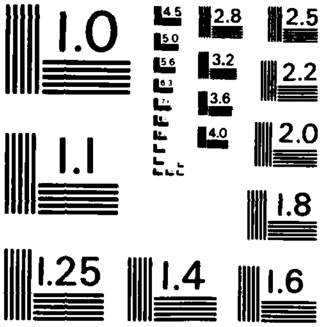
AD-A160 855 ACCESS TO INTER-ORGANIZATION COMPUTER NETWORKS(U) 1/3
MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER
SCIENCE D L ESTRIN AUG 85 MIT/LCS/TR-345
UNCLASSIFIED N00014-83-K-0125 F/G 5/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS
BEFORE COMPLETING FORM

1. REPORT NUMBER
MIT/LCS/TR-345

2. GOVT ACCESSION NO.
AD-A160

3. RECIPIENT'S CATALOG NUMBER
835

4. TITLE (and Subtitle)
Access to Inter-Organization Computer Networks

5. TYPE OF REPORT & PERIOD COVERED
Ph.D dissertation
August 1985

6. PERFORMING ORG. REPORT NUMBER
MIT/LCS/TR-345

7. AUTHOR(s)
Deborah Lynn Estrin

8. CONTRACT OR GRANT NUMBER(s)
DARPA/DOD
N00014-83-K-0125

9. PERFORMING ORGANIZATION NAME AND ADDRESS
MIT Laboratory for Computer Science
545 Technology Square
Cambridge, MA 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS
DARPA/DOD
1400 Wilson Blvd.
Arlington, VA 22209

12. REPORT DATE
August 1985

13. NUMBER OF PAGES
260

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)
ONR/Department of the Navy
Information Systems Program
Arlington, VA 22217

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release, distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
Unlimited

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Computer-communication networks, security and protection, network operations, electronic mail, public policy issues, organizational impacts, management of computing and information systems, system management.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

When two or more distinct organizations interconnect their internal computer networks they form an Inter-Organizational Network (ION). IONs support the exchange of cad/cam data between manufacturers and subcontractors, software distribution from vendors to users, customer input to suppliers' order-entry systems, and the shared use of expensive computational resources by research laboratories, as examples. This thesis analyzes the organization implications of using computer n'tworks for inter-organization communication, and the

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20.

technical implications of interconnecting networks across organization boundaries.

This thesis demonstrates the value of our bimodal approach to system design and analysis in which ~~we~~ ask, both how industry and organization contexts shape a new technology, as well as how a new technology affects the organization and industry contexts in which it is applied.

Keywords: [illegible] Information Systems [illegible]

Access to Inter-Organization Computer Networks

by

Deborah Lynn Estrin

Submitted to the
Department of Electrical Engineering and Computer Science
on August 21, 1985 in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Massachusetts Institute of Technology
Laboratory for Computer Science
Cambridge, Massachusetts 02139

This research was supported in part by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under Contract No. N00014-83-K-0125. Additional funding was received from discretionary funds provided to the MIT Laboratory for Computer Science by the IBM Corporation.

Access to Inter-Organization Computer Networks

Abstract

When two or more distinct organizations interconnect their internal computer networks they form an *Inter-Organization Network (ION).* IONs support the exchange of cad/cam data between manufacturers and subcontractors, software distribution from vendors to users, customer input to suppliers' order-entry systems, and the shared use of expensive computational resources by research laboratories, as examples. This thesis analyzes the organization implications of using computer networks for inter-organization communication, and the technical implications of interconnecting networks across organization boundaries.

We present a descriptive model of the effects of ION use. IONs change the economics of inter-organization communication. In particular, the speed and incremental cost characteristics support more intense communication, while the capabilities and automatic nature support a greater scope of information and resource sharing across organization boundaries. These enhanced communication patterns in turn allow participants to carry out more activities across their organization boundaries and with larger numbers of outsiders. At the same time, the ION-supported communication is more penetrating because outsiders access internal resources directly. In addition, when IONs are not universally accessible, communication is segmented between ION and non-ION organizations. These latter two characteristics introduce restrictions which detract from the expansive qualities of IONs. In particular, to compensate for increased penetration organizations may increase formalization of and controls on cross-boundary flows; while segmentation may lead organizations to narrow the range of favored interchange partners to those that are accessible via ION facilities.

We demonstrate the descriptive and predictive value of our general model in the domain of research and development laboratories. This domain provides evidence for our predictions of intensified communication of greater scope and penetration, as well as expanded numbers of cross-boundary activities and interchange partners. We attribute the absence of predicted restrictive behaviors to the absence of resource sharing.

Given our analysis of the organization context in which IONs are used, we demonstrate that such interconnections are not satisfied by traditional network design criteria of connectivity and transparency. To the contrary, a primary high-level requirement is access control, and participating organizations must be able to limit connectivity and make network boundaries visible. At the same time, these access control requirements are not satisfied by traditional computer security mechanisms. For example, this investigation of inter-organization networks makes clear that where traditional security mechanisms emphasize information flow, network environments are equally, if not more, concerned with command flow—i.e., invocation of services and applications. We develop a scheme based on non-discretionary controls that allows interconnecting organizations to combine gateway, network, and system-level mechanisms to enforce cross-boundary control over invocation and information flow, while minimizing interference with internal operations.

Access control requirements such as these impose new requirements on the underlying interconnection protocols. Just as internetwork access control requirements called for reevaluation of traditional computer security criteria and mechanisms, so cross-boundary connections call for reevaluation of traditional approaches to network interconnection. Consequently, we demonstrate the need for alternative interconnection protocols that support loose couplings across administrative boundaries and that accommodate the necessary control mechanisms. Message-based gateways that support non-real-time invocation of services (e.g., file and print servers, financial transactions, VLSI design tools, etc.) are a promising basis for such loose couplings.

The thesis demonstrates the value of our bimodal approach to system design and analysis in which we ask both *how industry and organization contexts shape a new technology*, as well as *how a new technology affects the organization and industry contexts in which it is applied.*

Keywords

computer-communication networks (C.2), security and protection (C.2.0), network operations (C.2.3), electronic mail (H.4.3), public policy issues (K.4.1), organizational impacts (K.4.3), management of computing and information systems (K.6), system management (K.6.4)

Acknowledgments

Jerry Saltzer, my thesis supervisor, introduced me to this intriguing topic, and then helped guide me towards the interesting problems, and away from the black holes. His commitment to support me in my inter-disciplinary studies made this research possible. His interest in my development as a researcher and teacher enriched my time at MIT tremendously. Finally, I am grateful for his prompt and ever-careful reading of earlier drafts.

Because of the interdisciplinary nature of this research, the contributions of my other committee members were particularly important. *Gordon Walker* provided invaluable inspiration and guidance in my study of inter-organization communication and relations. His gentle tolerance for my unorthodoxed training helped me to span the disciplinary boundaries. *Marvin Sirbu* shared his interests in the study of communications technology and policy and provided a model for how to integrate the two. *Jeff Meldman* helped me communicate many of my ideas for the first time, and then helped polish that communication in several successive drafts. I am also greatly indebted to the late *Ithiel de Sola Pool*, who welcomed me into The MIT Communications Policy Research Program.

In addition to my formal committee members, I wish to thank members of the Computer Systems Research group for their generous input, on paper and in the hallways: especially, Bob Baldwin, David Clark, David Feldmeier, David Reed, Karen Sollins, Juliet Sutherland, and Lixia Zhang. From the other side of campus, Diane Gherson provided an abundance of insight and scholarly criticism on the organizational aspects of my work. Tom Allen kindly reviewed my questionnaire design. Loretta Anania and Richard Solomon shared their interests in communications policy. From the other side of town, Jim Cash shared his experiences with inter-organization systems. From the other side of the country, Rob Kling and Isabel Valdes shared their perspectives on the implications of new technologies.

I contacted over twenty-five research laboratories during the course of this research and wish to thank them all for their participation; I refrain from mentioning them by name to protect their privacy. A very special thanks goes to Laura Breeden and Dick Edmiston of the CSNET Information Center at BBN for their generous support, information, and time.

Much thanks to those who build and support the wealth of computational resources with which this thesis was created. In particular, Arvind and Bob Iannucci generously allowed me to use their facilities for statistical analysis. Thanks also to the residents of the second floor for their hospitality while I used their Lisp Machines.

Many special people made life at the lab enjoyable:
The support staff—Rebecca Bisbee, Debra Fagin, and Muriel Webber.
The CPSR group—Steve Berlin, Walter Hamscher, Marie Macaisa, Ronni Rosenberg, Karen Sollins, and Joseph Weizenbaum.
The women—fortunately there are now too many to list individually!
Very special friends: Toby Bloom, for listening to my stories and reminding me that I was not alone; David Culler, for sharing hopes and fears as we ran along The Charles; Steve Heller, for keeping me laughing; and Karen Sollins, for her patient listening and advice.

I owe my strength and persistence to my parents, Thelma and Gerald Estrin. They were an endless source of encouragement and advice through these, and all the previous, years. My thanks also to Anna and Mayer Skrzypek for providing a warm home-away-from-home in Springfield.

Most of all I wish to thank my husband, Josef Skrzypek, for his support, encouragement, patience, and, of course, laughter. He inspired me to strive, and then, when the time was right, to finish so that we could return to our place in the sun!

Table of Contents

Table of Figures

Chapter One

Introduction to Inter-Organization Networks and the Thesis

Much has been written about the automation of factory and office functions. An important aspect of this automation is the ability to communicate and share resources between different physical machines, different administrative and production functions, and different geographic sites. This thesis focuses on computer-based communication and resource sharing that crosses *organization boundaries* as well. When two or more distinct organizations interconnect their internal computer networks to facilitate interchange, they form an *Inter-Organization Network (ION).* The interchange may be person-to-person communication; exchange of cad/cam data, software modules, or documents; input to an order-entry or accounting system; or use of shared computational resources.

The purpose of the research is twofold:

- To analyze the organization implications raised when computer-based communication media are used for inter-organization interchange—increased efficiency, capabilities, vulnerabilities, etc.

- To analyze the technical issues raised when computer networks cross administrative boundaries—security and network interconnection.

The first large scale, packet-switched, computer network, the Arpanet, interconnected computers in distinct organizations—namely, DARPA funded, research and development laboratories. However, the nature of the research relationships, DARPA's central role, and the explicit project goal of eliminating barriers to resource sharing, allowed the Arpanet in its early years to exhibit more of the characteristics of an *intra*-organization network. Even before the Arpanet, airlines used telex communications to coordinate reservation and flight information. Similarly, banks used telex and more recently data communications to support inter-bank transfers. The transportation industry has also made heavy use of telex and data communications to coordinate with one another and clients. Communication carriers have interconnected their networks for as long as international telegraph and telephone service has been available. More recently, firms in industries whose functions are less critically dependent upon inter-organization interchange and coordination—e.g., automotive, medical supplies, grocery—have established inter-organization communication links.

1.1 Contributions

The research contributions described in this thesis lie in three areas:

1. In the area of organization implications:

   - Analysis of the effects of this new communication medium on inter-organization interchange—efficiency, intensity, scope, penetration, and segmentation.

   - Analysis of the significance of these new communication characteristics for the management of cross-boundary activities—the number of interchange partners, the number of cross-boundary activities, restrictions on cross-boundary flows and interchange partners.

2. In the area of computer security and access control:

   - Characterization of security requirements that are not satisfiable using traditional non-discretionary control mechanisms—control of invocation, protection of invoked, accommodation of two-way communication.

   - Application of category sets and an intersect rule as simple mechanisms to address these requirements.

   - Design of access control mechanisms that allow strictly-internal applications to be unaffected by interconnection without requiring physical isolation from ION applications.

3. In the area of network interconnection:

   - Characterization of applications in which performance criteria alone and packet-level interconnection do not satisfy policy requirements.

   - Evaluation of high-level and visa-based interconnections in terms of implementation requirements.

- Proposal for message-based interconnection for loose couplings across organization boundaries.

The thesis as a whole represents a bimodal approach to the study of new technologies by asking *how industry and organization contexts shape a new technology*, as well as *how a new technology affects the organization and industry contexts in which it is applied*.

1.2 Summary of Thesis

The thesis is composed of three parts. The first portion of the thesis describes the context and implications of ION use. Chapter 2 sets the stage by characterizing inter-organization relations, and the traditional media used to support inter-organization communication. The subsequent chapter, Chapter 3, presents our model of how the use of IONs affects participants. The model begins with the technical characteristics of IONs and how these characteristics change the economics of inter-organization communication and interchange. Based on these technical and economic characteristics, we describe the behavioral changes that organizations are likely to make in their communication patterns and cross-boundary activities. In particular, we explain how IONs support intensified communication of greater scope, and expanded cross-boundary activities with larger numbers of outside organizations. We also describe the risks of ION use. ION communication is more penetrating because outsiders access internal resources directly. In addition, when IONs are not universally accessible, communication with non-ION organizations may be discouraged.

Given this understanding for the organizational context, the second part of the thesis analyzes and describes the design of network interconnections to fit the crossing of organization boundaries. A central concern of ION participants is protection of their organization boundaries in terms of access to information and resources. Chapter 4 introduces the access control issues using four real world examples which are referred to throughout the thesis. In addition to defining terms and concepts used in later chapters, this chapter reviews related work in network interconnection and computer and communication security.

Chapter 5 describes how the requirements for controlled external access can be met while minimizing interference with internal communication and operations. In particular, we describe how to use non-discretionary controls to isolate logical networks from one another while still allowing them to overlap. Although traditional non-discretionary control mechanisms are shown to be unsuitable, we specify an alternative set of non-discretionary mechanisms that each ION participant implements in the entry and exit points to its internal network, i.e., ION gateways. These mechanisms are designed to control *invocation* of computer-based resources instead of, or in addition to, information flow, and do not enforce strict confinement.

Chapter 6 analyzes the implications of the proposed approach for network interconnection. To implement these non-discretionary controls an ION gateway must have access to certain information about the logical characteristics of traffic; e.g., organization affiliation of source and destination. Most packet-level gateways do not have access to the information needed to make ION policy decisions. We describe a visa scheme for augmenting a packet level protocol in order to accommodate policy controls and compare it to the alternative of implementing higher-level gateways that actually terminate higher-level protocols. Our conclusion is that higher-level connections are preferable for many ION applications. We also conclude that these controlled, higher-level connections should be placed as close as possible to the administrative boundary being enforced. Finally, message-based gateways are suggested as being well suited to loose couplings desired for many inter-organization relationships.

The controls outlined assign categories or rights according to the organization affiliation of the source and destination. If the source and destination information can be falsified, then the controls are not effective. Chapter 7 add esses this issue and shows that Needham-Schroeder type authentication tools satisfy the authentication requirements outlined in the usage control model. The primary ideas presented are that internal authentication mechanisms need not necessarily be modified to comply with inter-organization requirements, and that multiple classes of authentication are desirable.

To conclude the technical discussion Chapter 8 describes the implementation of usage controls in an inter-organization network gateway. The most difficult aspect of implementing ION gateways is the association of communications with logical information. Aside from this difficulty the major implementation decision is whether to interconnect at the packet level and employ a visa scheme, or whether to interconnect at higher levels. The chapter evaluates these implementation issues for some of the examples described in earlier chapters. These examples provide insights into issues such as the distribution of control between an organization's internal gateways and its ION gateway, and the tension between supporting an open default for person-to-person electronic mail and a closed default for person-to-server invocation, when servers can be invoked via electronic mail messages.

Having discussed the design of IONs to fit organization boundaries we return to our discussion of how the technical characteristics of this medium affect the relationships and communication patterns among ION participants. Chapter 3 described our model in general terms. However, the implications of ION use are contingent on environmental factors and it is most useful to discuss IONs within the context of particular domains. Chapter 10 demonstrates the applicability of the general model to the study of ION use in distribution channels such as hospital supplies, airline reservations, etc.

The final chapters investigate more deeply the role of IONs among research and development laboratories. Chapter 11 characterizes R&D laboratories very generally and describes communication patterns and resource flows. We use the characterization of this domain and the general model described in Chapter 3 as the basis for our empirical study, described in Chapter 12. Based on almost 200 responses to an online questionnaire distributed to 25 commercial and university laboratories, we find strong evidence of increased intensity, scope, penetration, numbers of interchange partners, and cross-boundary activities. In addition, we find some evidence of segmentation. However, no increase in restrictions on communication or cross-boundary activities are evidenced; a finding that we attribute in part to the prevalence of reported person-to-person electronic mail, and absence of reported resource sharing. This chapter describes how the predictions were tested—the method and questionnaire used to collect data, the results of the data collection, and the implications of our findings.

Chapter Two

Inter-Organization Communication and Interchange

The first portion of the thesis describes the context and implications of ION use. This chapter sets the stage by characterizing inter-organization relations, and the traditional media used to support inter-organization communication. The next chapter presents our model of how the use of IONs affect inter-organization communication and interchange.

2.1 Organization Boundaries

This study of inter-organization networks focuses on communication and interchange between *distinct* organizations. In this study we consider two organizations distinct when they do not share a common authority with respect to primary budgetary or policy matters. Therefore, this discussion draws an organization's boundaries around its employees and resources. Furthermore, we assume that behaviors, not people, are organized and draw the boundary with respect to the roles that employees take on, not the individual people involved. [48] Our definition is loose because organizations are not neatly bundled, and consequently defining the term "organization" is largely a matter of analytic convenience. Examples of inter-organization relationships that fit this definition are customers/suppliers, manufacturers/subcontractors, joint venture participants, joint research collaborators, and companies that coordinate in order to serve common clients such as airlines, banks, insurors, and railways.

Organizations engage in many activities that blur this definition of organization boundaries. Vendors assign employees to work on the premises of major customers. Consortia and trade associations act as vehicles for sharing information and resources, and sometimes serve as super-organizations by creating common goals and policies under which members operate. Overlapping boards of directors and employee migration are less direct ways in which boundaries between otherwise distinct organizations are often blurred. At the same time, within single organizations (as defined above) there are often divisions and groups that relate to one another as if they were distinct organizations. For example, geographically-distant sites of a single company, distinct functional units such as research and manufacturing, or separate product-line divisions, often have independent budgets and significant local autonomy, even though they report to a common authority.

2.2 Inter-Organization Relations

Given this loose definition of organizations and organization boundaries, inter-organization relations can be characterized according to many parameters such as their function, power balances, etc. This thesis focuses on the economics of communication between organizations and the nature of inter-organization activities supported. Moreover, it focuses on *formal*, task-oriented communication and interchange among organizations, as opposed to informal, interpersonal communication among employees of the distinct organizations. This distinction between formal and informal is problematic in that both are important and interdependent; however it is a tractable place to begin our explorations and analysis. Examples of the formal communication and interchange addressed include:

- Exchange of purchase orders and invoices between customers and suppliers, as well as exchange of auxiliary product information and services.

- Exchange and sharing of design information and resources, and administrative coordination, between manufacturers and subcontractors or vendors.

- Exchange and sharing of design and development information and resources, and administrative coordination, between participants in a joint venture.

- Exchange and sharing of research information and resources, coordination of paper authorship, and administrative coordination, between researchers in a common discipline.

The transaction cost approach to the study of industrial activities, as developed by O.E. Williamson and others, [78, 79, 80] is particularly relevant given this focus. Building on work by Coase [14], and the assumption that economizing on transactions is the primary criterion for commercial organization. Williamson asserts that firms try to minimize production and transaction costs combined. He explains organizations' behavior with respect to other organizations in terms of transaction-cost efficiency and attributes the institutional arrangements effected between organizations to the type of transaction supported. He refers to these institutional arrangements as *governance structures.* One of Williamson's theses is that efficient governance structures vary systematically with the organizations' investments in durable, transaction-specific assets[1]

Williamson defines three critical dimensions of transactions: uncertainty of future exchange conditions, frequency of exchange (i.e., one-time, occasional, and recurrent), and specificity of investment in the exchange (i.e., non-specific, mixed, and idiosyncratic). He maps these conditions into three types of governance structures, market, trilateral, and transaction specific. Each of the three types corresponds to one of the three traditional types of contracting; market, trilateral, and transaction-specific [38]. Figure 2-1 illustrates the mapping between governance structures and transactions and gives an example of each. For each of the cells in the cross-classification table Williamson identifies the governance structure that is most efficient for that type of transaction. Both recurrent and occasional, non-specific transactions are associated with the classical governance structure, market. Occasional transactions of both the mixed and idiosyncratic type are associated with a trilateral structure in which a third party is engaged. Finally, recurring transactions of both the mixed and idiosyncratic type justify transaction-specific structures. Two types of transaction-specific structures are discussed. Bilateral structures are associated with mixed transactions, whereas unified (internal to an organization) structures are associated with highly idiosyncratic transactions. For example, for recurring, idiosyncratic transactions, market competition may be feasible at the contract-award stage. However, the subsequent relationship between buyer and seller transforms into a bilateral monopoly in which adaptation requires negotiation via an alternative governance structure.

[1] Transaction specific implies that the investment is not transferable to transactions with other organizations. [80]

		Investment		
		Non-Specific	Mixed	Idiosyncratic
Frequency	Occasional	Classical Market (purchase an office copier)	Trilateral (purchase a customized milling machine)	Trilateral (construct a plant)
	Recurrent	Classical Market (purchase paper for the copier)	Bilateral (purchase specially alloyed steel plate)	Unified (rail-transfer of coal from a mine)

Figure 2-1: Williamson's Determinants of Governance Structure [79].

Two additional areas of organization studies are of potential utility in characterizing inter-organization relationships: social networks and inter-organization relations. Social network research is used for identifying and then analyzing the communication patterns among a large group of communicating entities [72]. It is most useful for analyzing communications among organizations that do not have an explicit collective structure. In an ION the communications network is explicit at the formal, inter-organization level. In other words, the ION participants are mutually-aware and have distinct patterns of communications. Therefore, this stage of our research does not employ social network analysis.[2]

[2] Social network analysis should be well suited to future investigations of the less formal aspects of inter-organization communications; for example, a comparative study of ION and traditionally-mediated personal networks.

A large body of research in the area of inter-organization relations is also relevant to the research described in this thesis. For example, Marrett [39] suggested four dimensions of interorganization relations: formalization, intensity, reciprocity, and standardization. Formalization is measured by the extent to which exchange is given official recognition and the extent to which an intermediary coordinates the relationship. Intensity is measured by the size of the resource investment and the frequency of interaction. Reciprocity is measured by the extent to which elements are mutually exchanged and the extent to which the terms of interaction are mutually reached. And standardization is measured by the fixedness of units of exchange and the fixedness of procedures for exchange.

Over the past ten years, both Williamson's and Marrett's models have proven useful in empirical studies; for example see [74, 77]. The model of inter-organization networks described in the following chapter is structured along the lines of the transaction cost framework, but borrows from Marrett's characterization of inter-organization relations as well.

2.3 Traditional Communication Media

Any new technology should be analyzed in the context of the technologies that it augments and/or replaces. So, Inter-Organization Networks should be studied in the context of the traditional media that they augment, and in some cases, replace. Similarly, the adoption of this medium can be compared to the history of other new media which are now considered traditional—in particular telegraph and telephone.

The technology underlying the telegraph was first introduced in the late eighteenth century but was not used widely until the middle of the nineteenth century. For the first time information was transferred over distances far beyond that which could be achieved by human carriers. However, although Morse had developed a code of dots and dashes to represent the alphabet efficiently, the cost of transmitting one message at a time over many miles of telegraph line was still very high. Consequently, messages were written in cryptic language. J. Yates describes how the ability to communicate without delay over long distances allowed organizations to begin coordinating distribution and sales functions across geographic regions.[3] However, she also suggests that the motivation to use highly encoded—non-standard—language over this expensive communication medium meant that the cost of internal communications was reduced more than the cost of external communication—highly specialized telegraphic codes could be established *within* a single organization far more easily than across many distinct organizations. Based on this economic argument, Yates proposes that the telegraph encouraged manufacturers to forward integrate into distribution and sales instead of engaging independent distributors. Similarly, she suggests that the economics of the telegraph made it most appropriate for short routine messages and therefore favored the routine informal communication found within organizations over the more formal and protective, legalese used between distinct organizations. When the telephone entered the scene in the later nineteenth century it dampened telegraph developments. One of the main reasons for telegraph's decline was the very clumsy typewriter technology of that day for printing out telegraph messages. [52]

The telegraph and later TWX and Telex were always envisioned primarily as business communication media. Similarly, the telephone was perceived initially as a business tool more than an personal communication medium. For example, in 1879, 294 out of 300 telephones listed in the Pittsburgh directory were located in businesses; and all of the remaining 6 were used for conducting business from the home. [3] Even with the limitations of early technology, in particular the distance limitations (approximately 20 mile radius), the telephone allowed significant changes in the coordination of business activities because it was far less labor intensive than the telegraph and allowed true interaction between communicators. Examples of changes made possible by the telephone include: construction of skyscrapers—the telephone was used to coordinate construction at upper levels from the ground; messages and room service within hotels—previously messengers had to be available continually on every floor; coordination of railroad operations and coping with emergencies; and coordination between administration and manufacturing within the factory plant. [51]

[3] From a presentation at the M.I.T. Communications Forum, Fall 1984, entitled "Structural Effects of Communication Technologies on Firms: Lessons from the Past."

Eventually, the economics and utility of the telephone and the recognized *network effect*—in which the value of the service to all participants was greatly increased for each additional subscriber on the network—lead to increased residential development. Between the mid-1920's and mid-1970's the number of telephone calls per person per year grew from 200 to close to 1000, whereas the number of first class and air mail letters per person per year grew only from approximately 150 to 250. [51] Like telegraph, telephone charges were proportional to the amount of time spent on a connection. However, unlike telegraph, the source and destination of the communication were both human and both available at the time of the communication. Although this introduced the inconvenience of simultaneous presence, for many circumstances, the ability to respond immediately and even interrupt, could reduce significantly the amount of time used per completed interaction.

Computer-based communication is the medium of interest in this research. In some ways this medium represents a move back to the asynchronous mode of telegraph communications in which both parties were not, or did not have to be, present simultaneously. However, now, over 100 years later, other aspects of the technology combine to make this mode of communication more exploitable; in particular the quality of terminal equipment, the user interface, and the ability to automate the labor-intensive aspects. In addition to person-to-person communication via electronic mail, IONs can support online transaction processing and sharing of computer-based resources. We ask how this new medium affects the economics of inter-organization communication and interchange.

In recent years several studies have been conducted of electronic mail use within organizations.[4] Many aspects of this work are relevant to this study, in particular with respect to person-to-person electronic mail communications between employees of distinct organizations. For example, Rice and Case [55] describe the applications for which managerial and computer personnel perceive electronic mail to be appropriate; these perceptions have clear implications for inter-organization communication as well. At the top

[4] A summary of this research can be found in [56].

of the list for both groups (over 84%) were exchanging information, asking questions, exchanging opinions, and staying in touch. At the bottom of the list for both groups were exchanging confidential information, resolving disagreements, and bargaining (below 40%). In general the computer personnel found electronic mail more appropriate for a wider range of the tasks. This fact supports the notion that this new medium will be used as one of a multiplicity of media, and will not, or at least optimally should not, replace use of other media altogether; each medium appears to have its best set of uses. Kiesler and associates have found evidence for behavioral differences in the use of the different media; for example, that users of electronic mail express more extreme opinions about things and often relate more intimate information and questions than is typical of face-to-face, telephone, or written communications.

Several studies have characterized communication patterns over different media. For example, J.B. Goddard's comparative field data on telephone and face-to-face contact showed the following differences between the two channels: 87% of telephone calls were less than 10 minutes in duration, as compared with 19% of the face-to-face contacts; 83% of the telephone contacts were not arranged, as compared with 17% of the face-to-face meetings; 84% of the telephone calls covered only one specific subject, as compared with only 57% of the face-to-face meetings; and giving or receiving information or giving orders was the main purpose of contact for 50% of the telephone calls, compared with 23% of the face-to-face contact. [26] The characteristics of computer-mediated communications suggest that the breakdowns of usage will resemble those of telephone more closely than those of face-to-face meetings with respect to these parameters.

Picot et. al. measured the weight given to different evaluation criteria by users of various communication media, with the following findings. On a scale from very important (1.0), to less important (2.0), to unimportant (3.0), users ranked criteria in the following descending order: unambiguous understanding of context (1.1); speediness (1.2); certainty of exact wording, certainty of information reaching wanted receiver (1.3); availability of channel, capability of quick response, capability of quick feed-back, transmission of difficult content, short composition time (1.4); easy processing by receiver, short transmission time, resolving disagreement, capability of documentation (1.5); identification of sender, transmission costs (1.6); comfort, circular letters, transmission of small information volume (1.7); transmission of large information volume, protection from faking (1.8); confidentiality (1.9). [49] As the next chapter describes, computer-based communications media offer significant improvements in several of these criteria—speediness, availability of channel, capability of quick response and feed-back, short composition time, short transmission time, capability of documentation, transmission costs, comfort, circular letters, and transmission of small and large information volume. Note that these comparisons apply to person-to-person communications, not to resource-sharing or even formal business transactions.

Our discussion of IONs assumes that this new medium will be used in conjunction with traditional media. In addition, the discussion addresses a range of communication types, including person-to-person electronic mail, online transactions, and online access to computer-based resources. Unlike traditional communication media which support person-to-person communication only, IONs also support remote resource sharing. This latter function can be compared more directly to resource sharing arrangements such as joint ownership, equipment loans, off-site employees, etc. than it can be compared to traditional media per se.

The following chapter describes how IONs differ from traditional media for both communication and interchange. The model assumes interconnection across distinct organizations. It does not directly address the equally interesting question of communication and interchange across geographic boundaries but within a single organization. We begin with the technical characteristics of IONs and how these characteristics change the economics of inter-organization communication and interchange. Based on these technical and economic characteristics, we describe the behavioral changes that organizations are likely to make in their communication patterns and cross-boundary activities.

Chapter Three

Effects of ION use on Communication and Cross-Boundary Activities: A General Model

The thesis of this research is that IONs change the economics of inter-organization communication and interchange. We have developed a model that describes:

- The technical characteristics that underlie these changes in economics—speed, capabilities, cost, universality.

- Resulting opportunities for enhanced inter-organization communication—more intense communications of greater scope.

- Resulting opportunities for enhanced cross-boundary activities—more cross-boundary activities and with a larger number of outside organizations.

- Accompanying risks—more penetrating and segmented interchange, restricted sets of interchange partners, and more explicit administrative and technical controls on cross-boundary flows.

This model explains and supports design, deployment, management, and regulation of IONs. Examples from several domains are used to illustrate the model, including buyer-supplier relationships and peer relationships among research and development (R&D) laboratories. An empirical study of R&D laboratories is described in chapters 11 and 12.

Section 3.1 summarizes the model and addresses the issue of causality. Sections 3.2 through 3.4 describe our model of the technical and behavioral changes associated with ION use. Section 3.5 summarizes the model's predictive statements.

3.1 Overview of the Model

Like traditional inter-organization media such as telephone, paper, and face-to-face meetings, an ION is a medium for communication and interchange among organizations. However, because of its technical characteristics, an ION changes the economics of inter-organization communication and interchange. In particular, this new medium allows organizations to adopt new patterns of communication, such as greater frequency and scope, which reduce costs and enhance products or services. These new communication characteristics in turn allow organizations to expand their cross-boundary activities. At the same time, problematic effects of ION deployment can hamper communications and cross-boundary activities in ways not necessarily intended or foreseen by participants.

This model describes the opportunities for enhanced communication and interchange, increased cross-boundary activities, and restrictive side-effects, and the industry and organization factors that motivate ION participants to act upon the various opportunities presented. Our presentation of this three-stage model is summarized in figure 3-1.[5]

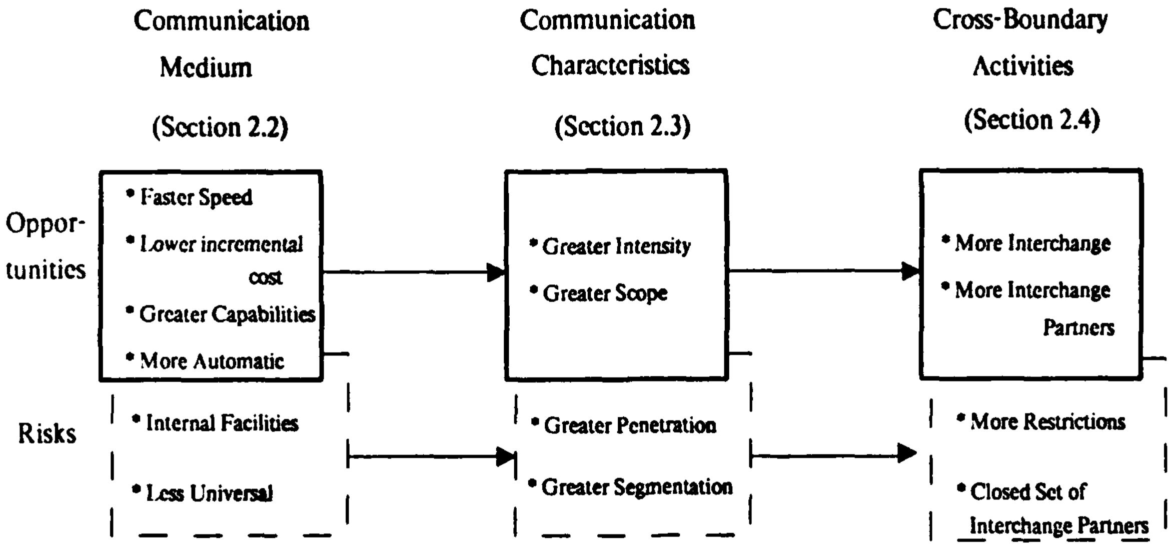


Figure 3-1: Effects of IONs on communication and interchange: overview and order of presentation.

[5] The unit of analysis of this model is a focal organization and one or more interchange partners. Accordingly, the characteristics of the communication medium, of the communications themselves, and of the cross-boundary activities are treated from the perspective of each ION participant.

The elements listed in the boxes are dimensions of *change* supported by ION use. The changes listed in the solid boxes are the *opportunities* that organizations can exploit using ION technology. Typically, these changes serve some organizational objectives such as reducing costs or increasing effectiveness. The *risks* (dashed boxes) are changes that may accompany the advantageous changes but which organizations may not have intended or even anticipated. However, in some cases, one or more of the ION participants may try to exploit these technical characteristics to its advantage. These disadvantageous changes are potentially more short-term in nature than the advantageous changes listed. Nevertheless, even short-term changes can have significant organizational and inter-organizational impacts.

Overall, our predictions of both advantageous and disadvantageous change are strongest with respect to IONs that support person-to-machine and machine-to-machine, in addition to person-to-person, communication, as compared with those that support person-to-person communication only. However, the model applies to computer-mediated person-to-person communication as well.

3.1.1 Causality

The model describes changes at three levels—communication medium, communication patterns, and cross-boundary policies. The lowest level, communication medium, describes the differences between ION technology and traditional communication technology. Although changes at this level support changes as higher levels, this does not imply causality. The desire for higher level changes in the communications themselves are the motivation for investing in and implementing a medium with different properties. The second level, communication patterns, describes behavioral changes of ION participants. Whereas the characteristics of the communication medium indicate which communication types the medium *can* support, this level describes the actual communication patterns—behavioral changes—that the ION *does* support. Finally, the highest level, cross-boundary policies, refers to the way in which activities carried out between organizations are managed. Once again, changes at this level may be enabled by lower-level changes, but the desired changes in cross-boundary policies are what influence demand for changes in communications and the communication medium.

Figure 3-2 illustrates this duality. Studies of change associated with other technologies, such as the telephone, have also called for "a logic more complex than simple causality—a logic that allows for purposive behavior as an element in the analysis." [51] The causality is neither that of Karl Marx in which the outcome is completely determined by the configurations of the technology and economics, nor is it simply that of Max Weber in which the outcome is determined by intervention of human will and values. The causality represented throughout this study, is bidirectional—changes in technical parameters alter the economics of communication and interchange (behavioral dimensions) and thereby support new forms of efficient behaviors, while desires for behavioral changes motivate the adoption and design of new media. The organization of this thesis reflects this perspective. We begin by discussing behavioral changes that IONs can support. We then analyze the technical designs that these behavioral changes motivate. And finally, we return to investigate the behavioral changes experienced in a particular setting.[6]

3.2 Communication Medium

When an organization adopts ION media, the new technology and procedures typically coexist along side the old. In fact, the new technology *can* be used in precisely the same manner as the old. However, IONs differ from traditional communication media (telephone, paper, face-to-face meetings) and offer significantly enhanced speed, capabilities, and cost-performance. The magnitude of the changes depends upon the design and investment by one or more of the ION participants.

Four technical characteristics of IONs differ most significantly from traditional media: speed, incremental cost, capabilities, and automatic response. These are the features of the new technology that motivate adoption and are the primary design parameters. In general,

[6] This view is similar to what J. Slack refers to as *structural causality*, [65] pg. 81.

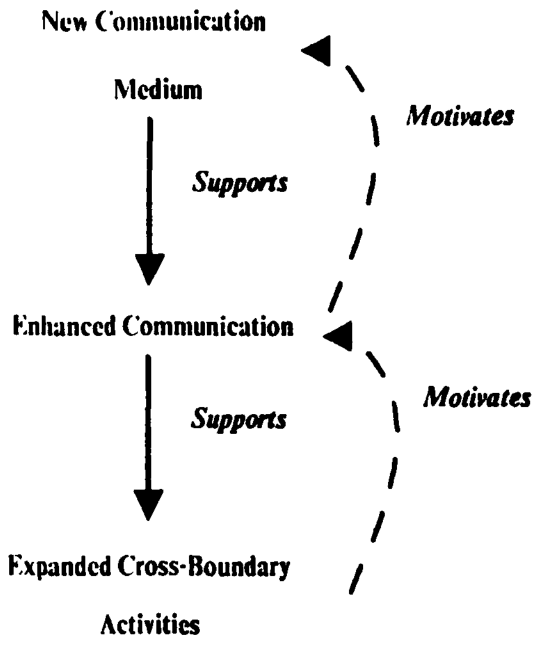


Figure 3-2:Causality as represented in the model.

IONs exhibit the following characteristics, as compared with traditional media.[7]

3.2.1 Speed

All ION application types—electronic mail, file transfer, data base query, and remote login—exhibit faster speed overall than traditional media; where speed includes the time to prepare, transmit, and process a message. It is easier to compare electronic mail to traditional media than it is to compare other ION applications because like traditional media, electronic mail supports person-to-person communication. Other ION applications that support person-to-machine or machine-to-machine communication are not directly comparable to traditional media; see section 3.2.3. Consequently much of the discussion below addresses person-to-person communication only.

[7] I qualify this statement because the characterisücs of an ION depend upon how it is designed. Therefore, an ION may conceivably support lesser characteristics instead of improved characteristics; for example, if users do not respond to electronic mail as readily as to a telephone call, turn around will be lengthened, not shortened.

Communicating a message between two or more persons involves preparing the communication for transmission, and then transmitting the desired information. The preparation time for a face-to-face meeting is the time needed to arrange the meeting. The transmission time is the duration of the meeting. In the case of telephone, the preparation time is the time it takes the originator to establish contact with the recipient(s) (including the highly-variable delays due to telephone tag). The transmission time is simply the duration of the telephone call once the parties have made contact. In the case of written communication (memos, letters) the preparation time is the time to compose and create the document and the transmission time is the time to transfer the physical paper from the sender to recipient. The originator of the message may prepare the document directly or may employ secretarial assistance. Written communications between organizations typically travels via US postal mail, express mail services, or facsimile type services. Finally, the preparation time for telex is that needed for composition and creation of the telex message (both by the originator and administrative support personnel), and the transmission time is that needed to transfer the message from the originator to the recipient via intermediaries such as telex operators (i.e., transmission time is *not* just the time to transfer the telex signals between telex machines).

Electronic mail messages are comparable to telex and paper mail in that preparation time is the time to compose and create the message, and the transmission time is the time to send a message from the originator to the recipient. The preparation time for electronic mail often is less than for telex and paper mail because originators typically have direct access to the electronic mail preparation system and need not employee administrative assistance if they do not want to. In addition, the computer-based editing tools often available facilitate the message creation process.[8] Although it is not a technical characteristic per se, the style of communication used via electronic mail is less formal than written memos, and less cryptic than telex (the latter because the incremental cost per word for electronic mail is not as high as for telex). Informality can reduce message composition and creation time because less care is needed to both content and form. Paper, telex, and electronic mail communications

[8] The difference between telex and electronic mail is mostly an artifact of the end-users having direct access to and better message-preparation tools on electronic mail terminals, than telex terminals.

are superior to telephone and face to face in terms of preparation time because of the highly-variable amount of time that it can take to set up telephone or face-to-face contact. On the other hand, psychologists have found that information is transferred among people at a faster rate via voice communication than via written communication only. However, this phenomenon is less relevant to exchange of formal or highly codified information. For most systems, the transmission time for electronic mail is somewhat faster (varies between instantaneous and one day) than for telex, and significantly faster than for paper mail. In summary, electronic mail messages (e.g., purchase orders, administrative, providing information, etc.) reach their destination faster, and therefore the *minimum turn-around time* between sending a message and receiving a response is shortened. However, to the extent human participation is required in the reply, turn around time is not deterministic because there is no guarantee that the electronic mail recipient will read the electronic mail message any more promptly than she or he would a telex or paper mail message.[9]

ION applications other than electronic mail support interactive access to computer-based resources. The speed of this access varies with the equipment used but in all cases is within the bounds of being considered interactive. This sharing of resources is difficult to compare to traditional media. The closest comparison is to physical exchange of data, programs, or equipment, or to human travel to another organization's facility for local use of resources.

3.2.2 Cost

The cost of a communication medium consists of fixed and incremental costs. The primary role of fixed cost is in the organization's decision to employ the medium; i.e, fixed cost determines what amount of communications is needed to justify the investment. Incremental cost influences the overall economic benefit of the medium, and is the primary factor in an organization's choice of communication patterns once the medium is employed.

The *incremental* cost of computer-based communication and resource-sharing typically is

[9] Some reports on electronic mail usage do claim that users read their mail and respond more promptly than they, or others, do to paper mail or telephone messages. However, much, if not all, of this could be attributed to the newness of the medium.

lower than comparable functions via traditional media. Once again it is worthwhile to discuss preparation and transmission costs separately. Preparation cost includes the human effort and time required to construct and assure delivery of a message; e.g., making telephone contact, writing a memo, arranging a meeting, or preparing an electronic invoice. As with the previous discussion of speed, only electronic mail is directly comparable to traditional media. Some components of preparation costs are proportional to the speed of preparation described above—in particular the labor cost of the originator is proportional to his or her message preparation time. Additional preparation costs are secretarial support employed and materials used in preparation—i.e., paper products, typewriter use, or electronic mail system use. Unfortunately, very little data is available on the incremental end-user and system costs of preparing electronic mail; in fact, little data is available on such costs for traditional media either. Moreover, the small amount of data that is available is of limited general use because so much of the measured costs are artifacts of the particular systems employed. [17, 47] Somewhat more information is available on transmission costs. In the remainder of this section we summarize some available data on preparation and transmission costs.

In 1982 Crawford summarized a study of the costs associated with electronic mail use at Digital Equipment Corporation. He compared the costs of two different internal electronic mail systems with telephone and inter-office memo preparation and transmission costs. Preparation costs for electronic mail were significantly lower than telephone or paper memo if the originator of the electronic message entered the text directly, without administrative assistance. Moreover, when additional copies of a message were required, the preparation costs for telephone increased linearly whereas electronic mail costs increased negligibly.[10] Including both preparation and transmission costs, electronic mail compared very favorably with inter-office memos but was not as low as telephone in some cases due to the fact that terminal and communication system costs were included for electronic mail. In addition, the significance of telephone transmission costs would increase significantly for inter-site and inter-organization communication. His findings are summarized in figure 3-3.

[10] This assumes no conference call facilities.

		Interoffice Memo	Phone Call	Elec Mail Orig:Mgr	Elec Mail Orig:Intrm
Preparation Costs	Originator (Labor)	$2.88	$1.80	$2.16	$2.88
	Secy/Operator (Labor)	$3.42			$1.08
	Non-productive	$0.25	$1.23		
Transmission Costs	Materials/Mail	$0.61			
	Communication		$0.82	$0.27	$0.27
	System			$0.83	$0.83
	Peripheral:				
	Equipment	$0.44	$0.46	$0.14	$0.14
	Communication			$0.83	$0.83
	Total Unit Cost	$7.60	$4.31	$4.70	$6.70

Figure 3-3: Comparison of message preparation and transmission costs for telephone, paper memo, and electronic mail; from [17].

Panko also studied the costs of message preparation and transmission. [47] He estimates approximately $9 (1977 dollars) for the cost of preparing a standard business letter; including author preparation and review, and secretarial time. His study of two 1977 electronic mail systems found the total cost of communication—including preparation time, terminal, computer system support, and transmission—to be $18/10 messages for the experimental Planet system and $70/10 messages on the experimental Hermes system. The costs were projected to be only $5 and $18, respectively, for a 1977 state-of-the-art system, $3 and $8 for a 1981 system, and $1 and $2 for a 1985 system. Unfortunately, it is difficult to judge the relevance of these numbers to state-of-the art systems. In addition, preparation time for open-form business correspondence (the subject of Panko's and Crawford's studies) might be significantly higher than for fixed-form correspondence which characterizes a large part of inter-firm communication.

A 1980 study by ADL for the grocery industry estimated the following costs for a system to support one kind of fixed-form correspondence, electronic purchase order and invoice exchange. [1] In one configuration, A communicates with B via B's service bureau. The costs to A for this arrangement include: per message and monthly service bureau fee, monthly phone line charges, monthly 2400 baud modem lease, and local call charges for transmitting batches of messages. This works out to a $175 fixed monthly cost plus $0.45 per message incremental cost to A. The ADL study compared these costs to traditional incremental costs: $0.35 for stuffing envelope, $0.35 for transcribing and keying in paper invoice or purchase order, and $0.15 for postage, totaling $0.85 per message incremental cost.[11]

Somewhat more data is available on the incremental transmission costs of electronic mail and traditional media, than for preparation costs. In general, the incremental cost of transmitting an electronic mail message is less than some media such as telex and, depending on message characteristics, telephone. However, the cost may be greater than for paper-based media. On the other hand, if we compare the incremental cost of transmitting a message at a given speed, or for processing messages automatically instead of manually, computer-based communications is lower cost for most message types. [17, 43, 47] Figure 3.2.2 is a summary of rates charged by public communication services—both traditional voice and paper-based media and electronic mail. However, the costs of *private* communication facilities used by most large organizations and even groups of organizations (e.g., AIRINC jointly owned by Airlines, Insurance Value Added Network (IVAN) jointly owned by insurance companies, etc.) are not represented. In his 1981 paper, Panko cites the

[11] In an alternative configuration for heavier users, A connects its computer directly to B's via Telenet. The costs to B for such a configuration include: $1300 per month for a host connection to Telenet or $85.00 per month for a terminal connection, $0.004 per 1,000 character message, and $0.05 for telephone charges to reach Telenet (1980 dollars).

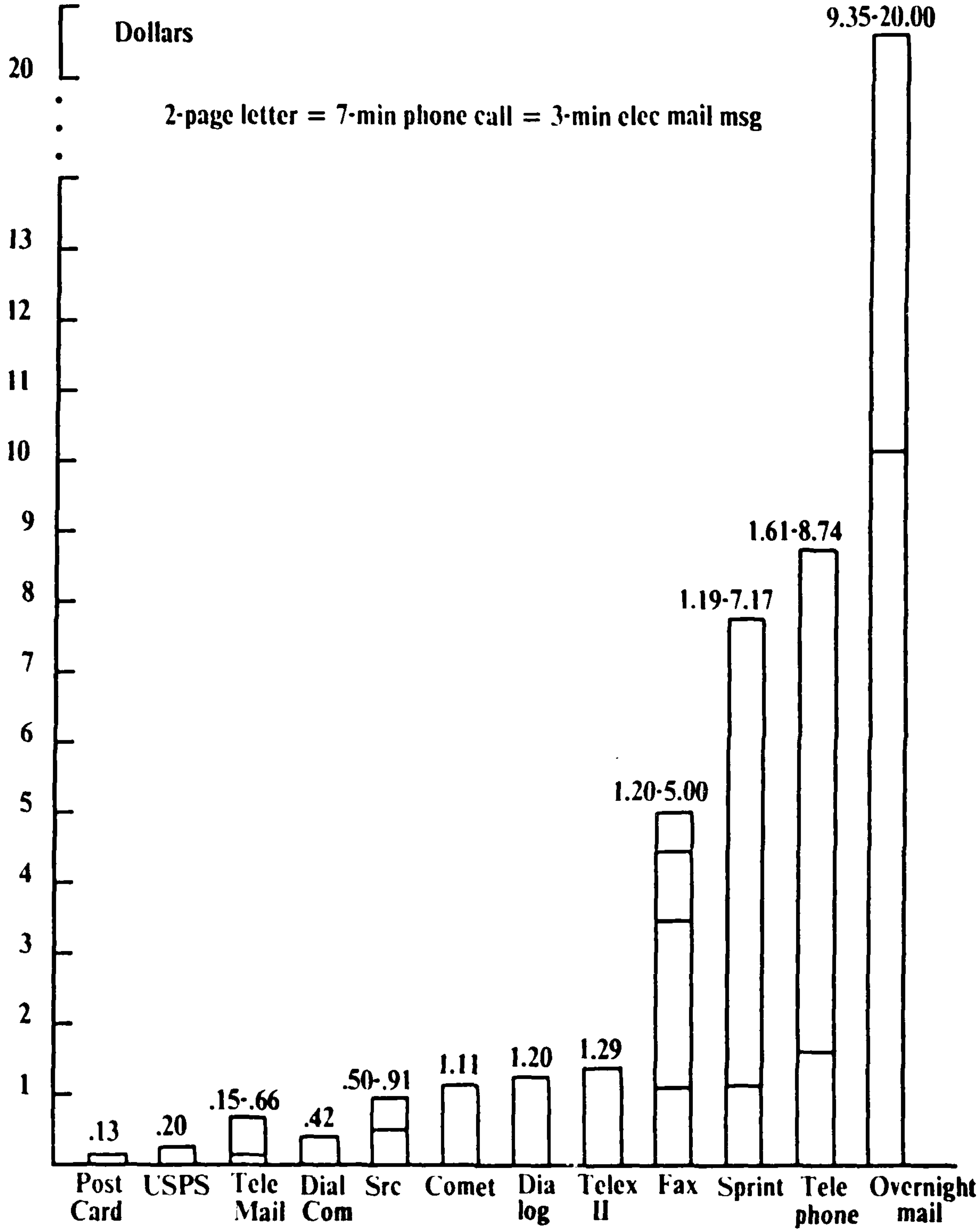


Figure 3-4: Transmission costs for different media. Data taken from [43].

hardware and telecommunications cost of Hewlett-Packard's internal electronic mail system at $1.06 million. Averaged across the 23 million messages sent in 1977, the average cost per message was $0.05. In general, it is difficult to obtain data on private network costs. Informal reports from other organizations with large, privately-operated, electronic mail networks confirm that their average transmission costs are significantly lower than those offered by commercial services; $0.10 (domestic) and $0.20 (international) per message are commonly-used estimates for large, high-volume, private networks.[12]

The incremental costs of ION applications other than electronic mail include transmission and processing. Transmission costs are the charges for telephone or packet-switched network usage. Telephone rates vary with distance and speed while packet-network rates include a fixed and volume-sensitive component. As with electronic mail, the incremental cost via privately operated facilities are rarely measured or publicly available but are considerably lower for high-volume networks. Processing costs include the overhead of the endpoint machines which similarly are rarely measured.

In general, the *fixed* equipment costs of electronic mail are higher than for traditional inter-organization communication media because the investment in telephone systems, mail rooms, and telex terminals were made long ago and were divided across the entire organization. The fixed cost of ION facilities varies widely according to capabilities, the existing equipment and expertise of the participant, the way in which the network is paid for, etc. The minimum fixed cost includes the hardware and software system used to connect the internal service to the ION, the communication equipment used to interface to

[12] Ullrich, Personal communication.

the ION, and any fixed communication or access charges.[13]

3.2.3 Capabilities

IONs support a wider range of capabilities for inter-organization communication and resource sharing than do traditional media. In particular, message-based IONs support remote, direct, and interactive communication and access to computational resources, databases, and information services.

There are four basic types of ION applications:

- Person-to-person electronic mail.
- File transfer.
- Database transactions.
- Remote login.

As described above, electronic mail does not represent qualitatively new capabilities as compared with traditional media; although, it is a faster, lower incremental cost, and often more convenient means of sending messages between people. The other three types of application support capabilities that are less comparable to traditional inter-organization communication media. Unlike speed and costs, capabilities do not lend themselves to quantitative metrics. In lieu of an appropriate metric, examples of each of the general application types will illustrate how these capabilities differ from traditional media.

One example of a file-transfer based application is software distribution. Computer system

[13] The rate structure for CSNET illustrates the tradeoffs between fixed and incremental cost (see Chapter 11 for a description of CSNET—a network connecting computer science research and development laboratories). In addition to an annual membership fee, charges depend upon whether a laboratory uses telephone facilities or X25-based packet network facilities. The fixed equipment cost for *Phonenet* access is approximately $1500. Using this equipment over local telephone, the estimated yearly incremental expense for electronic mail is $125 for light users (10 messages/day, 24 lines each), $250 for moderate users (38 messages/day, 36 lines each), and $625 for heavy users (75 messages/day, 50 lines each). In contrast, the fixed equipment cost for *X25net* access is approximately $15,000 while the estimated yearly incremental expense is only $75 for moderate users and $250 for heavy users. Furthermore, X25net access supports file transfer and remote login in addition to electronic mail.

vendors regularly release new versions of their operating system and various applications. Customers can be provided timely updates via an online file transfer system. Such a system can operate in a variety of ways. For example, when a new release is available, a computer on the vendor's premises could automatically dial and transfer the updates to designated machines at each of its customer sites. Alternatively, the customer's machine could dial up the vendor and request a file transfer of new software releases. This use of file transfer is most directly comparable to shipping a magnetic tape with the new release. File transfer applications might also augment or replace more traditional means of exchanging printed documents.

The airline reservation systems are a good example of a database transaction application. Travel agents communicate with the airlines' computer-based databases directly in order to obtain flight information and to make reservations. Similarly, the three major airline reservation systems communicate with one another to coordinate flights, and allow viewing of other airlines reservation data. The comparable capabilities using traditional media are telephone or paper communications with travel agents and the airlines directly. However, traditional media do not allow a travel agent, or a client, to consider as much information, from as many sources, in travel decisions, because of the variability of the information over time (e.g., seat availability) and the time needed to acquire it via traditional media. There are many other examples of customer-supplier, inventory-related applications for IONs (see Chapter 10).

Remote login provides the user with access to the full range of computational resources of the machine to which she or he is logged into; within the confines of access control mechanisms. Because of the generality of remote login, it may be used to support database access, electronic mail, or even file transfer. Furthermore, remote login can be used to access other types of applications such as a VLSI design simulation system, for example. Such general-purpose, remote access to computer-based resources is the least comparable to traditional media. The closest equivalent is sending a person to the site of the remote organization, or borrowing physical equipment.

Although person-to-person electronic mail does not introduce new capabilities, message systems can be used to *invoke* computer-based applications, peripherals, and various kinds of servers. One example is a long-standing application of the Arpanet. Researchers use electronic mail to send VLSI chip designs and commands to MOSIS, a facility that prepares chip layouts. Less sophisticated applications can be found in more general-purpose computing environments; for example, message-invocable name servers, file servers, program-tool libraries, and of course, mail forwarders.

In summary, IONs allow users in one organization to access information, manipulate data, and invoke computational resources, in a remote organization; capabilities which are not comparable to those of traditional person-to-person communication media.

3.2.4 Automatic Response

Access to a remote computer resource implies that the remote computer responds to requests or commands from outsiders automatically, without the participation of any human employee of the organization that owns the computer. Much of the efficiency of ION-based interchange arises because a human in each organization needs not be available at the same time to facilitate transfer of information or resources. Although this quality is clearly less applicable to person-to-person communications, even electronic mail that is read by a human in the destination organization, may support more direct, asynchronous access. [34, 23]

Traditional media support person-to-person communication. In all cases, an employee of one organization interprets and responds to communications from persons outside the organization. The response may be to pass the communication on to another employee in standard bureaucratic fashion, or it might be to routinely respond, perhaps by shipping out a requested item or initiating some other transaction such as a reservation or inter-bank transfer. However routine the response, employees typically are charged with some discretion over and responsibility for their actions. Using an ION that supports person-to-machine communication, the host, application, or peripheral in the remote organization may take action without involving a single employee of that organization; whether the action is generating an instruction to the shipping department to send 500 of part number 362f to a customer's address, for example; or updating the design file for a family of components, one of which is being designed by an outside party. In summary, the degree of change represented by the automatic nature of IONs depends on two factors: the extent of automatic processing done by ION-connected machines before an employee of the firm is involved, and the degree to which employee response to requests via traditional media are so routine as to be almost devoid of discretionary input. The types of automated and administrative controls that an organization might place on such automatic processing are discussed at length in subsequent chapters.

In addition to these direct changes in the communication medium, problematic increases in the internal value of accessible information and decreases in the universality of the communication medium often accompany ION use.

3.2.5 Access to Internal Facilities

Traditional inter-organization communication media connect persons outside the organization with persons inside. Often the internal employee has the assigned role of *boundary-spanner*, i.e., mediating access to internal information, resources, and people (e.g., customer representative, purchasing agent, etc.). Much of the value of computer-based information systems and networks within and between organizations is the ability to efficiently integrate related functions and streamline information flows. Likewise, one of the most significant motivations to interconnect is the elimination of intermediate time delays and labor costs. Consequently, the systems made directly accessible to outsiders via IONs often (and increasingly) support, or are connected to, related internal applications, e.g., inventory or engineering design databases. Moreover, these internal systems and applications often contain internally-valuable information or resources (i.e., proprietary, critical, limited, or costly) that in the past were accessed by outsiders only with the assistance of an internal employee.

IONs that support person-to-person electronic mail only do not support unmediated access to information or resources. However, often persons deeper within the organization (i.e., those who do not have official boundary-spanning roles) have more elastic demand for communications. If so, then the cost reducing characteristics of electronic mail contribute to deeper penetration by increasing the amount of external communication conducted by persons deeper within the organization. In other words, ION communication is more likely to involve persons who are not official *boundary-spanners*, than is communication via traditional media. On the other hand, much of this penetration may be attributed to the newness of the medium and therefore may be only a transient effect.

3.2.6 Universality

Theoretically, an ION facility could be used to interact with an unlimited number of interchange partners. However, an ION participant may be unable to use its ION facilities to communicate with more than one organization. Even if the participant can use the facilities with multiple organizations, it may be difficult for the organization to extend ION access to interchange partners outside of the initial set of ION participants without seeking agreement of all ION participants. The barrier to transferring ION facilities, i.e., the lack of universality, may be due to use of non-standard communication protocols or application procedures; the former prohibits using ION software or equipment, the latter prohibits using personnel training and know-how of ION procedures, e.g., learning effects. Universality may also be reduced due to the high fixed cost of ION equipment which may prevent some classes of organizations from participating. Finally, universality may be reduced intentionally through contract provisions.

In contrast, today telephone, telex, and paper mail are all highly universal media. Although special procedures and forms exist which are not universal, at least the underlying communication structure is common. At one time telephone and telex also lacked uniformity. However, the nature of the incompatibility and the implications were quite different. For example, in the early days of the telephone, there were many small independent telephone operators in addition to Bell. Although there were serious problems regarding interconnection and the ability to contact persons on the other side of town because of the multiplicity of systems, the barriers were largely administrative. Once the organizations agreed to interconnect (or were merged under the expanding Bell Telephone Company), connections were made without too much difficulty because the technical parameters were more similar across systems. Although over time different companies developed incompatible signalling schemes, all of the systems started with the original Bell patent and did not diverge significantly. In the case of IONs, a much larger and more diverse set of equipment and protocols exist. Therefore, interconnection and compatibility are more serious technical and economic impediments. On the other hand, computers support cost-effective translation among dissimilar protocols. Therefore, in the long run, third parties can offer interconnection, or protocol conversion, services if market players themselves fail to come up with a standard for clients. When this occurs the issue of universality will diminish. Nevertheless, the transient effects may have harmful long-term structural effects in terms of which organizations emerge as participants.

Message-based IONs are easier to interconnect than other types of IONs because of the greater homogeneity among message formats and transfer protocols, and, more importantly, the ease of protocol conversion (see chapter 8). Similarly, although protocols and formats for message-based invocation of computer-based resources and servers are less standardized than electronic mail transport and format, conversion among a small number of message formats and protocols is less expensive than is conversion among the same number of connection based (real-time) protocols, and experiences less performance degradation. In addition, the fixed cost of message based IONs is less than other types of IONs. Nevertheless, universality may still be lower than for traditional media due to fixed costs, barriers to adopting a new technology, and contract provisions. In the long run, development of third parties can increase universality by providing equal access and reducing the minimum fixed cost of participation.

3.2.7 Electronic Mail

Many IONs support person-to-person electronic mail only. Electronic mail applications do not introduce the same degree of change from traditional media that other ION service types do (e.g., file transfer, remote job entry, database query). In particular, the level of automatic reaction is far less since messages are interpreted and responded to by an *employee* of the remote organization. Therefore, the change in human oversight as compared with traditional media is not significant. However, electronic mail is sufficiently different from conventional communication media to make the general theory presented here useful. Three salient technical characteristics distinguish electronic mail from traditional media. Non-simultaneous presence increases the hit ratio[14] and reduces call set-up time. Nearly immediate delivery relative to conventional written forms greatly reduces minimum turn around time. And forwarding, receipt and delete, and distribution lists are all easier. See [56] for further discussion.

Table 3.2.7 provides a very rough comparison of the different media discussed—face-to-face, paper mail, telephone, telegraph, electronic mail, and full ION. Each of the six characteristics discussed in this section are outlined.

3.2.8 Exogenous Factors

All of these characteristics are subject to design by one or more of the ION participants. However, flexibility (i.e., cost of design changes) varies over the IONs development cycle, and is more available to some participants than to others. [4, 12]. Two general situations can be identified: symmetric control and ownership in which there is equal control among ION participants, and asymmetric control in which one party owns or otherwise controls design and modification of the ION. A special case of symmetric control is when a third party is employed. Depending upon the particular arrangement, third party service may not only facilitate balanced control over network design among initial ION participants, but may make it easier for new organizations to join the network at a later date. Shared services via third parties may aid universality in two ways: joining the network requires negotiation with only one entity, the third party, and the minimum cost of ION participation is potentially lower.

[14] Hit ratio is the number of successful connections over the number of attempted connections. In this case a connection is successful when the destination party is reached.

	Full ION	Email	Telegraph	Telephone	Mail	Face-Face
Speed	++	++	+	+	--	-
Incr. Cost	--	--	+	+	--	++
Capabilities	++	-	-	+	-	++
Automatic	++	+	+	-	+	-
Internal	++	-	-	-	-	-
Universality	--	-	+	++	++	++

Figure 3-5:Comparison of different media.

3.3 Communications

Use of a new communication medium does not necessarily imply changes in the communications themselves. An ION can support the same communication characteristics as were and are supported via traditional communication medium. Even if incremental costs and delivery time are reduced, organizations do not necessarily change their behavior, i.e., their communication patterns. However, if there is unmet demand, due to industry or organization factors such as pressure to speed up product turn around, organizations can use the new technology to support communications of greater intensity and scope, as described below. At the same time, increased penetration and segmentation often accompany greater intensity and scope, even if such changes are not intended by all ION participants. These latter characteristics are strongest when the ION supports person-to-machine and machine-to-machine, in addition to person-to-person, communication.

3.3.1 Intensity

Organizations with unmet demand for communications can take advantage of increased speed, reduced incremental cost, and automatic processing to *intensify* communications, in particular more frequent communication with reduced delay. The incremental cost for a given communication speed is reduced enough to make interactive communication and resource access across organization boundaries economically viable. Moreover, the automatic nature of the message processing in the remote location contributes to the overall increase in speed by eliminating the need for simultaneous presence and participation of a human being in the remote location.[15] As a result of the reduced delay between requests and responses, IONs can support more frequent communication in conjunction with finer grained coordination and interchange.

The ION participants may increase the intensity of communications in order to reduce costs or enhance products and services. For example, shorter turn around and overhead per-transaction allow customers of an electronic firm to order components more frequently but in smaller quantities and thereby reduce inventories.[16] The electronics supplier can use an ION to shift the costs of order-entry downstream to the customer.[17] The shift in inventory may also provide faster feedback on consumption patterns and thereby allow the supplier tighter product control and enhanced customer service. The ION may even allow suppliers to offer last minute, consumer-specified, product features such as system configuration of instrumentation systems.

In other domains, IONs may reduce costs for joint ventures and R&D collaboration, and both speed up and make more effective the coordination of technology transfer among participants. The electronics firm can use intensified communication to support joint

[15] In fact, both of the communicators, in both organizations, may be machines and not people. But for simplicity we discuss this process as if there is a human participant in one organization invoking services or information in a second, remote organization.

[16] This practice is referred to as *just-in-time* inventory management.

[17] I use the terms buyer and customer, and supplier and vendor interchangeably.

ventures and subcontractor relationships for those components where there is high technical uncertainty and therefore unmet demand for interaction. In particular, the new communication medium can cost-effectively support a larger number of, and more frequent, design updates between the parties—assuming other aspects of the *production* process are flexible.

3.3.2 Scope

As with intensity, organizations with unmet communication demand can use an ION's enhanced communication capabilities to support a wider *scope* of resource sharing and exchange of information. One type of scope increase is attributed to the increased speed of IONs, and the second to their capabilities and automatic nature. First, based on increased speed alone, an organization can export timely information before it becomes *stale*. To the extent that the utility of accessing computer based information and resources is sensitive to timeliness and convenience, traditional media prohibit sharing of some types of information and resources. In other words, information that before was not exchanged because of its perishable nature, can now be made available quickly enough to warrant purchase. As a result, communications include a wider scope of information and resource types. Second, the ION's automatic nature allows an outsider to access and invoke information services and resources without engaging a human intermediary. Therefore, ION participants may use the enhanced timeliness and convenience of the new technology to expand the range of resources and information interchanged.

For example, the electronics firm's customers may evaluate the component-inventory database when making manufacturing or purchasing decisions; whereas previously it was not economical or feasible for the suppliers to make as wide a range of information directly available to customers. A financial services firm can include information and services whose market value depends on the timely, interactive access offered by IONs. Similarly, a medical products supplier can introduce auxiliary services along with online order entry. From the buyer's perspective these services may enhance the product line and differentiate the vendor's products from those of its competitors (see section 3.3.4 below). [4] Similarly, a subcontractor can invoke the electronics firm's computer aided design (CAD) system to test the interaction of component specifications. And, research and development (R&D) laboratories can share expensive resources, such as supercomputers, since they can be conveniently utilized remotely.[18] These examples illustrate how organizations can increase the scope of their communications to enhance and differentiate products and services. [4, 12] Scope increases that allow suppliers to offer new types of services can in turn affect other industries; if the enhancements overlap with another industry, new ION-supported services may provide a substitute for existing sources of information or services (e.g., financial services firms and insurance companies).

3.3.3 Penetration

If organizations do increase the intensity and scope of their communications, two additional communication characteristics may be affected in problematic ways, namely, penetration and segmentation.

Automated communication media offer efficiency and functionality that were not available using traditional communication media. As a result, an organization can use the ION to efficiently provide an outsider with direct access to resources, information, and people that are located deeper within the organization.[19] In the process, this automated communication may reduce or eliminate human oversight from the accessing of the internal resources by outsiders. Together these technical characteristics—automated access and access to more internal information—can lead to behavioral changes, namely deeper *penetration* of outsiders into an organization. For example, whereas previously the electronics firm's customers would obtain projected price and inventory information from a sales department, an ION can allow customers to access the supplier's internal, online price list, without human mediation. Similarly, whereas previously the subcontractors would obtain part specifications from a person in manufacturing or engineering assigned to interface with

[18] This particular example is more generally applicable to IONs that support real-time communication.

[19] We think of a person being *deeper* within the organization, the fewer, and less significant are his or her dealings with persons *outside* of the organization.

subcontractors (an official boundary spanner). IONs make direct access to the electronics firm's internal engineering or manufacturing cad/cam database a more efficient and effective channel.

Typically, an ION participant does not have penetration as a design objective in the same way that it has intensity and scope. Rather, increased penetration is a possible result of seeking these direct design objectives. Although theoretically each ION participant can control the extent of penetration, typically this issue is not considered as explicitly as are intensity and scope. Because access to internal systems and the automatic nature of the channel are minimized, penetration is minimized for IONs that support only person-to-person communication. In addition, more sophisticated usage control mechanisms may reduce the extent of undesired penetration; see Chapter 4. In some cases, fear of this unknown technology may so dominate as to inhibit interconnection altogether. Alternatively it may simply dampen or delay behavioral changes such as expanded communication and cross-boundary communications.

As described in the previous section, person-to-person electronic mail does not bypass human oversight in the same way that other ION applications do. Nevertheless, to the extent electronic mail does not have the same set of customs associated with it that traditional media do, electronic mail may contribute to deeper penetration by supporting direct access to internal personnel, as opposed to official boundary spanners. In addition to faster turn around, non-simultaneous presence, and easy preparation, electronic mail is more often created directly instead of via a secretary, and has fewer associated institutional or cultural norms. [23] However, these changes are as much a function of new organization characteristics (the protocols of communication) as of new technical characteristics, and the impacts may be more transient as a result.

3.3.4 Segmentation

If an organization values the greater efficiency, intensity, and scope offered by IONs, it may be unwilling to substitute communications that rely on traditional media for those that have ION support. The organization's communications are thereby *segmented* into ION and non ION supported. If an ION communication medium is not universally accessible due to non-standard protocols or application procedures, or to restrictive contract provisions, this segmentation corresponds to the membership of interchange partners in the closed set of ION participants. If the fixed cost of ION participation is high, this segmentation corresponds to the size of the interchange partner and its ability to pay for ION access. However, as described in the previous section, in the long run, development of third parties may increase universality and thereby decrease segmentation.

For example, a computer manufacturer that uses an ION provided by one of its electronics suppliers may find that the ION provides inventory and administrative cost savings, or is simply more convenient for the purchasing agent. If the ION facility cannot be used to communicate with the customer's other electronics suppliers because it uses a proprietary, non-standard protocol and application interface, then from the customer's viewpoint, non-ION suppliers are less directly competitive for those communications that benefit from ION enhancements. In other words, if ION supported communications are preferred, the customer faces higher switching costs than were experienced when a traditional media was used (e.g., telephone, or standardized paper purchase order/invoices).[20] High switching costs can contribute to higher barriers to market entry as well as shifts in bargaining power of buyers and suppliers.

Even if the ION is standardized, if the fixed cost of access (equipment, software, etc.) is a barrier to participation by small organizations, the ION-supported interchange will include only larger suppliers. For example, there are several nation-wide R&D networks in the U.S. that are open to participation by most R&D laboratories (under certain restrictions such as no directly commercial use). However, these networks differ significantly from one another in capabilities and cost of access. Communication and collaboration that makes use of the more advanced capabilities available on the Arpanet (online access to computer-based resources), for example, is not as readily carried out with organizations that have access to

[20] Switching cost is the cost to a customer to switch from one supplier to another; in this case the cost of reduced efficiency or convenience of transacting with a non-ION supplier.

the lower cost networks only (i.e., CSNET, BITNET, UUCPnet—see Chapter 11 for further description of these networks.) Therefore, collaboration activities are segmented according to whether they can make use of the more sophisticated capabilities.

Even if the communication equipment itself is transferable, the ION may promote segmentation. Typically, the form of resources and information are more unique the deeper within an organization they lie. In other words, the information and resources are less standardized and therefore are less easily substituted with information and resources belonging to other organizations in the market. Segmentation will increase if an organization becomes dependent upon procedures, information, and resources that lie deeply within another organization, and that are not widely available from other sources. For example, if the computer manufacturer relies on detailed and timely information regarding electronic components availability for production planning, it will favor suppliers that provide such detailed information. Furthermore, if the information is of a special type or in a special form, the manufacturer may prefer to adopt a single source rather than deal with multiple formats and types of information, even if the information is accessible via standardized ION facilities. Similarly, an university research laboratory may make use of an experimental supercomputer being developed in an industrial laboratory. The university researchers may develop special software or techniques as part of a joint research project. If the supercomputer is not available from other laboratories, the university laboratory's interchange with the industrial laboratory is less readily substituted by interchange with other laboratories than if the relationship with this industrial lab were based strictly on exchange of technical reports, for example.

3.3.5 Exogenous Factors

Several industry factors influence whether and in which ways organizations choose to change their communication characteristics to make use of the new communication medium. In other words, these factors influence the level and elasticity of demand for communication. First, if there exist industry pressures to speed up turn-around time, then there is incentive to make use of ION features to speed up communications; if no such pressures exist, there may be no unmet demand and therefore no reason for an organization to upgrade its communication characteristics in response to a change in the communication medium. If a product has a short lifetime, or if the cost of holding inventory is high, for example, then demand is higher and an organization has incentive to change communication patterns by increasing the frequency and turn around with which it can order or sell products.[21] Similarly, if a joint venture or subcontracting arrangement operates in an industry characterized by high technical uncertainty, the participants can use increased communication intensity to shorten and make more flexible the production cycle. On the other hand, if an organization is already communicating intensively with outsiders, other factors—such as diminishing returns—may reduce the value of intensified communication.

Second, if the external communications are important to the function of an organization, there may be unmet demand for information and resource interchange between participants. For example, in the case of joint ventures and subcontracting arrangements, more tasks and activities can be coordinated efficiently across boundaries and thereby allow other constraints, such as location of resources and expertise, to determine how tasks are divided and allocated (these issues are discussed further in section 3.4). Similarly, if an organization can enhance a service by augmenting it with timely information or interactive access to additional resources, communications of increased scope will provide value to ION participants. In both cases, if the focal organization does not consider the interchange important to its central function, it may not be motivated to invest in new procedures.

Organization and industry factors influence how penetrating and segmented ION-supported communications are. Penetration is partly a function of an organization's internal computer and network facilities. If information, resources, and people deep within the organization are not accessible via computer-based communication, the ION would have no way of making them any more accessible to outsiders since the ION would not extend deeper into

[21] In an unpublished working paper, D. Gherson points out several product types that have this perishable quality—financial, airline reservations, etc.

the organization than traditional media do. Segmentation on the basis of interchange content was described in terms of penetration and consequently it too depends upon the extent of the remote organization's internal facilities. Industry and organization factors also influence whether an organization would benefit from increased segmentation, and whether the organization would actually have the foresight and power to impose it.

Although segmentation and penetration are treated as problematic changes, both may be intentionally imposed by one or more of the ION participants. Some organizations may find it particularly useful to design increased specificity into the ION, thereby increasing segmentation, tying in their interchange partners, and reducing competition. An industrial research laboratory might use this strategy to solidify its relationship with a university laboratory, to the exclusion of other commercial competitors. Industry factors determine whether or not there exist incentives to tie in interchange partners, while organization factors determine whether each ION participant is likely to recognize the opportunities. Finally, if other ION participants perceive increased segmentation as a threat, the outcome will depend on the structure of the industry.

3.4 Policies Governing Cross-Boundary Activities

Just as organizations do not necessarily modify their communication characteristics in response to changes in the communication medium, new communication characteristics may or may not lead organizations to change the way they manage and make use of cross-boundary activities. However, in addition to exploiting the new communication characteristics supported by IONs to reduce costs and enhance products or services, organizations can use the new communication characteristics to shift internal activities out across organization boundaries and to transact with a larger number of outside organizations. In other words, just as the new communication medium can support new communication characteristics, the new communication characteristics can support new patterns of cross-boundary activity. For this reason, the economics of IONs can affect the decisions that individual firms make about organizing production activities and managing inter-organization relations.

3.4.1 Cross-Boundary Activities

The greater intensity and scope of interchange means that some activities that previously were carried out most efficiently within the organization can now be carried out efficiently across the organization's boundaries. Williamson describes a continuum of ways in which production activities are managed, from internal to market. [80] He claims that a primary criterion for selecting the former over the latter is the relatively high cost of coordination in a market under conditions of high uncertainty. IONs support intense interaction between organizations, allowing them to coordinate adjustments quickly and efficiently; i.e., resembling coordination within a single organization more closely. Therefore, conditions that previously prohibited cross-boundary activity due to excessive coordination costs, can now be accommodated by virtue of ION-supported coordination. For example, technological and volume uncertainty increase the need for ongoing communication between computer manufactures and electronics suppliers. Previously, the expense and difficulty of intensive outside communication sometimes lead organizations to make some components internally instead of purchasing them, even if there existed production-cost advantages to outside production. The greater frequency, timeliness, and lower incremental cost of ION communication may cause buying a component (i.e., cross-boundary activity) to become a viable alternative to internal production.

Another type of increased cross-boundary activity is the introduction of certain kinds of new products. The greater scope of interchange supported by an ION allows an organization to offer as products internal information and resources that previously could not be made available to outsiders in a timely or economic manner. For example, some banks provide large customers with terminals that support direct access to internal portfolio management systems in addition to standard communications.

3.4.2 Number of Interchange Partners

Greater intensity and efficiency also allow a single organization to coordinate interchange with a larger number of organizations efficiently. Since the cost of preparing and executing communication is lower, the amount and frequency of communication can increase. This larger volume of interchange can spill over to increase the number of organizations that are contacted. In addition, because the minimum cost of querying an ION participant is lower than the minimum cost of querying a non-ION-connected organization, the number of organizations communicated with per transaction can be greater.

For example, if the ION decreases the incremental cost to the computer manufacturer of checking the price of its electronics suppliers, then the manufacturer can afford to survey a *larger market* before purchasing. Similarly, if an ION among research laboratories supports more intensive and efficient communications among researchers, each researcher and therefore each laboratory can exchange information and resources with a larger number of other researchers and laboratories.

3.4.3 Restrictions on Interchange

ION participants may perceive increased risk due to the reduced oversight and increased internal value of the information and resources accessed by outsiders. In order to cope with this new risk, organizations may apply restrictions to ION supported interchange. The restrictions may be as formal as a contract provision or technical and administrative controls on interchange, or as informal as limiting ION use to a few major interchange partners. Formal agreements specifying liabilities may reduce some risk. They may also inhibit the ease with which additional organizations are brought online. Because penetration and risk increase less for IONs that support person-to-person communication only, restrictions on interchange will also increase less for such IONs than for those that support machine communication as well.

Technical and administrative controls on cross-boundary flows may contain the extent of penetration and risk (see chapters 4 through 9). However, as part of the codification process needed to implement technical controls on ION flows, an organization may make more explicit what is and what is not allowed to flow across an organization boundary. Organizations often accompany computer-based automation with increased codification of rules, procedures, work and information flows, that were previously left vague. This impact is sometimes intended by the organizations, and sometimes is an unanticipated side-effect of defining a computer system to take the place of manual procedures. Therefore, in addition to the protective restrictive measures described above, the process of defining ION procedures may contribute to the restrictiveness of inter-organization interchange and relations. Moreover, codifying and implementing restrictions on external flows often impacts internal communication.

The fear of this *unknown technology* could overwhelm factors favoring expanded cross-boundary activities. The most extreme case is choosing not to interconnect at all. Even given interconnection, such fears could significantly dampen increases in ION mediated cross-boundary activities.

3.4.4 Restrictions on the Set of Interchange Partners

An ION participant may choose to minimize risk by using the ION only with a small set of select interchange partners. For example, a manufacturer may be able to efficiently coordinate interchange with a large number of subcontractors by using an ION for exchange of cad/cam data. However, to the extent the ION provides subcontractors with access to sensitive or proprietary internal information and resources, the manufacturer would be making itself more vulnerable. Consequently, the manufacturer may use the ION for a few subcontractors only. As with increased restrictions on interchange, because penetration is less, restrictions on the set of interchange partners will be much weaker, or even non-existent, for IONs that are used for person-to-person communication only.

Segmentation of communications according to ION support may result in interchange with a smaller, not larger, set of organizations. IONs can not contribute to greater numbers of interchange partners if the equipment or communications themselves are specific to a single or small set of organizations (due to non-standard protocols, contract restrictions, or high fixed costs). For example, although a customer can transact more efficiently with the electronics supplier that supports online order-entry, if the ION facilities cannot be used to communicate with other suppliers as well, the efficiency gained does not promote interchange with a larger number of suppliers. As described earlier, some ION participants may seek to impose such segmentation in order to tie in interchange partners and reduce competition. On the other hand, exploitation of segmentation will encouragement development of third parties which may increase universality of network access.

3.4.5 Exogenous Factors

Where there is unmet demand for increased cross-boundary activity (e.g., buying over making, or joint ventures over internal ones), organizations can use communications of increased scope to carry out greater numbers of activities across their organization boundaries. Similarly, where there is benefit to interchange with a larger number of organizations, organizations can use more intense communications to seek out greater numbers of interchange partners. Organizational factors influence the extent to which organizations recognize these opportunities for expanded cross-boundary activities. They also influence the organizations consideration of greater penetration and segmentation in decisions about how to manage production activities.

3.4.5.1 Production Cost Advantage

Communication or coordination costs are only one criterion according to which an organization decides whether to carry out an activity in the market or internally. The primary criterion is production costs. If other organizations have production cost advantages due to economies of scale or greater or unique expertise or resources, an organization has incentives to carry out an activity across its organization boundary rather than internally; e.g., if a supplier can produce a component more cheaply than the customer, the customer should buy the product rather than make it. On the other hand, if the production cost advantage of buying over making has always been very high, the decrease in communication costs may not affect the organization of production activities since external production would already be the primary mode of choice. Therefore, the reductions in communication costs offered by IONs will result in increased cross-boundary activity only when the relationship between communication costs and relative production costs change from a situation in which communication costs exceed production cost advantage to one in which production cost advantage exceeds communication costs. In some cases, basic limitations on production cost advantages, may impose an absolute ceiling on the effect of reduced communication costs; e.g., limitations on the amount of information and information sources that a researcher can assimilate meaningfully into his or her work. Similarly, the reductions in communication costs offered by IONs will result in a larger number of interchange partners, when the larger number offers production cost or quality advantages through increased variety, price competition, or bargaining power. Production cost advantage can be viewed as a firm's motivation to alter the way it currently organizes production activities.

3.4.5.2 Level of Decision Making

Increased penetration and segmentation supported by ION's may not be of concern at all levels of an organization. Therefore, the level of attention accorded to the interconnection determines the degree of risk that the organization will *perceive* and the extent to which it will place restrictions on ION-supported activities. Typically, higher levels of management are more concerned with managing vulnerabilities and dependencies; whereas, at lower levels, the primary concern is getting the job at hand done expediently. If higher levels of management pay attention to ION adoption, ION-induced penetration and segmentation is more likely to result in more increased interchange restrictions (e.g. restricted numbers of interchange partners, contracts, administrative and technical controls).

Similarly, lower levels are less likely to pay attention to the potential for taking advantage of new communication characteristics in the management of production activities. If the ION is dealt with only at the technical level of Data Processing or Telecommunications administration, organizations are unlikely to intentionally use an ION to achieve strategic objectives such as product differentiation, raised switching costs, or increased bargaining power. [12][22] If higher level management is involved in ION deployment, it may consider opportunities for strategic gain, such as manipulating specificity to tie in customers. An organization's role in initiation of the ION is related to the level of decision making. Initiators are more likely to have considered the implications of the ION before proposing

[22] In those organizations such as university research laboratories where the technical personnel also determine how to manage research activities, this statement does not apply.

the investment. Followers, are more likely to react on a strictly procedural level and thereby overlook higher-level, strategic implications. [12]

3.5 Summary

The model is of most use when interpreted in the context of a particular domain. However, below I summarize the general predictions set forth in the model. These predictions apply to those environments in which there is unmet demand for cross-boundary communications and activities.

1. Several predictions regarding communication characteristics can be made on the basis of this model.

    a. IONs will support communication of *greater intensity and scope.*

    b. If ION facilities are specific to a single or small set of interchange partners, an organization's communications will be *more segmented* than they were when only traditional media were used; i.e., the organization is not likely to substitute communications that rely on traditional media for ION-supported communications.

    c. The greater scope of information and resource sharing, and direct access to more internal resources and information, will result in communications that *penetrates deeper* into the organization. IONs that support person-to-person communications only will not increase the level of penetration as much.

2. Several additional predictions can be made about the way in which policies governing production activities may be affected:

    a. The expanded intensity and scope of ION-based communication will support an *increase in cross-boundary activity* in the form of vertical de-integration, joint ventures, or new products and services.

    b. IONs support interchange with a *larger number of outside organizations* than was engaged previously.
    However, if the ION cannot be used outside of a *closed set of organizations*, the number of interchange partners will be inhibited. The number may even decrease if the relative benefits of ION use, and cost of extending the ION beyond the initial set of participants, are both high.

c. Organizations will impose *restrictions on cross-boundary activities*, such as codification of cross-boundary flows, limited numbers of ION partners, or contract statements, in response to deeper penetration.

The model and predictions are summarized in the figure 3-6. The first part of the figure illustrates the direct opportunities introduced by the new medium. The second part illustrates problematic implications of exploiting these opportunities and ways in which organizations may respond, i.e., the indirect changes illustrated in figure 3-1. The thin-line arrows indicate opportunities offered to an ION participant by the characteristics at the tail of the arrow. The thick-line arrows indicate more direct implication, i.e., the characteristics at the tail of the arrows will bring about the changes pointed to under certain industry and organization conditions. The label above each arrow indicates which of the above predictions it corresponds to.

The three dimensions of policies governing cross-boundary activity are dependent variables—number of interchange partners, cross-boundary activities, and restrictions on cross-boundary activities. Segmentation and penetration are dependent variables with respect to intensity and scope and are independent variables with respect to number of interchange partners and restrictions. Intensity and scope are also independent variables with respect to number of interchange partners and cross-boundary activities. Finally, exogenous industry and organization factors are strictly independent variables in this model.

3.6 Conclusion

This chapter outlined our general model of how ION use affects participants' communications and cross-boundary activities. The model described changes on three levels—communication medium, communications, and policies governing cross-boundary activities. In general, ION use can support tighter coupling between participants but at the same time may introduce a new boundary between ION and non-ION organizations.

The general predictions set forth in the model are illustrated in the last portion of the thesis using examples of distribution channels (Chapter 10) and an empirical study of Research

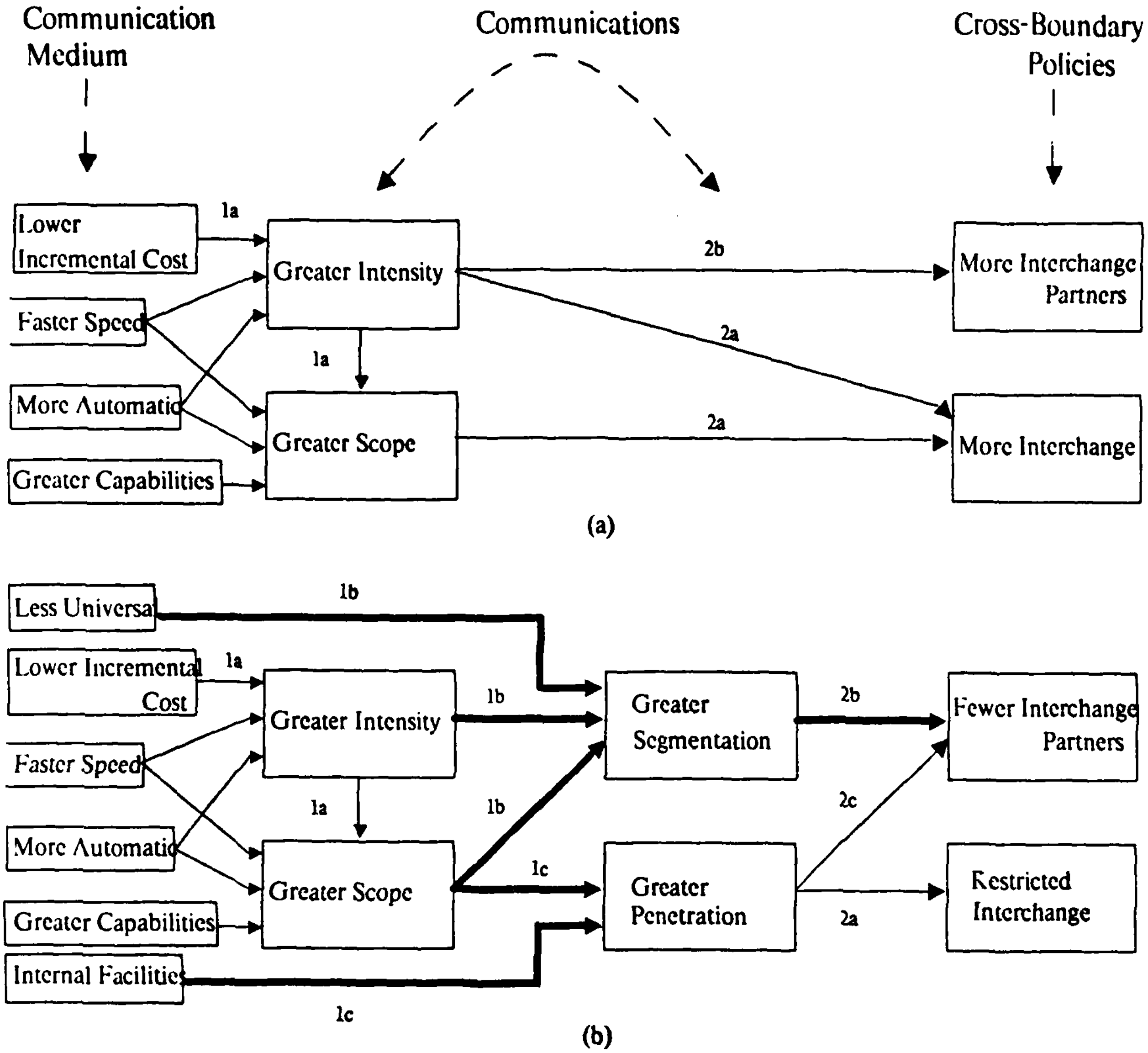


Figure 3-6: Model of ION Impacts. Part (a)
illustrates the opportunities introduced by IONs. Part (b)
illustrates some problematic implications of exploiting these opportunities.
The numbers refer to the predictions listed on the previous page.

and Development laboratories (Chapters 11 and 12). In the R&D environment we found greater intensity, scope, number of interchange partners, and cross boundary activities.

Some evidence of segmentation was also found. However, no restrictive behavior was indicated.

The next portion of the thesis, Chapters 4 through 9, investigates the technical characteristics of this medium, IONs. In particular, the desire to preserve organization boundaries, in terms of access to information and resources, raises new types of access control requirements. Hence, we develop mechanisms to embody organization boundaries in network interconnections.

The model described in this chapter serves as the context and motivation for the technical mechanisms developed. However, the technical issues also have direct implications for the model described above. In particular, the ability to implement these controls, and the costs of doing so, influence several parameters of the model. In terms of the medium itself, controls may reduce the range of capabilities and the automatic nature of ION services. Similarly, tighter controls are likely to reduce the increase in ION-supported penetration. At the same time, implementing technical controls may entail increased codification of communication flows and thereby increase the overall level of restrictions governing cross-boundary activities.

The interleaved organization of this thesis is representative of our treatment of causality. The organizational context is the basis for our technical design while the technical characteristics are central to our predictions of how use of IONs will develop.

Chapter Four

Usage Control Requirements in Inter-Organization Networks

As described in the previous chapter, a central concern of ION participants is protection of their organization boundaries in terms of access to information and resources. This chapter introduces the access control issues addressed throughout the next six chapters of the thesis. Section 4.1 describes the salient features of IONs. Section 4.2 introduces four real world examples which are used to illustrate points throughout the thesis. Section 4.3 begins the discussion of usage control requirements which is the subject of Chapters 5 through 9. Section 4.4 concludes this introduction with a review of related work in network interconnection and computer and communication security.

4.1 Classifications and Definitions

Computer networks in general, and IONs in particular, can be described on three levels—operational, logical, and physical (see figure 4-1). At the first, operational, level, an ION includes the administrative procedures and policies that govern use of the facilities encompassed in the ION; for example, the types of interchange, usage patterns, access rules, and accounting. This level is of most concern to the managers of the ION-supported interchange functions within the participating organizations and to the end users. Existing IONs include interconnections between airlines and travel agents, between banks, between insurance companies and agents, between research institutions, between medical-product suppliers and hospitals, between automobile manufacturers and parts subcontractors, etc. In each of these cases the interconnecting organizations want to enhance operations across their organization boundaries. The previous chapter focused on the organization implications of IONs and is tied most closely to this operational level.

At the second, logical, level, an ION is the set of accessible computer resources and applications formed via interconnection of facilities that are owned, operated, and/or used by two or more organizations. The logical ION excludes human decision making as part of the interconnection process and deals only with automatic procedures. It refers to all processes and applications that can be invoked automatically from another position in the ION. Participating organizations typically are concerned most with this level of the network. For each of the IONs mentioned above, a logical ION can be described; for example: the reservation systems of the airlines and the access equipment of the travel agents; the computers from and to which bank funds are transferred; the insurance companies' and agents' record management systems; general and special purpose computing and communications resources belonging to multiple research institutions; the medical-product suppliers inventory system and the hospitals' order-entry system; the automotive manufacturer's cad/cam and inventory systems and the subcontractor's cad/cam system, etc. The technical issues discussed in the following chapters are concerned most directly with this logical level.

At the third, physical, level, an ION is the transport mechanism and the supporting architecture (e.g., data format, coding, and exchange protocols) via which data are passed; this is the level commonly addressed by computer network designers. In the physical ION the interconnection of organizations' facilities need not manifest itself in the installation of a physical wire or switch, but only in an agreed-upon protocol for transferring and interpreting data.[23] Physical IONs that correspond to the arrangements described above are: travel agents using a specialized protocol over leased lines to communicate with an airline company's central computer; insurance companies and agents communicating through a third-party, value added network accessed via dial-up or dedicated telephone facilities or via a public packet-switched network; computer science research institutions using packet switched architectures over telephone lines; and the customer-supplier interchanges based on standardized or specialized protocols over dial-up and dedicated telephone lines, or magnetic tape transfers.

[23] For example, even magnetic tape transfers or automatic processing of telex messages qualify as automatic processing of external transactions; although in the case of tape transfer issues differ because transmission is not automatic, i.e., it requires human participation to transfer and down-load the tape.

Level	Function	Design
Operational	Policies and Procedures	Map organization policies to usage control requirements.
Logical	Applications	Define usage control mechanisms.
Physical	Network Architecture	Identify architectural support needed to implement mechanisms.

Figure 4-1: Levels of an ION

I distinguish among these three levels because although a given logical network can be supported by one of a number of physical configurations, and can be operated in a variety of ways, the design choices made at each of the three levels interact with one another. For example, policy requirements at the operational level imply implementation requirements at the logical level, which in turn imply design requirements at the physical level.

Many organizations have network connections to public carrier networks such as Telenet, Tymnet, and other packet-switched networks. Such connections, between a client organization's internal network and a public, packet-switched network, cross organization boundaries at the *physical* level, but not necessarily at higher levels. The client-to-carrier connection is not intended to support inter-working of the client's and carrier's computer based resources; i.e., there exists no logical network, by the definition given above. In fact, the client might use the public network to interconnect the geographically-distributed facilities of its own organization; in which case no ION exists. This thesis treats IONs at the logical and operational levels, primarily. Because the physical level is of interest only to the extent it affects higher levels, client-to-carrier connections are addressed only as used to support logical IONs among clients.

Existing IONs can be classified into two types—those that are dedicated to a single ION application, and those that support more general communication capabilities. At the physical level, the first type of ION is an interconnection between single computers, whereas the second is an interconnection between networks of computers. At the logical level, the first type is a system whose function crosses an organization boundary(ies). The interconnected facilities are dedicated to specific, well-defined, inter-organization interchange functions (e.g., a particular database transaction application such as airline reservations or order/entry). Because of their automatic nature, such interconnections can raise significant policy issues for the participants at the operational level. But, from a technical standpoint, usage control mechanisms can be treated as an extension of traditional, database-management and information system security, and do not impose on internal operations since the system is used for ION purposes only. In contrast, the second, more general, type of ION is composed of facilities interconnected to support generic inter-organization communications on top of which a multiplicity of user-defined applications may operate. This more general type of ION arises out of interconnections between networks of facilities of two or more organizations. By virtue of this interconnection a range of resources potentially are accessible to persons and machines within the other organization(s). However, at the operational level the participants may not intend that the entire set of internal resources in each organization form an integrated system or even be accessible.

This distinction between the two types of IONs can be described in terms of overlap between logical networks (see figure 4-2). We can model an organization's internal facilities as multiple logical networks operating on top of an internal physical network. Each ION participant's internal network consists of applications that pertain to strictly-internal operations. The logical ION consists of resources that the participants intend to make accessible to each other. The logical ION crosses organization boundaries and operates on top of physical networks belonging to multiple organizations.[24] In terms of this model, the

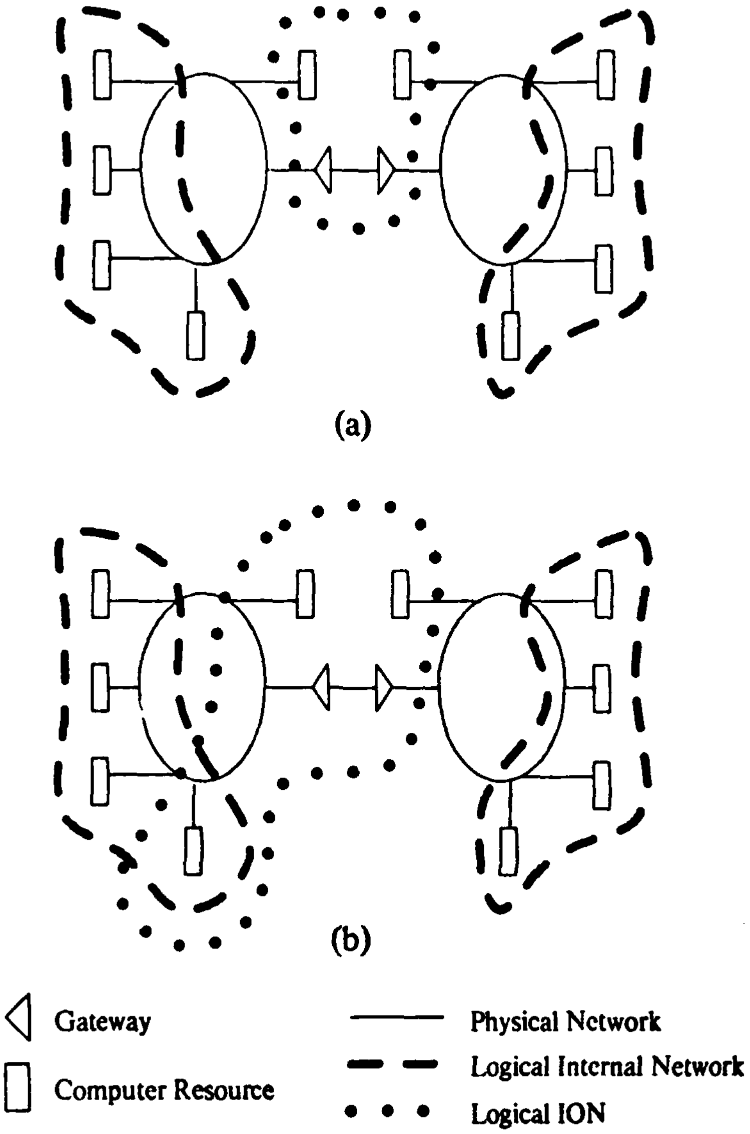
[24] If an organization supports multiple types of inter-organization interchange, electronically, each type constitutes a separate logical ION.

first type of ION is an interconnection in which the logical ION does not overlap with the participants' internal networks; i.e., the facilities accessed by outsiders are dedicated to that single function. The second, more general, type of ION is an interconnection in which the logical ION and internal networks do overlap; i.e., the facilities in the overlap are used for both internal and external applications. This potentially results in conflict between internal requirements for connectivity, transparency, and maximum performance, and inter-organization requirements for controlled access.

The remainder of the thesis focuses on the second, more general, type of ION. This more general case raises fundamental technical issues, not addressed by traditional intra-organization network interconnection, nor by traditional security mechanisms for shared systems. This emphasis is justified because in fact, organizations are using information technologies in support of a wiue range of internal and external activities, and are extending internal networks to support information and resource flows among these activities. As a result, there exist more internal computer-based resources that an organization might want to make accessible to an external interchange partner, and unrestricted external interconnection to one resource is more likely to imply access to other internal resources.

4.2 Examples

To make more explicit the discussion of usage control requirements in IONs this section describes the use of IONs by four organizations. Each organization described illustrates a different policy perspective and corresponding set of usage control requirements. The examples are real but the names have been changed to protect the proprietary concerns of the subjects. The examples are taken from the research community because that is where a number of sophisticated, internal and inter-organization networks are in use. Moreover, the relatively integrated nature of the internal, computer-based resources is representative of how many organizations are likely to use this technology in the future; based on the economics and technical characteristics of the technology and applications. The interconnections are described in this section, usage control requirements are discussed in section 4.3 and Chapter 5, and implementation issues are analyzed in Chapter 6 and 8.



(a)

(b)

Figure 4-2:(a) Non-overlapping and (b) overlapping logical networks.

4.2.1 MIT

MIT has extensive and varied internal computing resources, most of which are interconnected via local area networks. In addition, MIT has several external network connections, several of which are described here—the Arpanet gateway, public accounts on Multics, a dial-up gateway, and two dedicated connections to local companies.

Most Arpanet participants have individual hosts or groups of hosts, connected to the Arpanet directly via a special network interface. MIT has had such host-connections to the Arpanet since the network's inception, over fifteen years ago. However, several years ago the MIT Laboratory for Computer Science implemented a packet-level Arpanet gateway to provide all hosts on the MIT local networks with *direct* Arpanet access. The gateway supports any protocol that operates on top of the Internet Protocol (IP). [54]; in this case the most common applications used are mail transfer, file transfer, and remote login. MIT faculty, research staff, and students use the Arpanet for person-to-person communication via electronic mail, exchanging documents and software via file transfer, and accessing remote computers and applications via remote login.

MIT operates a Multics computer system. In addition to serving MIT users, Multics sells account time and space to non-MIT users.[25] Because Multics is connected to the MIT internal network and several external networks (i.e., Arpanet, Mailnet, Bitnet, and ScienceNet), it serves as a high-level ION gateway from the non-MIT organizations that have Multics accounts to MIT and the other external networks. The connection between non-MIT sites or users and Multics is achieved in one of two ways. The most common mode is for an individual with an account on Multics to call in via telephone or a packet-switched network and log in. A second mode is for a user to leave a daemon running under his or her Multics account that regularly wakes up and forwards mail or other kinds of traffic from Multics to the user's local machine, via a telephone or packet-switched network. Most Multics users have network privileges and thereby can communicate via electronic mail, file transfer, and remote login with other hosts at MIT and elsewhere. In yet another mode of access, users on BITNET and Mailnet exchange mail with MIT, and MIT-connected hosts, via Multics which is a node on each of these networks.

A third MIT gateway connects on one side to the public telephone network via dial-in ports and on the other side to MIT's internal network. This *dial-up gateway* is a packet-level

[25] There is a loose requirement for a user's work to be related to some interest of MIT in order for the user to be given a Multics account.

network entry point for both MIT and non-MIT users. To the end-user the gateway *appears* more or less like a terminal concentrator. The end-users connect to the gateway via the telephone network, and establish connections to their destinations as if they were connected directly to the same local network. However, unlike many terminal concentrators, the connections are established between the two endpoints directly—the gateway sees only individual packets and not the connection per se. The gateway routes packets according to the header addresses and some state information that maps dial-up-user-address to gateway port number. This gateway is used by a range of off-site users. Along with a terminal concentrator it provides users with remote access to hosts that do not have their own dial-in lines and it provides more flexible access to multiple hosts and peripherals. The gateway is also used for inter-organization communication. For example, an experimental community information service transmits data via the gateway to a local radio station for over-the-air distribution.

In addition to the three external connections described above, MIT has two low-level network connections to local companies. These gateways forward packets between the MIT local networks and the local networks of the two companies, respectively. Users and hosts on either side of these gateways can communicate with one another via electronic mail, file transfer, and remote login using this low-level gateway. At the communication level, the local companies look as if they are geographically-remote MIT sites because they are a part of the MIT network.

Within MIT there are several communities of users. The Laboratory for Computer Science and Artificial Intelligence Laboratory are the official users of the Arpanet and dial-up gateways. Multics is administered by an institute-wide information processing services center. Some machines used by the MIT administration contain private or sensitive information (student records, payroll, personnel, etc.) and are not even connected to the internal network because of the perceived risk of unauthorized access. However, in the future some of these hosts are likely to establish restricted connections to the MIT network. In addition, MIT is expanding its campus wide network to accommodate increased computer use in a wide range of departments, both for teaching and research. In particular, *every* student on campus will have access to some computer facilities in the coming year as part of a large-scale effort to incorporate computers in education, known a Project Athena. The diversity of this internal environment is very relevant to the external connections and their management, as will be described in later sections.

4.2.2 ABC Inc.

ABC Inc. is a large, US-based, computer manufacturer with sites all over the world. ABC makes heavy use of computers internally and supports world-wide inter-computer and inter-site communication for all research and development sites over an internal store-and-forward network. In addition, ABC sales, manufacturing, and other critical business operations depend heavily on direct access to business-oriented databases over a large private network of leased telephone lines, satellite links, etc. ABC also supports several inter-organization network connections, four of which are described below—access to two nation-wide R&D networks, links to subcontractors, and a value added network.

ABC operates gateways between its internal research and development network and two R&D IONs—BITNET and CSNET. BITNET participants are university computer centers and university computer science departments, primarily. The network supports mail and file transfer, and limited interactive communications. A participant joins the network by establishing a telecommunications link (leased line) to its closest BITNET neighbor and agrees to do the same for future members. CSNET is administered by NSF and connects computer science departments in universities and industrial labs. CSNET supports electronic mail throughout the network, and remote login and file transfer via a less widely used (more expensive) X25net service. Both networks have gateways to the Arpanet (described above) and USENET (a network composed of computer companies and universities which supports mail and file transfer). In addition, BITNET is connected to its European counterpart, the European Academic Research Network (EARN), making the ION, international as well. ABC's researchers use the BITNET and CSNET connections to facilitate joint projects, sponsored projects, collaboration, and informal communication, primarily with universities. The primary ABC gateway supports only electronic mail.

However, electronic mail is used to exchange software and data (e.g., software updates, and input data to simulation program), in addition to person-to-person communication (e.g., exchange of research ideas and progress reports, joint authorship of papers, and administrative scheduling).

ABC employs a separate ION connection to support a joint-development project between a west-coast subcontractor and an east-coast division of ABC. ABC connected the subcontractor's facilities to a west-coast node of ABC's R&D network, which then serves as a transit path between the ABC development group on the east coast and the subcontractor on the west coast. This connection supports mail and file transfer as well as some interactive communications.

Two other external connections are worth mentioning. A separate division of ABC operates a value added network over which they provide information and communications services to subscribers. This service network has connections to ABC's internal network, as well as to the subscribers' facilities. It differs from the two connections described thus far, and is of somewhat less interest, because the facilities made accessible to outsiders are strictly for external use, i.e., the logical ION and internal networks do not overlap. A second example is a proposed connection between ABC's network and the reservation system of the company's preferred airline carrier. This experimental connection would support direct online information and reservation services to end users throughout the ABC network.

The usage control concerns associated with these various connections will be discussed in section 4.3 and 8.5.

4.2.3 XYZ Inc.

XYZ Inc. is another computer manufacturer that uses computer communications internally and for communication with outside organizations. As with ABC, XYZ has a world-wide internal, network which supports mail and file transfer and remote login. The network is used for both R&D and management, manufacturing, and other business-related functions; the management and manufacturing and R&D communication systems are *not* as strictly separated as are ABC's systems. XYZ supports several external connections—R&D network, customer orders, customer program development. Additional connections are being planned such as online software distribution.

XYZ has a mail connection to Arpanet which serves the same purposes as ABC's described above. XYZ is also connected to USENET. The usage control issues differ across companies and are discussed separately in section 4.3. Because XYZ is a contractor for the DOD section that administers the Arpanet, it has hosts directly connected to the Arpanet. Like the mail gateway described, these hosts also forward mail automatically between any XYZ host and Arpanet host. In addition, remote login and file transfer capabilities are available to XYZ users that have accounts on the particular hosts; these capabilities are not available to other XYZ machines—the hosts do not act as gateways for these protocols, only mail. In addition, only outgoing file transfer and remote login is available, i.e., Arpanet users cannot establish connections to the XYZ hosts.

A second use of computer-based communications with outside organizations is the online product information and ordering system made accessible to customers. This system is not currently integrated with the rest of the internal network—orders are manually transcribed onto paper before being processed—but will be in the future. In addition, many orders currently handled via telephone are expected to be shifted to this mode. Some customer's are also given access to certain XYZ hosts to support program development while the customer is awaiting product delivery, or during repairs.

To support just-in-time inventory management, XYZ is also connecting its internal inventory and purchasing systems to its suppliers online order-entry systems. These connections support online ordering of components and materials, purchase orders, and invoices. These connections use a combination of XYZ and supplier-owned facilities. A particularly interesting example is communication between XYZ and printed circuit board shops that manufacture components of XYZ products.

Another external connection supports warehousing and delivery operations. XYZ provides computer systems to its warehouse operators (i.e., storage and trucking companies).

Typically, these systems are connected to a single XYZ host to support coordination of warehousing and shipping operations.

XYZ currently supports most joint ventures by providing the other company with a computer manufactured by XYZ. Thus far there has not been intensive interaction between XYZ and the outside contractors during the development period so online connections have not been established. In the case of contract employees, the individuals are simply given accounts on the appropriate system and temporarily treated as internal employees. XYZ plans to use IONs to support several applications in the future—for example, software and document distribution to customers, order/entry and coordination with suppliers, financial transactions with banks and credit union, and more extensive customer access.

4.2.4 QRS Inc.

QRS Inc. is a smaller computer manufacturer which, like ABC and XYZ, uses computer-networks heavily for intra- and inter-site networks, as well as connections to external organizations. Like XYZ, QRS is also an Arpanet contractor. However, it is connected to the Arpanet in a different fashion. QRS has a leased-line connection to a nearby university which in turn is connected to the Arpanet. QRS is effectively a subnet of the university's local network and in this way appears to be directly connected to the Arpanet. The Arpanet connection is used for electronic mail in the same way as was discussed above for ABC and XYZ. In addition, file transfer capabilities are used to distribute new software for testing and program updates to joint venture participants, research collaborators, and in some cases customers that have Arpanet access (mostly universities and government laboratories). A separate dial-up gateway supports electronic mail and file transfer communication with subcontractors and customers. QRS plans to extend this service to software distribution, bug reports, bug fixes, etc.

QRS uses its internal network to interconnect its R&D, manufacturing, and training centers which are located across the country.

4.2.5 Policy Perspectives

Each of the organizations described has a different policy perspective which shapes its requirements for usage control. The following sections illustrate the flexible technical and administrative mechanisms needed to address these diverse perspectives.

MIT is most concerned with supporting information and resource sharing, internally as well as with outsiders. Most usage control requirements at MIT arise in order to meet externally imposed requirements, i.e., controlling transit onto the external networks. In contrast, ABC is very conscious of controlling flows and is prepared to sacrifice benefits of external communication as well as increased internal regulation in order to protect its boundaries. XYZ and QRS are somewhere in between MIT and ABC. On the one hand, they share ABC's need to protect proprietary information and facilities. However, at the same time, they are more aggressive in exploiting the technology and in incurring the associated risks. Moreover, they share some of MIT's intolerance for impeding internal and external communication.

4.3 *Usage Control—Introduction to Issues*

Usage control issues in IONs vary widely depending upon the technical and organizational characteristics of the interconnected facilities and institutions. Nevertheless, this section begins to generalize and characterize these issues in preparation for subsequent discussions of appropriate technical mechanisms. The focus of discussion is the usage control issues that are unique to IONs. This section sets the stage for Chapters 5 through 8.

4.3.1 Assumptions

The conceptual model presented in subsequent chapters is based on assumptions about the technology and its capabilities. In its simplest state, an ION resembles a simple database system or service bureau that is used by multiple organizations. In its more sophisticated state, an ION is a collection of heterogeneous networks of heterogeneous computers carrying out a diversity of tasks in a decentralized manner. When not otherwise specified, a message-based, request-response protocol is assumed. Although existing IONs range from real-time communications to periodic exchanges of bulk data, many interchange types can be modeled as loosely-coupled message-based transactions.

This discussion also assumes that within an organization's internal network there exists a set of explicit and implicit policies and procedures that are considered adequate for the intended environment, namely, the members of the organization.[26] Often, the purpose of such internal networks is to facilitate communication and access to shared resources. Therefore, although individual hosts or servers connected to the network frequently include a protection system to isolate users, many services are treated as internal utilities that have limited or no protection. Often users perceive protection mechanisms as making use more cumbersome with little compensating benefit. It is even less common for data communications and processing facilities, in particular, network transport or electronic mail, to include logging or accounting mechanisms. Therefore, when such interconnected internal systems are made accessible to outsiders, there may be no existing means of treating external users differently from internal users other than by preventing access altogether. In any case, it is fundamentally difficult to convert from an environment composed of networks and resources in which the default is open access to one in which the default is closed; and the difficulty is increased the greater is the decentralization of management control over the resources. In other words, when an organization's internal network is exposed for the first time via an ION gateway, explicit design effort is needed if resource boundaries[27] are to be preserved in their pre-interconnection state. Even if internal security is high, many organizations would prefer to support some regions in which flow is less inhibited than in other regions.

The design of mechanisms that support *articulation* of policies can be separated from certification of the mechanisms' security. For example, specifying what type of information is needed to support a particular usage control policy is separable from questions of who

[26] This assumption is necessary in order to isolate issues regarding inter-organization networks from networks in general.

[27] The term resource boundary refers to the dividing line between facilities and information that are owned, operated, and accessed internally, and those that are not.

provides the information and whether it is forgeable or trustworthy. This research emphasizes the former. The rationale for this emphasis is *not* that certifiable security is unimportant. Rather, what is most different about IONs from traditional intra-organization networks and systems is the need to articulate and support new policies.

Security per se is addressed in two respects: as a primary motivation for some types of usage control policies (e.g., access control), and as a design parameter of supporting mechanisms. With regard to policy motivation, a significant difference between access control requirements for an ION connection and more traditional requirements is the greater acceptability in IONs of *detection* of abuse as opposed to a priori prevention.[28] The ongoing relationships among ION participants typically are such that there is significant disincentive to abuse the ION facilities, in the presence of detection capabilities, due to resource dependency, legal contract, or cultural standards. With regard to the design of supporting mechanisms, the security issues of enforcement and certification are not qualitatively different than they are in the case of internal networks, although the perceived need for such enforcement may be greatly increased. The primary security issue that must be addressed anew in IONs is the difficulty of authenticating information in the absence of a single, mutually-trusted mechanism to mediate, settle disputes, and provide authentication-related services such as key distribution (see Chapter 7).

4.3.2 Usage Control Requirements

As described in section 4.1, usage control issues differ for IONs that support outsider access to applications that are used only for external interchange functions, and those that support outsider access to internal communication and resource sharing. We explained the difference in terms of the overlap of internal and external logical networks. The well-defined applications of the more narrow type of ION can imply greater reactivity of internal resources to external inputs due to more concrete automatic processing of external communications. At the same time, this defined quality can support greater, and more

[28] The accuracy of this statement varies with the nature of the service type supported, information or resources interchanged, and the perceived threat of malicious attacks.

centralized, control over the extent of reactivity than in the case of the more general type of ION. In the narrower case, system security issues are intensified, but usage control policies can be satisfied, for the most part, by adopting or enhancing system security without infringing upon internal operations. In contrast, more general IONs raise network and resource control issues that differ from traditional security requirements. Each participant may want to implement multiple logical networks, some strictly internal and some that cross organization boundaries. If traditional access controls are implemented within each resource in such a way that all users (both internal and external) encounter equal scrutiny, conflict may arise between internal and external requirements (e.g., tolerance and need for cost and performance overhead of security measures). Alternatively, controls can be implemented in cooperation with other resources on the network so that internal users are treated differently from external ones. The benefit of the latter approach depends upon the value placed on minimizing usage controls encountered by internal users within each organization. These requirements are discussed in chapter 5.

In general, the function of ION usage controls is twofold: (1) to isolate non-overlapping logical networks that share a common physical network, i.e., build walls and gatekeepers around each logical network; and (2) to maintain the boundaries between overlapping logical networks by implementing usage controls within those resources that belong to multiple logical networks (i.e., resources in an overlap between walled domains). The first function is the more straight-forward of the two and resides in the domain of traditional lower-level network security. [53, 44, 75] In other words, any device that is physically connected to resources outside of a single logical network is responsible for maintaining the boundaries of that logical network.

A resource that resides within a single logical network can do whatever filtering is desired for the entities in that logical network. But, if a resource resides in multiple logical networks that have different usage control requirements, the resource must be able to discriminate between members of each logical network. For example, a device may define a barrier for one logical network (i.e., no information is intended to flow into or out of the device unless the information is going to and coming from other devices in the same logical network) while acting as a forwarder or shared resource for a second logical network. In order to discriminate in its provisior of forwarding services, the shared device must distinguish between the different sets of entities that access it via the common physical network. Traditional protection mechanisms are adequate within a resource dedicated to external functions. But, if the resource is used also for internal functions, performance overheads, restricted information flows, and disincentives to resource sharing may not be tolerable, and new mechanisms are needed. Within the internal network of an ION participant, implementation of more flexible usage controls involves two functions, *differentiating* between internal and external users within the network as a whole, and using this information to *discriminate* in the provision of services. In other words, how can one implement multiple logical networks on a single physical network[29] in such a way that minimizes imposition on internal users. These issues are analyzed in chapters 5 through 8.

The usage control requirements that arise in reaction to (or anticipation of) interconnection depend on the nature of the interconnecting organizations (hereafter referred to as participants). Formally structured organizations that manage resources conservatively and have proprietary interests to protect are unlikely to allow changes in external resource-accessibility to occur readily, assuming they are aware of the change. Such organizations are more likely to refrain from interconnection (the ultimate form of usage control) unless or until usage control mechanisms can be implemented to maintain existing resource boundaries. Alternatively, such organizations may adopt new usage controls that impose on internal procedures (such as increased internal access control or accounting) in order to accommodate interconnection without effecting a change in resource boundaries. On the other hand, loosely structured organizations that have ill-defined proprietary interests, and that manage resources more loosely, will likely accept some changes in resource boundaries more; in fact, they may be less tolerant of impinging on internal communications than they are of increasing external accessibility. Within the research and development community (which I describe in section 4.2) examples of the former are many industrial labs, whereas

[29] The single physical network might itself be composed of multiple local and long haul network facilities. But for the sake of this discussion I will refer to the entire internal facility as a single network.

examples of the latter are many university labs. Because requirements vary among organizations that are interconnected to one another, ION usage control mechanisms must support a range of participant-defined policies, and different coexisting policies for participants at either end of a connection. *The goal of this research is to define usage control mechanisms that will permit interconnection in such a way as to mitigate undesired changes in both external resource-accessibility and internal usage controls.*

In summary, most requirements for usage control in IONs are requirements to prevent outside users from obtaining all the rights and privileges of entities that reside within the network. The *undesired* situation can be modeled as a logical ION encompassing the *entire* internal network of the a participant. To control the rights of outsiders, a mechanism is needed to discriminate against externally originated messages and connections. In these terms, the fundamental usage control requirement is to isolate logical networks from one another. First, an organization must be able to isolate the ION from strictly-internal logical networks. Second, if an organization is connected to more than one ION, it must be able to isolate the various IONs from one another, i.e., to control transit across the two networks. In addition to restricting traffic flows, isolation entails insulating the logical networks from each other's policies. The following sections illustrate these three points with the examples presented earlier.

4.3.3 Isolating Logical Networks

A single institution may connect to several special-purpose networks, where each network represents an interest group (e.g., CSNET, SCIENCENET, EDUCOM, Supplier-networks, etc.). Membership overlap among the communities results in overlapping logical networks (i.e., organizations belong to more than one of the interest groups). Similarly, when multiple user communities within a single organization share communications and computer-based resources, the logical networks that correspond to the various user communities overlap (i.e., users share an organization-wide network as well as computational resources such as name servers and gateways). Chapters 5 through 8 describes mechanisms for isolating logical networks. As an introduction, this section illustrates the need for such isolation using the four examples described earlier.

As described in section 4.2, MIT is connected to several outside organizations via an assortment of ION connections, and the internal network itself is composed of several distinct user and administrative communities. Each of these internal and inter-organization user communities can be viewed as a logical network. The logical networks overlap since some users and computer-based resources (hosts, peripherals, servers, etc.) are a part of more than one user community. The usage control requirements encountered in this environment are not particularly stringent since university's have a public orientation and do not have proprietary concerns. Nevertheless, limited resources, personal privacy, and policies imposed by external ION participants, call for some isolation of the logical networks. For example, due to usage policies imposed by Arpanet administrators, and limited gateway capacity, the educational network developed for project Athena at MIT needs to be isolated from the logical ION that spans the Arpanet. However, the educational network overlaps and is connected to the research network in the MIT Laboratory for Computer Science. Mechanisms are needed to isolate these overlapping user communities. Similarly, some MIT resources that sit on the MIT internal network are not to be accessed by entities outside the MIT community (e.g., clipping service, high quality printer, gateways to other networks, etc.); other resources are *intended* for access by a restricted set of non-MIT users as well. MIT would like to enforce these policies with respect to Multics and dial-up gateway users in particular since they are less homogeneous than Arpanet users. At the same time, MIT is concerned about imposing boundaries on intra-MIT communication and resource sharing. Subsequent chapters describe mechanisms that balance these internal and external needs.

Both ABC and XYZ have a large internal user population that includes many computer-based services in addition to person-to-person electronic mail and remote login; for example, file servers, print servers, name servers, etc. Several of these internal services can be invoked via electronic mail, i.e., users send commands and data to the servers inside of specially formatted electronic mail messages. ABC and XYZ would like to support person-to-person communication with people on the external R&D network, but does not want to support external access to internal servers. Moreover, ABC does not want *all* ABC employees to use the ION for external communication. ABC wants the gateway to be used only for research-related activities on account of the highly proprietary proprietary nature of other business operations. QRS's Arpanet gateway supports file transfer and remote login as well as electronic mail. It must prevent external access to all internal systems that contain proprietary or unreleased software and documentation. These requirements can be viewed as the need to isolate the strictly-internal network and the ION from one another; where ABC's and XYZ's research mailboxes are part of the ION while their servers and non-research mailboxes are not.

The other external connections of ABC, XYZ, and QRS can be described in similar terms—the facilities included in the ION form a logical network which needs to be isolated from strictly-internal facilities. ABC wants to restrict the west-coast subcontractor from accessing any ABC facilities other than project-related east-coast facilities. Similarly, ABC wants to restrict employees from accessing the accounts and resources belonging to customers of the ABC value added network service, and from abusing the airline carrier connection. XYZ also wants to restrict vendors from accessing most internal information and resources since these same vendors may be competitors in other markets. Similarly, XYZ wants to restrict customer access to product information and ordering, and warehouse access to shipment schedules. Similarly, as QRS increases it use of its dial-up message gateway it will need to restrict the allowable destinations of customer's messages. In addition, QRS wants to isolate the strictly-internal logical network from the internal facilities made accessible to customers during training courses. All of these are examples of the need to isolate overlapping logical networks.

4.3.4 Insulating Participant Policies

The above examples refer to isolation in terms of allowing or disallowing traffic flow. One higher level motivation for this usage control is the need to insulate the ION participants from one another's policies.

Assuming for now that policy is uniform within any single logical network, different institutions have different policies regarding facility use, sharing, and gateway access. Similarly, different, but interconnected, interest group IONs have different policies regarding access, billing, etc. In order for these different institutions and interest groups to support their respective policies without imposing on one another's, participants must be able both to share facilities internally and still conform with the controlled connection desired by the external interest group network. In some of the examples described above the motivation for isolating the logical networks was based on the need to insulate the networks from one another's policies; e.g., MIT's Arpanet connection. Moreover, mechanisms used must entail minimum modification of internal procedures. For example, if external network traffic is billed on a usage-sensitive basis, an organization should be able to limit or control outgoing traffic without necessarily implementing accounting internally. An example of this problem is found on the Arpanet which does not implement usage-based accounting but which has a gateway to an X.25 public packet-switched network, Telenet, which charges on a per-packet basis.

Later chapters will show how each organization can implement its own ION gateway(s) to enforce policies and meet constraints imposed by the various IONs and internal network that connects it.

4.3.5 Transit and the pairwise connection problem

When an organization is connected to more then one ION, the issue of *transit* arises. Namely, the organization itself, or members of either ION, may desire to prevent transit from one ION to the other via the common organization (see figure 4-3).

We can use the case of CSNET to illustrate this problem. CSNET [35] is a network linking computer organizations engaged in computer science and engineering research throughout the US, Canada, and Europe. Membership is open to any university, corporation, government agency, etc., engaged in computer science research or advanced development. CSNET provides electronic mail, gateways to other networks, and a database of CSNET entries. CSNET also provides, at higher cost and effort, login and file transfer. Although some CSNET members attempt to constrain the overlap of ION and internal logical networks by forwarding mail to and from authorized registered personnel only, most member institutions forward mail to any mailbox within the organization. If these any of

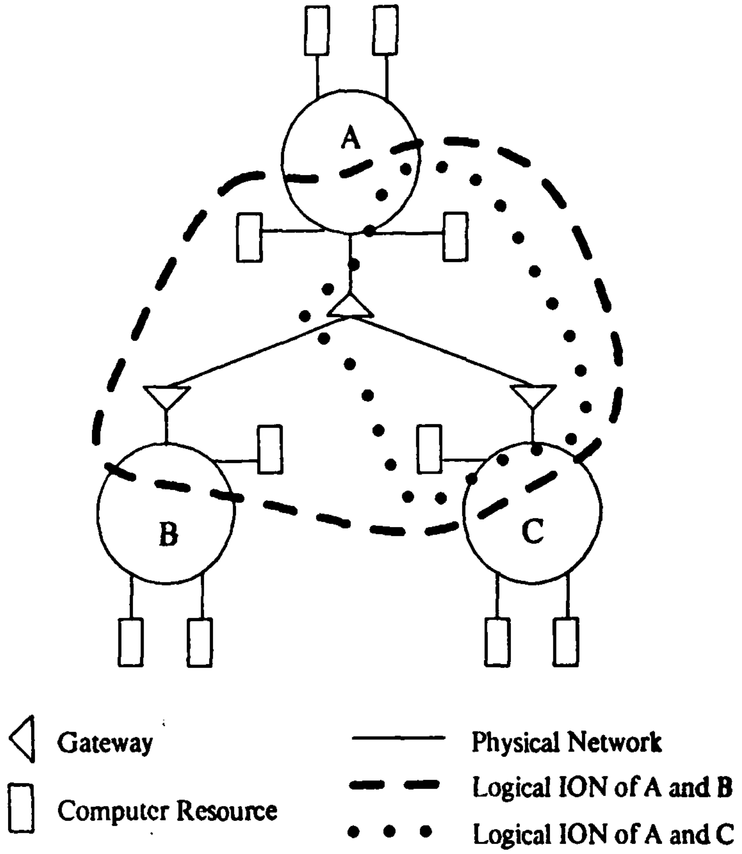


Figure 4-3: Network interconnection *without* controlled transit.

these mailboxes are forwarders and gateways, the logical ION completely encompasses the participants' internal electronic mail networks, as well as other networks to which the internal networks connect.

This cascading of networks, in which the logical ION encompasses the participants' internal nets including gateways, raises the problem of transit. For example, in order to control costs, preserve desired levels of service for CSNET members, and preserve the utility of paying CSNET membership dues, CSNET wishes to limit the amount of traffic that is originated by non-CSNET members and carried over CSNET facilities. On the other hand, it is not desirable to prevent or prohibit all forwarding because the value of CSNET to its members is proportional, in some sense, to the number of institutions that are accessible via CSNET; i.e., gateways are desirable. At a minimum, CSNET does want to prohibit communication between non-CSNET members over CSNET facilities, called transit, since no CSNET member benefits from such use.

Non-CSNET members gain access to CSNET via forwarding by CSNET members. To control undesirable forwarding, an appropriate tactic in this particular (research) community is to produce incentives for CSNET member hosts to not forward non-member traffic. One ad hoc mechanism is to state the policy and rely on peer pressure, since forwarding is detected when a message is read by the recipient. But this would have little effect on the transit problem since the non-member recipients and sources are not likely to be affected by peer pressure of this sort. A less ad hoc mechanism is to charge per message or set upper bounds on usage for each CSNET member host. This would force users to address the problem of implementing controllable forwarders and gateways so that they forward mail only from authorized machines within the member institutions but do not forward mail from non-CSNET members. In order to comply with such a policy the member hosts need a mechanism to tag and filter non-local from local traffic, i.e., to control the overlap between internal and ION logical networks. The following chapters describe usage control and interconnection mechanisms to address these requirements. One caveat regarding this approach is the tendency to discourage all forwarding of non-CSNET traffic due to cost, difficulty, and imposition of implementing flexible usage controls; even when it makes economic sense on a system-wide basis for CSNET members to receive and send some off-net mail via one another.

Note that CSNET differs from the two research networks, BITNET [22] and UUCP/USENET [46], in this regard. Neither BITNET nor UUCP/USENET charge for services and therefore do not face the burden of protecting their investment as a network.[30] But, although membership is free and largely unrestricted in BITNET and USENET, individual members may want to limit their forwarding burdens due to limited resources, both cpu time and leased line or dial-up capacity. Therefore, to varying degrees, individual members of the networks share CSNET's interest in controlling transit.

A somewhat different perspective on the transit issue is the *pairwise connection problem.* This issue arises when one organization interconnects to two other organizations that do not intend or desire to be connected to one another. Without usage controls in gateways that delineate and isolate logical networks from one another, such interconnection creates a path between the two unrelated organizations by default. One :xample is ABC's connection to Arpanet and UUCP/USENET by virtue of its CSNET connection, whereas the firm has not connected to ARPANET or UUCP/USENET directly. The question that arises is whether an organization must obtain the approval of the other organizations to which it is connected bc fore establishing new connections. If the answer is yes, an impractical situation arises in which pairwise agreements are no longer adequate. Another example is the connection between MIT Artificial Intelligence laboratory and a local company in support of joint research. The AI lab is in turn connected to the rest of the MIT networks and to the Arpanet, and the local company is in turn connected to a number of its customers. Due to the nature of the network interconnections, the logical ION encompasses all internal facilities, and therefore by default a connection exists between all of MIT's networked resources and this company's customers.

ABC and XYZ have additional concerns about transit because their internal networks span

[30] Joining BITNET involves acquiring a leased line to a nearby BITNET member and thereby picking up one's portion of the costs directly. BITNET communications software (RSCS) is an IBM product and is available for other types of machines at a small charge. BITNET services, i.e., BITSERVE, are developed in a cooperative manner with a large amount of direction coming from its birthplaces, City University of New York and Yale University. Similarly, an institution joins UUCP/USENET not by paying a fee or signing up with any central coordinator, but rather by finding an existing member to connect to and paying the telephone charges to transfer the traffic to that connected host. Again, the communications software is part of standard UNIX software and is available for other types of machines at little charge, and mailing list/bulletin board services are maintained in a distributed, cooperative manner.

national boundaries. The telecommunication administrations in Japan and some European countries have strict regulations about who can provide value added communication services. ABC's and XYZ's connections to customers and other types of organizations potentially supports communication among these external entities. Such transit could call into question the legality of the firms' internal networks in countries with more restrictive policies. For this reason, both ABC and XYZ must be able to control transit communication from external entities to other external entities via the internal facilities.

4.4 Related Work

The two most relevant areas of technical research are network interconnection and computer system security. Sample papers in each area are discussed in relation to the proposed research.

4.4.1 Network Interconnection

Over the past ten years research and development of computer network interconnection has advanced significantly [69, 13, 10]. Although interconnection of computer networks across organization boundaries raises issues discussed in this literature, the priorities differ. In particular, the emphasis of most network interconnection literature is on connectivity, performance, and transparency. These issues typically are of secondary concern in IONs. On the other hand, usage control is of primary importance in IONs but is not emphasized in most internetworking literature. In this section I review aspects of the internetworking literature that are significant to ION development: level of interconnect, gateway configuration, naming, usage control, standards for message transfer systems, and packet network interconnection.

A gateway can be configured as a front-end, protocol translator, host on two networks, or formal gateway with standardized end-to-end protocols [13]. Different configurations imply different levels of interconnection and different levels of homogeneity. For example, interconnection at the internet level[31] implies that the internal address structures of the

[31] *Internet* is the lowest level of the Department of Defense protocols employed on the Arpanet.

networks are uniform. Alternatively, if protocols terminate at the gateway, the address structure can be different in each network so long as a mapping is defined. If the gateway acts as a host on two (or more) networks it must encapsulate forwarded messages, be they packets or electronic mail. Finally, the feasibility of protocol-translation gateways depends upon the similarity between the design of the network protocols on either side. For example, the University College London (UCL) gateway described below interconnects at the transport level, and SNA-to-SNA[32] gateways do name mapping at the path control level [6].

Problems of naming in distributed systems are exacerbated by the crossing of organization boundaries. Sunshine proposes hierarchical naming to reduce the size of network routing tables. However, such a centralized approach is not applicable in many ION environments where the absence of a centralized rule-setter is an impediment to standardization. The SNA interconnect architecture provides name-mapping facilities in the gateway. This allows SNA networks with overlapping name-spaces to interconnect without affecting internal naming. This latter design addresses the concern for isolation desired in many ION activities. Sirbu and Sutherland also discuss naming issues in directory systems for interconnected, heterogeneous systems [64]. Related to naming are issues of addressing and routing. Source routing, in which the source of a message specifies the complete route to the destination,[33] appears particularly suited to environments in which access control and restrictive routing are desired [68]. In addition, source routing reduces the need for global agreement on network names.

Although many research papers mention the need for authentication, access control, and accounting, only a few systems actually implement such controls. Two exceptions are the facilities developed at University College, London (UCL) and Cambridge University which implement access control mechanisms; these systems are described in sections 6 and 8

[32] SNA stands for System Network Architecture, IBM's network protocol and architecture.

[33] This is in contrast to hop-by-hop routing in which each gateway or node uses a routing table to determine the appropriate action given a packet's destination.

[11, 81, 18]. Most existing access control schemes for interconnected networks maintain a table of authorized users in the gateway. Many rely on a trusted entity to authorize or authenticate users and maintain the access table [13, 81, 11]. Independent of how authorization is administered, existing access-table schemes support a flat protection space, in which a user has either full access to the network, or no access at all.

Symbolics' Chaosnet design includes a secure subnet feature which classifies subnets as either trusted or untrusted. Servers on one subnet can use this information, held in address tables, to discriminate in providing service. Chaosnet also distinguishes between communities in an internet via the definition of Namespaces. A closely-coupled community shares a single namespace, because within a namespace, routing updates and other control information are broadcast frequently to all hosts and servers. Across different namespaces, detailed information about the internals of other namespaces are more transparent. [70]

Most network interconnection designs ignore accounting mechanisms. Often, implementors assume that each network will charge for gateway traffic as it does for host traffic [69]. However, there may not be an accounting mechanism in place for charging internal hosts. Therefore, an ION gateway that admits outsiders for the first time might introduce the first demand for internal accounting. Even if internal accounting does exist, it may be necessary to account differently for incoming and outgoing traffic than for strictly-internal traffic. Consequently, there would still be a need to modify internal accounting to discriminate between internally and externally generated traffic. Interconnection of public networks raises a specific accounting requirement, revenue sharing. Interconnections between public carriers are more formal and static than are interconnections between private networks. Many solutions can be borrowed from interconnection of public switched telephone networks; however, these solutions do not satisfy the dynamic requirements encountered in private interconnections.

Several international standards activities address network interconnection. The CCITT has developed a standard for interconnection of public packet switched networks, X.75. The standard is not intended for use by private networks [13, 27] and it does not specify security features. NBS, IFIP, and CCITT are developing message transfer system standards for mail and document interchange [63]. On a similar vein, IBM has developed a Document Interchange Architecture [62]. Meanwhile, the American National Standards Institute, based on earlier designs by the Transportation Data Coordinating Committee [71], developed a Business Data Interchange standard which specifies formats for invoices, purchase orders, and other business forms [21]. These application-level standards will facilitate the ION process in general and could make adoption of standardized ION protocols, as opposed to specialized ones, more likely.

4.4.2 System Security

Several environmental factors distinguish ION usage control requirements from typical system-security requirements. First, a single trusted third party often is not available across multiple organizations. Therefore, mechanisms must accommodate negotiation among mutually suspicious entities, or at least reliance on the minimum common mechanism. Second, there is no single point of mediation for all operations in a network. This issue is less troublesome for IONs that are dedicated to a single interchange function and therefore have a single application interface which can act as a central mediation point. The third difference is the relative difficulty of articulating usage control policies that both reflect organization requirements and can be mapped into technical mechanisms. Certification of these mechanisms as *secure* is a somewhat separable concern. Finally, investment in a priori security enforcement can be traded off for reliance on a posteriori detection in conjunction with legal contracts. In this section I discuss models of system security that contribute to the development of a usage control model for IONs. I conclude with a discussion of network security. The usage control requirements encountered in IONs are elaborated in the next chapter 5.

Articulation and models of computer system security are discussed in a large body of literature most of which will not be discussed here. Summarizing the literature, Saltzer and Schroeder [60] classify security violations as unauthorized information release, information modification, or denial of use. For both intra- and inter-organization networks, these violations should be extended to protection of resources in addition to information. Saltzer and Schroeder also categorize functional levels of information protection: unprotected, all or nothing, controlled sharing, user-programmed sharing controls, and strings on information (information flow control). Most existing IONs reside at level one or two of this classification. This research effort will define mechanisms to support levels three through five.

Saltzer and Schroeder also discuss two types of sharing controls, list-oriented and ticket-oriented. *List-oriented* refers to protection systems in which each object has associated with it a list of authorized users. *Ticket-oriented* refers to systems in which each user maintains an unforgeable bit pattern (i.e., ticket) for each object that it is authorized to access. This dichotomy applies to usage control mechanisms in an ION environment. For example, the Cambridge system mentioned above is ticket oriented, whereas the UCL system is list oriented. A proposal by Mracek [42] uses a combined approach analogous to the international practice of issuing and checking visas (see chapter 6).

In information systems in general, policy is best separated from mechanism. Policies are high-level statements indicating which resources are to be secured, from which kinds of usage, and by which users. Mechanisms are the automatic procedures and administrative controls used to implement the policies. Different organizations require different policies. Therefore, an information systems mechanism should support a range of policies without extensive alteration. In datab.se systems that support a diversity of data, users, and applications simultaneously, articulation of security requirements is particularly complex. At the same time, the more complex the underlying mechanisms are, the harder it is to achieve performance and certifiability. Consequently, because mechanisms define the range of supportable policies, there is a tradeoff between flexibility of policy on the one hand, and performance and certifiability of mechanism on the other.

Control policies can be classified as either *discretionary* or *non-discretionary.* Discretionary policies allow the owners, and sometimes the users, of a resource to specify who may have access to that resource. Non-discretionary policies enforce restrictions on access that are beyond the control of resource users, or even owners. A typical mechanism used to support discretionary policies is an owner-specified access list attached to a file. In contrast, a typical mechanism used to support non-discretionary policies monitors all read and write actions and disallows flows between users or files of different classification levels; where each user, file, and resource in a system is assigned a classification level. The latter is called non-discretionary because it disallows flows independent of user or owner defined restrictions such as access lists.

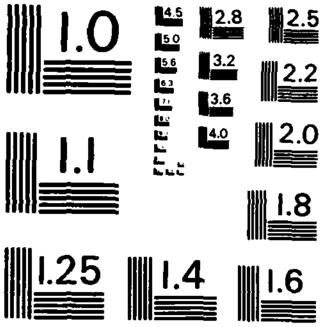
Numerous designs for network security employ encryption. Popek and Kline [53] discuss the use of encryption for protection of data and authentication. Voydock and Kent [75] also discuss requirements and mechanisms for protecting data in networks, with emphasis on encryption as the tool. Needham and Schroeder [44] developed specific protocols for using encryption to authenticate users (see chapter 7). In a paper design, Rauthier [58] extended their work to accommodate an ION environment where there is no single, trusted third party. Encryption has been used to control access to the data transport facility in a local network by verifying sequences of packets at the link layer of the communication protocol [7]. Gifford [24] designed a method of cryptographic sealing to protect and self-authenticate objects. The cryptographic keys represent desired access and ownership policies. Karger describes a proxy login scheme whereby hosts differentiate non-local users from local users and access control is based on user ID as well as the network path via which a user accesses the host [33].

The systems that are accessible via IONs are equipped with a wide range of security and access control mechanisms. These mechanisms were designed primarily for use by persons within the organization. As discussed in the following section and chapters, the existence of connections to the outside world changes the ground-rules on which the original design and implementation decisions were made. Building on the system security literature, I describe and analyze access control policies and mechanisms suitable for this environment.

4.5 Roadmap

This introduction to IONs is followed by descriptions of how to address these usage control requirements using non-discretionary controls (Chapter 5). Subsequent chapters analyze the implications of the proposed approach for network interconnection (Chapter 6), authentication (Chapter 7), and implementation (Chapter 8).

AD-A160 855 ACCESS TO INTER-ORGANIZATION COMPUTER NETWORKS(U) 2/3
MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER
SCIENCE D L ESTRIN AUG 85 MIT/LCS/TR-345
UNCLASSIFIED N00014-83-K-0125 F/G 5/1 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Chapter Five

Non-Discretionary Controls for Inter-Organization Networks

This chapter describes a conceptual model for implementing usage controls in Inter-Organization Networks (IONs). After reviewing usage control requirements in networks that cross organization boundaries, a proposal is outlined for adapting traditional, non-discretionary controls to support usage control in IONs.[34]

5.1 Summary of ION Usage Control Requirements

When two or more distinct organizations interconnect their internal computer networks to facilitate inter-organization interchange, they form an Inter-Organization Network. The interchange may be person-to-person communication via electronic mail; exchange of cad/cam data, software modules, or documents via file transfer; input to an order-entry or accounting system via a database query and update protocol; or use of shared computational resources via an asynchronous message protocol or remote login. In most inter-organization arrangements, the set of resources that an organization wants to make accessible to outsiders is significantly smaller than the set of resources that it wants to remain strictly-internal (i.e., accessible to employees of the organization only). In addition, because the potential user is a person (or machine) outside the boundaries of the organization, the damage associated with undesired use can be high. Because of these characteristics, IONs have unique usage-control requirements.

Unlike traditional simple security requirements, the goal is not simply to prohibit access by outsiders; some outside access is explicitly desired. The goal is to support access to certain

[34] This chapter is based on a paper of the same title published in the Proceedings of the 1985 Symposium on Security and Privacy, Oakland, CA April 22-24, 1985. IEEE Computer Society.

4.5 Roadmap

This introduction to IONs is followed by descriptions of how to address these usage control requirements using non-discretionary controls (Chapter 5). Subsequent chapters analyze the implications of the proposed approach for network interconnection (Chapter 6), authentication (Chapter 7), and implementation (Chapter 8).

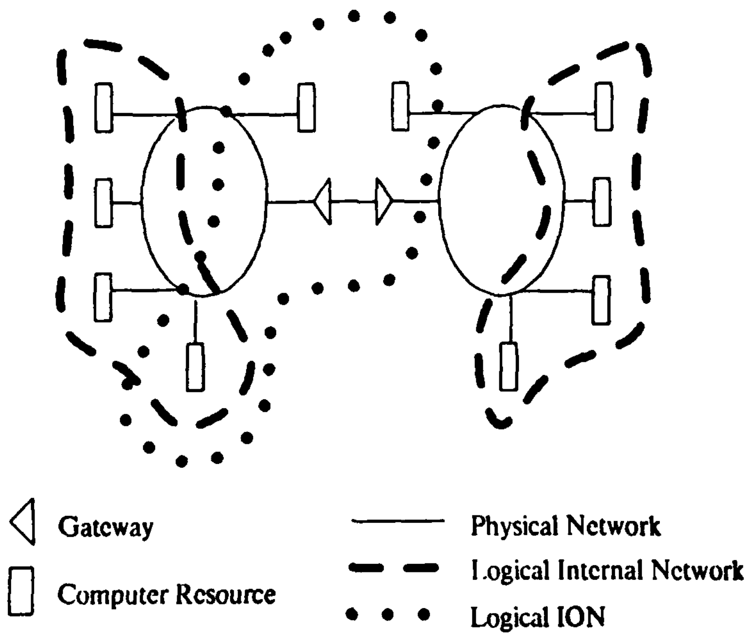
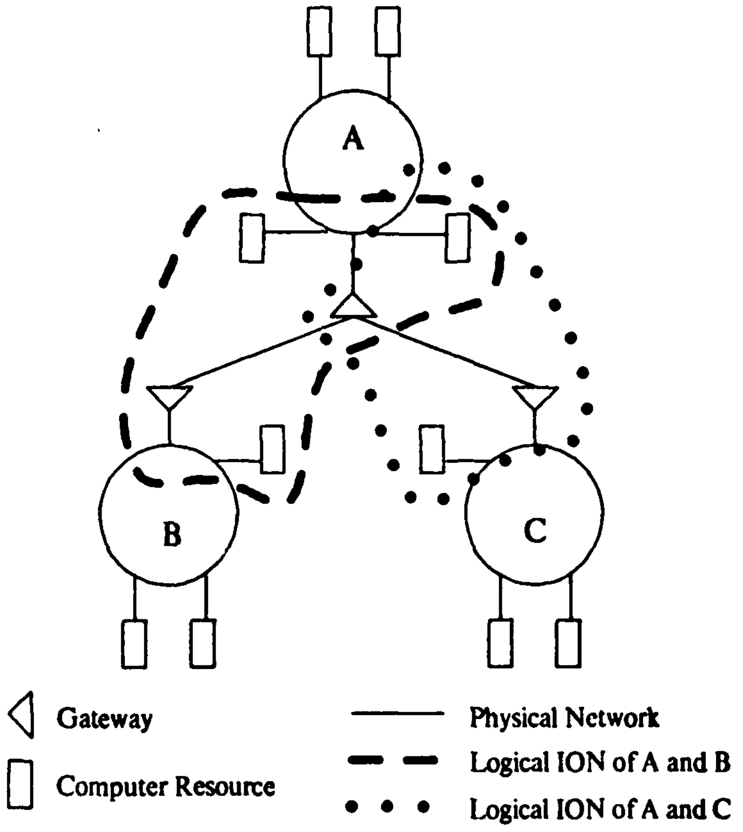


Figure 5-1:Overlapping Logical Networks: The ION shares physical resources with the two organizations' internal networks. However, at the logical level, the ION is isolated from the strictly-internal facilities.

machines, services, and processes, while preventing access to all other internal facilities. In addition, because the function of the internal network predates and dominates that of the ION, interconnection must not interfere with internal operations. Therefore, it is not acceptable that ION facilities be physically isolated from all strictly-internal resources for this would interfere with internal communications and resource access. We want to implement *logical networks* that can be isolated from one another yet share physical resources (see figure 5-1).[35] Similarly, when two organizations interconnect, it may be inappropriate to impose a connection between the other organizations to which each was interconnected previously. In other words, the new ION may overlap physically with the existing IONs, but it must not form a transit path between those organizations that desire to remain isolated from one another (such as *B* and *C* in figure 5-2).

[35] The term *logical network* refers to a collection of computational resources and applications that communicate with one another. Logical networks operate on top of physical networks which are composed of communication links and switches.



Figurc 5-2:Overlapping IONs: The ION between A and B shares physical resources with the ION between B and C. However, at a logical level the two IONs are not connected to one another, i.e., B cannot communicate with C via A.

5.2 Constraints on the Solution

ION participants typically want to make only a subset of their internal resources accessible to outsiders; and in most cases, the default condition for external access is *no* access. There are two *obvious* ways to support access to certain resources while preventing access to all other resources. The first is to physically isolate those resources that are to be made externally accessible from those that are to remain strictly internal (see figure 5-3 (a)). The second is to increase access controls on *all* systems on the *internal* network so that no system allows external access unless it is explicitly approved to do so (see figure5-3 (b)). Both solutions support controlled interconnection, however, the constraints described below make both approaches unacceptable as general solutions.

First, in most cases, the function of the internal network predates and dominates that of the inter-organization network. Moreover, typically, the purpose of an internal network is to facilitate communication and resource sharing. Increased internal usage controls that are tailored to restrict outsiders may interfere with this objective. In addition, the administration of most networks is intentionally decentralized. Consequently, it is very difficult to assure conformance with new policies such as accessibility of internal resources to outsiders. Internal networks also grow incrementally by adding connections to other internal networks as well as single machines. Therefore, it is hard to determine whether such additions introduce resources into the internal network that do not conform to network-wide policy. Finally, in order for resource owners or users to enforce a security policy they must be educated as to its purpose and operation. Educating all resource owners and users in a decentralized network is hard to accomplish once, let alone every time an external link is established.

A solution based on physical isolation may be acceptable for some special cases, but, given these constraints, it is not a general solution because it imposes excessive restrictions on communication and integration between externally-accessible and internal systems. For example, XYZ Inc. provides MIT with access to some research facilities. Physical isolation would imply that these internal research facilities could not be integrated with the XYZ's internal development system. Similarly, if to protect itself from customers a supplier had to physically isolate customer-accessible online order-entry system from the internal inventory system, the supplier would forego one of the main benefits of online order-entry—the potential for integration of order processing and inventory control. The second obvious solution described above is also unacceptable as a general solution. The constraints summarized imply that strictly-internal resources which have nothing to do with the interconnection not be required to take any action such as modifying security mechanisms, in order to be protected from external access. A requirement to take explicit security action when an external link is added violates several of the constraints listed. First, the access

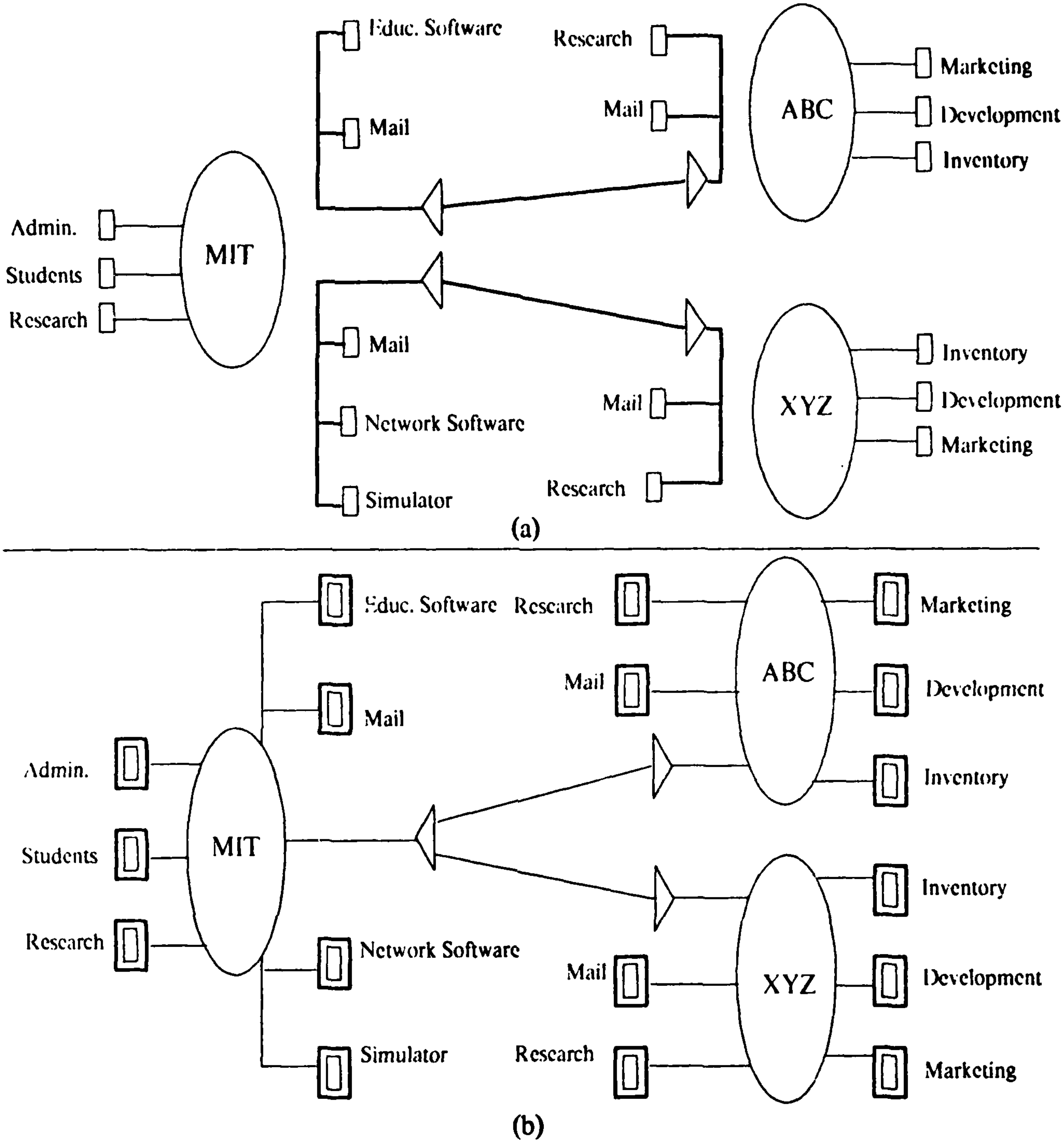


Figure 5-3:(a) Physically isolated logical networks. (b) Modified access controls on all internal systems.

controls implemented may impose on internal communication and resource sharing. Second, modifying all internal systems is an exceedingly costly proposition in most reasonable sized organizations. Finally, even if such a cost could be justified, given decentralized management of internal facilities and/or interconnections, it is not feasible to assure conformance of all systems with new interconnections.

If internal security levels are high, all users have limited capabilities, and therefore the extent of damage that would result from treating external users as internal is contained. Nevertheless, there remain two reasons why internal security measures must be augmented in the presence of inter-organization interconnection, even if the existing internal access control measures are conservative and non-discretionary, to begin with. First, in most environments, internal needs are best met by open internal access to *some* shared internal resources which nevertheless should not be accessed by outsiders; in the same way that small office supplies often are freely accessible to employees. Second, the design of a security mechanism depends critically upon an accurate model of the user population. External connections that are implemented incrementally under decentralized management may undermine the assumptions on which some internal security mechanisms were developed previously. Requirements for increased internal security raises issues for divisions within a single organization. Divisions that wish to communicate and share resources but that wish to remain autonomous and control access to local resources encounter tensions between connectivity and autonomy or liability that are analogous to the general ION issues described here.

In summary, only the administrators of the external link (i.e., the ION gateway) and the internal resources that are made explicitly accessible should be required to take security action; in accordance with organization-wide policies or guidelines, perhaps. Owners of all other internal resources should be assured that their facilities are not accessible to outsiders. In other words, the management of a strictly-internal resource should not have to rely on its own discretionary action for restriction of external access to its facilities. *This requirement suggests the use of non-discretionary access controls to isolate strictly-internal resources and networks from the ION without relying on the discretion or explicit action of strictly-internal resource owners.*

5.3 Non-Discretionary Controls

There are three essential differences between the non-discretionary access controls called for here, and those traditionally employed in military security systems [32, 36]. First, in the case of military systems the most common use of non-discretionary controls is to restrict the flow of information from higher classification levels to lower ones.[36] [5] In IONs, of equal or greater concern is preventing outsiders from *invoking* proprietary, expensive, or scarce resources that are supposed to be strictly internal. In traditional terms, control of invocation concerns unauthorized disclosure, modification, and denial of *resources*; whereas, information flow control concerns only unauthorized disclosure of *information.* Although many commercial and government institutions are extremely concerned about the outgoing flow of information, in this paper we focus on invocation control because it has received far less attention in the past.[37]

Second, the non-discretionary invocation controls that have been developed are designed to protect the integrity of the *invoker*, not the *invoked* [8, 37]. For example, the integrity rating of a program indicates the level of assurance that the program does not contain any trojan horses. Based on these ratings, the simple integrity policy allows a user to invoke programs of *equal-or-greater* integrity only.[38] In contrast, we are trying to protect each ION participant in its role as service *provider*, not *user*. To do so, we must protect the provider from unauthorized disclosure, modification, and denial of resources. Therefore, we want a policy that prevents a program from being invoked by a user that does not have an adequate

[36] No read up, by a lower classification level of a higher one; and no write down, by a higher classification of a lower one.

[37] One form of information control that we do address explicitly is information flow that is not mediated by an employee of the organization (i.e., extraction). Such flows require invocation of a file transfer or database or other computer-based service by an external entity and therefore are covered by invocation controls. IONs also raise concerns about the outgoing flow of information that *is* mediated by employees. For example, automatic distribution lists remove direct employee discretion from the process of generating outgoing mail. In addition, organizations often are concerned about excessive dependency on resources that are not controlled by the organization itself. However, these concerns are more traditional in nature and can build on traditional mechanisms. Extension of the mechanisms described here to information flow control is disc ıssed briefly in section 5.5 of this paper.

[38] The simple integrity policy is described as the mathematical dual of the basic security policy by Biba in [8].

integrity rating. The invocation policy should allow a user to invoke services of *equal-or-less* integrity only. Rotenberg [57] was also concerned with protecting information providers but did so only in the form of controlling information flow, i.e., unauthorized disclosure of information. He assumed that all services necessarily returned information, and that information flow controls would prevent the returning of information to unauthorized users. In current-day network environments there exist facilities that do not necessarily return information or that do so only after the resources have been expended or an irreversible action has been taken (e.g., gateways, print servers, robotic devices, order-entry systems.) In this environment, control of invocation is needed in order to protect the owners of such services. A related issue is that in the communication applications addressed here the distinction between object and subject is not meaningful because both participants in a communication take on both roles. Consequently the distinction between clearance and classification is not useful.

Most systems that enforce non-discretionary policies enforce confinement between categories of information. In other words, information can flow from a source to a destination only if the destinations category set contains *all* of the elements contained in the source's category set. The third distinction between traditional non-discretionary controls and those proposed here is that organizations would like to support *overlapping* logical networks (see figure 5-1). In order to do so, the non-discretionary controls enforced on network communications should implement a relaxed rule, namely, that information can flow from a source to a destination so long as the source's and destination's category sets *overlap*, i.e., have a non-empty intersection. As is described later, this intersect restriction on network communications would then be complemented with traditional system-level controls in those systems made accessible. These issues are summarized in table 5-4.

Based on these characteristics and requirements, we suggest that special network entry points, *ION gateways*, implement non-discretionary invocation controls. ION gateways are logical gateways that mediate and control the forwarding of messages from outsiders into

Traditional	*ION*
* Information Flow	* Invocation
* Protect Integrity of Invoker (equal-or-greater rule)	* Protect Invoked (equal-or-less rule)
* Sensitivity Levels	* Category Sets
* Subjects and Objects	* Communicating entities
* Enforce Confinement	* Overlapping category sets

Figure 5-4: Comparison of traditional non-discretionary controls and the requirements encountered in IONs.

the internal network.[39] Each organization operates its own ION gateway. Therefore communication between any two ION participants involves two ION gateways. In addition, ION facilities which can be communicated with by outsiders must implement discretionary or non-discretionary controls to protect other non-ION resources. Finally, because organizations communicate with multiple external organizations, and these inter-organization relationships are not hierarchically related to one another, the access rights should be based on category sets (compartments) and not sensitivity levels.

The participating organizations can tailor the strength of the gateway's implementation to suit their security requirements. These requirements will vary with the value of the online information and resources, as well as the nature of the inter-organization relationships. One way to formulate the general requirement is to require that the level of monitorability and accountability equal that of telephone and paper communication.

In the following sections we describe two example IONs and discuss how non-discretionary controls can be implemented in the ION gateways and ION facilities without modification to strictly-internal facilities.

[39] ION gateways may be composed of multiple, physically distributed components, e.g., a packet forwarder, policy filter, authentication server.

5.4 Examples

The following two examples illustrate how non-discretionary invocation controls could be used to protect the resources of interconnected organizations. The first example is from the perspective of MIT The second is from the perspective of ABC Inc. These examples are *representative* of existing activities. However, the details have been changed somewhat to illustrate several points in a single example; therefore the examples are *hypothetical*, not actual, cases.

MIT has connected some of its internal computer facilities to those of two of its industrial sponsors, ABC and XYZ who happen to be competitors of one another (see figure 5-5). The connections are intended to support exchange of software modules, access to some unique computational resources, and electronic mail. ABC has access to a host on which an MIT research group is developing educational software. XYZ has access to a different host on which another MIT research group is developing network software. XYZ also has access to a design simulation program developed by yet another group of researchers at MIT. In addition, both ABC and XYZ have access to electronic mail communications. For the sake of this example, we assume that both ABC and XYZ invoke the various services (i.e., file servers for software distribution, simulator, and mail distribution) by sending appropriately-formatted messages through the gateway, and the servers return the requested data via the same gateway to the requesting organization. Aside from these ION resources, MIT has other strictly-internal computer-based facilities: administration, student accounts, other research projects, gateways to other networks, etc.

In this example there are *five* logical networks that need to be isolated from one another; where logical network refers to a set of computer resources that are intended to communication and interwork. The two logical IONs are shows in figure 5-5, one between MIT and ABC, and the second between MIT and XYZ. In addition, each of the three organizations has a logical internal network which each organization should be able to isolate from the IONs. Note that there is no logical ION between ABC and XYZ because none of their facilities are intended to communicate or interwork. In order to isolate ION from strictly-internal facilities, and the XYZ ION facilities from the ABC ION facilities, MIT can implement the following controls:

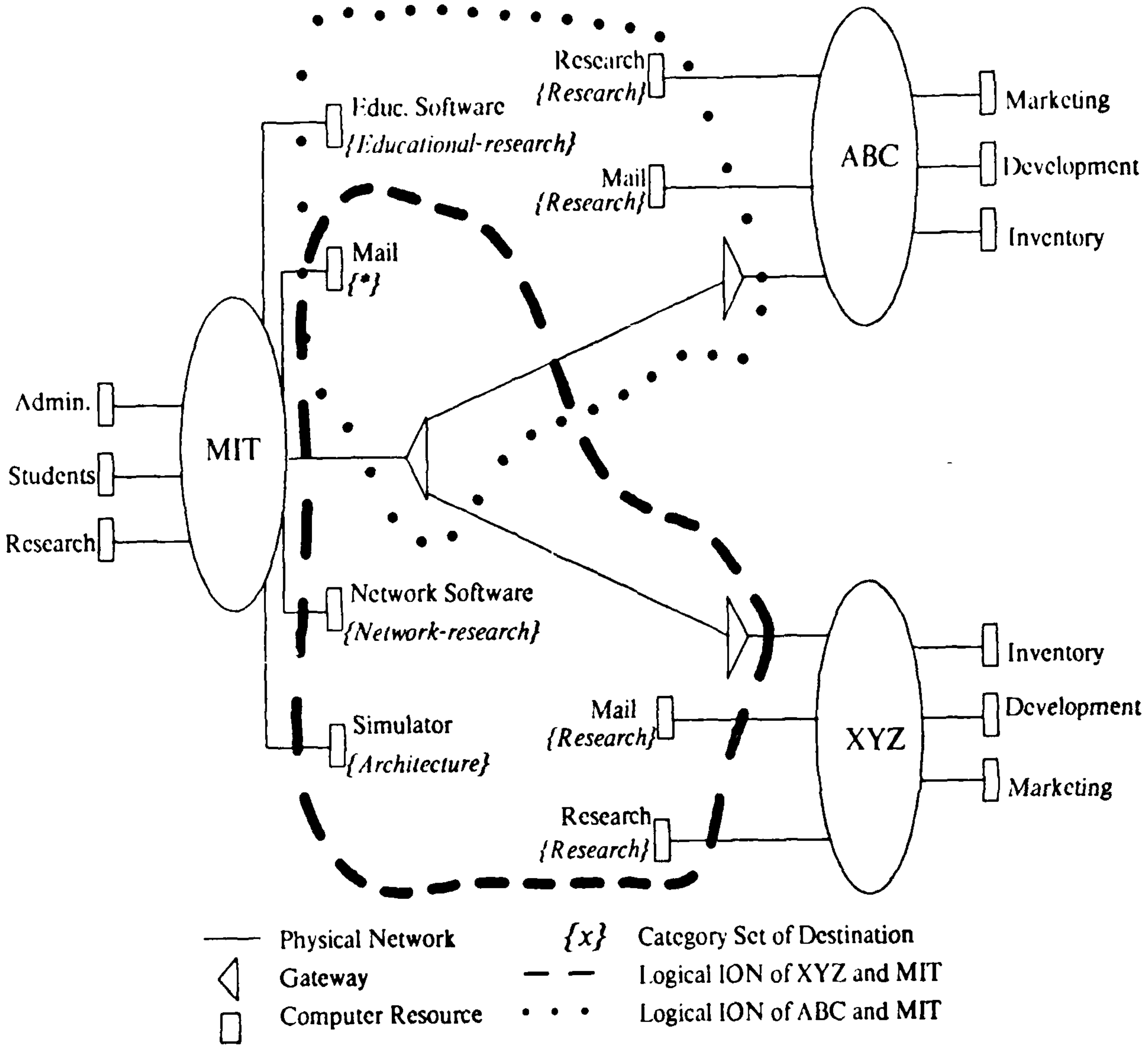


Figure 5-5: Example of an Inter-Organization Network: One ION exists between MIT and XYZ and another ION exists between MIT and ABC. Both IONs overlap physically yet are isolated logically from the internal networks of the three organizations.

1. Implement a single ION gateway and prohibit direct connection of all internal machines to outside organizations. Equip the gateway with an authentication mechanism to certify the source of each message.

2. Assign appropriate category sets to each of the ION facilities, and no category sets to strictly-internal ones. MIT assigns the category set {Educational-research} to the host used for development of educational software, the set {Network-research} to the host used for development of network software, {Architecture} to the design simulator, {*} to its electronic mail system to indicate all. and {Strictly Internal} to all other internal systems. See figure 5-5. If an internal facility is not registered at all the gateway assumes that it is not accessible via this entry exit point. The category information is assigned to internal resources but the information is maintained in MIT's gateway, see figure 5-6.

3. The ION gateway checks the category set of the source, $\{Ci\}_s$, and of the destination, $\{Ci\}_d$, of each message and forwards the message to the intended destination *If and only If* $\{Ci\}_s$ *Intersect* $\{Ci\}_d$ does not equal nullset, {} (referred to as the *Intersect* rule).

4. Equip the internal ION facilities (software distribution servers, electronic mail server, and design simulator) with discretionary or non-discretionary controls to enforce application-specific controls (e.g., restrictions based upon the dollar amount of a purchase order or the filename of a cad/cam file request), isolate non-ION files and processes, and prevent transit between the ABC ION and the XYZ IONs.

Similarly, ABC and XYZ each label their own research hosts and inventory systems with the category set {MIT} only, and implement gateways with message authentication and the *Intersect* rule. *Note that each organization assigns category labels to incoming messages for interpretation by its own internal facilities. Therefore, although naming must be consistent within each organization, it need not be consistent throughout the ION as a whole.*

Our second example is from the perspective of one of the computer manufacturers, ABC, that connects its internal network to a nation-wide network of computer research and development (R&D) laboratories. Informal, person-to-person research communication transpires with a large subset of all the organizations on the network. In addition, there are two universities, MIT and Northeastern (NU), with which ABC is conducting two separate joint studies, one with each of its major research divisions. In conjunction with these

Entity	Category Set
ABC	{Educuational-computing}
XYZ	{Network-research, Architecture}
Educ. Software	{Educuational-computing}
Network Software	{Network-Research}
Mail	{*}
Simulator	{Architecture}

Figure 5-6:MIT's gateway table containing category set information.

studies, ABC supports some file transfer and remote job entry with these two organizations only. To support such tailored connections, ABC assigns the category sets {MIT} and {NU} to Division 1's and 2's respective R&D systems, and all three organizations assign the wild-card category set {*} to their respective mail servers. See figure 5-7. The mail server is thereby made accessible to all network members, whereas the joint-development facilities are made accessible to the select parties only. As described above, the gateway authenticates messages, implements the *Intersect* rule, and ION facilities are equipped with discretionary or non-discretionary controls to isolate non-ION processes and files. For the most part, the two universities, MIT and NU, are not concerned about protecting internal resources. One exception is that MIT has another gateway to a special network, Bitnet, which in turn connects to the European Academic Research Network (EARN) that interconnects European research institutions. In order for MIT to remain on Bitnet, it must guarantee that no non-university parties send mail or other traffic to international destinations over the subsidized network.[40] For this purpose, MIT implements the *Intersect* rule in its gateway to the R&D net, assigning wild card category sets to entire regions of its internal

[40] This restriction to university parties is necessary in order to conform with policy requirements of the European PTT's.

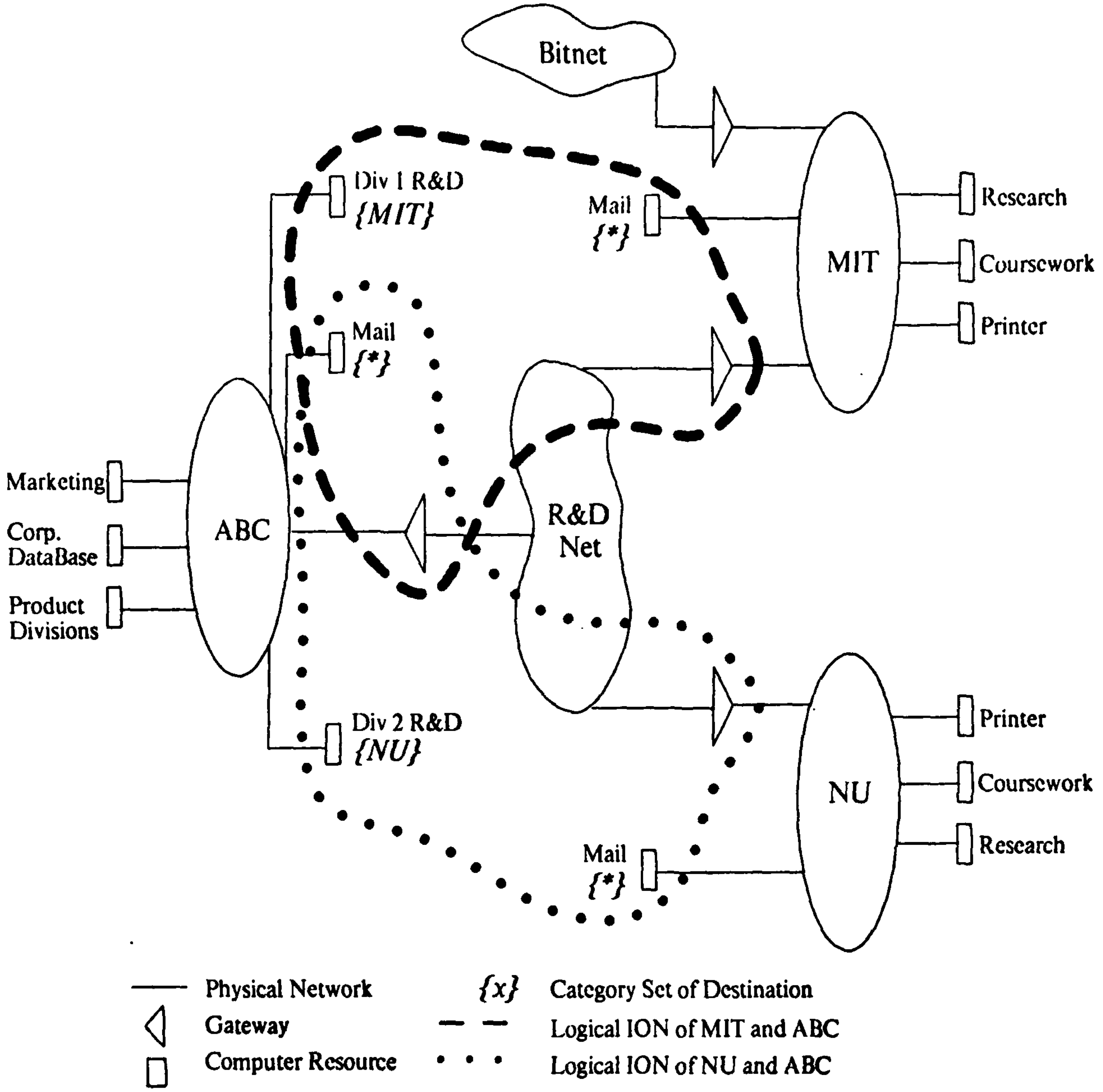


Figure 5-7: Example of an Inter-Organization Network: One ION exists between ABC and MIT, and another ION exists between ABC and NU. Both IONs overlap physically yet are isolated logically from one another and from a third ION, Bitnet, to which MIT is connected.

network that it wants to be globally accessible, but assigning the Bitnet gateway the strictly-internal category only.

5.5 Implementation

In general, the following controls should be implemented by each ION participant. Detailed description is provided in Chapter 8.

First, an ION participant must define categories and assign to each ION machine or process the set of categories (compartments) for which it is to be used. For example, the organization may define a category for each ION in which it is a participant.[41] If a process is intended for external access by members of a single ION (e.g., an order entry system for customers), then the process is assigned only that category. If the process is to be accessed by members of multiple IONs (e.g., a mail server), then it is assigned multiple categories. Strictly-internal services are assigned only the *strictly internal* category since they are not intended for any ION access. External entities are assigned the category associated with their particular ION. If the gateway supports multiple types of invocation (e.g., connection-based remote login, message-based server requests, electronic mail), each ION or subgroup may be assigned multiple category sets, where each set applies to a different type of invocation.

Given these category sets and assignments, the organization allows external invocations to enter the internal domain only through specified gateways; similar to the notion of entry points [57]. Each gateway forwards (routes) an invocation to the indicated destination if and only if the *Intersection* between the category set of the external invoker, $\{Ci\}_s$, and the category set of the destination ION process, $\{Ci\}_d$, is not nullset, {}.[42] Strictly-internal entities (processes, programs, mail-boxes, machines, etc.) are assigned a category set of

[41] If different external entities in an ION are to be given different capabilities, then subgroups are defined and each subgroup is assigned its own category set.

[42] Note that the gateway's control policy does not require that the category sets of the invoker and invoked be exactly the same, i.e., it does not necessarily enforce confinement. Each ION application enforces additional controls which restrict flows across categories; see below.

{strictly internal} only , and are therefore not accessible to any external groups. In addition, if an internal process or machine is not registered, it is not accessible via the ION gateway. The mechanism can be implemented using a non-discretionary access control list for each ION process. where the list contains groups that are allowed to access the process.[43] In addition to maintaining category information for each outside ION or subgroup, the ION gateway must know where to route incoming and outgoing messages based on the destination address, as do traditional gateways.

Finally, the ION participant must implement discretionary or non-discretionary controls within ION facilities. ION facilities must enforce application-specific controls that the gateway can not enforce itself (e.g., restrictions on the dollar amount of a purchase order or the filename of a cad/cam file request). In addition, ION facilities that reside in overlapping logical networks must prevent information or invocations from flowing between ION and non-ION entities, and between different ION entities. A range of traditional system security mechanisms can be employed. *System level controls complement the controls on network communications, i.e., at the network boundary.*

The result of these mechanisms is that no external invocations can be sent to entities that are not explicitly registered as accessible to outsiders. In addition, the ION participant can specify which categories of external users may access each ION process.

5.5.1 Information Flow

In many cases, information flow controls on outgoing traffic may be needed as well.[44] If an organization is unable or unwilling to rely on existing policies to discourage employees from exporting confidential information via the ION, the organization may require additional information flow controls. For example, some features of computer-based communications

[43] Note that internal invocations originate within the organization's domain, do not enter through gateways, and therefore are not subject to these non-discretionary controls.

[44] For many inter-organization networks invocation control is needed for incoming traffic whereas information flow control is needed for outgoing traffic. Therefore, the two do not conflict with one another as they often do when both apply to traffic flowing in the same direction [37].

remove direct employee discretion from the generation of ION messages, such as automatic distribution lists. A user who sends a message to a distribution list typically does not know which individuals are on the list; the user knows only that they share a common interest. If one of the addressees on the list is located outside of the organization, an employee may export information without realizing it and therefore without considering relevant company policies. In addition, an ION participant may enforce controls on outgoing flows if it must conform to policies imposed by other ION participants (e.g., Bitnet's no transit requirement described earlier in the MIT-ABC-NU example), or if it must pay for outgoing flows on a usage-sensitive basis.

Some information flow controls can be implemented using category sets and the *Intersect* rule: internal user A can send a message/file X to external user/resource B if and only if $\{C\}_a$ and $\{C\}_b$ have a non-empty intersection. However, more elaboration is needed and capabilities will vary with the type of control mechanisms available internally. For example, if internal systems implement non-discretionary controls that mark objects with security labels, the gateway can control outgoing information flow based on the security level of the message content as well as the category set of the message creator. Because each organization implements its own gateway, each can integrate existing, internal, labeling systems into its ION gateway.

In summary, the gateway authenticates, labels, and maintains information on category sets while most of the rest of the internal systems can go on unchanged. Because of the gateway's central role, there are a number of important design issues which require further elaboration but which we will only mention here. One is that the gateway and ION entities are programs that must be "trusted" by the organization that owns them. However, each ION participant can make its own decision as to the investment and trust that it will place in its ION gateway. The general requirement is to raise the level of monitorability to that of non-ION communication channels (e.g., telephone and paper). Therefore, organizations' security requirements vary according to the value of online information and resources, as well as the nature of the inter-organization relationships. The second issue is that the gateway must authenticate the source of a request/message in order to properly evaluate its category set. A range of tools, of varying strength, can be used; from third-party authentication servers to one-time encryption keys. Although participants need not use a common authentication scheme internally, they must agree on the authentication protocol used by their respective gateways (see Chapter 7).

5.5.2 Confinement and the Role of System-Level Controls

The above discussion and examples illustrate how a non-discretionary policy based on a relaxed *intersection* rule can be used to isolate logical networks without imposing either physical isolation or increased access controls on all internal systems. Before concluding, it is important to emphasize the role of system-level controls in those systems that an organization does include in the ION. First, as was stated earlier, it is essential to the security of critical strictly-internal systems to implement controls in all ION systems to prevent external traffic from traveling via the ION system to strictly-internal systems or systems belonging to other IONs. Although such system-level controls depend on traditional security techniques, the task may be very difficult, depending upon the application involved. In particular, the larger in number and more varied are the tasks performed by the system, the harder it is to certify or audit system-level controls. At the same time, if the system is used for strictly-internal purposes as well, the overhead experienced by internal users must be minimized (see section 5.2).

Despite these difficulties, an intersection rule on communication flows could be combined with confinement rules at the ION-system level to achieve confinement in the larger network system. For instance, in the first example (see figure 5-5), MIT provides both ABC and XYZ with access to its electronic mail delivery system. At the communications level, an intersection policy is enforced so that both customers can communicate with the order-entry system. As indicated in figure 5-6, ABC has category set {Educational-research}, XYZ has category set {Network-research,Architecture}, and the mail system has category set {}. If a traditional confinement policy were implemented, XYZ would be unable to communicate with the simulator because the simulator's category set is not a superset of XYZ's category set. Moreover, the confinement rule would prevent the mail system from sending online messages back to either of the companies because neither of the companies' category sets is a superset of the mail system's category set. By equipping the ION gateway with a non-empty intersection rule instead of a strict confinement rule, the desired communication between customers and the order-entry system is achieved. However, because the rule governing communications has been relaxed, the order-entry system must take responsibility for preventing flows across, what are intended to be, *isolated* logical networks—that between MIT and ABC and that between MIT and XYZ. In other words, the mail system needs to enforce the traditional confinement policy to prevent information and invocation flow (message based) between ABC and XYZ entities. Although the security of such a scheme depends critically upon the security of the system-level controls employed, the approach described throughout this chapter structures the problem so that security risks can be isolated and managed by the organizations involved. In other words, although we have not eliminated the problems that are common to all computer system security (e.g., certifiability, overhead cost, etc.), we have developed an approach to access controls in IONs for which the security risks are as tractable as traditional computer system access controls. And we have done so without violating organization constraints such as minimizing interference with internal resource sharing.

5.6 Conclusion

In conclusion, initial analysis suggests that category sets and non-discretionary control mechanisms can be adapted to satisfy usage control requirements in inter-organization networks; namely, to isolate the strictly-internal facilities from the ION facilities, and distinct IONs from one another. This approach has implications for network interconnection, in particular the level of interconnection. Further research is needed to understand the range of applications for which the proposed modifications might be suited, the implications for non-discretionary security models, and appropriate authentication schemes. The following chapter investigates the implications of the proposed mechanisms for network interconnection protocols.[45]

[45] I thank Bob Baldwin, Carl Landwehr, Steven Lipner, David Reed, Suzanne Sluizer, and Juliet Sutherland for insightful comments and suggestions on drafts of this chapter.

Chapter Six

Implications for Network Interconnection

As the previous chapters described, when organizations establish inter-organization network connections and extend their networks internally, they require new usage control policies and mechanisms to cope with the increased heterogeneity of the user population. For example, consider the case of a university computer science department, such as MIT's, that is connected to the Arpanet. In the past, the user population that could access the department's Arpanet-connected machines was small and the department required no special measures in order to adequately comply with the Arpanet policy that the Arpanet be used only by computer science researchers. However, as MIT extends its computer networks out from the computer science and engineering department to the rest of the campus, the user population that can access the Arpanet-connected machines is no longer small nor homogeneous. In addition, the university has established external network connections with local industry. In this case, the potential user population of the computer science department's facilities includes not only members of other departments, but members of other organizations altogether. In this new environment the computer science department may have to introduce control mechanisms to restrict access to the Arpanet gateway or Arpanet-connected hosts in order to adequately comply with Arpanet policy.[46] Such control mechanisms would have to discriminate between various segments of the user population; in this case these segments are logical groupings of users or hosts according to organization or department affiliation.

A second issue that arises when interconnection reduces the homogeneity of a network concerns *efficient* use of network resources. Typically, the larger and more heterogeneous is the user population, the less tightly coupled are the applications that operate across the

[46] A related policy requirement with similar technical implications is that the Arpanet not be used as a transit path between two points, neither of which is itself a legitimate Arpanet node.

entire network. However, many applications that assume tight coupling among users operate under the assumption that this tight coupling spans the entire network, even in the presence of extensive interconnection. Applications that use broadcast communication protocols are the best example. I will use Xerox's Grapevine system illustrates this point. [9] Grapevine is a distributed database service that provides mail, naming, and other information to users and applications. Fundamental to Grapevine is the manner in which it keeps the distributed data repositories up-to-date. The updates can take up significant network resources. If two organizations that each run Grapevine interconnect their networks in order to support some inter-organization application such as electronic mail, the Grapevine updates will travel across both networks.[47] However, this up to date naming information may be far less appropriate to the loosely coupled relationship of the two organizations than it is within a single organization. And given the gateway bottleneck through which all inter-organization traffic must flow, the updates may place a significant burden on network resources.[48] For this reason, some types of communication traffic should not be forwarded onto external networks. More efficient use of network resources would be possible if broadcast information such as minute-by-minute Grapevine updates would carry information in the packet header indicating that the packet is intended for logically-local destinations only. Logical locality is emphasized since it is what determines the appropriate degree of coupling for this application.

A second example is the use of Address Resolution Packets (ARPs) to locate hosts. Some networks broadcast ARPs over the entire network in order to locate a particular machine.[49] Consequently, when two networks that use ARP are simply interconnected, all ARPs flow across both of them. If these two networks belong to two distinct organizations, it may not

[47] Primarily because ION connections and the like were not an explicit consideration during the design process.

[48] In addition, such inter-organization connections are often more transient than intra-organization connections. Transient connections are the exception internally and the tightly coupled applications may not be designed to adapt easily.

[49] CMU is one example. Their use of ARP is described in [41].

be cost effective to broadcast all ARPs across the boundary.[50] However, without information about the logical affiliation of a packet source and destination, there is no easy way to use a broadcast-based search mechanism locally without having it propagate across the ION gateway as well. An application level example of the same phenomenon is a network-wide broadcast search for a service of some kind, e.g., a high-quality printer. In this case it is likely that the user is interested in printers belonging to his or her organization only.

The following sections focus on the issue of controlling ION flows to meet policy, as opposed to efficiency concerns. However, similar mechanisms and issues pertain to the latter concern as well and therefore are relevant to intra-organization connections as well.

6.1 Non-Discretionary Controls

An *uncontrolled* connection between two distinct organizations implies that the organizations are willing to trust one another and all organizations to which each interconnects. Under some circumstances this level of trust is appropriate due to the nature of the organizations (e.g., low risk), their relationship (e.g., not competitive), or existing contract provisions (e.g., liability for violation of connection). However, under many conditions the level of trust implied may be inappropriate and inconsistent with other aspects of inter-organization relations and interchange.

The proposed approach to the problem of usage control in IONs is to implement non-discretionary controls in all entry and exit points to each organization's internal network, i.e., all ION gateways (see Chapter 5 for discussion). Any ION communication involves at

[50] Problems related to broadcast of ARPs have been experienced at MIT. MIT's Lab for Computer Science's local network is connected to the AI laboratory's Chaosnet. The AI lab Chaosnet is in turn connected to a local company's, QRS's, Chaosnet. QRS's Chaosnet is in turn connected to many of its customers' hosts or networks. A bug in the machine of one of these customers caused large amounts of ARP traffic to be generated to the extent that it flooded a MIT LCS local network and caused several network-attached personal computers to cease functioning. Although this same problem would have occurred if the defective machine had been on the same local network as the personal computers, it is unfortunate that by virtue of interconnecting an organization makes itself so dependent on the correct operation of another organization's machines.

least two ION gateways—one belonging to the source's organization, and one belonging to the destination's organization. Each organization implements its own ION gateway and selects the control mechanisms to fit its own policy requirements. Examples of policies that an organization might enforce using such controls are:

1. Accept incoming traffic only if it is from an authorized outside entity and is destined for an internal system or gateway that has been explicitly registered as available to such outside access. Access may be refused to an external user either because of the user's organization or group affiliation, or because of the type of access requested.

2. Forward outgoing traffic only if it is from an authorized internal entity and is destined for an authorized external network. External network access may be refused to an internal user either because of contract provisions that restrict the use of the external network, or because of usage fees charged by the external network. Information flow may also be restricted on the basis of internal sensitivity classifications.

All entry points to an organization's internal network must be treated as ION gateways and equipped with controls. For example, as described earlier, MIT sells time on one of its timesharing systems to a wide range of users—small local companies, international research centers, government personnel, other universities, etc. Because these non-MIT users can access the MIT network to which the system is connected, the system itself acts as a gateway and must enforce controls consistent with MIT's policies; for example, restricting access to other gateways and certain internal resources (e.g., printers, scarce computational resources, etc.). If controls are desired in connections that cross organization boundaries, all entry points, both dedicated gateways and hosts that provide outside access, must be equipped to address incoming and outgoing traffic. The benefit of the non-discretionary approach proposed is that systems used strictly by insiders need not be modified, nor made aware of the presence of new interconnections. However, those systems that are made accessible must employ mechanisms to enforce application-specific controls (e.g., which files can be accessed, which programs can be run), and to isolate ION processes from non-ION processes.

To implement these non-discretionary controls an ION gateway must have access to certain information about the logical characteristics of traffic; e.g., organization affiliation of source and destination, type of service, amount of resource requested, etc.[51] According to this information the gateway determines which categories of internal information or resources the external entity may access. In other words, in addition to the traditional bindings between user or service and node, node and network attachment point, and network points and path [61], the ION gateway needs a binding between user or service and organization affiliation. Domain style naming might be used to capture this notion of affiliation. [40] However, as is described in the following section, it is not possible to evaluate the domain affiliation of a packet based solely on the network number that it carries in its header.

If the logical information required for policy decisions is available, then the non-discretionary controls can be implemented by assigning category sets to incoming and outgoing traffic, according to logical characteristics of the traffic and enforcing invocation and information flow controls accordingly. [20] The next section describes issues associated with low-level connections (packet-level), for which this information is not directly available.

6.2 Packet-Level Interconnection

As with any gateway, an ION gateway can be designed to operate at one of several levels. Most gateways can be classified as either high or low level. A high-level gateway is an end-point in a message- or connection-based communication session, such as file transfer, remote login, or electronic mail. A low-level gateway forwards packets between machines that are the endpoints of higher-level message or connection-based communication sessions, but the gateway itself is not an endpoint. Low-level gateways may operate on individual packets (datagrams) or virtual circuits, depending upon the protocol design of the interconnected networks; I refer to both as packet-level gateways. High-level gateways operate at *session*, *presentation*, or *application* layers, whereas low-level gateways operate at *transport* or *network* layers of the International Standards Organization, Open Systems Interconnect (ISO-OSI) reference model. [30]

[51] For simplicity, much of this discussion focuses on organization affiliation of source and destination and mode of access. Similar arguments apply to other types of information.

As is discussed below, most packet-level gateways do not have access to the information needed to make ION policy decisions. This is not necessarily inherent to this level of connection, but is a result of the competing requirements that constrain the design of low-level protocols.

6.2.1 Network Numbers

The organization affiliation of source and destination is fundamental to many, if not most, conceivable usage control requirements for IONs. Consequently, an ION gateway must be able to identify the organization affiliation of the traffic destination and source. Given this information, the ION gateway can assign categories and determine the rights of that source and destination.[52] The source and destination in a packet header appear in the form of network numbers. This section describes some of the problems of relying on these numbers for identification of organization affiliation.

Networks interconnected at the packet level (e.g., Internet Protocol (IP) level in the DARPA TCP/IP family of protocols) must coordinate the assignment of network numbers in order for packet addresses to be meaningful throughout the internet. In addition, network numbers provide information about proper routing of a packet to its destination, e.g., which subnet on which network a particular host sits. This routing information pertains to the physical location of the destination. When networks cross organization as well as geographic boundaries, *logical* information is desired in addition to *topological* information. In other words, policy control mechanisms need to know the organization domain to which a message is being sent, and from whence it came, in addition to the physical locations.

Currently, network numbers in the Internet are allocated to sites by a centralized number czar. Each site may then allocate numbers to hosts and even subnets that lie within its topological network. Most of these hosts and subnets are within the confines of a single organization, but some are not. For example, MIT has direct network connections to several

[52] Many policies may require more information than just source and destination affiliation. But for simplicity I focus on this information to illustrate the argument.

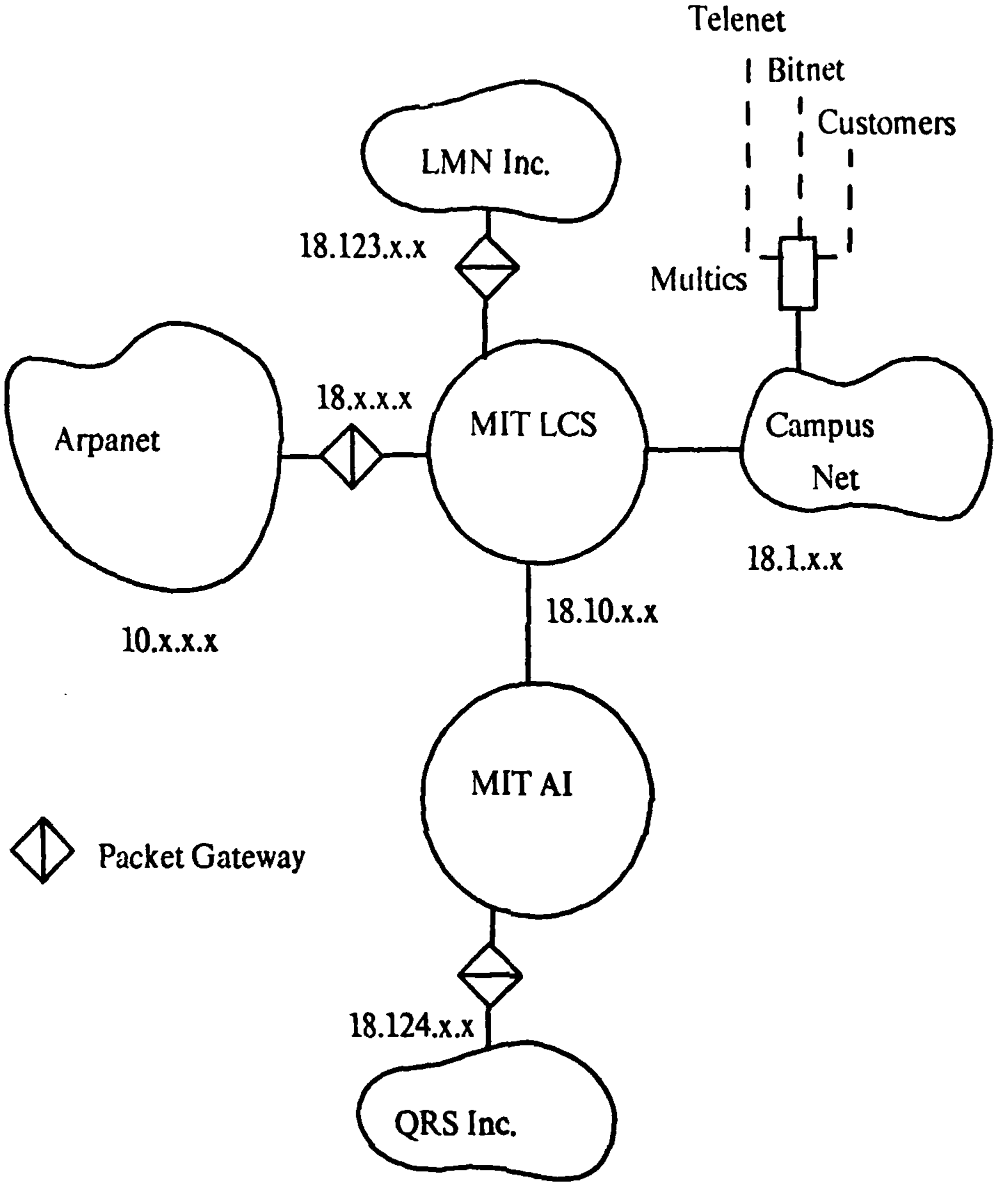


Figure 6-1:A simplified depiction of MIT internal networks and several external networks to which it is connected. The Internet network numbers are listed near the gateway to each network.

local companies; see figure 6-1. The network numbers of destinations in these companies look like the network numbers of other MIT subnets because they contain topological information for routing purposes. In order to discriminate between subnets and hosts that are part of MIT's logical network (i.e., actually belong to MIT) and those that lie outside of the logical network (i.e., facilities in the local companies which are accessible but do not belong to MIT), the gateway must be able to bind the source and destination network numbers in the packet header to the organization affiliations.

These issues were not among the many considered during design of the Arpanet/Internet protocols. At that time, the primary concern was to achieve connectivity and transparency and make network boundaries disappear.[53] Therefore, it makes sense that providing information needed to enforce organization boundaries was not a design requirement. Even if it had been a consideration, the number of competing requirements and constraints on the low-level protocol would probably have led the designers to leave such application-specific information to higher levels. In particular, because routing table size is limited, there is pressure to be able to make routing decisions on the basis of a packet's destination subnet number.

One might try to use the network and subnetwork numbers as a hint to organization affiliation. However, because of the decentralized manner in which networks and subnetworks may establish their own interconnections, over time these topological numbers may not map into meaningful logical groupings. For example, MIT might implement a filter in the Arpanet gateway to reject outgoing messages with source addresses other than MIT's Arpanet network number, 18. However, if an MIT department or laboratory connects some local company to the department or laboratory local network, according to current practices, that company is assigned a subnet number within net 18. Therefore, the filter would not catch transit traffic sent from that company to non-MIT Arpanet sites. In addition, several MIT hosts do not sit on network 18.

Of course it is possible to identify the various subnet numbers that are assigned to non-MIT entities and add such information to the gateway filter. However, this is not a general

[53] In many ways the Arpanet is not a typical ION. In the past, the Arpanet was intended to encompass the internal networks of the participating organizations. Only recently, as the participating organizations have extended their internal networks to other internal communities, is the Arpanet community manifesting many of the issues attributed here to IONs.

solution because such interconnections are established in an incremental and decentralized manner, and therefore there is no good way of tracking these exception cases without centralizing the interconnection and number allocation process in some way. One approach might be to establish guidelines that set aside blocks of numbers to be used for non-MIT sites. However, because of the nature of the namespace, it is hard to know a priori how many such numbers to set aside, and exactly what groupings one will want to be able to distinguish between, i.e., MIT/non-MIT is only one relevant distinction. If such guidelines do not exist, and the connection is not centrally managed, it is not feasible for the gateway to maintain a list of allowable host and/or subnet addresses with which to implement packet-level controls.

An example of a packet-level ION gateway that implements usage controls is the University College London (UCL) network connection to the Arpanet. [11] The UCL network employs two gateways to the Arpanet. One connection forwards packets via a private satellite network to the Arpanet. The second connection forwards packets via an X.25 connection over public packet-switched networks. The two separate gateways are needed because of the different protocols used, and the division satisfies policy requirements. Due to PTT regulations, only Ministry of Defense traffic can be sent via the private satellite path, while civilians (such as many university researchers) must send traffic via the public-network path. Because only routing information is available at the IP level, the restriction is enforced by making UCLnet appear as two separate networks, UCLnet and PSSnet. This is achieved by splitting the namespace in two and assigning addresses to MOD and civilian hosts accordingly. Because there is a small and fixed number of user groups (i.e., two) the mechanism works. In addition, higher-level controls enforce restrictions on invocation of higher-level network applications, i.e., mail, file transfer, and remote login (see the description in section 6.4). A similar mechanism could be employed by MIT to restrict access to the Arpanet. For example, most student access to computational facilities and the MIT network will occur via the MIT subnets that belong to project Athena.[54] The Arpanet gateway could simply reject all packets originating from those subnets, but in so doing it

[54] Athena is a university-wide experimental project in the use of computers in education.

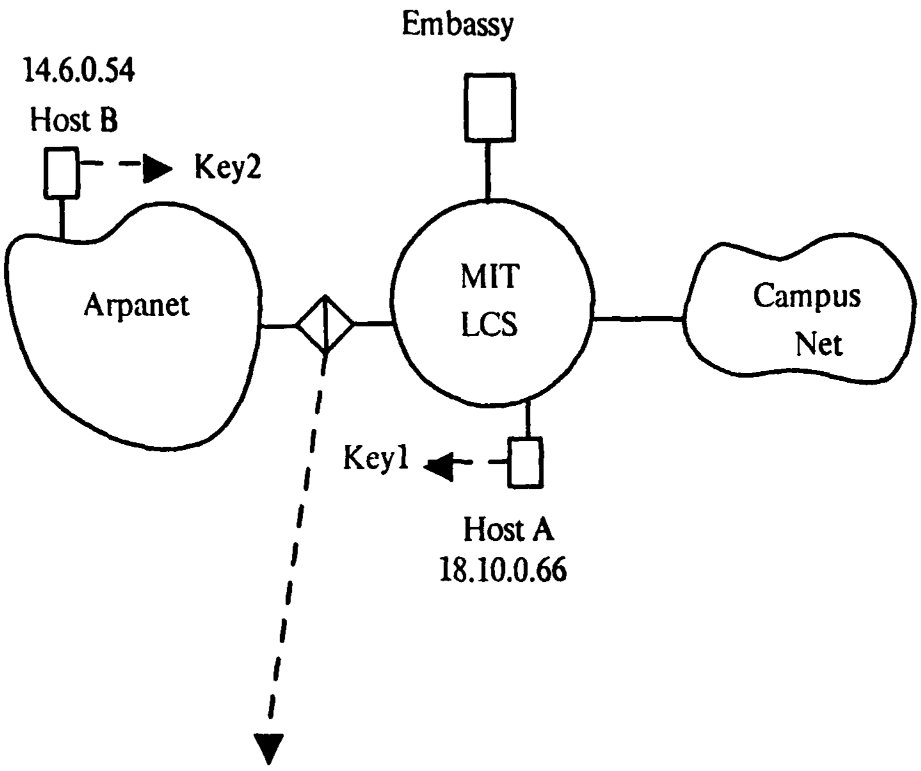
would preclude transit by legitimate research hosts that are physically located on Athena subnets.

In conclusion, traditional packet level protocols do not carry the information in each packet header that an ION gateway needs to make policy decisions. The following section describes one scheme for augmenting a packet level protocol in order to accommodate policy controls.

6.3 Visa Scheme

Given that logical information generally is not deduceable from packet headers alone, we can adopt an approach first suggested by D. Reed and documented by J. Mracek. [42] This scheme, depicted in figure 6-2, requires that the source carry out a higher-level dialog with a policy server in the destination network in order to authorize a particular conversation (e.g., mail, file transfer, etc.). The policy server passes the authorization information to the packet-level gateway along with a means of authenticating the authorized traffic (e.g., an encryption key). The scheme is referred to as a visa scheme because gateways are analogous to border crossing stations, access control servers to embassies, and the keys to visas.

In this scheme, in order for a source host to send a packet or set of packets via an ION gateway, the source must obtain a key from the access control service (ACS) of the network that it wishes to enter or leave. If the source passes the ACS's policy filter, the ACS gives the lower-level gateway a source network id, key pair with which to authenticate packets from the authorized source as they pass through the gateway. The same key is given to the source. The key may simply be a ticket appended to each packet header or it may be an encryption key used to calculate the packet checksum. For example, using an encryption key, the ION gateway records which keys correspond to which network numbers. The gateway looks up the key corresponding to the source network-number of incoming packets and calculates the checksum using the key. If a key is found and the checksum is properly computed then the gateway knows that the packet has been authorized by the local ACS. If the organization's policy requires that the gateway discriminate according to the destination



ID	Visa
18.10.0.66	Keyl
14.6.0.54	Key2

Figure 6-2: Example of a packet-level ION gateway using a visa scheme.

of each packet in addition to the source, the ACS can include this in the authorization information as well by passing the ION gateway a source, destination, key triplet. The visa may be granted to many different units of authentication so long as the packet forwarding gateway has access to the additional information; e.g., source-destination-service type, number of packets, etc. Visas can not be granted according to the application supported, or user ID, because this information is not available to the gateway in the IP packet headers.

This scheme offloads to the ACS the mapping of traffic affiliation to access privileges by requiring the source to have a higher-level dialog with the ACS before its packets are able to enter or leave the network. Once the ACS provides the low-level gateway with the access and authentication information, the low-level gateway operates as if it were itself equipped with the category information and a way of mapping traffic information to that category information. In effect, by granting a visa to a source or source-destination pair, and informing the gateway of the granting, the ACS wraps a set of individual packets into a logical unit which is then subjectable to policy control in the packet-level gateway.

A visa scheme has been proposed to control incoming traffic to a dial-up, packet-forwarding gateway to the MIT network. This gateway is connected on one side to the public switched telephone network, and on the other to an MIT local network. Although a single physical gateway is used, MIT would like to apply different access policies to the different groups that use it. Some MIT resources are intended for access by members of the MIT community only (e.g., gateways to other networks, a New York Times clipping service, high-speed printers, etc.). Other resources are intended for access by some non-MIT users as well. In order to implement non-discretionary controls as proposed in Chapter 5, the gateway would operate as follows. When a user calls the gateway, the gateway associates the call with a particular port and accepts packets from the user only if they are addressed to an ACS. The external user carries out a high-level dialog with the ACS and authenticates itself. After authenticating the user and identifying the internal facilities that the user is authorized to access, the ACS sends the gateway a key and a list of destination addresses to which the particular user should be allowed to send packets. In addition, the ACS sends the same key to the user. The gateway associates both the list of destinations and the key with the port assigned to the user. The key is used as a *connection-authenticator* for the duration of the connection. In order for the gateway to accept a packet through the port, the checksum of the packet must have been calculated including the connection-authenticator that is currently associated with the port. When the user first dials up the connection-authenticator is zero and the user can send packets to only a single destination, the ACS.

If access policies are relatively stable, an even simpler visa scheme can be used which would reduce the need for frequent higher-level dialogs with the ACS, and would therefore reduce the overall performance requirements of the ACS itself. For example, assume that an ION gateway needs only distinguish between those hosts that belong to a particular research community and those hosts that do not. In this case it would be adequate to regularly distribute a key to all eligible research hosts and the ION gateway. The gateway could use the key to identify packets belonging to authorized research hosts, as in the above description. The research hosts would no longer require a dialog with the ACS for each external communication. Consequently, the ACS could even operate offline or via electronic mail.

A packet-level gateway together with an ACS can effect higher-level controls. However, if the ION is intended to support only a small number of higher-level applications, if the policies can not be expressed in terms of information available in packet headers, or if performance benefits of packet-level interconnection are not significant, it makes sense to consider interconnection at a protocol level at which the policy-related information is available directly. The following section describes higher-level interconnection. The two schemes are contrasted in more detail in Chapter 8.

6.4 Higher-Level Interconnection

In the introduction I gave two motivations for treating entities on the other side of an ION gateway differently from those within an organization's internal network. First, policy concerns may require that non-local users be restricted from using some internal resources and other gateways. Second, efficient network use suggests that information needed for local, tightly-coupled applications, not be broadcast through an ION gateway across which applications are more loosely coupled. In the discussion of network numbers I explained that the logical information needed to implement intelligent filtering in an ION gateway is not available at the packet level. The visa scheme described above can effect higher-level control for many simple usage control policies. However, when policy decisions are dependent on higher-level information that cannot easily be bound to packet-level information, higher-level connections may be more practical. The primary advantage is that the information needed to evaluate policy, such as organization affiliation, service type, size of request, etc., is available to higher level protocols directly; i.e., these protocols deal with aggregated units of traffic that contain more semantic information in the headers and control fields (e.g., electronic mail messages, remote login or file transfer connections, etc.). A packet-level gateway with an ACS can effect higher-level controls, as described. However, these controls must be representable in the form of a key or ticket. In addition, the visa scheme requires modification of lower-level communication protocols because of the need to alter calculation of the packet checksum.

Even with a higher-level ION gateway, some controls are best implemented in the endpoint applications themselves; in particular, controls that discriminate according to the content of a message, e.g., the size of a purchase order, or the name of a file requested. In addition, these applications-level controls may be required to isolate the ION processes and applications from the non-ION ones. In the remainder of this section I will continue to focus on higher-level, communication protocol controls, under the assumption that some application-level controls are implemented in the endpoints.

Higher-level gateways terminate the higher-level communication protocols and thereby gain access to more information about the application of the connection. Depending upon the level of connection and application, this information may include the logical affiliation of source and destination, the actual service being performed, and the amount of communication resources requested, for example. Although a key security issue arises with regard to the authentication of this information, the point is that the information is available for evaluation, and authentication mechanisms can be employed as needed. For example, Harvard University is connected to the Arpanet via a packet-level gateway. Harvard would like to allow any university member with a computer account to send electronic mail via the Arpanet gateway, but at the same time it wants to provide file transfer and remote login to select groups of users only. Currently Harvard is able to control remote login use because for internal resource control purposes, it does so anyway within the internal network. Remote login is a restricted command on all internal hosts and because only certain users can use it internally, only those users can use it through the gateway. However, file transfer is not a restricted command; it is too common and useful a facility to even consider restricting internally. As a result, there are no controls on doing file transfer via the gateway. Because the gateway is a packet-forwarding gateway, and such information is not deduceable directly from packet headers, controls that discriminate according to service type (i.e., remote login, file transfer, mail) and host are difficult to implement on such a gateway. If the gateway operated at a higher-level, it would be a more simple and modular task to restrict file transfer use to authorized users, because the connection would carry information about the affiliation of source and destination as well as the communication mode.

This discussion of level of interconnection is concerned most directly with what Sunshine refers to as service level and implementation approach. [69] Service level refers to the communication mode supported in the gateway, e.g., datagram, virtual circuit, file transfer, remote login, mail, etc. He distinguishes between two implementation approaches, *endpoint* (where the source and destination each act as an endpoint in the communication mode and each gateway passes lower-level information), and *hop-by-hop* (where each intermediate gateway acts as an endpoint of the communication mode as well). I refer to the former as lower-level, or packet-level, and the latter as higher-level, interconnection.[55] With respect to traditional interconnection concerns alone (not inter-organization concerns), Sunshine finds the hop-by-hop (higher-level) approach more appropriate where backward compatibility of protocols and immediate needs predominate and where user awareness of crossing network boundaries is acceptable. He finds the endpoint approach (packet-level) preferable when robustness and generality are important and there is more basis for agreement and conformance to standards. Sunshine's conclusions lend support to the argument in favor of higher-level (i.e., hop-by-hop) ION connection. When an organization interconnects to the outside, backward compatibility with internal protocols and procedures is still of primary importance. Similarly, expediency is often a key criterion for the interconnection and it is

[55] Application level refers to the endpoints in the application itself and therefore application-level controls are even higher than the higher-level communication controls discussed here. For example, where a high-level gateway would forward a file transfer request without looking at the content of the request, an application-level gateway would interpret the request itself.

often desirable for the connection to be less than transparent so that insiders are conscious of their actions when communicating with outside entities.

Two types of higher-level gateways can be distinguished—connection based and message based. Some applications require that the gateway set up a real-time connection between machines on either network. Message-based applications require only that the gateway forward messages between two applications. Due to their store-and-forward nature, message-based gateways are sometimes referred to as relays. However, message-based gateways operate at a higher protocol level than packet gateways, which also operate on a store-and-forward basis. A message is a complete semantic unit whose content, and address, a remote process, application, or person can interpret. Packets are small, equally sized components of a connection or message. Any single packet may have no semantic meaning by itself and typically carries only low-level addressing information. The most common type of message-based gateway is a person-to-person, electronic mail relay.

In both message- and connection-based high-level gateways, the unit of transfer and therefore of control makes accessible more policy-related information than does an individual packet. Connection based protocols must perform under tighter real-time constraints and therefore are more difficult to implement. Connection-based gateways establish connections between entities on either side of the gateway and then ship undifferentiated packets back and forth via that connection. Consequently, connection-based gateways *may* introduce more vulnerabilities than do message-based gateways. Even if controls are added to the connection set-up process, unless controls are applied on a per packet basis, there is more chance for a connection that was approved initially to be used for some undesirable purpose. A message-based gateway applies controls to each message that passes through it while avoiding the cost of applying controls to each packet. However, if a machine automatically processes the messages (even to simply display or format it) similar attacks may be made by embedding executable commands or special control characters within the text of the message. Even so, the staged delivery of messages makes it easier to guard against such attacks by filtering traffic. On the other hand, message-based gateways are ill-suited to delay-sensitive applications. Message-based gateways are discussed further in section 6.4.2.

There are many examples of higher-level gateways in use today. A few examples are described below:

- As illustrated earlier, a host connected to two different networks can act as a high-level gateway between them. For example, the host might forward electronic mail between users on either network using the host's local mail facilities. Similarly, a remote user logs in to the host (perhaps via the internal network and public-network gateway of the user's own organization), and from there might establish another connection to some other host on the internal network. In both cases, the host acts like an endpoint in the mail or remote login protocol.

- ISI operates an experimental mail gateway called *Intermail.* This program accepts mail from Arpanet users and forwards it onto one of several public mail networks, including MCIMail and IEEE Compmail. Messages must conform to a special format in which the destination network, and destination address, are listed as the first lines of the message contents itself. This information is interpreted automatically by the Intermail program, a new message is constructed with this information as the header, and the remaining message contents as the contents of the new message, and the message is forwarded on to the appropriate network. [19]

- IBM's SNA interconnection technique terminates the communication protocol at the gateway between SNA networks. To the end user, the gateway is transparent, but at the connection level, two connections exist—one between each gateway and each endpoint. Name translation is also done in the gateway in a manner transparent to the end user. [6]

- Cambridge university uses an X.25 gateway to connect its local area network to a public packet-switched network. The gateway implements access controls by checking all connections against a database of authorized users. Similarly, accounting information is collected in the gateway. [18]

- The University College London network, described earlier, supports mail, file transfer, and remote login to the Arpanet in the U.S. Each mail user must register his or her mailbox with the mail server in order to send and receive mail; alternatively an entire host may be registered. File transfer and remote login services require that the user log in to the terminal gateway using a personal or group password. Mailbox registration is achieved via the terminal gateway as well. The control information is used to conform to the Ministry of Defense and PTT imposed policy requirements described earlier, as well as for accounting purposes.

- The UUCP based network operates at a higher-level than packet forwarding. [29] Electronic mail, mailing list digests, and sometimes files are transferred between hosts or networks of hosts via telephone connections. It is relatively easy to add filters to the forwarding of UUCP network traffic. Filters may determine such things as which mailing lists are distributed, and which types of services are provided to each of the neighboring UUCP sites.

Even at higher levels the distinction between topological (routing) and logical (organization affiliation) information may be blurred. If an organization supports transit, and guidelines are not set for the structure of source and destination names, address and routing information can be confused with logical information. For example, if person Smith at QRS Inc. sends mail to Jones at XYZ Inc. via MIT's network, then if XYZ receives the source address as Smith.QRS%MIT, the XYZ ION gateway must figure out that the source's logical affiliation is QRS and not MIT. At the same time XYZ must be able to determine how to return a message, namely via MIT. UCLnet has experienced this problem in its mail connections to the Arpanet. As described earlier, mail from Ministry of Defense and civilian users must be treated differently. Some mail is forwarded to the Arpanet from other civilian research networks and hosts that are connected to UCL. The addresses assigned to mailboxes on these hosts are constructed so that the mailboxes appear to lie within UCL. The mail gateway relies on a list of registered users to filter mail, in part because the organization affiliation of a user is not necessarily evident from the header. [16] However, in general textual mailbox names carry more semantic information and are taken from a larger namespace than network numbers. Therefore textual mailbox names can be constructed in such a way that the correct affiliation can be interpreted using easy to follow guidelines.[56]

6.4.1 Placement of Controls and Higher-Level Connections

Previous sections have argued that based upon the need for non-discretionary controls in network connections that cross organization boundaries, higher-level connections should be considered. However, there remains the question of *where* the controlled, higher-level

[56] Arpanet Domain Name format is one possibility although it was not designed for this purpose and therefore may not be practical. [40]

connection should be placed. This issue is less straight forward than it might appear because it is sometimes unclear where the administrative boundary that requires enforcement lies. In particular, if an organization and its computing facilities are divided internally into subgroups, there exist administrative boundaries *within* the organization's internal network as well. There are two general situations in which the placement of the controlled gateway may be unclear. First, if one internal subnet has an external connection which entities on a seconc internal subnet are not to use, it is unclear whether controls should be placed in the external connection or in the connection between the two internal subnets. Second, if an internal application (or set of applications) is used only by external users and is the only internal application that external users access, it is unclear whether the controls should be placed at the application level in the endpoint systems or at the communication level in the ION gateway. The two design criteria that guide placement of controls are: location of the information needed to make control decisions, and proximity to the administrative boundary between internal and external facilities.

The appropriate location of a control function depends in large part upon the information needed to make the control decisions. For example, assume that the ION gateway forwards messages and communication requests. If usage control policy discriminates according to the source or destination of traffic, the controls can be implemented in the ION gateway since the gateway has access to source and destination information in message headers or connection set-up requests. On the other hand, if the policy discriminates according to content (e.g., the amount of a purchase order or the particular file requested), it is better done by the endpoint (i.e., destination) which has access to such application-specific information. However, as we have argued, controls on external use should be placed where they will cause minimum interference with internal use. In other words, controls should be placed as close to the administrative boundary as possible so that the controls will impose on internal operation as little as possible. Consequently, in some cases where an end-to-end arguments might appear to favor of placing controls on external use in the end application, [59] concerns over interference with internal use of the application and access to the policy-relevant information favor placing control in the gateway that connects external users to the internal network.

Although usage control design often must balance these two criteria, in two degenerate cases, both criteria can be satisfied simultaneously: namely, when an ION gateway is dedicated to a single application, i.e., outsiders use it to access only a single application, and when the application(s) accessed by outsiders are used only by outsiders, i.e., not by internal users. In the first case, all controls can be placed in the gateway since it is application-specific and need not support more general communication. This solution accommodates end-to-end concerns and minimizes interference with internal operations at the same time. For example, if the ION gateway supports outsider access to a mail system only, the gateway can stage the mail and implement all necessary communication and application-level controls. In the second case, all controls can be placed in the end application since it is used only by outsiders. This solution accommodates end-to-end concerns without interfering with internal operations since there is no internal user of the application. For example, if the ION gateway supports outsider access to a VLSI design, service system that is used only by an organization's paying customers, all controls can be placed in the VLSI design system without infringing upon internal users. The exceptional characteristic of the first case is that the application-level access can be extended out to the gateway since only one internal application is accessed. The exceptional characteristic of the second case is that the administrative boundary can be said to exist between the application and the rest of the internal network, instead of at the ION gateway per se, since no internal entities use the ION application.

In other situations, there is more of a trade-off between application level controls and interference with internal usage. For example, many university computer science departments are facing the following dilemma. These computer science departments have connections to the Arpanet (or one of several other research networks) which they have made available to all users of department facilities, routinely. Recently, many of these computer science departments have connected their respective internal network(s) to a much larger and diverse population of users across their respective university campuses; and sometimes to other non-university organizations as well. These connections introduce the need for usage controls in order to meet Arpanet policy requirements such as that the Arpanet be used for research purposes only. The question is, should usage controls be placed in the gateway between a Computer Science department's internal network and the Arpanet, or in the gateway between a department's internal network and the other university networks? The gateways are not application specific, nor are the applications external-user-specific. Therefore, neither of the degenerate cases exists. Nevertheless, by considering the two design criteria in turn, the appropriate decision can be determined. First, because both gateways (i.e., CS-Arpanet and CS-Campus) support access to multiple applications, end-to-end argument implies that the controls in the gateway restrict communication-level functions, not application-level functions. Application level functions must be restricted in the end applications that are made accessible to outsiders. Second, because the policy requirements that must be enforced in these cases are imposed by Arpanet management, the communication-level controls should be placed in the CS-Arpanet gateway and not in the CS-Campus gateway.[57]

6.4.2 Message-Based IONs

A message-forwarding gateway is a special type of high-level gateway This section is devoted to message-based interconnection because of its suitability to the loose coupling often desired across organization boundaries.

Organizations can augment traditional person-to-person communication media such as telephone, postal mail, telex, and face-to-face meetings, by interconnecting their internal electronic mail systems. In addition, messages can be used as the transport method for sending commands to, and receiving responses from, computer-based services. For example, the *Electronic Data Interchange* standard (ANSI X.12) specifies a set of formats and commands related to purchase orders and invoices. [31] Using this standard a machine can automatically interpret and act upon messages, and thereby convert a *passive* communication channel into an *active* one. In a similar way, message protocols can be used to retrieve documents or programs, and to subscribe to and deliver information services. For example, Arpanet users invoke a fabrication system called MOSIS (developed by

[57] Additional controls may be placed in the latter gateway but these would address intra-campus requirements, not Arpanet-imposed requirements.

Information Sciences Institute at USC) via electronic mail. [45] Clients use specially formatted messages to enter designs into the fabrication queue, request status information, or check that a design is acceptable to the system. MOSIS interprets the design messages as textual descriptions of the geometry of masks for IC fabrication and the physical products are eventually shipped to the client via air carrier. More recently, Sciencenet, a network proposed to support supercomputer access, proposes to support some batch, message-based, service access. As end-users increase their computing activities and autonomy with personal hardware and software, electronic mail can provide a convenient substrate via which to offer end-user developed servers. For this reasons alone, the number of message-invocable servers within organizations is likely to increase across a broad range of organization functions. Even in those environments whose services are invocable via connection-based protocols only, message-based gateways may support loose external couplings by acting as front-ends. For example, an outside user would send a formatted message to the gateway which would interpret it and apply access control filters. Assuming the gateway found the request to be legitimate, the gateway would carry out the users request by setting up a a connection to the indicated server and generating the indicated commands. To the end-user the service would appear to be message-invoked; however, the ION gateway is acting as a front end to those services for which it is programmed to understand messages. Although the asynchronous nature of message-based services is not transparent to the end application or user, speed and turn-around can be tailored to support semi-interactive applications.

Several characteristics make message-based interconnection appropriate for inter-organization communications and interchange. First, many ION applications are more loosely coupled than are internal applications and can be supported adequately by asynchronous message-based communications. Second, protocol conversion, which is often necessary between distinct organizations that have not coordinated equipment selection, is easier because it need not be achieved within real-time performance constraints. For the same reason, overall gateway implementation is less complex. Finally, message-based communication accommodates filtering and access control more easily because it provides higher-level information than packet-based communication, and yet does not impose real-time constraints as do connection-based protocols. In addition, if internal users invoke most internal servers via connection-based protocols while external users have access to message-based protocols only, then only those internal servers that support message-based invocation will be accessible to external users.

However, message based interconnection alone does not solve the problem of ION usage control. Non-discretionary controls are still needed in the gateway as described earlier. The need for such controls is particularly great in those environments that contain message-invocable services; as compared with environments where message-based communications is used for person-to-person communication only. Moreover, because message-based front-ends are relatively easy to write, message-invocable servers will probably grow in number within organizations; i.e., as a means for end-users to construct and share their own servers. Although we claimed above that controls are somewhat easier to implement, message-based gateways may raise conflicts not encountered with connection-based gateways. In particular, conflict will occur if the same communication protocol and gateway are used to forward person-to-person and person-to-machine messages. As describe earlier for XYZ Inc., an organization may want to support an open default for person-to-person electronic mail, and a closed default for person-to-machine messaging. If the default is open so that all internal hosts have electronic mail capabilities, then all internal, message-invocable servers must know about ION connections and be equipped with appropriate defenses. However, this may be unrealizable in an environment where both servers and ION connections are established incrementally and without centralized administration. If the default is closed, then all internal persons or hosts must be registered explicitly; which may or may not be acceptable to the organization. If a host is registered as a unit, so that all users potentially have ION mail access, a mail-invocable application on a registered host must still take action to protect itself from external invocation. However, if host administrators have to take explicit action to register their hosts as accessible, it may be feasible to require that they make their local user population aware of the need to defend message-invocable applications.

Whether message-based IONs support person-to-person electronic mail or more general message protocols, problems of mail forwarding must be addressed. Cohen and Postel have described *mail tunnels* via which mail can travel in encapsulated form between networks, automatically forwarded from one to the next; the Intermail system mentioned above is one example. [15] Many mail systems automatically forward mail to another host if they receive a message for a machine with a known network address. This feature is quite useful at times but flies in the face of both packet-level and mail-level controls. To a packet level gateway the traffic will appear to be from the host that did the forwarding, and not from the originator of the message. Similarly, if the mail is encapsulated and a mail-level gateway checks the outermost header only, the mail will appear to be from the forwarding host. In order to regulate transit from internal systems without ION access, to the ION via internal systems that do have ION access, a mail gateway must examine encapsulated mail headers, or the organization must take some action internally to restrict off-net forwarding. In fact, taking action with respect to the internal systems may be justified in this case because the forwarding host is acting as an high-level ION gateway. Therefore organizations could manage this problem by requiring that any host that automatically forwards mail call itself a gateway and conduct checks and controls in accordance with ION gateway guidelines.

6.5 Conclusion

This chapter has argued that to implement non-discretionary controls in ION gateways, certain logical information must be available. Two types of interconnection and their suitability to ION applications were discussed. The conclusion was drawn that higher-level connections are preferable for many ION applications. And where lower-level connections are adopted for performance or generality, a visa scheme can be used to support many simple usage controls policies. In addition, these controlled, higher-level connections should be placed as close as possible to the administrative boundary being enforced. Finally, message-based gateways are well suited to loose couplings desired for many inter-organization relationships. The next chapter describes security, and in particular, authentication requirements in IONs and following that, Chapter 8 analyzes the implementation issues and trade-offs of ION gateway design.[58]

[58] Discussions with J.N. Chiappa, D. Clark, D. Reed, and L. Zhang contributed to the ideas presented here. In addition, D. Feldmeier and L. Zhang provided detailed comments on an earlier draft.

Chapter Seven

Security and Authentication in Inter-Organization Networks

7.1 Introduction

Previous chapters did not address vulnerabilities introduced by ION communication that *is sanctioned* by official policies. In other words, discussion has focused mainly on how to tailor IONs to include only those resources desired by the organizations and did not address the risks associated with converting a manual information or resource channel into an automated one. In many cases the risks may be fundamentally altered in the absence of extremely costly measures, where cost is measured in both dollars and convenience of use. On the other hand, in many applications it is adequate to accept the risks if they can be identified and assessed a priori and monitorability and auditability are adequate.

Two aspects of ION communications that are inherently more risky than traditional voice or paper communications are the difficulty of authenticating information about the source and destination of communications, and the ability to cause internal mechanisms to behave contrary to the intended procedure. In general, human boundary-spanners are more cost-effective at both tasks than are machines because of the flexibility with which humans can combine multiple sources of information and detect suspicious or unordinary events. Often these procedures are not conscious or are at least hard to codify. The majority of this chapter discusses authentication in IONs. The last section addresses the second issue which resembles the Trojan Horse phenomenon.

7.2 Authentication Requirements

The controls outlined in previous chapters rest upon assigning categories or rights according to the organization affiliation of the source and destination. If the source and destination information can be falsified, then the controls are not effective. More accurately, they are not effective if they can be falsified *without detection* before or after the fact. Therefore, environments in which the risk of falsified source information is significant will require mechanisms to authenticate that a communication is from the entity it claims to be from. A range of authentication mechanisms can be used. In some cases the gateway may rely on the correct operation of the communication system, i.e., that the communication system does not accept mislabeled communications from sources and does not change the information in transit. Such reliance may be reasonable if the organizations employ a third party whose reputation depends upon proper operation and whose own accounting interests rely on proper identification of sources. Similarly, the gateway may rely on the fact that the source will not be able to receive replies to its messages if the source address is not correct. If falsification is not detectable through malfunction of this kind, the gateway may require a more explicit authentication mechanism. For example, a predefined password, ticket, or key may be used to authenticate a source to the gateway. Finally, ION participants can employ a third party mechanism to dynamically authenticate a source using session keys or ticket mechanisms. These alternatives are discussed in the following sections. It is also necessary to authenticate that the destination is who the source claims it to be. The latter function is more easily left to the reliable operation of the communication delivery system. Since the delivery system for incoming communication belongs to the organization, it is more reasonable for it to be trusted. In the case of outgoing communication, the authentication mechanisms described above must be employed to achieve the same level of trust.

An interesting problem that is not solved by any of these authentication mechanisms is message encapsulation and forwarding as discussed in Chapter 6. Message forwarders can cause a message to appear to have a different organization affiliation than that of the message originator. However, we assume that message forwarders must take responsibility for their actions and if messages are encapsulated in such a way as to mask the name of the originator, it is the responsibility of the forwarder to declare itself as a mail gateway/tunnel and conform to organization guidelines (e.g., implement usage controls). [15] Therefore, as part of the guidelines specifying the ION usage controls that each ION gateway must implement, an organization must make clear that automatic forwarding of mail off of the internal network constitutes gateway behavior.

The remainder of this chapter describes authentication requirements and protocols for IONs. Needham-Schroeder type authentication tools are shown to satisfy the authentication requirements outlined in the usage control model. The primary ideas presented here are that internal authentication mechanisms need not necessarily be modified to comply with inter-organization requirements, and that multiple classes of authentication are desirable.

In order to enforce the desired policies and controls, and to comply with contract agreements, interconnecting organizations must be able to authenticate one another. The main purpose of authentication in this domain is to assure *accountability* should some behavior transpire that is in violation of contracts.

There are two types of authentication required:

- First, when one organization contacts another for the first time, the organizations must authenticate that each is legitimate. For example, when a new client contacts a vendor, the vendor typically checks the client's credit rating just as the client has checked the vendor's credibility in the market. In this case, the new computer-based transaction mechanisms should allow organizations to assess one another via third parties in the same formal way that is done currently via telephone and paper.

- Second, each time an organization contacts another, it must authenticate that it is the organization that it claims to be. When an established client contacts a vendor to reorder some item by telephone or paper mail, both parties typically have informal or formal procedures for assuring each other that they are who they claim to be. For example, purchasing agents recognize one another's voices, or rely on the letterhead of invoices and letters, or call back the requester at the claimed organization. In this case, the new mechanisms must substitute for what are often informal procedures via telephone or postal mail.

In both cases, different levels of authentication are appropriate for different organizations and types of transactions. For example, the larger the proposed purchase, the more confident the vendor will want to be that the customer has an adequate credit rating and that it is who it claims to be. Similarly, the larger the purchase, the more confident the client will want to be that the vendor will be able to uphold its end of the agreement—delivery date, quantity, quality, service, etc.

The goal of this discussion is to specify how two or more organizations can make use of trusted third parties to authenticate one another without having to modify internal systems and protocols, with the exception of the ION gateway. The methods proposed fit well with the model of usage controls described in earlier chapters. The discussion begins with a list of assumptions about organizations' and third party facilities. Initiation authentication is then described, followed by transaction authentication, and multiple levels of service for both types of authentication.

7.3 Assumptions

Several assumptions are made about the ION participants' internal facilities. First, each organization,[59] A, has an internal authentication server, AS_a, that it, A, trusts to authenticate individuals *within* organization A. Contracts between ION participants specify that an organization, A, is responsible for the integrity of the information provided by AS_a. A second assumption is that an organization should be able to participate in ION authentication using existing internal authentication mechanisms. Although the organization might choose to beef up such mechanisms in the presence of new liabilities for correctness, it remains an internal decision. Finally, organizations are assumed to have a known and small number of ION gateways (we will assume 1 for simplicity). All packets that enter or exit an organization's internal network must pass through one of these official ION gateways. Much of the function described below for the ION gateways can be offloaded to special policy servers to improve the gateway's packet-forwarding performance. However, for simplicity all functions will be treated as a part of a single logical gateway, even though they may be physically separated.

Several assumptions are made with respect to third party services as well. First, each organization has many ION-supported relationships each of which is governed by a separate contract. If no third party is employed, authentication must be handled on a pairwise basis. Since authentication fundamentally depends on sharing a secret, each organization would

[59] An organization is defined here as a set of entities that are willing to trust and be represented to other organizations by a single authentication server and gateway.

have to keep track and guard as many secrets as there are organizations it communicates with. The benefit of employing a third party is not the traditional space considerations, but rather the liability associated with guarding each of the secrets. In addition, minimizing the number of organizations that one trusts with a secret makes it is easier to certify that the secret is being kept. Also, if the communicating organizations are competitors or otherwise mistrustful of one another, the third party can act as a buffer between them. The function of the third party is twofold. The first is to provide information about organizations to one another when they interact for the first time. The second function as ION Authenticator is *to certify that a particular transaction/connection/message/*packet is from the *organization* that it says it is from. It is left to the source organization's AS to certify that the packet/message/connection is from the claimed individual, i.e., x, within the organization, A. Our final assumption about third party services is that they are available to authenticate organizations (ION participants) to one another. Different levels of service (of *guarantee*) are available for different types of organizations, transactions, and relationships. Any two (or more) organizations that want to be able to authenticate messages from one another must agree on a single mutually-trusted third party.[60]

7.4 Initiation Authentication

If transactions are carried out *online* it makes economic sense for organizations to be able to initiate relationships with one another online as well. For example, a computer manufacturer may buy a certain chip by sending online price queries to a collection of suppliers and initiating a purchase with the lowest bidder. In this case, the selected supplier will want to check the credit rating of the new client just as it does when a first-time purchase is proposed over the phone or on paper.

In the paper and voice world a wide range of requirements and corresponding procedures exist for evaluating the legitimacy or credit of a new client. We will discuss this further in section 7.6. For now, we will assume that the third parties that a supplier traditionally

[60] Actually, the scheme below could be extended to allow the participants to use different third parties [58] but for simplicity we will assume that they agree on a single one.

checks with are accessible online. !f they are not, then the supplier must use traditional media for evaluating new clients.

7.4.1 Protocol

The general approach to initiation authentication is described in terms of a new client, A, proposing to purchase something from a supplier, B.

When the supplier receives a message it checks the source listed on the order against a list of known entities, i.e., initiated clients. If the source does not appear on the list, the supplier sends an authentication request to one of several third parties employed for this purpose. A may send a suggested third party's name along with the original message if A anticipates the need for initiation authenticat..n. Along with the name of the claimed entity, A, B includes the criteria according to which the AS should evaluate A, e.g., credit rating. B may set the criteria according to the destination of the message (i.e., the level of risk or value of information or product control residing at the destination), or the size of the request.[61] If the source is not registered with any of the third parties employed by the supplier the purchase order may be rejected or a message returned saying that registration with third party X is required. It is then up to the customer to reinitiate the purchase after establishing its identity with X. If the third party does have the client registered, the third party returns its evaluation of the client (e.g., credit rating, or perhaps just an assurance that the client is a real company) to the supplier. The supplier adds the client to its list of initiated clients along with the evaluation. The supplier also records the name of the third party that was able to provide the information about the client. From this point on the client is initiated until the supplier decides to recheck the evaluative information.

Following is an example of a dialog that could be used to implement initiation authentication as described above:

[61] In the latter case, gw_b would have to pass the purchase order to some service in order to determine the appropriate evaluation criteria since it is based on something other than the source and destination of the message, which is the only information the gateway has direct access to.

1. A ---> B: purchase order

2. B ---> AS: A, evaluation criteria

3. AS ---> B: A, evaluation

4. { B ---> A: m, register with AS }

5. B adds A, evaluation, AS address to known-entity list

At this stage the organization that the purchase order *claims* to be from is initiated as a legitimate entity to do commerce with. However, the supplier still needs to know that the purchase order in fact came from that organization. In addition, in the future, when the initiated client sends other purchase orders, the supplier must be able to authenticate that the purchase orders are from the claimed client for which the supplier maintains credit rating information, etc. What is needed is a mechanism for authenticating that a particular transaction is from the claimed party. The following section describes an approach to *transaction authentication.*

7.5 Transaction Authentication

Assuming that a client has been initiated and is now a registered client with the vendor, each transaction must be authenticated. This section outlines the approach and describes a simple protocol for transaction authentication and implementation issues.

7.5.1 Protocol

A protocol for ION authentication will be described for two organizations, A and B, who want to authenticate messages from one another. However, we assume that both organizations communicate with many other organizations as well so that the approach must scale well. After each organization has registered itself and a secret key with a common third party, a Needham and Schroeder protocol is used to authenticate the organizations and provide communicating pairs with session keys so that they can authenticate messages from one another [44].

Before describing the protocol, we should emphasize why a third party is employed in this dynamic phase of authentication. As long as each organization is maintaining information about the other, each pair of communicating organizations could exchange a secret key with which to authenticate one another. Our rationale for employing a third party is that there is significant overhead in protecting a secret. Given that organizations have many correspondents (i.e., other organizations that they transact with), it is significantly more manageable for an organization to safeguard a single key to communicate with a third party than it is to safeguard n keys, one for each of its n correspondents. Note that the concern is not for space, since as mentioned, some contract or other information is already stored for almost every correspondent. Rather, the concern is for the nuisance associated with safeguarding secrets. For this reason, a third party is employed for transaction authentication.

The protocol begins when an individual x in organization A sends a message to y in organization B; x and y may be people or machines. The message header lists the source and destination organizations and individuals. All messages travel in and out of A and B via gw_a and gw_b, respectively. If B considers there to be no need (i.e., no risk associated with open access to y), it may forward the message to y unauthenticated. However, if B wants wants to control external access to internal resource, y, then for this discussion we will assume that B uses non-discretionary controls and assigns category labels to incoming messages, as was described. Because B assigns a category label according to the source of the message, B wants to authenticate the source, i.e., make sure that the source listed in the message header really generated the message. Functionally, this means that the organization listed in the header will take responsibility for the message.

To authenticate the source organization, B sends a message to the third party that it has listed as the one to use to authenticate messages from A; we will call this third party authenticator AS_{ab}.[62] B asks AS_{ab} for a key with which to authenticate A and subsequent

[62] We assume that during initiation authentication described above, the two parties identified a mutually trusted third party.

messages sent by A during this session. B also returns the message, m, to gw_a saying that authentication is required. When gw_a receives the returned, unauthenticated message from gw_b it asks its internal AS_a to authenticate x. B also authenticates y through a conversation with its internal AS_b.

AS_{ab} sends gw_b a session key, E_{ab}, along with the session key encrypted in A's private key, E_a; included also is a timestamp and an identifier of B. The entire message includes a timestamp or nonce and is encrypted under B's private key, E_b.[63] B then sends A the session key encrypted in A's secret key. B does not have A's secret key, but was given the encrypted session key by AS_{ab}. B is guaranteed by AS_{ab} that only A will be able to read this message. Similarly, A is guaranteed by AS_{ab} that any message identifying B along with a session key encrypted under A's secret key must have originated with AS_{ab} and that only B has been given a copy of the session key. A and B now each have a copy of the session key and are guaranteed by AS_{ab} that any message encrypted under that key can be read by the other organization, only. Finally, to protect against replays by an intruder, A and B carry out a simple handshake, e.g., exchanging the current date and time.

Both gateways store the session key and gw_a resends the message, m, from x encrypted with the key. Both gateways encrypt all subsequent communication between x and y with the session key until the session ends or either party decides to reauthenticate. gw_b is assured that any messages arriving under that key came from gw_a and gw_a relies on internal authentication to assure that the message came from party x within A. Similarly, gw_a is assured that only someone in B can receive the message, since only gw_b can decrypt the message, and gw_b relies on its internal authentication to assure that the message goes to y, only.

The dialog that corresponds to this protocol is listed below.

1. x-->GW_b: m

[63] Both organizations' addresses and private keys have been stored with AS_{ab} previously when A and B registered with AS_{ab}. AS_{ab} uses these secret keys to authenticate the organizations.

2. GW_b-->AS_{ab}: (B,A)

3. GW_b-->GW_a: m, error-unauthenticated

4. GW_a-->AS_a: x and GW_b-->AS_b: y

5. AS_{ab}-->GW_b: $E_b(A,E_{ab},E_a(B,E_{ab},T))$

6. GW_b-->GW_a: $E_a(B,E_{ab},T)$

7. GW_a-->GW_b: $E_{ab}(I)$

8. GW_b-->GW_a: $E_{ab}(I-1, J)$

9. GW_a-->GW_b: $E_{ab}(J-1)$

In summary, using their secret keys (e.g., E_a and E_b), each organization can authenticate itself to the trusted third party in order to request a session key. The gateways use this session key to authenticate the source and destination organizations of each message. The organizations take responsibility for authenticating the destination within their respective organizations, based on existing internal authentication mechanisms. Consequently, AS_{ab} is liable if organization A or B is incorrectly authenticated, whereas AS_a and AS_b are liable if x or y are not who they claim to be. This characteristic is significant because it allows an organization with tight physical security to dispense completely with internal authentication if it so chooses.

7.5.2 Implementation

The following changes are required to implement this protocol among organizations with heterogeneous internal networks:

Third party:

1. A method for distributing keys between organizations and trusted third parties is needed so that the trusted third party can authenticate the organization.

ION gateway:

1. The gateway must maintain a list of trusted third parties so that when an unauthenticated message arrives from another organization, the gateway knows where to go to request authentication. The gateway must also store the private key used to authenticate its organization to trusted third parties. In addition the gateway maintains the known-entities list which includes evaluation information and mutually-trusted third party for each initiated organization.

2. Encryption in the gateway. *No* internal entities need to encrypt messages for the purpose of authentication. Each gateway must store the session keys and associate them with the appropriate incoming and outgoing packets; e.g., by assigning the source, destination pair and the key to a virtual or physical port.

3. The gateway must be able to ask the authentication mechanism to authenticate the source of an outgoing message (i.e., generated internally).

Note that the individual persons or machines that originate messages need not be concerned with this procedure other than responding to authentication challenges from the internal AS. The gateway handles external authentication requests, retransmission of the first message in a session, as well as all encryption.

Several of the functions that logically are done in the gateway when a session is first authenticated may be *offloaded to different hardware* in order to improve the efficiency of forwarding packets that belong to ongoing sessions. However, if the level of authentication is such that sessions consist of one message only (e.g., authenticating electronic mail), there is little savings. On the other hand, if each packet in a mail, remote login, or file transfer session is authenticated individually, the overhead may be great and warrant offloading. Therefore, the appropriate engineering depends on the level of interconnection, i.e., whether the gateway is a packet forwarding gateway or an application level gateway in which application protocols are terminated.

To offload this function to a server, the protocol would be modified as follows. When the first packet in a session arrives it is assumed to be unauthenticated and is forwarded to the ION policy server which sits in the destination organization (B in the above example). The policy server carries out the protocol listed above for the gateway (gw_b in the above example). The gw automatically forwards all unauthenticated incoming packets to the policy server during this dialog with the third party ION authenticator (AS_{ab}). Once the source organization is authenticated and the session key is obtained, the ION policy server sets the port in the gateway to authenticated and sets the session key. From then on packets arriving to that port in that key will be forwarded to the destination(s) for which they were authorized (determined by the rights assigned to the source organization, see [20]), until the session is closed or until either side decides to reauthenticate. In either event, the policy server resets the port and session key entries. The policy server could also handle the initiation protocols for authenticating new clients.

As organizations adopt more sophisticated internal authentication mechanisms, such as badge readers, the discrepancy between internal and external authentication levels will grow. If internal facilities and applications assume that authentication involves the sophisticated internal mechanism instead of a less sophisticated external mechanism, it may be inappropriate to tell the application that an ION user is "authenticated" in the same way that internal users are. The ION gateway must compensate for the lower level of authentication of external users by taking responsibility for their authentication, or by adopting additional mechanisms; assuming the internal application cannot be updated to accommodate multiple levels of authentication.

7.6 Multiple Levels of Service

Different types of transactions require different degrees of confidence in the credit or authenticity of the client. And, different strengths of authentication require different types of equipment and facilities. When the highest level of authentication is not available, some lower level of authentication may be adequate. If a purchase order arrives for $10,000 worth of goods, the supplier must be relatively confident that the client is legitimate and in fact made the order, before the order is acted upon; the cost associated with incorrect authentication is high. However, if a smaller client sends a purchase order for $100 worth of goods, relatively little authentication may be necessary and the facilities needed for the protocol described above may not be available. Therefore, it would be nice to support intermediate services, i.e., multiple levels of service.[64]

One method for offering a "second-class" authentication scheme is to rely solely on initiation type authentication. as described below.

7.6.1 Protocol

The protocol begins when A sends a message to B. We assume that A has no encryption capabilities at all.

Initiation authentication is only slightly affected by the lack of encryption capabilities. If A is not on B's known-entity list then B contacts a (set of) third party(ies) to authenticate the existence of organization A and to evaluate it. Assuming B contacts a third party that does have A registered, that third party returns to B values of the requested evaluation criteria along with a flag indicating the level of authentication that A can support; for example, first-class to indicate that A has encryption capabilities and can carry out the protocol described earlier, and second-class to indicate that A has no capabilities and must rely on passwords sent in the clear to authenticate itself to the third party.

Transaction authentication can no longer rely on a Needham-Schroeder protocol since A has no encryption or decryption capabilities. Therefore, when B asks the third party to authenticate a particular transaction or message from A (either the first transaction or later ones), the third party informs B that only second-class transaction authentication is available. One procedure that the third party could use in the absence of encryption would be to ask the source of the message to B (presumably A) to resend the password that it submitted upon registration. If the resent password matches A's registered password, the third party could send a message to B indicating that the third party believes the source of the message is in fact A. Similarly, the third party could authenticate B and inform A that the third party believe that the destination is B. In both cases, the third party must include the authentication level rating, second-class. A and B can then decide whether to accept or

[64] This feature was suggested by J. Saltzer.

reject this level of authentication for the proposed transaction. The primary risks are that there is no session key for the parties to authenticate themselves to one another directly and there is no control over an impostor intervening in the transaction after it has begun. In addition, passwords are subject to intervention because they are sent to the third party in the clear.

For certain types (low risk) of transactions and communications, this limited level of assurance may be acceptable, and preferable to no authentication at all to the extent casual impostors are detected or discouraged. However, it is vital that both parties keep track of the level of authentication in use. For example, if in the middle of a transaction A proposes to increase a purchase order by an order of magnitude, B should know that only second-class authentication is being used and reject the suggestion if it sees fit.

7.7 Internal Authentication

Before concluding it is worth emphasizing that although organizations do not have to modify their internal authentication mechanisms in order to support ION authentication, inter-organization connects can heighten the need for reliable internal authentication mechanisms. The mechanisms described in this chapter allow an organization to decouple internal and external authentication. However, if an organization's internal authentication mechanisms are weak or non-existent, the ability to authenticate external entities leaves several problems unaddressed.

First, using the protocols described above, an organization is liable for requests that its gateway allows to flow to the outside world, i.e., by passing a message the gateway has asserted its belief that it was generated by an authorized program or user. Just as an organization needs to guard against employees writing fraudulent or excessive paper purchase orders, so it must guard against fraudulent or excessive online purchase orders. Many organizations have extensive authentication procedures in place with respect to traditional paper purchase orders; certainly more extensive than it has for electronic mail. An ION participant must recognize the need to augment traditionally-weak electronic mail authentication with something better suited to the application.

A second concern that impinges on internal authentication requirements regards incoming communication. For example, an internal order-entry system that accepts online orders from customers via the gateway relies on the gateway to authenticate that the originator of the order belongs to the organization that it claims to belong to. However, if the level of trust on the internal network is low, the order-entry system will need to authenticate that the requests that claim to be from outsiders actually are coming from the ION gateway. If this authentication does not take place then an internal user could spoof the order-entry system by sending a message that claims to be from the ION gateway.

In summary, offloading responsibility for external authentication to the ION gateway is desirable because it decouples external from internal authentication. However, such decoupling may also be dangerous if the organization does not carefully reevaluate assumptions made about the trust-worthiness of the internal environment. In the absence of external interconnection, the cost of unauthenticated communications may have been deemed less than the expense of authentication mechanisms. Inter-organization connections change the parameters of the equation and should lead ION participants to reevaluate their internal as well as their external authentication mechanisms.

7.8 Subverting Access Controls

The first security issue outlined above is similar to the trojan horse phenomenon; namely, that special commands or control characters can be embedded in a communication and cause the receiving machine to behave in a way in which its operating organization did not intend it to behave. For example, control characters can be embedded in a text message which is then sent for display on the screen by a person's mail-reader. If the recipient's display terminal interprets these control characters as pseudo-commands (e.g., to clear the screen, interrupt the current program, interpret the following text as a command, etc.), the text message can turn an otherwise passive communication medium into an active one. In other words, the remote machine can be made to take some action aside from delivering and displaying text to a human. A similar problem arises with some text formatters that interpret a piece of text as a command if it is preceded by special characters. As with the display system that interprets control characters as commands, the formatter interprets these special characters as commands and transforms a passive channel into an active one. In order to guard against such misuses, the gateway must filter messages for control sequences. This requires that the gateway be able to recognize a control sequence. Therefore, generic control characters and commands can be filtered at the gateway, while system-specific ones may require filtering at the endpoint.

7.9 Conclusion

In summary, organizations can initiate relationships with one another using third parties to authenticate one another's identity and desired credit information, can carry out transactions using third parties to authenticate that the transaction request travels from and to the claimed party, and finally, both of these activities can be carried out at the appropriate authentication cost level.

Chapter Eight

Implementation of ION Gateways

The implementation of usage controls in an inter-organization network gateway consists of two parts: the protocol for forwarding packets, messages, or connections via the ION gateway; and the association of these packets, messages, or connections with access rights. In general the ION gateway implements controls related to communication characteristics, i.e., source, destination, and mode of communication. Application-specific controls are left to the endpoints. If a gateway supports communication with only a single application it could implement the application-specific controls as well; if in addition the application supports only outsiders, the controls could be implemented in the application instead of the ION gateway. This chapter focuses on a solution to the more general case in which the ION gateway interconnects the internal network to multiple external organizations, and supports communication with multiple applications which have internal, as well as external, users.

8.1 Protocol

The protocol used by internal hosts and users to communication with outsiders via the ION gateway depends upon the communication service supported (e.g., mail, remote login, file transfer) and upon the level of interconnection. We discuss three types of protocols via which hosts can communicate with the outside world: packet-level, message-forwarding, and connection-terminating.

If the ION gateway operates at the packet level, an internal host communicates with the outside by setting up connections directly with the external destination host. The ION gateway simply routes the packets to the network designated in the packet header. Policy controls operate based on information available in the packet header or using a visa scheme as described in chapter 6.

Message protocols, by definition, imply asynchronous interaction between source and destination: person-to-person electronic mail is a typical example. A message-based gateway stages messages between internal and external host. Internal machines send mail destined for external machines to the gateway, and vice versa. The gateway applies policy controls and sends the messages on to the appropriate external network. An internal user may specify an external destination to an internal machine by explicitly listing the gateway as part of the destination address (e.g., Smith at ABC via ION gateway, or Library at MIT via ION gateway), or the internal machine could be programmed to recognize names of external destinations and forward such mail to the ION gateway (e.g., Smith at ABC and Library at MIT are forwarded to the ION gateway if ABC and MIT are not recognized internal host names). In either case, internal hosts forward the mail to the ION gateway using the same mail transfer protocol used to forward mail among internal hosts. The ION gateway implements policy controls on a per-message basis using information in the message-header. An internal host that is connected directly to an external network, as well as the internal network, can be configured to serve as this type of mail gateway.

A connection-based ION gateway operates at a higher protocol-level than packet-level and supports highly-interactive services (e.g., remote login and interactive file transfer) which are not easily supported by message-based protocols. As with message-based gateways, a connection-based gateway acts as an endpoint in the higher-level protocol and terminates the connection. Internal hosts set up sessions with the ION gateway which in turn sets up sessions with external hosts, assuming policy controls permit the requested communication. An internal hosts that is directly connected to both an external network and the internal network can act as a this type of higher-level gateway, or a dedicated gateway machine, may be used. A multi-network host gateway enforces controls on all commands that generate network traffic based on the initiating user's identity, the particular command, and the indicated destination. A dedicated gateway filters all communication requests based on connection set-up control information, i.e., source, destination, and application type. As in the message-gateway case, users may specify explicitly to their local machines that a session be established with the gateway, or the local host may detect that the destination is external (based on the destination host name) and initiate a session with the gateway automatically.

In all three cases, the usage controls implemented involve a common set of functions which are described below. The protocol alternatives are compared and contrasted in section 8.4.

8.2 Evaluating Access Rights

When a packet, message, or connection request arrives (depending upon the level at which the gateway operates), the ION gateway must associate the communications with access rights. The first step in this process is to identify the communication characteristics that are relevant to the policy decision. The second step is to identify the category sets that correspond to these characteristics. The final step is to compare the category sets against the policy algorithm's definition of a permitted operation. These three steps are each described in turn below.

8.2.1 Associating Communications with Access Rights

Many ION usage control policies can be expressed in terms of the mode of access (e.g., electronic mail, file transfer, database query, remote login) and the organization affiliation of source and destination. Some control policies may require logical information other than mode of access and organization affiliation of source and destination for evaluating access rights; e.g., the amount of communication resources requested, or the time of day. However, the logical information focused upon here (source, destination, mode) satisfies a significant portion of high-level gateway usage control policies that are not application specific. This *logical* information about a packet, message, or connection is available to most higher-level communication protocols. In the approach outlined in chapter 5, the ION gateway uses this logical information to determine which destinations a packet, message, or connection may travel to, i.e., the access rights of the communication. In order to enforce non-discretionary policies in this way, the ION gateway must be able to evaluate the logical characteristics of communications based on the packet, message or connection-request headers. If only one outside organization uses the gateway, there is a default association between the external sources and organization affiliation. However, if the gateway is used by more than one organization, association must be achieved according to the name information provided.

In many cases associating communications with logical information is the most difficult aspect of implementing controls in an ION gateway. The solution is constrained by existing protocol and naming semantics which are not easily changed and most often were not designed with usage controls in mind. Internet Domain-style naming provides a means of structuring names so that organization affiliation can be easily determined for higher-level communication such as electronic mail and remote login connections. [40] However, Internet Domains were established to correspond to name management authorities which do not necessarily correspond to access control authorities.[65] Finally, depending upon the application, the gateway may need to employ authentication mechanisms such as those described in the previous chapter to authenticate the logical information which the gateway uses to make its policy decisions.

8.2.2 Information Management

Once the gateway has associated a packet, message, or connection with logical information, it must determine the corresponding category sets in order to enforce a non-discretionary policy such as that described in chapter 5. The gateway may maintain the category information locally or may request it from a separate policy server, as suggested in chapter 6 The following discussion assumes that the information is held in the gateway; however, the form and management is the same in either case.

The category information can be represented as a table. Depending on the symmetry of the policies applied to incoming and outgoing communication, one or two tables may be maintained. If the policies are asymmetric, an external host may be assigned one set of categories when it acts as a source of incoming communications and another set when it acts as a destination of outgoing communications. For example, if an organization is concerned about invocation of internal resources but is not concerned about flow of information to the outside, internal hosts may be assigned a restricted set of categories for incoming communications and an unrestricted set for outgoing.

[65] Note that domain-style naming alone does not offer a means of associating lower-level packets with access rights since packets carry numerical destination addresses only.

Traditional non-discretionary controls apply to two generic access types, read and write, and use the same category sets for both. However, an ION participant may want to discriminate among several different communication modes, e.g., mail, database query, file transfer, remote login, etc. Each mode represents a different logical network. If a gateway supports a single communication mode such as mail, then each external group or entity can be assigned a single category set. If the gateway supports multiple modes of communication to any host or group, then different category sets may apply to the different modes of communication, i.e., to the different logical networks. The same general non-discretionary framework can be used. However, a host that can be accessed using multiple communication modes may be assigned multiple category sets, one for each communication mode. These multiple entries represent the fact that there are multiple logical networks—one corresponding to each mode of communication.

Category information may be aggregated so that all hosts in an outside organization are assigned the same category set, or it can be very detailed so that each individual host is assigned an individual category set. The former imposes a greater burden on the association process described earlier, i.e., associating information available in the packet, message, or connection header with organization groupings. The second scheme requires maintenance and searching of a much larger amount of information and is equivalent to maintaining a set of capabilities for each user. In many cases, organizations treat outsiders according to their organization affiliation and role more so than on an individual basis. Individual identity is more relevant to the end applications accessed. Therefore, generally, the gateway should control access on the coarser level of organization, and leave more refined distinctions to the end applications that are being made accessible. In summary, aggregating category information is generally more appropriate where sources and destinations are grouped and assigned category sets according to organization affiliation, or subgroup affiliation within an organization (e.g., according to project or division).

An organization or subgroup is assigned categories on a need-to-know basis. Categories or compartments represent the finest grain of access desired. Some compartments may include only a single machine or mailbox, while others may include entire subnets. The fine grain is necessary because according to the intersect rule described earlier, if a resource is assigned a category, anyone belonging to a group that has that category in its category set can access the resource, for the specified communication modes.

For the sake of illustration, assume that XYZ is a customer of ABC. ABC could assign sales a single component so that XYZ would have access only to those resources that pertain to sales. Meanwhile, MIT is engaged in a joint research project with the research division of ABC and has access only to research-related resources. Finally, geographically-remote employees of ABC are permitted to communicate with all internal resources via the gateway. An example is a sales agent located at a customer site.

8.2.2.1 Assigning Categories

Assignment of categories and category sets is critical to any non-discretionary control system. Because it is so critical, assignment should be adaptable to changing environments and policies. The maintenance of all such information in the ION gateway(s), as opposed to in numerous hosts distributed over the entire internal network, facilitates this adaptability.

Each ION participant must classify outside entities into groups such that all members of a group have uniform capabilities with respect to internal facilities. A participant can then assign a single category label to each group. If an outside entity is a member of more than one group, its category set should include multiple categories. A group may include an entire outside organization, a division of an organization, or several organizations. The gateway includes a particular group's category label in the category set of each internal facility that that group is allowed to communicate with. In order for this method to be meaningful, the gateway must be able to evaluate and authenticate the group membership of messages and communication requests.

8.2.3 Control Policy

Earlier chapters argued that most usage control policies can be expressed in terms of non-discretionary access controls. In particular, the following non-discretionary rule was

			Resource
Organization	Categories	Mode	Limit
XYZ	Inventory	Query	x/day
MIT	Rsearch, Development	Mail	*
MIT	Research	*	*
Remote employees	*	*	*
Mail	Research	*	*
R&D Hosts	Research	*	*
Customer accounts	Internal only	*	*
Field service support	Internal only	*	*

Figure 8-1: Category set information maintained by ABC's ION gateway. * indicates all.

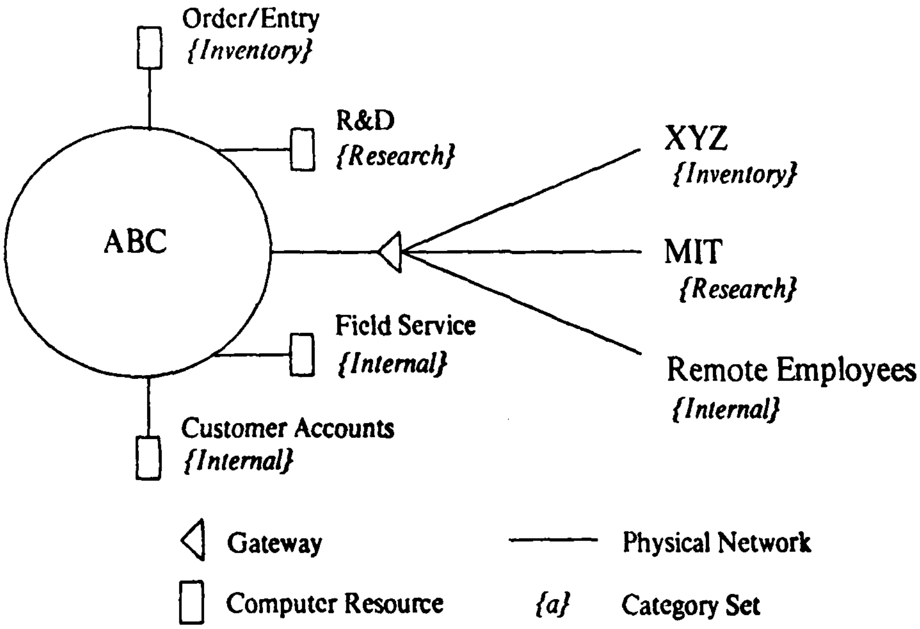


Figure 8-2: The ION that correspond to the above category information.

proposed: a source may send a message or establish a connection to a destination if and only if their respective categoty sets overlap, i.e., share some element in common. In terms of the evaluation procedure described above, the gateway uses the logical information associated with packets, messages, or connections to identify the category sets that correspond to the source, destination, and mode of the communication. The gateway forwards the traffic if the category sets of source and destination overlap for the mode. The gateway rejects the traffic if category information is not found for either source or destination, or if the category sets are disjoint.

8.3 ION Application-level Controls

Some usage controls are best implemented in the ION applications themselves because that is where the access-related information resides.

Table 8-3 illustrates several ION applications and the information needed to evaluate communication or transactions requests. The examples listed are ones in which the information needed for at least some of the policy decisions is available to the ION application, and not to the type of high-level ION gateway described above. In the first example, the operator of the online order-entry system provides each outsider with access to inventory data about select product lines only. Similarly, the operator limits the quantity and dollar amount according to credit rating and dates. In the Cad/Cam system each subcontractor can access a small set of components only, and can access only certain versions of those files. These restrictions depend in part on the type of contract and mode of access. The algebra system restricts users according to licensing and cpu time, the software distributor restricts updates to customers according to the customer's contract, different field service is provided to different customers depending upon their contract provisions and system configuration, and the information retrieval system filters according to the user's name and the particular file and record requested.

The controls required in the ION application are relatively standard. Discretionary or non-discretionary control mechanisms can be employed. The choice depends upon the nature of the policy and the mechanisms available in the ION application's environment. Aside from

ION Application	Non-Discretionary Control Criteria
Purchase order	Product type
	Dollar amount
	Quantity
	Date
	Credit rating of source
Cad/Cam or VLSI	Component type
	File
	Contract type
	Operation (read/write/append/modify)
Algebra system	CPU time
	Licensed
Software updates	File
	Operation
	Contract
Field service	Files
	Contract
	Configuration
Information retrieval	Records
	Files
	User name

Figure 8-3: ION applications and information needed to evaluate access.

the application-specific controls, the primary requirement introduced by the ION is that the ION applications prevent external communications from propagating to internal, non-ION applications and machines. For IONs in which the ION applications are not used for internal purposes, application-level controls can be relied upon more heavily because they do not interfere with internal operations.

Application-level controls designed before, or without cognizance, of external connections must be reevaluated if the application is to be made accessible to outsiders. For example, ABC Inc.'s international sales database system implements application-level controls which assume that terminals have unique IDs and can be associated with fixed physical locations. Therefore, terminal ID is used as a basis for access decisions to the extent physical security is trusted. In addition, some newer applications assume that users must authenticate themselves via a badge reader in order to use a terminal. When external users access these applications via an ION gateway these applications may wrongly assume that a user can be associated with a particular physical terminal or that the user has authenticated him/herself with some internal systems badge reader. A properly designed ION gateway should authenticate external users to meet the needs of internal, ION applications. However, in addition, owners of applications that are made accessible to outsiders must reevaluate the assumptions under which the existing application-level controls were designed.

8.4 Comparison of Visa and High-Level Gateways

Based on the implementation issues presented, this section outlines the trade-offs between visa and high-level gateway architectures described in Chapter 6. The two approaches differ from one another with respect to several of the ION gateway tasks described above; in particular, associating communications with logical information. The approaches also differ with respect to several performance parameters.

A high-level gateway can associate communications with logical information directly, or it can call an ACS. The visa scheme must employ a high-level dialog with an ACS to associate packets with logical information. Similarly, high-level gateways evaluate each connection or message according to programmed control policies. They may or may not apply some check to each successive packet in a connection. In addition to employing an ACS to apply a high-level control algorithm to each connection or message request, a visa-based gateway always checks each and every packet against the visa.

A more significant limitation of visas for packet forwarding gateways is that they must make decisions based on information in the packet header, which usually contains source and destination addresses only. The packet-level gateway must be able to evaluate the legitimacy of each packet based solely on the packet header and the visa. It is difficult, and sometimes impossible, to represent complicated policies in this manner. For example, there is no way for a packet-level gateway to discriminate on the basis of mode of access (e.g., mail, file transfer, remote login, etc.) because no information about higher application levels is available in the packet headers. Consequently, even if mode of access is indicated in the visa, there is no way for the gateway to verify that a particular packet is supporting one mode of access and not another since this information is not carried in the packet header. The same problem arises if the gateway needs to discriminate on the basis of user ID. For this reason, higher-level gateways are better suited to implementation of some types of policies.

The two schemes are comparable in terms of several cost and performance criteria—storage and trusted components—but differ significantly in terms of others—end-user performance and protocol modification. Storage requirements are the same for both, although a high-level gateway may store control information locally or in an Access Control Server (ACS) and a visa gateway by definition stores it in an ACS. In addition, the visa gateway stores locally a small number of currently-in-use keys, whereas the high-level gateway maintains more state information about the connections passing through it. In both cases, the amount of storage used for access control information depends on the grain of control, i.e., user, host, network, organization. The two approaches are also similar in terms of the number and extent of components that must be trusted. In both cases, security depends upon the authentication of header and connection request information, the evaluation program in the gateway and ACS, and the ability to subvert the access control mechanisms used to approve connections or messages. The latter risk is somewhat higher for gateways that do not authenticate each packet.

The two schemes differ most in performance overhead and the modification required of existing protocols. Each of the methods exacts a performance cost. The visa gateway is costly because of the required dialog with the ACS and the checks applied on a per packet basis The high-level gateway is costly because protocols are terminated and because the gateway must be programmed with each higher-level protocol that it supports. The tradeoff depends much on traffic patterns; in particular, the number of packets per session or message, the volume of traffic, and the number of communication service types. On the other hand, protocol conversion is hardest for the lower-level protocol because of the tighter real-time constraints; for the same reason protocol conversion is harder for connection-based gateways than it is for message based.

A second significant difference between the two methods is that visa gateways require that all internal systems that use the ION add the visa to the header or checksum calculation. This requires that each machine modify its low-level communication protocols. In contrast high-level gateways require that application-level procedures be changed; or, in some cases, only that name tables be updated (see section 8.1). Although the latter is less transparent to the end user, the cost and inconvenience of software modification is avoided. This cost can be quite high if it implies incompatibility with existing and future equipment. On the other hand, an additional cost associated with higher-level gateways is the need to program the gateway separately for each higher-level protocol that the organization wants to support; in contrast, the packet-level gateway supports all higher-level applications.

8.5 Evaluation of Mechanisms for the Examples

This section illustrates the implementation issues and tradeoffs by evaluating gateway designs for some of the cases introduced in Chapter 4. Most of the mechanisms proposed below do not actually exist; those that are in place are indicated.

8.5.1 MIT Dial-up Gateway

MIT currently employs two packet-level gateways—the dial-up and Arpanet gateways. The dial-up gateway is a candidate for the visa scheme as described in Chapter 6. The characteristics of this gateway and its application are such that the costs associated with visa schemes are minimized. First, the policies desired are quite simple and can be expressed in terms of packet header information. If MIT wanted to restrict access according to the application type (mail, file transfer, remote login) or according to an individual user's identity, this would not be the case. Moreover, the size of the external user population is rather limited so the expense of including visas in checksums is limited to that small population. Because MIT is concerned with controlling incoming traffic via this gateway, and not outgoing traffic, it need not be concerned with the much larger internal user population. The only serious cost is implementing the ACS and investing in encryption hardware for the gateway and all outside users of this gateway.

The function of the gateway controls is to specify the set of resources that each external user can send packets to. This set is determined by the group to which the user belongs. If the user is a member of MIT who happens to be geographically located elsewhere, that user should be able to send packets anywhere on the MIT network. If the user is a member of another organization that has been granted access to host *x*, that user should be able to send a packet to host x only. MIT could implement user-dependent policies in the shared gateway by assigning each group of external users and each MIT resource (host or device) a category set. The gateway would associate communications with logical information by calling an ACS where such category information would be held. Based on the category information obtained, the gateway would apply the intersect rule—dial-up user *U* can send a packet to MIT host (or device) *H* if and only if the intersection of *U's* category set with *H's* category set is non-empty.

The gateway would operate as follows:

1. A user dials up the gateway and the gateway associates the user with a particular port.

2. The user authenticates itself to the gateway as belonging to a particular category (e.g. MIT, ABC, WMBR). The port is thereby bound to policy information.

3. The gateway looks up the user's category set and assigns the set to the port that the user was given at dial-up time.

4. All subsequent packets that enter through that port, and that belong to the authenticated user, are forwarded to their indicated destination if and only if the destination's category set includes at least one of the elements in the port's category set (i.e., non-empty intersection).

5. When the user terminates the conversation, the category set associated with the port is cleared before the port is made available to other callers.

The organization or group affiliation of a user would be determined through a conversation with a higher-level, ACS, as described in the previous chapter.

8.5.2 MIT Arpanet Gateway

According to the procedure for implementing usage controls in an ION, described earlier in this chapter, incoming and outgoing traffic are bound to policy information about source, destination, and mode; category sets are identified and compared; and if they overlap, the packet is passed. In the Arpanet gateway a less general mechanism is acceptable because of the relatively simple policy required—prevent flows from educational users to Arpanet. The gateway can assume a fixed association between organization affiliation and category set, i.e., research has full, and education has no, Arpanet access. The gateway's task therefore simplifies to binding traffic to organization affiliation and determining access based on a simple pass-research-traffic-only rule. However, because the gateway operates at the packet level, this task is difficult as described in the previous chapter 6. In order to bind packets to organization affiliation, the gateway must either use a visa-type scheme, or maintain a table locally that maps subnet and host numbers onto organization domain, e.g., Education, Research, and Arpanet. Unfortunately, unlike the dial-up gateway, the concern is with outgoing traffic and the size of the user population is very large. Therefore the cost of implementing a visa scheme—encryption facilities and modified network software for all users—would be very large. However, the second approach, of maintaining the association of network address to organization affiliation locally, is only viable for small user populations.

The MIT-Arpanet gateway provides a good example of the difficulties associated with implementing controls at the packet level. A higher-level connection would greatly facilitate the binding of traffic characteristics to policy information. In addition, controls applied on a per-message or connection basis, instead of per packet, reduce overhead. On the other hand, the major cost of a high-level gateway is the performance experienced by the end users because protocols are terminated in the gateway. A compromise between the generality and performance of a packet-level gateway and the convenience of a higher-level gateway is a visa scheme as described for the dial-up gateway and in the previous chapter. Alternatively, interactive protocols such as remote login could operate via a visa gateway while mail could operate via a high-level, mail-forwarding gateway. This would allow those hosts that communicate with the outside via mail only to avoid updating their communication protocols to accommodate the visa scheme. These hosts would send all external mail via a designated mail-forwarder that would in turn implement appropriate policy controls at the mail level.[66] Only those hosts that use remote login and interactive file transfer would need to update their protocol software to use the visa gateway. The designated mail-forwarder would also update its lower-level protocol to make use of the visa gateway. Moreover, only some of the hosts that use connection-based, Arpanet applications need direct Arpanet access. The remaining, more casual users, could use the directly-connected hosts as high-level gateways. The Arpanet gateway itself could then be simplified to a small fixed table of approved network IDs.

An alternative to putting controls in the Arpanet gateway at all is to put controls in the gateways connecting the Athena educational network to the MIT research network. The advantage of this approach is that each Athena gateway serves a much smaller community than is served by the Arpanet gateway. Consequently, the binding of packet source and destination numbers to organization affiliation is easier since the gateway needs distinguish between a smaller population of sources. Nevertheless, there are several problems with this approach that suggest that the controls belong in the Arpanet gateway.

1. Regulating the flows between MIT and external domains should not impose regulation on the flows within MIT. It is in the interest of MIT to facilitate communication and integration of the research and education communities.

2. Policy information and mechanisms must be easily checked, updated, and modified. Although MIT can establish guidelines for how ION gateways must

[66] As stated earlier in this discussion, we assume the existence of some administrative policies and procedures requiring that external network connections be registered or approved in some fashion. Consequently, we assume that all external mail would be sent via known ION gateways.

be managed and the controls they must implement, it is not practical or desirable to require that each internal subnet gateway be modified every time another ION gateway is established or modified, or each time an external policy requirement changes. Therefore, the mechanisms that enforce a particular policy should be located as close as possible to the administrative boundary that imposes the policy requirement.

3. As the number of external network connections increase, the policy requirements of the different IONs may conflict with one another and the gateways between mit-subnets will not be able to satisfy all of them without compromising capabilities with respect to others; or the rules will become burdensomely complex.

8.5.3 Multics

Multics' role as a vehicle for external communication is by definition a high-level gateway—users must establish high-level connections with Multics before communicating with other MIT entities. Each Multics user has an account ID. In addition, each user is given a project classification which could correspond to organization affiliation, e.g., UCLA, CERN, DCA, Proteon. MIT internal facilities include MIT hosts, printers, and other servers, in addition to the several gateways to external networks, i.e., Arpanet, Bitnet, Usenet, Telenet gateway, Telex gateway. The modes of communication supported between Multics and other MIT sites via the local network include mail, remote login, and file transfer.

As described earlier, the non-discretionary policy that we want to enforce in the gateway is not the traditional security policy. We want to implement a policy that allows invocation or information to flow from A to B if the category sets of A and B overlap, i.e., have a non-empty intersection. This rule does not achieve strict confinement and isolation of categories and therefore can not be implemented under the traditional security policy. Consequently, it can not be implemented under Multics with Access Isolation Mechanism (AIM) which enforces the traditional security policy. The remainder of this discussion will assume therefore that AIM is turned off and an alternative set of non-discretionary control mechanisms are activated, which are similar to AIM but with a different non-discretionary policy.

For each communication request Multics has access to the source, destination, and mode information needed to determine category sets. For example, mail messages include source and destination text names which can be associated with organization affiliation if guidelines are followed in assigning project IDs in Multics. Similarly, connections established for file transfer or remote login require explicit network commands using the textual name of the source and destination sites which can be evaluated and associated with to organization affiliation. In addition, the mode of communication invoked is apparent from the command of the user. In other words, all packets sent over the internal network can be associated with a command and a particular user, organization, or group.

Figure 8-4 illustrates the category information that Multics could maintain to implement non-discretionary controls. The categories defined for this environment are commercial, university, international, and government (* indicates all of the above). Multics looks up the category sets associated with the individual source and destination textual names. If the individuals are not listed, it looks up the names of groups (e.g., organizations) to which the source and destinations belong, respectively; according to the textual names. If category information is not found for either source or destination, individual name and group name, then the communication is rejected.[67] An internal site without any registered category information is also treated as inaccessible to outsiders. Similarly, an external site without any category information is treated as strictly external (i.e., not accessible to, and not able to access, insiders). When the category set assigned to a MIT facility differs for different communication modes, two entries are listed; alternatively, two separate tables could be maintained, one for each communication mode.

8.5.4 MIT Administrative Mail

If a gateway is used to connect to only a single organization then the association between source ID and organization affiliation is fixed. Similarly, if the gateway is used for only a single mode of communication, then the association between category sets and mode is

[67] Alternatively, the gateway may query a policy or ACS before rejecting traffic.

Organization	Categories	Mode
MIT	*	*
Proteon	Commercial	Mail
DCA	*	Mail
UCLA	Research	Mail, File
CERN	International	Mail
CERN	Science	*
French PTT	International	*
LCS hosts	Research	*
EECS hosts	Educational	Mail
Athena hosts	Educational	Mail
MIT admin	MIT-only	*
UUCP gw	Commercial,Research Gov't,International	*
Arpanet gw	Gov't,Research	*
Bitnet gw	Research,Educational	*
ScienceNet gw	Science,Research	*

Figure 8-4: Example of category tables for a Multics system connected to multiple networks.

fixed. If an ION gateway has both these characteristics, its construction is simplified immensely. One example of such an application in the MIT environment is described below.

The headquarters of MIT's Laboratory for Computer Science wants to be able to send and receive electronic mail from other laboratory members. However, the personnel and budget information maintained on the headquarters computer is too sensitive and the operating system's security is too weak to make direct network connection tolerable. By inserting a mail-forwarding gateway between the headquarters' and the laboratory's local area networks, headquarters can achieve interconnection at the mail level without exposing its computer system to an environment against which it can not protect itself adequately.

Because the headquarter's mail gateway is "hard-wired" to a single outside organization, i.e., the rest of the laboratory, and because it communicates in only one mode, i.e., mail, a single category set can be applied to all communication. In short, this single-organization, single-application ION gateway requires none of the category information described above. All it needs to do is attend to the security issues of whether in fact only a single organization is able to communicate with the gateway and whether in fact that organization is able to communicate with the gateway in mail mode only. A prototype of a mail-forwarding gateway implemented on an IBM PC-XT is described in [25].

8.5.5 R&D Mail Gateways

Both ABC Inc. and XYZ Inc. would like to support mail and some file transfer and server access to universities with which they collaborate. However, much of the information and resources kept on their respective internal R&D computers is quite proprietary. Moreover, many industry competitors are also connected to these shared research networks. Some internal information and resources can be made accessible to outsiders on a special contract basis. Other information and resources can not be made accessible at all without endangering trade secret rights. An additional constraint on ION usage is imposed by the operators of the shared research networks. Namely, the participants must not allow other non-participants to which they are connected to use the participant as a transit path onto the shared research networks.

Currently ABC and XYZ Inc. use high-level mail gateways to interconnect their internal R&D networks to CSNET. Given the importance of protecting proprietary information in these environments, mail is the only widely-available service type that the organizations support. ABC Inc. is more protective of these resources and implements controls that allow only registered users to send and receive CSNET mail; users register with their local site manager and are permitted access if they can justify an R&D related need. XYZ Inc. implements no controls in the mail gateway so that all internal users may send and receive mail. As network use increases, and the table of individual users grows, ABC could move towards a per-group-access scheme, in place of the current scheme that registers individual users. In that case, it would need a mechanism for associating header information with group affiliation, and category information similar to that described for Multics (See figure 8-5.

Organization	Categories	Mode
Corporate R&D	Corporate	*
UCLA	University	Mail,File
CERN	International	Mail
MIT	University,Athena	Mail
MIT	Athena	File
Corp R&D hosts	*	*
Athena hosts	Athena	Mail
UUCP gw	University	Mail,File
Telex gw	Corporate	Mail
Bitnet gw	University	*
CSNET gw	Corporate	*

Figure 8-5: Example of category tables for a a corporate R&D organization connected to multiple research networks.

Both ABC and XYZ have internal servers that are invocable via electronic mail. ABC Inc. is protected from unwanted invocation because only registered mailboxes can receive mail—the default is no access. In contrast, all message-invocable servers on XYZ's internal network are vulnerable because the default is open. XYZ desires to maintain the open default for person-to-person communication and therefore can not implement a closed default for servers that are accessed via the same protocol and naming structure. One solution is for the gateway to tag all incoming messages from outside the organization and require that the message-invocable servers internally check such tags if they contain sensitive information or services.

A similar gateway for connection-based communications could be implemented for ABC's connections to subcontractors and XYZ's various customer and vendor connections. The design tradeoffs are similar to those described for the MIT Arpanet and Multics examples of high and low-level connections, respectively. Alternatively, many of these applications may be supportable with a message-based invocation paradigm. The advantages of this approach were described in section 6.4.2.

8.6 Conclusion

The most difficult aspect of implementing ION gateways is the association of communications with logical information. Aside from this difficulty the major implementation decision is whether to interconnect at the packet level and employ an ACS and visa scheme, or whether to interconnect at higher levels and employ structured naming. Each approach is well suited to different environments and may be used in conjunction with one another in some cases. Finally, some loosely coupled connections may be best served with message-based high level interconnections.

Chapter Nine

Technical Conclusions

The preceding chapters characterized IONs, and the security and network interconnection issues raised therein. In the area of security and usage controls we characterized a set of non-discretionary control requirements that were not addressed by traditional non-discretionary controls (i.e., invocation control, protection of invoked parties, two-way communications applications) and applied category sets and a simple intersect rule to address these requirements. This approach was integrated in a design that allows strictly-internal applications to be logically isolated from external interconnections without requiring physical isolation, and demonstrated the approach with design studies of existing IONs. The technical mechanisms are described in the context of administrative guidelines and controls under which they operate. In the area of network interconnection we identified and characterized a class of applications for which performance criteria alone are not adequate for selecting the interconnection method and evaluated higher-level and visa-based interconnections as alternatives to packet-level interconnection. Message-based, high-level gateways were emphasized as a means of achieving loose couplings across organization boundaries.

The concepts presented can also be applied to *intra*-organization applications. Most large organizations are composed of subdivisions which exhibit many of the characteristics of separate organizations. When these subdivisions interconnect their respective computational resources and communication networks to support interchange, they can be considered to form an ION; even though a common administration exists, and accountability among participants is greater than in the case of distinct organizations. Even when the organization is not formally subdivided, communities and regions of users typically exist. Each community may not perceive a need for controls on internal use. However, as was described at the beginning of Chapter 6, efficient network usage may favor differentiating between these communities. Searching for a resource is an application in which the end user may desire to confine the search to a local environment, both geographically (e.g., a near-bye printer) and organizationally (e.g., one that my account number is good for). In addition, recognizing these regions can allow an organization to allocate communication resources more efficiently; for example by employing lower cost interconnections at those points where traffic is lower. One example of a lower cost interconnection is message-based interconnection described in chapter 6.

The first part of the thesis asked the question "how is inter-organization interchange impacted by the use of interconnected computer networks?" Subsequent chapters addressed the dual question "how should and can usage controls and network interconnection protocols be adapted to the requirements that arise when network interconnection supports inter-organization interchange?" We now return to the first question and describe the use and impacts of IONs in particular domains. In the following chapter we illustrate the components of the model for distribution channels. Then in Chapters 11 and 12 we describe in significantly more detail the use of IONs in R&D.

Chapter Ten

IONs in Distribution Channels

Having discussed the design of IONs to fit organization boundaries we return to our discussion of how the technical characteristics of this medium effect the relationships and communication patterns among ION participants. Chapter 3 described our model in general terms. However, the implications of ION use are contingent on environmental factors and it is most useful to discuss IONs within the context of particular domains. One domain in which IONs are and will increasingly have a significant impact is distribution channels. Moreover, the general model described has much to offer this domain in both descriptive and predictive power. Examples of distribution channel applications include: airlines to travel agents, hospital suppliers to hospitals, subcontractors to manufactures in automotive, integrated circuit chip, and computer systems design. In this chapter we illustrate how the model applies to such relationships. We describe general characteristics of distribution channels, the changes in communications and cross-boundary activities predicted by the general model, and two examples of customer-supplier IONs. In chapter 11 we describe in far more depth IONs in a second domain, Research and Development Laboratories.

10.1 Distribution Channels

Distribution channels are composed of manufacturers, wholesalers, retailers, and consumers. Among these participants, the several functions are carried out: [67] carry inventory and physical distribution; selling; after sale service; and extending credit to customers. The use of IONs may result in a reallocation of functions among distribution channel members.

The individual customer-supplier relationships within a distribution channel are inherently asymmetric. Suppliers provide non-liquid assets (goods and services) in exchange for highly-liquid assets (currency) from customers. Consequently, the predicted impacts of IONs on customer-supplier relationships are a tailored subset of the impacts described in the general model.

Porter's competitive analysis framework for value added chains is quite germane to the study of distribution channels. He characterizes the competitive environment according to the following parameters: buyer power, supplier power, substitution, entry, and intra-industry rivalry. Cash describes how IONs can impact these characteristics of distribution channel relationships. [12] Bargaining relationships among buyers and suppliers may be affected through changes in selection criteria, threats of backward and forward integration, switching cost changes, and product differentiation. Entry barriers may be increased through economies of scale, increased switching costs, product differentiation, and restricted access to distribution channels. Rivalry among competitors may also be affected by restricted market access, changes in cost effectiveness, and product and service differentiation. Finally, substitution may also be affected due to redefinition of products and services.

10.2 Implications of ION Use in Distribution Channels

As described in the general model, suppliers can use increased efficiency of ION-based communications to reduce operation costs of order entry and other communication-intensive customer service. In addition, faster feedback on customer orders can support tighter production control. Suppliers can also use broader capabilities to enhance existing products and services by introducing information, maintenance, and other online services that were not feasible previously. Meanwhile, given lower per-transaction costs and more information from suppliers, customers can reduce inventory costs and increase flexibility by ordering in smaller lots with shorter lead time—just-in-time inventory management. These behavior changes correspond to increased intensity and scope of the communications. Increased scope can also include new information and design or product selection support facilities, and new value added products altogether. At the same time, IONs can make communications more segmented for the customer if the ION cannot be used to communicate with competing suppliers.

Almost by definition, suppliers do not willingly decrease the number of organizations (customers) with which they transact, nor do they shift activities out of the market. Therefore, the primary change in cross-boundary activities that a supplier might use an ION to support is an increase in the number of customers reached and in the market share. The increased efficiency and capabilities of the ION can increase the effectiveness of a sales force and therefore may allow the supplier to support a larger market. In addition, the supplier can design the ION to prevent its use for communication with competing suppliers.[68] As a result of this reduction in universality, customers may find it relatively more costly or inconvenient to transact with other competing suppliers, thereby increasing the ION-providing supplier's sales per customer. In other words, customers may segment their transactions into those that are and those that are not ION-supportable. To the extent they favor the latter, the customers may find themselves transacting with a smaller number of suppliers, i.e., only those that have ION access. Suppliers may also increase restrictions on customer relations by asking customers to sign agreements regarding usage of the system or of facilities provided by the supplier. If fixed cost is high, a supplier may subsidize the ION facilities for large customers only and therefore other changes attributed to the ION would apply to those large customers only.

Another form of increased cross-boundary activity is the introduction of new products that previously were not feasible or economical to distribute in the market. Suppliers can use ION facilities to distribute information and resources as stand alone or auxiliary products that previously could not be distributed in a timely or cost-effective manner. Furthermore, IONs reduce communication and coordination costs and thereby allow a manufacturer to support direct service to consumers. Such forward integration in the distribution channel, i.e., bypassing distributors, is another form of shift in market governance, but in the other direction, from external to internal (vertical integration).

In contrast to suppliers, customers do vary the number of suppliers transacted with and do

[68] The non-universality of an ION facility can be designed into the ION or it can result due to lack of planning (i.e., due to lack of an available standard).

make choices between internal and market governance. Therefore, a primary opportunity introduced by IONs is to increase reliance on cross-boundary activities (e.g., buy over make, or joint ventures) and purchase from a larger number of suppliers, if the ION is not specific to a single supplier. Customers may also increase the restrictions imposed on supplier relationships to cope with new risks introduced by the ION. For example, in the case of a cad/cam connection between manufacturer and subcontractor (customer and supplier, respectively), risk is high due to the internal value of computer-based information and resources. Therefore, the customer has greater requirements for restrictions on subcontractor interchange, and might limit the number of ION-supported subcontractors. On the other hand, in the case of traditional product purchase (e.g., medical supplies, grocery stock, etc.), there is less concern about the proprietary nature of online information and therefore less perceived risk, allowing the customer to use the increased efficiency of the ION to survey a larger market, if non-standard ION facilities are not imposed by the supplier.

The eventual impact of IONs on cross-boundary activities depends upon the characteristics of the particular industry; the unrealized production cost advantages of cross-boundary activities, the nature of the internal facilities made accessible, and the level of decision making attention to ION issues. For customer-supplier IONs, the result can be complementary benefits for both customer and supplier, or it can be a shift in bargaining power or cost burden from one to the other. The outcome is determined largely by which party initiates and retains control over the ION.

10.3 Examples

More work is needed to elaborate the model's analysis and predictions for distribution channels. I describe the major impacts reported in trade literature and by Barrett and Cash and others for two cases, the hospital supply and the airline industries. These examples are outlines for studies that could be carried out to test the model. The examples do ***not*** represent systematically collected data!

10.3.1 Hospital Supplies

American Hospital Supply Corp. (AHSC) ervices close to 3000 hospitals with automated purchase orders, inventory control, and record keeping. [50][69] The technology is relatively simple, terminals connected over telephone lines to a single time sharing system.

The speed and turn-around of online purchase inquiries and orders is improved as compared with traditional paper and telephone media. Similarly, customers experienced lower incremental cost per order in terms of preparation time, and for the supplier in terms of eliminated data entry. Capabilities include order-entry and inventory support. This major industry supplier does not employ standard facilities and therefore the facilities cannot be used with other suppliers.

Enhanced speed and capabilities allow orders in smaller lots and thereby reducing inventory for customers. One customer, based on the new 24 hour delivery period, reportedly reduced inventory from 80 to 33 day supplies. Intensity is increased. Scope increased to a less quantifiable extent in the form of new services and information related to products. Penetration also increased somewhat since the supplier has more market data on the customers habits and inventory needs. Segmentation is higher due to the combined effects of ION advantages and lack of universality. For example, in one case buyers so preferred ordering via the online system that they did so even when prices were lower elsewhere. [4]

The size and volume of the suppliers market is increased whereas the size of customers' markets is decreased, or at least no greater due to non-universal facilities. Cross-boundary activities are not necessarily increased for customers since in this industry, most customers do not consider the make option on products because production cost advantage overwhelms the transaction costs independent of the ION. Customer's market size was reduced because the system convenience discouraged seeking of competitive bids as time-consuming and expensive. The ION has supported new product offerings by the supplier in the form of ancillary information and services. No known restrictions or controls have

[69] Information for this section was also obtained from Mr. E. Doerhoefer in an MIT Laboratory for Computer Science Seminar in 1983.

increased. In any case, the risk is relatively low given the nature of the information and resources accessed. The supplier introduced and controlled the network and treated it at a relatively high level of management as a strategic tool. The customer reacted and only upon encountering problems did the issue elevate beyond a procedural development.

10.3.2 Airline/Travel Industry

Another example of a distribution channel ION is the connections between airlines and travel agents. The suppliers (airline companies) use online reservation systems to provide a larger range of information in a more timely fashion. Consequently the travel agent potentially can access more flight information, more easily and economically, and in addition can scan a larger number of airlines when satisfying a clients request. However, most travel agents use systems that belong to a single airline company. Although these systems provide information about other airlines, the systems are somewhat biased in favor of the ION providing airline. Consequently, the number of airlines actually considered for a transaction may be limited if not decreased as a result of these biased online systems.

In addition to improved speed and turn around, capabilities are increased because of the greater amount of more detailed and timely information available from a wide range of services. However, the various forms of bias built into the airlines reservation systems reduces universality since the purchaser does not have equal access to all suppliers.

In addition to increased intensity, scope is greater as a result of new types of information and services (e.g., hotels, cars, travel packages). Segmentation is greater to the extent non-universality in the form of system biases are effective in making travel agents and customers favor the ION-supplying carrier. Penetration is not increased significantly since the nature of the information and resources accessible are not of significantly greater internal value or internal nature than non-ION access.

Customers reduce their effective market to the extent non-universal facilities leads to greater segmentation. Once again the production cost advantage of the airline makes customers' make/buy decisions irrelevant. However, for the supplier, it is possible to view the move away from travel agent mediation to direct end-user access to reservation systems as a form of forward integration. In addition, new services offered by carriers are made possible by the ION and can be viewed as increased cross-boundary activities of a sort (e.g., travel packages). Restrictiveness is increased along with segmentation in the form of travel agent contract conditions (e.g., that a certain percentage of reservations be booked via the ION). Controls are not necessarily increased since the internal nature of access resource, and associated risk are low. As with the previous example, the supplier initiated and controls the network and treated it at a strategic level whereas most travel agents reacted at a procedural level.

10.3.3 Additional Examples

Other industries also provide interesting examples of ION impact which should be investigated further.

- Grocery: The electronic data interchange standard for invoices and purchase orders supports enhanced speed, turn around, and incremental costs but capabilities initially are not changed. Universaiity is relatively high because standardization preceded ION adoption. Based on these characteristics intensity is likely to be greater but scope and penetration are not affected significantly. Segmentation is low given the standardization process but participation rates are not known. There are significant opportunities for expanded numbers of interchange partners (expanded competition?), and moreover, for forward and backward integration in the distribution channel. [1]

- Banking: Electronic funds transfer among banks and online financial services to customers are two forms of IONs. The speed and incremental costs of EFT supports more intense inter-bank coordination. However because of the extensive regulation and rigid institutional structures the implications for inter-bank activities requires special investigation. Online services to customers—in particular, large business customers—resembles other customer-supplier connections. Enhanced speed and capabilities support more intensive communication of greater scope, while the non-standardized equipment and interfaces increase segmentation of ION and non-ION supporting banks from the client's perspective. Depending on the extent of this segmentation and other market factors, the client may or may not change the number of banks dealt with. Similarly, the extent of enhanced capabilities and scope may or may not lead to new banking activities.

- Automotive Dealers: Online connections between manufacturers and dealers support speedier communication at lower incremental cost, as well as new capabilities of accessing online automobile information. In addition to intensified communication, greater scope can be supported including customer-defined car configuration and locating of unavailable parts, warranty information, and even accounting assistance for dealers. Dealers are often dedicated to a single manufacturer anyway so the effect of segmentation is not detectable.

- Automotive Subcontractors: Online exchange of design specifications between auto manufacturers and their parts vendors supports greater speed at lower incremental cost, as well as capabilities to exchange machine readable cad/cam designs and updates. In addition to more intensive interaction in the design process there is potential for penetration and segmentation to increase significantly. However further investigation is needed in both these areas. Small vendors without the cad/cam capabilities may be affected if manufacturers narrow their purchases to online accessible vendors only; however, third party services can alleviate this barrier.

10.4 Issues

In these market environments IONs raise several additional questions:

1. Is it feasible or appropriate to automate non-routine, in addition to routine transactions? And what is the impact of automating routine transactions only?

2. How real is the threat of suppliers tying in their customers through non-standard facilities? How rapidly will third parties and standards obsolesce these tactics ?

3. Does the tighter coupling between neighbors in a distribution channel make one organization's information and policies more relevant to the strategy of its neighbors?

4. What will be the role of third parties in IONs? What qualities favor the use of private and internal facilities over, shared, third party facilities?

The following chapters investigate more deeply the role of IONs in R&D relationships.

Chapter Eleven

The Use of IONs
Among Research and Development Laboratories

The previous chapter described how the use of IONs can affect inter-organization communication and interchange in distribution channels. This chapter describes ION use in another domain—research and development (R&D) laboratories. The next chapter describes an empirical study conducted in this domain to test many of the propositions and predictions set forth in the general model.

There are several reasons for choosing R&D as a starting point for evaluating the general model. In addition to the accessibility of a large number of organizations that use IONs, the users and applications are varied and range from simple administrative scheduling to more sophisticated resource sharing. This wide range of applications allows investigation of the contingencies described in the model within a single study. In contrast, most other domains currently exhibit only a single ION application, e.g., airline reservations, banking funds transfer, or customer order-entry. Therefore, although there are several reasons why the R&D organizations studied are *poor* models for more competitive market environments, the diversity and integrated nature of the applications makes R&D a rich domain in which to begin such investigation.

One of the limitations of conducting our study in this domain is that R&D laboratories have uncommonly porous organization boundaries—communication and resource sharing are vital to the survival of any laboratory. Similarly, competition is a less dominant factor in dealings among R&D laboratories—even among commercial labs— than in traditional market activities. Similar to non-profit organizations and public welfare agencies, etc., R&D relationships are non-economic in character. [76] They are not governed by markets based on prices, nor are they internal. Consequently they do not fit traditional Williamson and other economic models as well. A third factor is the sophisticated nature of the computer usage in the R&D laboratories studied.

The remainder of this chapter characterizes R&D laboratories very generally and describes communication patterns and resource flows. The parameters of the general model are evaluated for this domain and the IONs used among R&D laboratories are described—historical, functional, and technical aspects. Based on this characterization of the context and the networks, the general hypotheses put forth in the previous chapter are presented for this particular domain. These hypotheses form the basis for the study described in the next chapter 12.

11.1 Characteristics of R&D Laboratories

This section briefly outlines the general functions, organization structure, resource and information flows, and communication patterns typical of R&D organizations.

11.1.1 General Functions

The functions performed by the R&D laboratories studied range from theoretical research to advanced experimental design of new hardware and software structures, to development and engineering of information systems; all of the laboratories focus on computer-related areas. The "end-products" of the research are not the end-products of the companies nor universities, rather they are the research components. These products and services are in the form of reports, tools and techniques, solutions to problems, etc. The market for these products is composed of other R&D laboratories, product development groups, and large commercial and government users.

The degree of risk or uncertainty in research is universally high but varies with the maturity of the area as well as its location in the basic-applied spectrum. Most research laboratories cover a large variety of subjects, and therefore participate in several markets. Resource availability also varies across the organizations studied but is often lower in universities. Similarly, the number of employees and operating budgets vary widely but are often higher in industry.

11.1.2 Structure of R&D Organizations

In university research the number of supervisory levels is small. Typically, a faculty member directs a small staff, and is quite autonomous. Consequently, administrative intensity is relatively low and the faculty member, who is also the researcher, holds most of the decision making power. Similarly, within a university department, job titles and research areas are loosely defined according to the general area of research.

Commercial labs remain less formally structured than other aspects of the commercial world, but are more bureaucratic than universities. Usually there are at least two management levels and a wider variety of job titles. Departments are also more strictly defined. Consequently, administrative intensity is higher than in universities and less of the ultimate decision making power belongs to the researchers themselves. However, actual distribution of power between management and research staff varies significantly from one organization to the next. Another common difference between the two types of laboratories is that universities typically have only one geographical operating site; whereas commercial laboratories often have at least two. Geographic dispersion can influence communication patterns significantly and encourage the use of computer network technology *within* the organization for inter-site coordination.

11.1.3 Resource and Information Flows

Resource and information flows from commercial to university laboratories include: equipment, research funds, access to computational resources, manufacturing and engineering resources and expertise. In contrast, flows from university to commercial laboratories more often include: resources of a very experimental nature, people and expertise in the form of consultation or training. The intensity and quality of these flows varies across organizations and pairs of organizations. In addition, information and resource flows can interact with one another—increased information flows may enhance a laboratory's ability to assimilate expertise from the outside and thereby reduce the need for activities such as joint ventures which make accessible outside personnel and physical resources.

11.1.4 Communication Patterns

In applying our model to the R&D domain, we refer to Allen's characterization of R&D organizations, communications among them, and in particular the differences between commercial and academic laboratories. [2]

Allen illustrates that no R&D laboratory can be completely self-sufficient, that every laboratory must "import information from its environment in order to sustain life." No matter how resource and expertise rich a laboratory is, the nature of scientific and technological development is such that it does not grow in a single location without input from related activities elsewhere. Given that all R&D laboratories operate as somewhat open systems, Allen distinguishes between two aspects of the process of importing information: acquisition from the outside and dissemination within the organization. The concern of this research is on inter-laboratory flows and therefore it focuses on acquisition.

Allen studied the time spent with outsiders, ideas acquired from outsiders, other problem solving functions involving outsiders, the relationship between project performance and reliance on outsiders, and the difference between university and commercial laboratories. The findings that are most relevant to our study of IONs in R&D are summarized below:

- Sources: Vendors made up the largest group of outsider sources. Second was unpaid consultants such as researchers from government, non-profit, and university laboratories; these relationships were largely informal and brief. Competitive pressures precluded most inter-industry consultation. The third and rarest source was paid consultants.

- Project engineers spent 5% of their work time and 33% of their communication time with outsiders.

- A greater number of ideas for solution of problems was acquired from outsiders than from all other sources combined. Allen noted that this preference for outside sources is not necessarily based on the unavailability of information internally.

- Other problem solving functions served by outsiders were generating criteria for problem solution, setting the limits of acceptability of solutions, and testing alternative approaches.

- Project performance was found to be inversely related to outside communication. However this relationship was not thought to be causal. Poor performers made much of poor use of outside communication, thereby skewing the data. In addition, the problems were due to the poor quality of these communication channels. This issue is described further below.

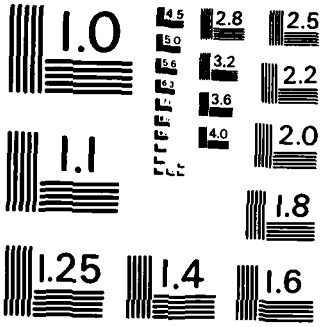
Allen also investigated the characteristics that most differentiate commercial from university laboratories. He found outside channels of communication to be consistently more effective for university than for commercial laboratories. He attributed this difference to two factors. First, most commercial laboratories are somewhat typical bureaucratic structures, i.e., hierarchically organized, clear division of labor, work procedures, differential rewards by position, etc; whereas universities are composed of autonomous departments and relatively independent faculty. Formally structured bureaucracies tend to have formally structured boundaries. These boundaries serve to isolate the organization from the outside. In the process, the isolated organization develops specialized approaches and expressions that help communication ideas rather efficiently within the organization. However, this same specialization makes it harder to communicate effectively with outsiders. The second difference is that the proprietary interests of the commercial laboratory leads it to demand the employees' loyalty; whereas a university faculty's loyalty is first to his or her colleagues in the discipline, wherever they may be, and second to the university. In other words, researchers in the commercial laboratory are first a part of the social structure of the lab and therefore that is where the primary communication transpires. University faculty are first a part of the social structure of their discipline so their primary communication transpires with others in the discipline, across university boundaries.[70] These factors together result in reduced and weaker outside communication for commercial than for university laboratories.

Allen also found strong evidence of a *gatekeeper* role in commercial laboratories. A gatekeeper is an individual in the organization who disseminates information acquired from the outside to other researchers inside the organization. Allen found that the gatekeeper role is particularly relevant to commercial laboratories and not so much to universities. This

[70] Allen argues that in fact, the relevant organization boundary in the case of university researchers is not the university but the discipline.

AD-A160 855 ACCESS TO INTER-ORGANIZATION COMPUTER NETWORKS(U) 3/3
MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR COMPUTER
SCIENCE D L ESTRIN AUG 85 MIT/LCS/TR-345
UNCLASSIFIED N00014-83-K-0125 F/G 5/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

gatekeeper phenomenon is relevant to the study of IONs because it affects the number of outside communicators found in commercial labs as compared with the number found in university laboratories. Because the gatekeeper plays a dissemination role, the information he or she collects is more likely to spread to others than is that of a typical communicator in a university. Consequently the number of people in a commercial laboratory who have contact with outsiders may be considerably smaller than the number of people in a university laboratory who have outside contact; and yet the information may be disseminated more widely and directly within the commercial organization.

11.2 Traditional Media, Communication, and Cross-Boundary Activities in R&D

In preparation for applying the general model of change associated with ION adoption to R&D laboratories, this section describes the communications and cross-boundary activities typical of R&D labs in the absence of IONs. The description focuses on differences between university and commercial labs with respect to the dimensions of the model.

R&D labs traditionally communicate via written documents, written correspondence, face-to-face contact, and telephone; telex has some application internationally. The communication and interchange patterns of inter-lab communication vary with the organizations, research projects, researchers, phase of endeavor, and many other factors. This section makes no attempt to characterize the absolute intensity, scope, etc. of inter-lab communications. Instead, this section uses the differences between commercial and university laboratories as a vehicle for summarizing the pre-ION state. The differences are generalized statements; exceptions exist to each one of them. The importance is that the general cases differ in interesting ways.

Commercial laboratories typically carry out a smaller proportion of their projects across organization boundaries than do university laboratories because of the need to retain control over research products in order to benefit from them commercially. Information is exchanged with outsiders only at very early stages of research, before it is ready for exploitation in a product, and after product release. In contrast, universities are public institutions and within the limitations set by individual researchers, there are no formalized restrictions on information exchange at all stages and aspects of research. Similarly, commercial labs engage a more restricted set of interchange partners because of the need to establish ownership terms of agreement, and codification and restrictions on information and resource flows is greater. One way in which external communication in a commercial lab is controlled, *implicitly*, is through the gatekeeper roles described by Allen and summarized in the previous section. In commercial labs, the number of individuals that communicate regularly with outside entities is proportionally smaller than in university labs. These *gatekeepers* then disseminate the information within the commercial organization. Consequently, the portals to the outside world are fewer and somewhat more controlled than in university labs. At the same time intensity may *appear* to be reduced—because of the smaller number of communicators—but in fact because of the fanout from the gatekeeper to lab members, the difference in intensity may appear larger and more significant than it is. In general, in commercial laboratories, work related communication and interchange with outsiders requires management approval of one sort or another. Different companies place responsibility at different levels of management but it is universally higher than for universities. At the same time, in commercial labs, the value of communication with other laboratories is recognized by higher levels of management, particularly in areas experiencing rapid technological change. In addition to limiting penetration of external communications, these restrictions can limit the scope of information and resources flows, as well as the intensity. On the other hand, the more restricted is the set of partners with whom joint research is carried out, the more segmented the interchange is likely to be between members of that small set, and other research organizations with whom contact is less serious. Consequently, commercial labs have more segmented interchange than do university laboratories because the latter do not have as much in the way of special arrangements surrounding their dealings with outside organizations; dealings with outside organizations is one of the defined roles of the university.

University labs do not have as a dominant goal the protection of commercial interests and therefore engage in more cross-boundary activities. In fact, as Allen and others have shown, a university researcher's affiliation to his or her "invisible college" (i.e., members of the researcher's discipline) is greater than the affiliation to his or her university and therefore promotes cross-boundary activity. Similarly, the set of interchange partners is larger since the primary criteria for selecting a correspondent on a particular research topic typically is area of research, more so than organization affiliation. In this absence of policy restrictions on information and resource flows, cross-boundary activities and the number of interchange partners are limited by the incremental overhead of coordinating and assimilating interchange with each outside organization or person. The greater amount of cross-boundary activity entails a greater intensity and scope of communication and interchange to support it. For many cases, the wider range of interchange partners and reduced formality implies that communications is less segmented for universities than for industry labs. University collaboration typically does not entail any formal arrangements, unless there is a commercial research lab involved, and therefore there is less distinction between an organization's communication with current collaborators and potential collaborators and among different collaborators. To the extent a university laboratory makes use of information or resources that lie deep within an organization, i.e., to the extent penetration is deep, the information or resources used are likely to be rather unique to the outside organization and can therefore lead to segmented communication. Because university laboratories are less restrictive of penetration, they are somewhat more likely to have communication that is segmented on this basis.

Although as described, university researchers tend to be less protective than researchers in commercial laboratories, some fields have a tradition of being quite protective of their work in progress, even in the university. The field of research studied and described in Chapter 12, computer science and engineering, has a tradition of greater openness than many fields in the more traditional sciences. Consequently, the distinctions between university and industrial laboratories described apply to the particular R&D laboratories studied more so than to other fields of science.

A final point is that because university laboratories have fewer institutional restrictions on interchange, university researchers are more likely to maximize their interchange, *within* *limits of how much they can assimilate.* Researchers in industrial laboratories are more likely to operate below their personal saturation points on account of institutional restrictions.

This is a very rough and over-generalized picture of the environment into which IONs are introduced. The new economics of communication and interchange could allow university laboratories to increase cross-boundary activities because of the reduced overhead of coordinating and assimilating a wider range of external inputs. The new economics could also encourage commercial labs to reduce some restrictions so as to take advantage of the increased, and lower-cost, benefits available. At the same time, the deepened penetration of ION-supported communication could cause commercial labs to intensify restrictions or could cause universities to find their interchange more segmented between different external entities. The remainder of this chapter addresses IONs in R&D and outlines several predictions. Many of these predictions are evaluated empirically in Chapter 12.

11.3 Characteristics of R&D IONs

This section describes existing Inter-Organization Networks (IONs) in the R&D domain, both their common characteristics and the ways in which they vary. It begins with a brief history of the primary R&D networks. The Arpanet was developed in the early 1970's by the Department of Defense (DOD) to link university, government, and industrial computer science research centers that were funded by DOD. It was partly a project in communications and partly an effort to share expensive or scarce computational resources. Arpanet supports electronic mail, file transfer, and remote login to all sites. In the early 1980's CSNET was started by NSF to extend connections to commercial and industry computer science research centers that were not DOD contractors and therefore not Arpanet members. CSNET is a logical network that operates on top of Arpanet, telephone, and X.25 facilities. CSNET supports only electronic mail to phonenet sites (i.e., those that rely on telephone access) and electronic mail, file transfer and remote login to X25net and Arpanet sites. A third research network is called BITNET. Historically BITNET connected university computer centers that happened to be IBM shops. It now includes a more diverse set of university sites. Because of this history BITNET uses IBM protocols to support file transfer, mail, and some remote commands; although software has been developed to allow non-IBM machines to participate as well. A fourth research network, called UUCPnet or USENET, has grown in the most decentralized manner of all the research networks. However, unlike BITNET, a very large number of members are commercial firms of varying sizes. UUCPnet is composed of Unix machines and is used mostly for mailing lists. Although UUCPnet also carries person-to-person electronic mail, many if not most organizations experience long, variable delays in receipt and delivery. Consequently users do not choose this network to support conversations that are sensitive to delays and variability. For both UUCPnet and Bitnet an organization need only have a system that can speak the communication protocols (for UUCPnet any Unix system and will do, and for Bitnet, most IBM machines and many Digital machines), and arrange with a geographically-nearby organization that is already on the network to poll the new member organization in order to send and receive communications.

The technical/performance characteristics of these different networks vary across the different networks, and sometimes across the different participants. *Speed* varies widely from 300 to 56,000 baud lines, and from 3-day to less than one minute turn around for mail messages. File transfer and remote login, where supported, are almost transparent. *Incremental costs* are ill measured, however many users perceive lower message preparation time relative to most other media. *Fixed Costs* include the communication equipment and transmission, as well as software and maintenance costs. Electronic mail service requires significantly lower investment than do more delay sensitive services such as remote login. *Capabilities* vary systematically—most places have either electronic mail only, or they have the full range of services. The most common service is electronic mail. *Automatic response* varies with the capabilities supported. Where electronic mail refers to person-to-person communication only, the change in automatic response is not significant. Remote login, file transfer, message-invocable services and other person-to-computer or computer-to-computer applications do introduce significant change. Message-invocable services are more common in environments that do not have connection-based internal networks. However, even in connection-based networks, message-based invocation will become more common because of the ease and low cost of implementation. The IONs described typically connect directly to one host in each laboratory. That host is then connected to the laboratory's internal network, if one exists. In this way, *access to internal information and resources* is provided to outsiders. However, the degree of access varies between universities and industry because of the latter's concern for proprietary information and resources. The access also varies among universities and industry according to the sophistication of internal facilities. The more widely computers are used throughout the organization, the greater the chance that valuable information will be made accessible to the outside via the ION. IONs are less *universal* than are other communication media. However, universality is higher in R&D networks than in many other domains because of the quasi-standardization that took place a priori. There are only three or so protocol families used on the various R&D networks and means of exchanging electronic mail exist between all of them. Nevertheless, incompatibilities across the four basic networks, Arpanet, CSNET, Bitnet, UUCPnet, do reduce universality as compared with telephone, telex, or postal mail. As a result, all sites cannot experience equal benefits of ION interchange with all other sites. Each network provides full access to its own capabilities and only mail access to the others. Moreover, Arpanet is the only one of the networks that supports remote login. The network membership is based on whether a laboratory has DOD funded research and whether it is rich in resources.[71]

11.4 Hypotheses for IONs in R&D

The general model predicts that organizations will change their external communication and cross-boundary activities in ways that are consistent with their differing organization characteristics. However, in the longer term, to the extent IONs change the economics of communication and interchange, organizations may make decisions that involve altering the nature of their organization policies and boundaries. Based on the general model described

[71] A second source of non-universally is the access to unique remote resources that IONs support. For example, a university laboratory may be able to make effective use of a very expensive simulation resource in a large company. The ability to do so remotely and conveniently may encourage researchers in the university laboratory to incorporate use of this facility into their research plans. At the same time, because the resource is unique to the particular company involved, the interchange may be more company-specific, i.e., less universal, than other research arrangements that involve more traditional forms of information and resource sharing.

in the previous chapter, this section predicts the effects of ION use among R&D laboratories. The predictions are broken down into two parts—characteristics of the communications and of the cross-boundary activities. Each of these two parts discuss the dimensions introduced in the model.

11.4.1 Communications

Intensity of communication will increase for both university and commercial laboratories based on greater speed and lower incremental cost of IONs.

Scope of communication will also increase for both types of laboratories based on the greater speed and capabilities supported by IONs. Increased scope will be limited in commercial laboratories because of the need to limit information flow for commercial exploitation.

Penetration will increase for commercial laboratories, within the bounds needed to protect proprietary information, resources, and trade secrets. Because such bounds do not exist for university researchers, and because the role of universities is to act as public institutions, university R&D labs may *exhibit full penetration before the ION.*

Segmentation will increase for both types of laboratories when all former interchange partners do not have access to one of the research networks. Increased segmentation may be more visible for university laboratories because industrial labs have more limited joint-research relationships, as a matter of course.

11.4.2 Cross-Boundary Activities

Cross-boundary activities will increase for both university and commercial labs based on the greater intensity and scope of communication that can be supported at the same cost. However, the increase for commercial laboratories will be limited by counterbalancing needs to retain commercial control of research products, while the increase for university laboratories will be limited because of the very public nature of university research activities to begin with.

Set of interchange partners will also be greater for university and commercial laboratories because of the reduced overhead of coordinating collaboration. However, commercial laboratories will still limit this set in order to manage their external dependencies; i.e., to control research products and commercial operations.

Restrictions on interchange and cross-boundary activities and partners will increase for commercial labs because of the penetration potentially supported by IONs. The public nature of universities reduces the risk due to penetration and therefore reduces the need for such restrictions.

In general, the absence of policy-driven restrictions implies that commercial laboratories will experience greater increases in cross-boundary activities and interchange. On the other hand, to the extent university researchers were already operating near saturation with the old media, they too will limit the extent to which they expand their cross-boundary interactions. The medium only eases some economic and technical constraints, it does not expand the researchers ability to cope and assimilate information into mental activities. In contrast, commercial labs have policy requirements that are in tension with some of these expansions. At the same time, the commercial labs are operating much farther from saturation point (largely because of these restrictions) and therefore can expand more without encountering this limitation. This qualification is related to the discussion of production cost advantage in the previous chapter.

11.4.3 Electronic Mail in R&D

As described earlier, electronic mail is a somewhat special case because it is the one ION application discussed that supports person-to-person communication as opposed to person-to-machine or machine-to-machine communication. At the same time, it is the most pervasive of all the applications. For the most part the predictions outlined above hold for electronic mail. However, some of them need amending in order to best reflect IONs that support only person-to-person electronic mail.[72]

[72] As more organizations make use of message-invocable servers to support cross-boundary resource sharing, the distinction become more difficult.

Scope will still increase but because electronic mail does *not* introduce fundamentally new technical capabilities, as do other ION services, it will not increase as much. The increase that will occur is due to the ability to import and export more perishable information based on the improved speed and incremental cost.

The increase in *Penetration* will also be less for electronic mail because it does not support unmediated online access to computer based resources. Because email is a person-to-person communication medium, the automatic response is not greatly increased, and oversight is not fundamentally different than traditional media. However, some increase in penetration is likely because of the greater ease of sharing information in general, which implies internal information as well.

Segmentation will still increase but because electronic mail is more standardized and easier to interconnect across different standards, segmentation due to non-universal access will be reduced. Similarly, because penetration does not increase as much, that aspect of segmentation will be dampened.

Restrictions on interchange is another dimension that will not change for electronic mail alone as for other applications because of the limited penetration experienced. Restrictions will still increase where an organization does not trust its ability to educate its employees adequately to the new medium, or where the organization is unable to differentiate between person-to-person electronic mail and person-to-machine (server) invocation messages; the latter introduces the automatic nature of other ION applications.

11.5 Conclusion

The next chapter describes an empirical study of IOns in R&D. The predictions proposed in this chapter reflected a direct mapping of the general model onto the R&D domain. We found strong evidence of increased intensity, scope, penetration, numbers of interchange partners, and cross-boundary activities. In addition, we found some evidence of segmentation. However, no increase in restrictions on communication or cross-boundary activities was found.

Chapter Twelve

Empirical Study of ION Use by R&D Laboratories

To investigate the predictive and descriptive value of our model, we conducted a study of ION use by R&D laboratories. The study included two types of organizations whose organizational boundaries differ—industrial and university laboratories. Comparison of the two groups' behavior illuminates the influence of organization boundaries on the use of IONs. Each laboratory had a connection to at least one of the networks described earlier—CSNET, Arpanet, UUCPnet, and BITNET. After collecting background information from a liaison in each organization, we distributed an online questionnaire. The questions addressed the characteristics of inter-organization communication and interchange described in the model. As described in detail in section 12.3, strong evidence was found of increased intensity, scope, penetration, cross-boundary activities, and number of interchange partners. Segmentation was indicated only weakly. We found statistically significant behavioral differences between university and industrial laboratories for scope and penetration, and scope did not include expanded exchange of resources, only information. Furthermore, we found no increase in protective behaviors.

This chapter describes how the predictions were tested—the method and questionnaire used to collect data, and the results of the data collection. After discussing the implications of our findings and the limitations of the study, we reevaluate the theoretical model and propose follow-up work.

12.1 Testing the Model

Our predictions regarding ION use in the R&D domain were outlined at the end of the previous chapter. To test these predictions empirically, we investigated seven conceptual variables. Following is a description of each of these variables and how it was represented and measured.

Intensity is represented by frequency of communication. Frequency was measured as the number of times per week that a respondent communicated with a person or machine in an outside organization: 0, 1, 2-5, 6-10, or more times per week. Change in intensity is represented by reported increase, decrease, or no change in the frequency of communication.

Scope is represented by the range of information and resource types exchanged with outside organizations. Information types measured were: research ideas, research results, joint authorship, information for solving a particular problem, information about tools and techniques, administrative scheduling, and other. Resource types measured were: software, computer resources, remote applications, databases, and other. Change in information and resource types is represented by reported increase, decrease, or no change in the amount of each type reported.[73]

Penetration is represented by the classes of information and resources exchanged. Information classes measured were: publicly available, available in internal documents only, related to unpublished research, related to unreleased system or product, proprietary, and other. Resource classes measured were: widely available, limited, costly, critical for internal operations, proprietary, and other. Change in information and resource classes is represented by reported increase, decrease, or no change in the amount of each class reported.

Segmentation is represented by the change in communication with non-ION organizations. Change in communication with non-ION organizations was measured by reported increase, decrease, or no change in the amount of information and resource types exchanged.

The number of *cross-boundary activities* is represented by the number of research projects that involve information or resource sharing with external organizations. Change is represented by reported increase, decrease or no change in the number of research projects that involve outside input since ION adoption.

[73] See section 12.4.2 for discussion of flaws associated with our measurement of this variable.

The number of *interchange partners* is represented by the number of external organizations with which information or resources are shared. Change in the number of external organizations is represented by reported increase, decrease, or no change in the number of organizations communicated with regularly since ION adoption.

Restrictions on interchange is represented by contracts between the organizations and by factors that inhibited more extensive ION use. Contracts included informal, consulting, grant, and joint venture. Change in the restrictiveness of contract agreements was measured by the difference between the ION contracts reported and the contracts governing non-ION-supported interchange: no difference, more explicit, more protective, more exclusive, more open-ended or ill-defined, other. Inhibiting factors were measured by reported limitations: destinations inaccessible, inconvenient, poor performance, confidentiality of information, company policy, none, other. Additional information on organization-wide restrictions (administrative and technical) was collected from the liaison in each organization.

Several other variables were measured which are related less directly to the model: number of people and change in the number of people communicated with in each outside organization; change in use of traditional media with each outside organization; and the information and resource types exchanged within the respondent's organization.

The primary contingency variable is sitetype. Each respondent is associated with either an industrial or university laboratory and the results are aggregated within these two types. Other context variables measured were: years of ION use, job title, area, and frequency of computer use.

12.2 Method

An ideal study of the behavioral changes that accompany ION use would measure communication and cross-boundary activities before and after ION adoption and would cover an environment in which the applications are advanced enough, in particular, integrated, to reflect future uses of this technology. At the same time, an ideal study would look at multiple organizations to increase the external validity of the results. The study conducted investigates a group of organizations that use IONs in interesting and diverse ways but for which no pre-adoption study was possible. Consequently we rely on retrospective data to measure the behavioral changes. On the other hand, the large number of participating R&D laboratories does allow cross-organizational comparisons.

We addressed a questionnaire to researchers in twelve commercial and fifteen university laboratories One hundred and ninety-two persons responded; all but a few responses were returned online. Of the 192 respondents, 73 indicated no work-related ION communication and therefore did not answer most of the questions. These individuals who indicated no work-related use were left out of the final data analysis.

Data was collected in two stages. First background data on each of the participating organizations was collected from liaisons. The data was collected through a combination of electronic mail correspondence and telephone conversations. The liaison in each organization was the person officially responsible for administering the ION connection for that organization. Table 12-1 lists the frequencies of 5 background variables, broken down by the type of site—industrial or university laboratory. The first variable is the type of ION service supported—electronic mail only, electronic mail and file transfer, or a full range of electronic mail, file transfer, and remote login. Most of the 12 industrial sites supported electronic mail only. In contrast, most of the 15 university sites studied supported the full range of services (remote login, file transfer, and electronic mail). Note that these numbers are representative only of the population included in our study, not necessarily of the larger population of computer science/engineering R&D laboratories. The two site types also differed significantly in the number of years they had been connected to the network—for example, seven of the industry sites, but only one of the university sites had been connected for one year or less.

We grouped responses from the various sites into university and industry responses. The number of respondents per site did vary significantly as can be seen in Appendix B. Tests showed that the responses from those sites that had very large numbers of respondents did *not* differ from other sites. However, we cannot completely eliminate the possibility of organization effect. See section 12.4.2 for discussion of why the numbers of responses per site varied so.

The nature of the organizations' respective internal facilities did not vary significantly across the two site types. Most of the industrial and university sites had internal networks connecting multiple hosts to the ION. In addition, several of these laboratories supported rather extensive internal services. However, a few of the industrial sites and one of the university sites had only stand alone host(s) connected to the ION. Of the sites with internal networks approximately half consisted of large numbers of heterogeneous hosts, peripherals, and services. The rest had smaller networks consisting of a few to a dozen machines.

Very few sites (one of the industrial sites and two of the university sites) reported any explicit technical access controls on gateway access—most relied on limited access to network-connected facilities or limited capabilities supported through the gateway (i.e., electronic mail). Similarly, none of the sites charged end-users for network usage.

Most of the sites were connected directly to at least two of the four research networks identified (CSNET, Arpanet, BITNET, and UUCPnet). University sites had on the average more connections than did industrial sites.

After providing background information, the liaison in each organization distributed our online questionnaire to end-users of the ION. The questions were closed ended and were divided into four multiple-question parts. We pretested the questionnaire in two sites and a small number of ambiguities were rectified. The final questionnaire, as it appeared on line, is found in Appendix A. A short explanation was included with the questionnaire requesting online or hard copy response. In many of the organizations the questionnaire was posted on an electronic bulletin board. In the other organizations, the questionnaire was distributed as electronic mail to mailing lists of individuals. The distinction between mailing list and bulletin board distribution turned out to be critical since persons are more inclined to pay attention and respond to personal electronic mail than they are to a bulletin board

	Site Type	
	Industrial Lab	*University Lab*
ION Capabilities		
Elec. Mail Only	9 (75%)	1 (8%)
File Xfer, Elec. Mail	0 (0%)	1 (8%)
Remote Login, others	3 (25%)	11 (85%)
ION GW Controls		
None	11 (92%)	11 (85%)
Registered Hosts	0 (0%)	2 (15%)
Registered Users	1 (8%)	0 (0%)
Years on ION		
One or fewer	7 (50%)	1 (8%)
Two or more	5 (42%)	12 (92%)
Internal Facilities		
Single host only	3 (25%)	1 (8%)
Network	2 (17%)	3 (25%)
Enhanced services	7 (58%)	8 (67%)
Size of internal net		
Small	4 (33%)	2 (15%)
Medium	2 (17%)	5 (39%)
Large	6 (50%)	6 (46%)
# ION connections		
One	4 (33%)	0 (0%)
Two	4 (33%)	6 (46%)
Three	4 (33%)	5 (39%)
Four	4 (33%)	2 (15%)

Figure 12-1: Background information for each site:Capabilities of ION connection, Access controls on ION gateway, Number of years on network, Nature of internal facilities, Size of internal network, Number of ION connections.

posting. Unfortunately, most of the liaisons did not have appropriate mailing lists and therefore resorted to posting on electronic bulletin boards.

Each of the questions relates to one of the variables presented in the model and described in the previous section of this chapter. In the first section on external information and resource sharing:

- Question (a) asks about the number of organizations communicated with via the ION.

- Question (c) asks about the change in this number.

- Question (d) asks about the change in the number of projects with external input.

- Question (e) asks the extent to which the changes indicated in (c) and (d) were attributed to ION use.

In the second part of the questionnaire—communication and access patterns—responses are in terms of individual outside organizations identified in question 1(f):

- Questions (a) and (b) ask about the number of people communicated with in each organization and the change in this number since ION use. (This variable is not a direct part of the model.)

- Questions (c) and (d) ask about frequency of communication and change in *frequency—corresponding to intensity.*

- Question (e) asks about the change in use of traditional media. (As with (a) and (b), this variable is not a direct part of the model.)

- Questions (f) and (g) ask about the types of information and resource exchanged, and the change in the amount—scope.

- Question (h) investigates segmentation in the form of altered communication with non-ION organizations.

- Question (i) asks about the range of information and resource types exchanged within the respondents internal organization.

- Questions (j) and (k) investigate penetration in terms of the classes of information and resource types exchanged and change in the amount.

- Question (l) asks to what extent the changes indicated were attributed to the ION.

The third set of questions focuses on contracts and restrictions:

- Questions (a) and (b) investigate restrictions and policies governing ION-supported relationships in terms of contracts between the respondent's organization and the outside organizations, and the difference between these contracts and those governing non-ION mediated relationships.

- Question (c) asks about other factors that inhibit ION use.

Finally, the fourth set of questions collected background information on the respondent:

- Questions (a) and (b) ask about the total number of projects and outside organizations, non-ION as well as ION.

- Questions (c) and (d) ask about the respondent's work area and job title.

- Question (e) asks about the frequency of computer use.

Respondents who answered zero to the first question (1a)—i.e., during an average work week they did not exchange work-related information or resources via the ION with any outside organizations—were asked to provide background information only (section 4) and were not included in the data analysis.

Each response was encoded into 166 one- or two-digit variables. The large number of variables is due to the fact that each respondent provided information on up to four outside organizations and each information and source type and class were coded separately as one or zero. If a respondent answered for only two outside organizations, for example, missing data codes were entered for the remaining two organizations. The coding scheme is summarized in figure 12-2.

For each respondent we took the average of his or her responses across the one to four external organizations for which he or she provided data. For example, an average change in frequency was assigned the value SUM(change−in−freq1, change−in−freq2, change−in−freq3, change−in−freq4)/f, where f is the number of organizations for which the respondent provided change values, i.e., 1, 2, 3, or 4. Of the industry respondents, approximately half reported on one or two outside organizations and most of the rest reported on three (21%, 27%, 33%, 18%). Of the university respondents over three quarters of the respondents reported on three or four outside organizations.

Questions	*Response*	*Code*
1a,1b,2a,4a,4b	Number between 0 and 98	0 to 98
1c,1d,2b,2d,2e,2g,2h,2k	Less/Same/Greater	0/1/2
1e,2l	None/Some/Quite a bit/Very Much	0/1/2/3
1f	Univ/Industry/Gov't/Bboard/Committee	0/1/2/3/4
2c	0/1/2-5/6-10/More	0/1/2/6/8
2f,2i,2j	Yes/No for each type/class	0/1
3a	None/Informal/Consulting/Joint Devel/Other/Grant	0/1/2/3/4/5
3b	None/Explicit/Protective/Exclusive/Less/Other	0/1/2/3/4/5
3c	None/Proprietary/Inaccessible/Performance/Other	0/1/2/3/4
4c	Sftwr/Hdwr/Theory/Systems/Other	0/1/2/3/4
4d	Mngr/Faculty/Rschr/Tech Staff/Scientist/Other	0/1/2/3/4/5
4e	Rarely/Weekly/Daily/Always	0/1/2/3

Figure 12-2:Coding scheme for questionnaire data.

12.3 Results

This study tests whether certain behavioral changes accompany ION use. Chapter 11 concluded with a set of predictions for R&D laboratories which were derived from the general hypotheses set out in Chapter 3. The test of these hypotheses is the proportion of respondents reporting change and the variation in this proportion across the two types of sites—university or industry. The results are presented in three parts—context variables, hypothesis variables, and the interaction between variables—and are broken down into industry and university respondents. The implications of the results are discussed in section 12.4.

To summarize, most respondents who use the ION for work-related communication indicated increased intensity, scope, penetration, numbers of interchange partners, and cross-boundary activities. However, only reports of increased scope and penetration varied significantly across university and industrial respondents. Contrary to our predictions, reports of increased penetration were *not* accompanied by reports of increased segmentation or restrictive measures. Since fewer than 25% of the respondents shared any kind of computer-based resources—almost all of the interchange consisted of person-to-person information exchange—increased segmentation and restrictions may not have been needed.

12.3.1 Context Variables

All of the respondents reported using a computer daily. Most of them work in software or systems related areas; only a few in theory or hardware. However, they were more or less evenly divided among management, faculty, research, and technical staff positions (only a small number referred to themselves as scientists). The frequencies of these context variables are grouped by site type in figure 12-3.

Over half of the industry respondents reported on communications with government and university research laboratories only. In contrast, three-quarters of the university respondents reported on communication with at least one industrial laboratory. Although our results, reported below, did not vary significantly with the type of outside organization, the numbers are too small to draw any conclusions.

12.3.2 Hypothesis Variables

The results reviewed in this section are summarized in figure 12-4.

Most respondents reported more intense (frequent) communications in connection with ION use (82% of 94). Although a greater percentage of university than industry respondents reported an increase, the difference was not statistically significant (93% university, 77% industrial, $p < .16$).

Most respondents also reported increased exchange of a wide range of information and

	Site Type	
	Industrial Labs	***University Labs***
Years of ION Use		
Less than one	5 (8%)	2 (7%)
One to Two	17 (26%)	2 (7%)
Two to Three	17 (26%)	6 (21%)
Three to Four	7 (11%)	5 (17%)
More than Four	20 (30%)	14 (48%)
Area of Work		
Software	61 (69%)	19 (51%)
Hardware	6 (7%)	1 (3%)
Theory	6 (7%)	6 (16%)
Systems	9 (10%)	10 (27%)
Other	6 (7%)	1 (3%)
Job Title		
Manager	22 (25%)	2 (5%)
Faculty	0 (0%)	22 (60%)
Research Staff	22 (25%)	7 (19%)
Technical Staff	37 (42%)	1 (3%)
Scientist	5 (6%)	2 (5%)
Other	1 (1%)	0 (0%)

Figure 12-3: Context variables grouped by type of site:years of ION use, area of work, and job title.

resource types. The most commonly exchanged information types were research ideas, information for solving a particular problem, information about tools and techniques, administrative scheduling, and software (see Appendix D). Overall, a greater portion of university than industry respondents reported exchanging a research ideas, research results, and joint authorship comments. Furthermore, a greater percentage of university

	Site Type	
	Industrial Labs	*University Labs*
Frequency of communication		
Less	2 (3%)	0 (0%)
Same	13 (20%)	2 (7%)
Greater	50 (77%)	27 (93%)
Types of information and resources		
Less	0 (0%)	0 (0%)
Same	13 (20%)	1 (4%)
Greater	51 (80%)	26 (96%)
Number of outside org.s		
Less	0 (0%)	0 (0%)
Same	5 (9%)	1 (3%)
Greater	54 (92%)	27 (97%)
Number of projects with outside input		
Less	0 (0%)	0 (0%)
Same	21 (37%)	6 (22%)
Greater	34 (63%)	20 (78%)
Classes of information and resources		
Less	2 (3%)	0 (0%)
Same	18 (28%)	2 (7%)
Greater	45 (69%)	27 (93%)
Communication with non-ION org.s		
Less	13 (22%)	9 (36%)
Same	44 (75%)	15 (60%)
Greater	2 (3%)	1 (4%)

Figure 12-4: Responses to hypothesis variables grouped according to site type—university and industrial labs.

respondents than industry respondents reported an increase in this measure of scope (96% university, 80% industrial, p < .1). Almost all of the exchange reported for both site types was information sharing between people. Less than 25% of all the respondents reported sharing any kind of computer-based resources even though 80% did so *within* their respective organizations.

Almost all of the respondents reported communicating with a greater number of outside organizations. Moreover, there was no difference in the percentage of university and industry respondents reporting an increase (92% university, 97% industry, $p < .67$). Over half of the respondents also reported communicating with a greater number of people within each organization (59% university, 61% industry, $p < .29$).

A smaller but substantial percentage of the respondents indicated an increase in the number of projects involving outside organizations (i.e., cross-boundary activities). A somewhat larger percentage of university respondents than industry respondents indicated such increase (78% university, 63% industry, $p < .28$).

Most respondents also reported increased exchange of some classes of information and resources. Moreover, a substantially larger percentage of university respondents than industry respondents reported an increase in this measure of penetration (93% university, 69% industry, $p < .04$).

Despite the reported expansion in inter-laboratory communication and interchange, our predictions of increased segmentation was not supported as strongly. Only a small number of respondents reported any decrease in communication with non-ION organizations and the difference between university and industry respondents was not significant (36% university, 22% industry, $p < .4$). We found even less indication of restrictions on interchange and cross-boundary activities. Almost no university respondents, and less than half of the industry respondents, reported any kind of formal contracts governing their ION-supported relationships. Furthermore, almost all respondents indicated that there was no difference between the contracts governing ION-supported relationships and those governing traditionally-supported relationships.

12.3.3 Interaction of Variables

We tested for variance of these variables with context variables other than site type and with one another and found almost no statistically significant correlations. The only significant correlations found were for industry respondents. The frequency of communication, number of cross-boundary activities, and number of interchange partners varied significantly with the number of years of ION use. This correlation was not apparent for university laboratories; in fact, there was some indication that the percentage of respondents reporting increases *decreased* as the number of years of ION use grew. This finding is probably due to the difficulty of recalling communication patterns before ION use, or the complete *absence* of communication before ION use—i.e., for those persons who had used ION facilities since their first day on the job.

The following section discusses the implications of these results.

12.4 Discussion

After reviewing the hypotheses supported and refuted, this section discusses the problems and limitations of the conceptual model, research design, and data collection instrument, and theoretical implications of the overall findings.

12.4.1 Hypotheses

As reported above, the data supported our hypotheses of increased intensity, scope, number of outside organizations, number of cross-boundary projects, and penetration. Moreover, our measures of increased scope and penetration showed a significant difference between the two site types; supporting our hypothesis that the more bureaucratic and protective organizational boundaries would inhibit change for industrial laboratories. However, although all other reported changes were greater for university laboratories than for industrial laboratories, the differences were not statistically significant. Two other hypotheses clearly were not supported. Most dramatically, the prediction of increased restrictions, in the form of contracts and technical access control mechanisms, was refuted. In addition, very few respondents indicated an increase in segmentation. This subsection discusses each of the hypotheses, both the results and the limitations of the measurements and theory.

The data confirmed our predictions that the greater speed and lower (end-user) incremental cost of ION communication would foster more intense communication. No difference between industrial and university respondents was found. In chapter 11 we suggested that the level of communication engaged in by university researchers might be closer to saturation than that of industrial researchers before ION adoption and consequently university researchers might not expand their communications as much. The only possible evidence of saturation found was the decrease in increased communication for those university respondents who had been using the ION for the longest periods of time. On the other hand our questions were not fine grained enough to be test this issue conclusively; we asked only whether there was an increase, not what the size of that increase was.

The measured increase in information and resource types exchanged supports our hypothesis of increased scope. In addition, a significantly greater percentage of university respondents than industry respondents reported increased scope. This relatively lower increase for industry sites demonstrates the restrictive effects of bureaucratic boundaries suggested in our model. Although increased exchange of *resource* types is a particularly interesting form of increased scope, (most resources cannot be exchanged via traditional media), practically no sharing of computer-based resources was reported. This constitutes a negative finding in and of itself with respect to the model's scope hypothesis. We predicted that sharing of computer-based resources would contribute to the scope of inter-laboratory interchange. Despite these limitations, the wide range of information types exchanged, the increased amount of these types, and the increased number of people communicated with in each outside organization, does indicate a wide range of information exchanged.

The number of *cross boundary projects* increased equally for both university and industry sites. However, overall, a smaller percentage of respondents reported an increase in this variable than in the other variables. One of the reasons that we predicted an increase in cross-boundary activities was the predicted increase in scope of ION-supported communications. To the extent increased scope was more limited than expected because of the absence of resource sharing, it is consistent with the model that a smaller increase was reported. A second factor that might have dampened the increase in cross-boundary projects is the greater institutional overhead associated with changing the management of a project, for example, to include outside entities. For these reasons, this variable, as well as segmentation and restrictions discussed below, are likely to change more slowly than those variables over which researchers exert more individual control (i.e., frequency, scope, and even penetration). As reported these impediments to change appeared lower for university than for industrial sites. Although the difference was not statistically significant, it is consistent with the differences between university and industrial laboratories described in chapter 11. In addition to these reasons for why increases in cross-boundary projects might actually be dampened, our *measurement* of this variable may have suffered from a basic mismatch in the units of analysis of the questionnaire and model. In particular, projects are likely to vary more for a group as a whole than for an individual. For example, when the number of projects grows, the size of the research staff may be increased so consequently individual researchers may not experience an increase in the number of projects.

The number of *interchange partners* increased for a very large portion of respondents, and increased equally for both types of sites. The model predicts increased numbers of interchange partners based on the increased intensity supported. The model also predicts that for some sites this increase would be limited by increased segmentation and the need to cope with risk of increased penetration. As described above, segmentation did not increase significantly and therefore would not be expected to inhibit growth in the number of outside organizations. Similarly, no increase in protective or restrictive measures was indicated.

The measured increases in *classes* of information exchanged indicated increased penetration for both university and industrial respondents. However, a significantly smaller percentage of industrial respondents than university respondents indicated increased penetration. According to our model, this predicted difference reflects the industrial laboratories' more bureaucratic and protective boundaries.

Segmentation was hardly supported by the very weak report of decreased communication with non-ION organizations. Moreover, the reports were equally low for both industrial and university sites. However, this finding is consistent with the overall model, given the other findings in this study. The model predicted segmentation would increase due to non-universal facilities and penetration. In fact, the ION use reported consisted almost exclusively of person-to-person electronic mail. Electronic mail is more universal, more easily interchanged and intermixed with traditional media, and less penetrating than resource sharing capabilities. Therefore, although electronic mail alone introduces some bases for segmentation, the limited segmentation measured is consistent with our model given the absence of resource sharing. Unfortunately, there remain two reasons why the *measurement* of such segmentation may have been dampened. The first is the mismatch in the unit of analysis described above. Namely, change in communication with non-ION entities is more easily detected on an aggregate level across a group or laboratory than on an individual basis. Often researchers do not make independent decisions about whom to communicate with—i.e., certain outside organizations are part of their home organization's community and others are not.

The hypothesis of *increased contractual and technical restrictions* on cross-boundary flows was clearly refuted. No evidence was found of administrative policies to cope with risks of ION-supported penetration. Contracts of all kinds were rare and almost no change in contracts was indicated. Several respondents even commented that the contracts with ION partners were *less* restrictive and less explicit. In addition, most of the organizations studied had minimal or non-existent gateway access control mechanisms or policies in place. The prediction of increased restrictions was premised on an increase in external access to internal facilities. Many of the organizations' internal facilities are quite sophisticated, integrated environments which we predicted would amplify issues of penetration and risk related to ION connections. Although restrictions were not anticipated for universities, given their public nature, the proprietary nature of industrial facilities did lead us to predict restrictions on cross-boundary flows. However, because most of the interconnections reported supported person-to-person electronic mail only, these valuable internal resources were not accessible to outsiders anyway. One of the bases for predicting an increase in penetration and restrictive measures, was ION-supported online invocation of another organization's internal resources. Given that most of the ION use measured was information exchange via electronic mail, little such access actually occurred, and therefore it is less surprising that a more significant increase in penetration and compensating restrictions was not detected. Apart from these explanations for our negative finding, the mismatch in units of analysis described earlier may have dampened our measurement of this variable since individual researchers are less cognizant of administrative mechanisms governing their external relations.

In summary, it is consistent with our model that some of the predicted changes in cross-boundary activities were not found. Namely, because almost all reported communication consisted of person-to-person information exchange, there were no increases in capabilities or automatic response, or decreases in universality of the communication medium. Consequently, the communication patterns did not increase as much in scope, penetration, or segmentation, and no restrictive measures were imposed on cross-boundary activities.

12.4.2 Problems and Limitations

The most problematic aspect of testing the model is that it makes predictions about *change*. Two major difficulties in measuring change were encountered. The first is that because no pre-post study was possible, we had to rely on retrospective data. Although the questions addressed formal communication for which such subjective data is somewhat more reliable, we are still limited by the subjectivness of retrospective data. For example, many of the variables for which the model predicts an increase due to ION use, would naturally increase over the course of a researcher's professional life, independent of a change in media. We conducted one test for the effect but found that although many of the ^hange variables did vary with the number of years, none of these variations was statistically significant (with the exception of number of outside organizations). Because the model predicts an increase in the rate or amount of change, a control is needed in order to isolate the effects of the ION. We attempted to address this need by including data only for those respondents who explicitly attributed some change to the ION. A related problem was that some respondents we.e unable to comment on changes in behavior resulting from ION use because they had always used the ION to carry on professional communications and therefore had no point of reference; in general, or with the particular outside organization reported,

A second problem in designing this study was that of unit of analysis mentioned in the previous section. The unit of analysis addressed in the questionnaire is the individual researcher and his or her cross-boundary activities; whereas the theoretical model addresses organizations as aggregate units. For many of the variables measured, this mismatch in unit of analysis was acceptable. However, some of the variable measures such as changes in contracts and cross-boundary projects probably suffered significantly because an individual researcher may not have the perspective of the larger group of which s/he is a member, and therefore may not be aware of higher level factors such as contracts and aggregate projects. Moreover, it is reasonable to expect that the number of researchers in a group would be increased along with the number of projects or cross-boundary projects, in which case individual researchers would not necessarily experience a change. Several of the respondents also commented on the fact that often they are not conscious of the organization affiliation of an individual and therefore had trouble answering questions that addressed an aggregated external unit—the organization—instead of the individual.

In our measures of increases and decreases in the various information types and classes exchanged space limitations forced us to aggregate the indicated changes across all types. It would be interesting to measure increased scope and penetration in more detail to investigate whether the various types of information and resources are affected differently by ION use.

A final issue of concern is the online distribution and response to our questionnaire. L. Sproull conducted a study comparing the responses to an online questionnaire to the responses to a traditional paper questionnaire. [66] Her criteria for rating the different media were respondent access, willingness to respond, and comparability of the data collected. Sproull found that overall response rate was ten percent worse (lower) for online questionnaires than for traditional, while the time to answer was fifty percent better (lower) for online than for traditional. Content of responses to factual objective questions were comparable while responses to less objective questions were more extreme. Some of the difference in response rate may be do to the fact that the paper questionnaires were followed up by telephone to encourage response, and this was not the case for the online response. In any case, Sproull concluded that the lower cost of online questionnaires made them an attractive alternative. Because the questions in our online questionnaire were primarily factual, and because online was the most effective (and perhaps the only) way to reach the intended audience, we feel confident that our results did not suffer from online response per se.

Online distribution was nevertheless problematic. Because of the diversity of mail reading programs and editors used by would-be respondents, we could not design the questionnaire to take advantage of computer-based editors; e.g. an interactive questionnaire. Consequently, although online editing is more convenient for most of the individuals in this population, the questionnaire itself would have been easier to design and fill out on hard copy.

A *second problem* related to online distribution was the difficulty of determining the number of persons receiving questionnaires. As a result, the total population size was not known and no meaningful measure of response rate can be calculated. Most of the organizations do not monitor use of the ION connection and have no idea how many researchers actually use it. Similarly, the bulletin boards are seen by an unknown number of users and even the mailing lists often had an unspecified number of addresses on them.[74] Consequently, we cannot determine how representative the self-selected respondents actually are of the total ION-using population.

Another issue related to the mode of distribution was the significantly larger number of responses from those sites that distributed the questionnaire via mailing lists compared to

[74] Mailing lists can contain pointers to other mailing lists so the only way to count the number of individuals is to trace the entire tree. If different mailing lists are owned by different people, this is not always possible by the person at the top of the tree.

those that posted it on electronic bulletin boards. In addition, the response appeared to vary depending upon the nature of the mailing list used—in particular, the size, frequency of use for other purposes, and position in the organization of the originator of the message. One site in particular had a mailing list composed of all persons who had ever sent a message through the ION gateway. Probably because of the appropriateness of this list, the fact that it was rarely used, and that the distributor of the questionnaire was known to be the administrator of the gateway, this site generated a very large number of responses.

Finally, several issues were not addressed in the general model, nor in this study. First is the strategic role of personal communication. Particularly among researchers, the use of electronic mail to enhance personal networks was not studied and there is good reason to expect a significant role. A related issue is the saturation point of many of these variables, and the reasons for the saturation, i.e., the limits to useful increase in communication between researchers. We did not address computer based communication within the boundaries of the organization. Inter-site and inter-division communication, in particular, have undoubtedly been impacted by the use of internal networks.

12.4.3 Normative Implications for R&D Laboratories

Before discussing the more general theoretical implications of our findings in section 12.4, this section identifies some normative implications for R&D laboratories. We will discuss both opportunities and dangers of which ION participants should be aware. The items discussed below were indicated by the study but require further investigation. Our objective is simply to flag issues of concern to R&D laboratories and future ION studies.

Above all else, participation in R&D networks is important because it is becoming a determining element of participation in the R&D community. Although it is difficult to codify the significance of this parameter, some ongoing technical discussions already occur only on the network, and some informal contact may be practically maintained only via the network.[75] Our study does not address this informal role of the R&D networks, but

[75] The IONs under discussion are Arpanet, CSNET, BITNET, and UUCPnet

comments of respondents and discussions with site liaisons do indicate that this is a primary function of the network services.

The study clearly illustrated that IONs can support intensified communication and expanded cross-boundary activities. The question remains as to what the implications are of these changes for the R&D organizations themselves. For example, although the use of traditional media was reported to decrease in conjunction with ION use, there was indication of an overall increase in communication activities, not just a shift from one medium to another. The reported increase in the number of people communicated with in each outside organization is one example of this trend. However, in addition to an overall increase in communication levels, introduction of this new medium will involve some reshifting of communication among media. Different media are more appropriate for different types of tasks (both in terms of cost and function) and some formal evaluation and guidelines for when to use which media may be in order.

Most of our discussion has addressed the effects of increased communication on relations with outside organizations. However, such increases will affect internal communications as well. For example, as larger numbers of individuals communicate with outside organizations directly, the nature of internal information dissemination and the role of gatekeepers and boundary-spanners will change.

Almost all of the communication reported by respondents involved exchange of information, not resources. Consequently this study provides no data on the implications of intensified ION-based resource sharing. However, it does suggest that there are more impediments to incorporating shared resources into operations than there are to incorporating increased amounts and sources of information. One such impediment is the absence of application level standards. At the same time, what makes IONs so powerful, potentially, is the ability to share resources. In other words, although we have found evidence of the barriers to resource sharing, we should not conclude that resource sharing will remain minimal or unimportant. Furthermore, if and when resource sharing does increase, issues of penetration, segmentation, and restrictive actions which were minimally indicated or refuted in our study will require reevaluation. Evidence of greater formality with respect to fund or resource flows is consistent with the findings of several studies in other domains. [28, 73, 76] In fact, the absence of resource sharing could be indicative of the ultimate form of restriction, i.e., abstention.

In general, the formal inclusion of outside sources of information and resources into projects needs to be reevaluated in light of the change in economics of cross-boundary activities. At the same time, policies governing these relationships will require reevaluation. Although these policies may have to be adapted to cover risks introduced by ION-mediated interchange, it will be of equal importance to assure that these policies are consistent with other aspects of both external relationships and internal operations. For example, in some cases ION use could be rejected because it would require a level of codified information flow. and risk-management that is perceived to impose on internal or external communications. The flexibility of the technical mechanisms employed can significantly affect this dynamic; see Chapter 5.

Assuming that the levels of risk are not so excessive as to preclude ION use, the type of coping mechanism that participants should consider include educating more personnel to act as boundary spanners. [2] A similar educational requirement is the need to caution boundary spanners and organizations as a whole against inefficient criteria for selecting interchange partners, or sources of information. For example, convenience to the purchase agent or researcher of accessing online sources may lead individuals to obtain information or orders from an online supplier instead of from an "off-line" supplier of superior price-performance.

Finally, universities have a different, and in some ways conflicting, set of interests than do industry laboratories. Universities play a public role to which accessibility and multiplicity of information sources and sinks is essential. In a larger sense, there has been general concern in recent years that university relations are narrowing as a result of special and sometimes binding industry contracts, and that this trend has some unhealthy implications for academia. Because of their public role, it is in the interest of university laboratories to maximize the number of entities on the network, so that reliance on the network does not result in a narrowed community. Similarly, easy information flow within and and without is particular important to universities. Consequently, ION participants should not be forced into imposing on internal or external communication with increased codification and controls (see 4).

12.4.4 Theoretical Implications

Thus far we have discussed the findings of our study in terms of the hypotheses set out for R&D laboratories. In this section we discuss the implications of our findings for the more general structure of our theory.

In the spirit of Williamson's transaction cost framework, our model of ION impacts propŏsed a set of relations between communication medium and communication patterns, and between communication patterns and cross-boundary activities. This study provided evidence for the existence of these connections since a large portion of the respondents reported expanded communications and cross-boundary activities, and attributed at least some of the expansion to ION use.

Our theoretical model also suggested a connection between the direct improvements that motivated ION adoption and problematic side-effects of ION use. Generally, these problematic side-effects were related to codification and formalization of inter-organization flows as a result of penetration-related risks. Our findings did not show any increase in restrictions or codification of cross-boundary flows and activities. However, we believe that this negative finding reflects an absence of the resource sharing that we predicted would lead to these problematic side-effects. Because the ION communication studied turned out to be almost all person-to-person electronic mail, we were not able to investigate the connection between opportunities and problematic side-effects for resource sharing. In other words, although this study suggests that for person-to-person information exchange alone, the restrictive effects suggested in the model are not active, we should not generalize to other types of IONs.

This finding does suggest that communication patterns can change without being reflected in the formal structures governing the relationships. At the same time we posit an untested hypothesis that such is not the case for resource sharing—in other words, that the restrictive actions suggested in the model (i.e., contracts, technical restrictions, etc.) *would* be exhibited for IONs that support new kinds of resource-sharing patterns. Walker and Townsend [76] and Van de Ven et. al. [73] findings support this hypotheses as to the difference between information and resource sharing. Two highly related aspects of ION-supported resource sharing leads to this distinction. The first is the difference between resources and information; and the second is the difference between person-to-person communication and person-to-machine or machine-to-machine communication. In short, information exchanged between persons typically is less codified, less specific, and more fungeable than information exchanged between machines because of the inherent flexibility of human intelligence. At the same time, people in an organization can be charged with responsibility for monitoring, overseeing, and applying discretion to incoming and outgoing communication. In contrast, sharing computer-based resources implies that at least one of the two correspondents is a computer that reacts automatically. These connections can supports a wide range of capabilities beyond person-to-person communication and information retrieval, but at the same time the rules of operation for a computer system must often be specified much more explicitly than those for a human being. Furthermore, a computer system cannot be held responsible for the same type of discretionary actions that a human employee can be. Moreover, because sharing computer-based resources is qualitatively different than traditional communication, it raises many more unknowns and fears for participants which itself exacerbates inhibitions. In summary, we propose that the coupling between the opportunities of IONs and the problematic side effects is much stronger for resource-sharing than for information-sharing alone.

12.4.5 Generalizability

One way to assess the generalizability of our study is in terms of the environmental conditions under which our hypotheses and findings hold. In particular, chapter 3 suggested that the production cost advantage of external communication and production, and the level of decision making were the two primary environmental, or exogenous factors. However, our study suggests that the two primary factors are:

- Whether the function supported by the ION is person-to-person information exchange or computer-based resource sharing, or both.

- The type of competitive relationship—peer coordination, peer competition, or distribution channels.[76]

In the previous section we discussed the difference between information and resource sharing and the implications for our model. Given these differences, the results of this particular study apply most directly to information sharing contexts. Similarly, the findings are most relevant to peer coordination relations such as government agencies and joint ventures. As described in the previous chapter, the R&D environment differs from more traditional market environments in several ways: communication and resource sharing are vital to the survival of all R&D laboratories; competition is a less dominant factor since the R&D activity is removed from the organizations' end-products or business; R&D organizations do not operate in traditional price-based markets; and R&D laboratories house many sophisticated users of computer equipment. The findings can be generalized in part to other domains but require further investigation in more competitive environments. Moreover, we studied R&D laboratories engaged in computer-related research. Not only is there a tradition of freer information sharing in this discipline than in many more traditional fields of natural sciences, but the use of computers internally is exceptionally high.

Looking at the individual hypotheses, neither intensity, nor number of interchange partners are affected significantly by the two primary environmental factors identified-resource sharing and competitive relationship. Consequently, our findings should apply directly to other domains such as distribution channels discussed earlier. On the other hand, resource sharing is an explicit component of increased scope. Although we observed scope increases

[76] In addition, several other environmental factors influence the level of demand for ION benefits and should be investigated in future studies: the advantages of speeding up production processes; the importance of external interchange; the role of unstructured person-to-person communication; the nature of computer use internally.

for the information-only IONs, we expect the extent of such changes to be far greater when resources are also shared. Similarly, although increases in cross-boundary activities were observed for the information-only IONs, we expect resource-sharing IONs to encourage a far greater range of activities at the same time that they introduce significantly more risk. Consequently, when participants share resources as well as information, communication patterns and the management of cross-boundary activities will undergo more significant changes. Examples of ION applications that are resource-based and therefore likely to experiences amplified changes are the exchange of cad/cam designs and design tools between manufacturers and subcontractors, or shared use of expensive processing equipment by members of a joint venture.

Penetration, segmentation, and restrictions are all three directly affected by both resource sharing and the organizations' competitive relationship. Penetration was found to be increased through information sharing but the nature of automatic, online access to resources is qualitatively different than penetration based on person-to-person information exchange alone. Segmentation is less relevant to person-to-person information exchange than to resource sharing since information can be easily translated between formats even if the transmission medium is non-standard. Resource-sharing protocols, i.e., protocols for people and machines to interact with other application programs and machines, are far more complicated than protocols for transferring textual messages between people. Consequently, resource-sharing protocols are far less standardized and an interface to one system is less likely to support access to another without significant expense. Segmentation is also affected by the inter-organization relationship. In particular, in distribution channels the incentives to impose or oppose segmentation (switching costs) are much higher than for other types of relationships. Finally, restrictions on cross-boundary flows and relationships follow from penetration and segmentation and therefore are tied directly to these factors. Therefore, although we did not find segmentation or restrictive behavior in the R&D environment, we cannot generalize to other domains where resources are shared and/or where relationships are more explicitly competitive.

12.5 Future Work

This initial study has confirmed the descriptive value of the general model and the overall approach. The data provides a good base on which to begin more in depth studies of ION participants and applications.

A particular study that could illuminate some of the issues addressed here would investigate networks that connect different geographical and functional sites of a single organization. This comparison would complement the two types of organizations studied here by contrasting issues related to organization boundaries. For example, it would be interesting to see if the more porous boundaries that exist between divisions or sites within a single parent organization resemble the porous boundaries of universities; and if so, if the behavioral patterns with respect to ION use also match.

Further work is also needed to understand the differences between information and resource exchange. One way of approaching this problem is to compare email-only IONs to IONs that support access to online computer-based resources. In addition to investigating the distinction between information and resource sharing, it would be useful to isolate the effect of automation, i.e., automatic response to external invocation as compared with traditional human response which is maintained in the case of electronic mail. Theoretical issues should also be investigated further to examine whether the relationship between the three stages of the model—medium, communication patterns, and cross-boundary activities—differs for information and resource exchange.

Changes in inter-organization communication and interchange patterns is likely to affect internal communication as well. One example is the possible reduction in the roles of gatekeepers and boundary-spanners should larger numbers of internal employees communicate with outsiders directly. This change has implications for dissemination and interpretation of internal information, as well as need for education of a wider range of internal employees about rules and regulations governing external communication.

Perhaps of most importance, future studies should examine IONs in other, more traditional, market environments, where the relationships and governing rules are more codified. In this spirit, our previous discussions used examples from other domains, in particular, distribution channels. Of particular interest is to determine whether the coupling between formal governance structures and communication behaviors is any stronger in other environments: and to identify the critical organizational and technical characteristics that affect this coupling. For this reason, the relationships we would like to investigate further are those that: support more than person-to-person information exchange via electronic mail, involve more risk for participants because of the value of internal information and resources, and are not necessarily supported by third party. Finally, we would seek to conduct such a study at a different unit of analysis—the organization and projects conducted by the organization, as opposed to the individuals within the organization.

Chapter Thirteen

Conclusions and Implications

The preceding chapters described the policy and technical implications of inter-organization computer networks, which we refer to as IONs. We described how IONs change the economics of inter-organization interchange and how interconnection across organization boundaries affects the design requirements for network access controls and interconnection. This chapter outlines the central conclusions of the thesis and discusses implications for public policy and the study of new technologies.

13.1 Conclusions

The findings of this research lie in three related areas: organization implications of new communication technologies, computer security and access controls, and network interconnection.

We described how this new medium, i.e., IONs, changes the economics of inter-organization communication and interchange and supports communications of greater intensity and scope. These new communication characteristics allow participants to carry out greater numbers of activities across their organization boundaries and to do so with greater numbers of interchange partners. At the same time ION communications is more penetrating and segmented and therefore encourages participants to impose restrictions on cross-boundary flows and interchange partners. These effects are most pronounced when participants share resources, as well as information, via the ION. Our study of R&D laboratories showed that when IONs support electronic mail only communication is enhanced and cross-boundary activities are expanded, but restrictive reactions to the more penetrating and segmented communications are not very evident. As laboratories begin to exploit the opportunities of resource sharing, the risks will increase along with the enhancements.

Based on the penetrating communications supported by IONs, we analyzed the unique control requirements of networks that cross organization boundaries. Access controls for IONs must address invocation as well as information flow, must protect the invoked as well as the invoker, and must address two-way communications, all without imposing on internal communication and resource sharing. Traditional non-discretionary control mechanisms do not satisfy these requirements. However, we presented alternative non-discretionary gateway mechanisms, based on category sets and an intersect rule, which do allow strictly-internal applications to remain unaffected by interconnection without requiring physical isolation from ION applications. These access control requirements and proposed mechanisms indicate that packet-level interconnection is not always appropriate or adequate. ION policy requirements are better met by high-level and visa-based interconnections. We concluded that, in many cases, message-based interconnection is the most-appropriate means of implementing loose couplings across organization boundaries.

Our analysis and design of ION mechanisms was greatly enriched by our bimodal approach which asked *how industry and organization contexts shape a new technology*, as well as *how a new technology affects the organization and industry contexts in which it is applied.* The following sections conclude by discussing the implications of this research for public policy and study of new technologies.

13.2 Implications of IONs for telecommunications policy

This research has value beyond that of design, implementation, and management of IONs. As demonstrated in our model and study, IONs can be used to support new modes of production that were not economic previously. In this way IONs exemplify the central role of telecommunications and computing technologies to industrial behavior and policy. This coupling has implications for telecommunications policy. In the past, inter-organization telecommunications consisted of person-to-person information exchange via telephone and telex. Today, a wide variety of machines and human users in distinct organizations communicate with one another via interconnected computer networks. Just as the development of private networks had serious implications for the public telecommunication infrastructure, so does the development of these IONs. By granting network access to external entities, organizations are using their private networks to support their "public" communications. As a result, it is increasingly unclear where the ION participants' networks end and the public network begins. Questions arise as to who will own, operate, and have access to, these communication facilities and services which mediate growing numbers of business activities. Use is not widespread enough to say which modes of service provision will dominate. Some IONs may be operated by closed consortia such as AIRINC, the Insurance Value Added Network, and CSNET. Other IONs may be private offerings, such as the individual airline reservation systems and supplier-owned online order-entry systems. Alternatively, third party value added network operators such as MCI, Telenet, GEISCO, IBM, and ATT may enter this potentially lucrative market. However, even if it is too early to state conclusively what the long-term implications are for public telecommunications, it is clear that the telecommunications policy issues are related to the industrial issues.

Telecommunications and computer technologies are used increasingly as vehicles for enacting and supporting corporate strategy. Consequently, telecommunication policy, industrial policy, and organization policy, increasingly impinge on one another. Given this interaction, one way to anticipate the effects of IONs on the public telecommunications infrastructure is to examine how and why organizations are using this medium, and what their needs and constraints are. To this end, this thesis has described how the use of IONs can impact inter-organization activities.

13.3 Research in the design and use of new technologies

We have summarized the contributions of this thesis to our understanding of network interconnection and security, the effects of computer-based communication on organization boundaries and activities, and implications for telecommunications policy. At a more general level, this research demonstrates how studies of system design and system adoption enrich one another. In particular, although our models of the technical mechanisms and organization implications in many ways are dissimilar in structure and language, they are linked in two critical ways. The most explicit connection is the issue of penetration. The organization model investigates how using this medium can affect organization boundaries, while the technical model investigates how to design the medium so that it can be tailored to support different types of organization boundaries and relationships. The second connection relates to causality. To some extent, the organization model assumes technical characteristics of the medium to be fixed, while the technical model assumes organization requirements to be fixed. By investigating the technical characteristics in depth, we are able to differentiate the fundamental from the artifact. Similarly, by investigating the use of IONs in several domains we identify those requirements and constraints that are common across types of organizations and relationships and those that vary. We can thereby evaluate which technical parameters must be variable. Our bimodal approach to the study of system design and use reflects our view of causality in which organization and societal contexts influence and structure the development of the technology, while application of the technology affects the organization and social contexts in which it is used.

In conclusion, this study of IONs served as a rich example of:

- A systems design effort that *necessitated* investigation of implicit and explicit assumptions about the organization context in which the systems are used.
- An analysis of the organization implications of a new technology that was based on a careful specification of the technology's fundamental technical characteristics.

We hope that this approach can serve as a model for analysis and design of other new technologies.

References

1. Anonymous. Electronic Data Interchange for the Grocery Industry. Feasibility Report. 83156, Arthur D. Little, Cambridge, MA, 1980.

2. Allen, T.. *Managing the Flow of Technology.* M.I.T. Press, Cambridge, MA, 1977.

3. Aronson, S. Bell's Electrical Toy: What's the Use? The Sociology of Early Telephone Usage. In *The Social Impact of the Telephone*, I. de Sola Pool, Ed., M.I.T. Press, Cambridge, MA, 1977, pp. 15-39.

4. Barett, S. *A Framework for the Analysis of Automated Inter-Organizational Information Sharing Systems.* Ph.D. Th., University of Arizona, Department of Business Administration, June 1983.

5. Bell, D., LaPadula, L. Secure Computer Systems. Technical Report ESD-TR-73-278, The Mitre Corp., Bedford, MA, June, 1974.

6. Benjamin, J., Hess, M., Weingarten, R., Wheeler, W. "Interconnecting SNA Networks". *IBM Systems Journal 22*, 4 (1983), 344-366.

7. Berson, T. Local Network Cryptosystem Architecture: Access Control. In *Advances in Cryptology. Proceedings of Crypto82*, Chaum, D., Rivest, R., Sherman, T., Eds., Plenum Press, New York, 1983, pp. 251-258.

8. Biba, K. Integrity Considerations for Secure Computer Systems. Technical Report ESD-TR-76-372, The Mitre Corp., Bedford, MA, April, 1977.

9. Birrell, A., Levin, R., Needham, R., Schroeder. M. "Grapevine: an exercise in distributed computing". *Communications of the ACM 25*, 4 (April 1982), 260-274.

10. Boggs, D. "Pup: An Internetwork Architecture". *IEEE Transactions on Communications COM-28*, 4 (April 1980), 612-623.

11. Braden, R., Cole, R. Some Problems in the Inter-connection of Computer Networks. In *Pathways to the Information Society: Proceedings of the 6th International Conference on Computer Communications*, Williams, W., Ed., North-Holland, 1982, pp. 969-974.

12. Cash, J. Information Systems That Change Company and Industry Boundaries. Graduate School of Business Administration, Harvard University, Boston, MA 02163, May, 1984.

13. Cerf, V., Kirstein, P. "Issues in Packet-Network Interconnection". *Proceedings of the IEEE 66*, 11 (November 1978), 1386-1408.

14. Coase, R. "The Nature of the Firm". *Economica 4* (1937), 386-405.

15. Cohen, D., Postel, J. Gateways, Bridges, and Tunnels in Computer Mail. In *Local Networks, Strategy and Systems, Proceedings of Localnet '83*,
Online Ltd., Northwood, UK, 1983, pp. 109-123.

16. Cole, R., Higginson, P., Lloyd, P., Moulton, R. "International net faces problems handling mail and file transfer". *Data Communications* (June 1983), 175-187.

17. Crawford, A. B., Jr. "Corporate Electronic Mail—A Communication-Intensive Application of Information Technology". *Management Information Systems Quarterly* (September 1982), 1-13.

18. Dallas, I. Implementation of a Gateway between a Cambridge Ring Local Area Network and a Packet Switching Wide Area Network. In *Pathways to the Information Society: Proceedings of the 6th International Conference on Computer Communications*, Williams, W., Ed., North-Holland, 1982, pp. 137-142.

19. DeSchon, A. MCI Mail/Arpa Mail Forwarding. Technical Report ISI/RR-84-141, USC Information Sciences Institute, August, 1984.

20. Estrin, D. Non-Discretionary Controls for Inter-Organization Networks. *Proceedings of the 1985 Symposium on Security and Privacy*, Silver Spring, MD, 1985, pp. 56-61.

21. Foley, J. "Business Data Usage of OSI". *Proceedings of the IEEE 71*, 12 (1983), 1442-1445.

22. Fuchs, I. "BITNET—Because It's Time". *Perspectives in Computing 3*, 1 (March 1983), 16-27.

23. Geller, V., Lesk, M. The Computer as a Communication Modality. Working Paper. AT&T Bell Laboratories, Murray Hill, NJ 07974., 1984.

24. Gifford, D. "Cryptographic Sealing for Information Secrecy and Authentication". *Communications of the ACM 25*, 4 (April 1982).

25. Gillies, D. Improved Network Security with a Trusted Mail Relay. S.B. Thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, June, 1984.

26. Goddard, J. Office Linkages and Location. In *Progress in Planning*, Diamond, D., McLoughlin, J., Eds., Pergamon Press, Oxford, 1973, pp. 111-231.

27. Grossman, G., Hinchley, A., Sunshine, C. "Issues in International Public Data Networking". *Computer Networks*, 3 (1979), 259-266.

28. Hall, R., Clark, J., Giordano, P., Johnson, P., Van Roekel, M. "Patterns of Interorganizational Relationships". *Administrative Science Quarterly 22* (September 1977), 457-472.

29. Horton, M. Standard for Interchange of USENET Messages. Request for Comments RFC 850, USC Information Sciences Institute, June, 1983.

30. ISO. Directives for the Technical Work of ISO. International Standards Organization, Geneva, Switzerland, 1982.

31. Jones, T. "Paving the way for universal document interchange". *Data Communications* (July 1982), 123+.

32. Karger, P. Non-Discretionary Access Control for Decentralized Computing Systems. S.M. Thesis. Massachusetts Institute Technology, Dept. of Electrical Engineering and Computer Science, May, 1977. Also available from the M.I.T. Laboratory for Computer Science as TR-179.

33. Karger, P. Security in DECnet: Authentication and Discretionary Access Control. Technical Report TR-121, Corporate Research Group, Digital Equipment Corporation, Hudson, MA 01754, January 8, 1982.

34. Kiesler, S., Zubrow, D., Moses, A., Geller, V. "Affect in Computer-Mediated Communication: An Experiment in Synchronous Terminal-to-Terminal Discussion". *Human-Computer Interaction 1*, 1 (1985), 77-104.

35. Landweber, L., Solomon, M. Use of Multiple Networks in CSNET. *IEEE COMPCON Spring '82*, IEEE Computer Society, February, 1982, pp. 398-402. San Francisco, California.

36. Landwehr, C., Heitmeyer, C., McCleen, J. "A Security Model for Military Message Systems". *ACM Transactions on Computer Systems 2*, 3 (August 1984), 198-222.

37. Lipner, S. Non-Discretionary Controls for Commercial Applications. *Proceedings of the 1982 Symposium on Security and Privacy*, IEEE Computer Society, Oakland, California, April, 1982, pp. 2-10.

38. Macneil, I. "Contracts: Adjustment of Long-Term Economic Relations Under Classical, Neoclassical, and Relational Contract Law". *Northwestern University Law Review 72*, 4 (1978), 854-905.

39. Marrett, C. "On the Specification of Interorganizational Dimensions". *Sociology and Social Research 56* (October 1971), 83-99.

40. Mockapetris. P. The Domain Name System. In *Computer-Based Message Services.* Smith. H., Ed., Elsevier Science Publishers B.V., North-Holland, 1984.

41. Mogul. J. Internet Subnets. Request for Comments RFC 917. USC Information Sciences Institute, October, 1984.

42. Mracek, J. Network Access Control in Multi-Net Internet Transport. S.B. Thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, June, 1983.

43. Nakamoto, S., Ching, C. Electronic Mail: Aa Comparative Evaluation. Information Text Series 021, University of Hawaii, Department of Agriculture and Resource Economics, Hawaii Institute of Tropical Agriculture and Human Resources, Univeristy of Hawaii, Honolulu, HI 96844, 1984.

44. Needham, R., Schroeder, M. "Using Encryption for Authentication in Large Networks of Computers". *Communications of the ACM 21*, 12 (December 1978), 993-999.

45. Newell, A., Sproull, R. "Computer Networks: Prospects for Scientists". *Science 215* (February 12 1982), 843-852.

46. Nowitz, D., Lesk, M. A Dial-Up Network Of Unix^{TM} Systems. Bell Laboratories, Murray Hill, New Jersey 07974, August, 1978.

47. Panko, R. "The Cost of EMS". *Computer Networks*, 5 (1981), 35-46.

48. Pfeffer, J., Salancik, G.. *The External Control of Organizations* . Harper & Row, New York, 1978.

49. Picot, A., Klingenberg, H., Kranzle, H. Organizational Communication: The relationship between Technological Development and Socio-economic Needs. In *Information Technology: Impact on the way of life*,
National Board for Science and Technology, Ireland and Commission of the European Communities, FAST Programme, Dublin, Ireland, 1981, pp. 114-132.

50. Pillsbury, A. "The Hard-Selling Supplier to the Sick". *Fortune* (July 26 1982), 56-61.

51. Pool, I. de Sola (Ed.). *The Social Impact of the Telephone.* MIT Press, Cambridge, MA, 1977.

52. Pool, I. de Sola. *Technologies of Freedom.* Belknap Press, Cambridge, MA, 1983.

53. Popek, G., Kline, C. "Encryption and Secure Computer Networks". *Computing Surveys 11*, 4 (December 1979), 331-356.

54. Postel, J. Internet Protocol—DARPA Internet Program Protocol Specification. Request for Comments RFC 791, USC Information Sciences Institute, September, 1981.

55. Rice, R., Case, D. "Computer-Based Messaging in the University: A Description of Use and Utility". *Journal of Communication 33*, 1 (1983), 131-152.

56. Rice, R. and Associates. *The New Media.* Sage Publications, Beverly Hills, CA, 1984.

57. Rotenberg, L. *Making Computers Keep Secrets.* Ph.D. Th., Massachusetts Institute Technology, Dept. of Electrical Engineering and Computer Science, February 1974. Also available from the M.I.T. Laboratory for Computer Science as TR-115.

58. Routhier, S. An Improved Authentication Server For Inter-Computer Communication. S.B. Thesis, Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science, June, 1983.

59. Saltzer, J., Clark, D., Reed, D. "End-to-End Arguments in System Design". *ACM Transactions on Computer Systems 2*, 4 (November 1984), 277-288.

60. Saltzer, J., Schroeder, M. "The Protection of Information in Computer Systems". *Proceedings of the IEEE 63*, 9 (September 1975), 1278-1308.

61. Saltzer, J. On the Naming and Binding of Network Destinations. In *Local Computer Networks*, Ravisio, P.C., Hopkins, G., Naffah, N., Eds., North-Holland Publishing Company, New York, 1982, pp. 311-318.

62. Schick, T., Brockish, F. "The Document Interchange Architecture: A member of a family of architectures in the SNA environment". *IBM Systems Journal 21*, 2 (1982), 220-244.

63. Schicker, P. "Conference Report, IFIP WG6.5 Workshop on International Computer Message Services". *Computer Networks* (February 1983), 47-57.

64. Sirbu, M., Sutherland, J. Naming and Directory Issues in Message Transfer Systems. In *Computer-Based Message Services*, Smith, H., Eds., Elsevier Science Publishers B.V. (North-Holland), 1984.

65. Slack, Daryl J.. *Communication Technologies and Society: Concepts of Causality and the Politics of Technological Intervention.* Ablex, Norwood, New Jersey, 1984.

66. Sproull, L. Using Electronic Mail for Data Collection in Organizational Research. Committee on Social Science Research in Computing Working Paper Series, Department of Social Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, December 13, 1984.

67. Stern, L., ElAnsary, A.. *Marketing Channels.* Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

68. Sunshine, C. "Source Routing in Computer Networks". *ACM SIGCOMM Computer Communication Review 7*, 1 (January 1977), 29-33.

69. Sunshine, C. "Interconnection of Computer Networks". *Computer Networks*, 1 (1977), 175-195.

70. Symbolics Inc. Networks. Documentation for Release 6, Volume 9. Document #996095, Symbolics Inc., Cambridge, MA, March, 1985.

71. Transportation Data Coordinating Committee. The United States Electronic Data Interchange (EDI) Standards. Office of Facilitation, Research and Special Programs Administration, U.S. Department of Transportation, Washington, D.C., May, 1979.

72. Tichy, N., Tushman, M., Fombrun, C. "Social Network Analysis for Organizations". *Society of Management Review 4*, 4 (1979), 507-519.

73. Van de Ven, A., Walker, G., Liston, J. "Coordination Patterns Within an Interorganizational Network". *Human Relations 32*, 1 (1979), 19-36.

74. Veith, R. *Multinational Computer Nets*. Lexington Books, Lexington, MA, 1981.

75. Voydock, V., Kent, S. "Security Mechanisms in High-Level Network Protocols". *ACM Computing Surveys 15*, 2 (June 1983), 135-171.

76. Walker, G., Townsend, E. Interorganizational Effectiveness, Efficiency and Network Structure. Working Paper 1349-82, Sloan School of Management, M.I.T., 50 Memorial Drive, Cambridge, MA 02139, August, 1982.

77. Walker, G., Weber, D. "A Transaction Cost Approach to Make-or-Buy Decisions". *Administrative Science Quarterly* (September 1984), 373-391.

78. Williamson, O.. *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York, 1975.

79. Williamson, O. "Transaction-Cost Economics: The Governance of Contractual Relations". *The Journal of Law and Economics 22* (October 1979), 223-261.

80. Williamson, O. Mitigating Contractual Hazards: Bilateral Equilibration. *Transaction Cost Economics Workshop*, University of Pennsylvania, Philadelphia PA, Febrary 23, 1982.

81. Yudkin, M. Resource Management in a Distributed System. *Proceedings of the Eighth Data Communications Symposium*, IEEE Computer Society, October 3-6, 1983, pp. 221-226.

Appendix A

Online Questionnaire

The following questionnaire was distributed online at all sites.

Online Questionnaire: Inter-Organization Networks

Computer-based communication and resource sharing ACROSS organization boundaries are the focus of my doctoral research in the MIT Lab for Computer Science. As a test case, I am studying the effects of INTER-ORGANIZATION NETWORKS on communication amoag Research Laboratories. I am seeking responses to the following questionnaire. The 5 multiple-part questions are all short answer or multiple choice. And as you will see, because I am primarily interested in detecting patterns of change, the questions do NOT require ultra-detailed answers. Please do take the few minutes to respond; it takes most people between 5 and 15 minutes. All information will be treated confidentially. You may respond online (to estrin@mit-xx) by inserting your responses after each question or by numbering your responses. Or respond on paper by printing the questions double spaced and writing in or numbering your answers; send to Deborah Estrin, NE43-508, MIT.

NOTE: I refer to electronic mail, file transfer, remote login, database, and other computer-based communication mechanisms as INTER-ORGANIZATION NETWORK (ION) FACILITIES. Telephone, face-to-face meetings, and postal mail are referred to as TRADITIONAL MEDIA. EXTERNAL ORGANIZATIONS are government, university, or industrial laboratories outside of your university.

1) EXTERNAL INFORMATION AND RESOURCE SHARING:

a) During an average work week, with about HOW MANY EXTERNAL ORGANIZATIONS do you exchange work-related information (e.g. research ideas, tools and techniques) or resources (e.g. equipment, software, data bases, computer services) VIA ION FACILITIES ? If the answer is 0, please skip to question 4.

b) About HOW LONG AGO did you first begin using ION facilities to communicate with these and other external organizations? (number of months or years)

c) Since you began using ION facilities, is the NUMBER OF ORGANIZATIONS with which you share information or resources less, the same, or greater than it was when you used only traditional media?

d) Since you began using ION facilities, is the NUMBER OF RESEARCH PROJECTS that involve information or resource sharing with external organizations smaller, the same, or larger than it was when you used only traditional media?

e) To what extent do you attribute the changes indicated in (c) and (d) to the use of ION facilities? (not at all, some, quite a bit, very much)

f) Identify the individual organizations with which you exchange work related information or resources via ION facilities most intensively; select no more than 3 or 4. Assign a code letter to each one (i.e., a,b,c) and indicate whether each is a university(u), government(g), or industrial(i) lab. Bulletin boards and distribution lists do NOT qualify as organizations per se; please do not include more than one of these among the 3 or 4 organizations.

2) COMMUNICATION AND ACCESS PATTERNS:

Respond to the following questions by listing each organization's code letter (assigned above) followed by the appropriate answer for that organization.

a) Approximately HOW MANY people in EACH of these organization do you communicate with via ION facilities during an average work week?

b) Since you began using ION facilities, is the NUMBER of people that you communicate with per organization less, the same, or greater than it was when you used only traditional media?

c) HOW OFTEN do you communicate with people or machines in EACH of these organizations via ION facilities during an average work week (0 times, 1 time, 2-5 times, 6-10 times, more)?

d) Since you began using ION facilities, is the FREQUENCY of communication with each of these organizations less, the same, or greater than it was when you used only traditional media?

e) Since you began using ION facilities, do you communicate with each of these organizations via TRADITIONAL media less,the same,or more than you did when you used only traditional media?

f) For each of these organizations, which of the following INFORMATION and RESOURCE TYPES do you exchange via ION facilities ? INFORMATION: (1)research ideas (2)research results (3)joint authorship comments (4)information for solving a particular problem (5)information about tools and techniques (6)administrative scheduling (7) Other, please specify. RESOURCES: (8)software (9)computer resources (10)remote applications (e.g.,Macsyma, VLSI tools) (11)database (12)Other, please specify. (List each organization's code letter followed by the appropriate numbers.)

g) For each of these organizations, indicate if the average amount of EACH INFORMATION and RESOURCE TYPE exchanged per week is less, the same, or greater than it was when you used only traditional media.

h) Since you began using ION facilities to communicate with these outside organizations, has your communication with outside organizations that are NOT accessible via ION facilities changed? Indicate if the average amount of EACH INFORMATION and RESOURCE TYPE exchanged with the non-ION organizations is less, the same, or greater.

i) Which of the information and resource types do you exchange with people INSIDE your organization via internal computer facilities ?

j) For each of the external organizations that you communicate with via ION facilities (identified in 1f), which of the following CLASSES of INFORMATION and RESOURCES do you exchange via ION facilities ? INFORMATION: (1)publicly available (2)available in internal documents only (3)related to unpublished research (4)related to unreleased system or product (5)proprietary (6)Other, please specify. RESOURCES: (7)widely available (8)limited (9)costly (10)critical for internal operations (12)proprietary (11)Other, please specify.

k) For each of these organizations, indicate if the average amount of EACH information and resource CLASS exchanged per week is less, the same, or greater than it was when you used only traditional media.

l) To what extent do you attribute the changes indicated in (b),(d),(e),(g), (h),(k) to the use of ION facilities? (not at all,some,quite a bit,very much) If appropriate, provide a separate response for each of the 6 questions (b,d, e,g,h,k).

3) CONTRACTS AND RESTRICTIONS

a) What kinds of AGREEMENTS exist between your organization and each of the individuals or organizations that you communicate with via ION facilities? (none,informal,consulting contract,joint development contract,other specify)

b) Indicate if these agreements differ from the agreements governing relationships that use only traditional media (no difference,more explicit conditions,more protective,more exclusive to other organizations,more open-ended or illdefined,other please specify) ?

c) Indicate if any of the following factors significantly INHIBIT your using ION facilities more extensively (destinations inaccessible,inconvenient, poor performance,confidentiality of information,company policy,none, other please specify)

4) BACKGROUND:

a) About HOW MANY RESEARCH PROJECTS are you working on currently that involve regular contact with persons in organizations outside of your own company/university ?

b) During an average work week, with about HOW MANY EXTERNAL ORGANIZATIONS do you share work related information or resources (via either traditional or ION facilities) ?

c) Which aspect(s) of research/development do you work in, primarily? (software, hardware, theory, systems, applications, other please specify)

d) Which job category do you belong to, primarily? (manager,faculty, scientist, research staff, technical staff, other please specify)

e) How often do you use a computer of some kind in conjunction with your work? (daily,several times a week,once a week,monthly,other,please specify)

5) COMMENTS:
If you use ion facilities in interesting ways that the above questions have not touched upon, please describe them here.

Thank you very much for your time! Deborah Estrin

The following table outlines the mapping of questions to hypothesis variable.

Question	Hypothesis variable
1a	number of interchange partners
1c	change in number of interchange partners
1d	change in number of cross boundary projects
2c	intensity
2d	change in intensity
2f	scope
2g	change in scope
2h	segmentation
2j	penetration
2k	change in penetration
3a	restrictions
3b	change in restrictions
1e,2l	change attributed to ION
	Context variable
1b	years of ION use
1f	types of outside organizations
2a	number of people per outside organization
2b	change in number of people
2e	change in use of traditional media
2i	info and resource types exchanged internally
3c	inhibitions to ION use
4a	*total number of projects with external input*
4b	total number of outside organizations
4c	area
4d	job
4e	computer usage

Appendix B

Numbers of respondents per site

Industrial Labs	No ION Use Reported	ION Use Reported
1	3 100% of row 2% of column	0 0% of row 0% of column
2	4 80% of row 3% of column	1 20% of row 1% of column
3	7 88% of row 6% of column	1 13% of row 2% of column
4	0 0% of row 0% of column	1 100% of row 2% of column
5	3 100% of row 2% of column	0 0% of row 0% of column
6	4 100% of row 3% of column	0 0% of row 0% of column
7	3 75% of row 2% of column	1 25% of row 2% of column
8	3 60% of row 2% of column	2 40% of row 3% of column
9	1 100% of row 1% of column	0 0% of row 0% of column
10	4 57% of row 3% of column	3 43% of row 5% of column
11	44 49% of row 35% of column	45 51% of row 68% of column
12	13 81% of row 10% of column	3 19% of row 5% of column

p<.05
chi-sq = 19.8
d.f. = 11

This table lists the number of respondents per industry site. Respondents are divided according to whether or not they reported using the ION for work-related communications.

University Labs	No ION Use Reported	ION Use Reported
13	1 100% of row 1% of column	0 0% of row 0% of column
14	4 80% of row 3% of column	1 20% of row 2% of column
15	1 100% of row 1% of column	0 0% of row 0% of column
16	4 50% of row 3% of column	4 50% of row 6% of column
17	1 100% of row 8% of column	0 0% of row 0% of column
18	7 88% of row 6% of column	1 13% of row 2% of column
19	1 100% of row 1% of column	0 0% of row 0% of column
20	1 100% of row 1% of column	0 0% of row 0% of column
21	3 75% of row 2% of column	1 25% of row 2% of column
22	1 100% of row 1% of column	0 0% of row 0% of column
23	4 80% of row 3% of column	1 20% of row 2% of column
24	3 100% of row 2% of column	0 0% of row 0% of column

(continued)

This table lists the number of respondents per university site. Respondents are divided according to whether or not they reported using the ION for work-related communications.

		No ION Use Reported	ION Use Reported	
University Labs (cont.)	25	4 80% of row 3% of column	1 20% of row 2% of column	
	26	1 100% of row 1% of column	0 0% of row 0% of column	

	27	1
		100% of row
		1% of column
			0
			0% of row
			0% of column
				p<.90
				chi-sq = 7.7
				d.f. = 14

This is a continuation of the table listing the number of respondents per university site.

Appendix C

Cross-Tabulations of Variables by Site Type

(A)

Frequency of communication with outsiders.

	Indust	Univ	
Less	2 3%	0 0%	2 2%
Same	13 20%	2 7%	15 16%
Greater	50 77%	27 93%	77 82%

	65
	69%
		29
		31%
			p<.16
			chi-sq=3.7
			d.f.=2

(B)

Types of information and resources exchanged.

	Indust	Univ	
Less	0 0%	0 0%	0 0%
Same	13 20%	1 4%	14 15%
Greater	51 80%	26 96%	77 85%

	64
	70%
		27
		30%
			p<.09
			chi-sq=2.8
			d.f.=1

(C)

Number of outside organizations.

	Indust	Univ	
Less	0 0%	0 0%	0 0%
Same	5 9%	1 3%	6 7%
Greater	54 92%	28 97%	82 93%

	59
	67%
		29
		33%
			p<.67
			chi-sq=0.2
			d.f.=1

(D)

Number of projects with outside input.

	Indust	Univ	
Less	0 0%	0 0%	0 0%
Same	21 37%	6 22%	27 32%
Greater	36 63%	21 78%	57 68%

	57
	68%
		27
		32%
			p<.28
			chi-sq=1.2
			d.f.=1

(E)

Number of people contacted per outside organization.

	Indust	Univ	
Less	0 0%	1 4%	1 1%
Same	27 42%	10 36%	37 40%
Greater	38 59%	17 61%	55 59%

	65
	70%
		28
		30%
			p<.29
			chi-sq=2.5
			d.f.=2

(F)

Use of traditional media for outside communication.

	Indust	Univ	
Less	32 49%	18 67%	50 54%
Same	25 39%	7 26%	32 35%
Greater	8 12%	2 7%	10 11%

	65
	70%
		27
		30%
			p<.31
			chi-sq=2.4
			d.f.=2

(G)

Classes of information and resources exchanged.

	Indust	Univ	
Less	2 3%	0 0%	2 2%
Same	18 28%	2 7%	20 21%
Greater	45 69%	27 93%	72 77%

	65
	69%
		29
		31%
			p<.04
			chi-sq=6.5
			d.f.=2

(H)

Communication with non-ION organizations.

	Indust	Univ	
Less	13 22%	9 36%	22 26%
Same	44 75%	15 60%	59 70%
Greater	2 3%	1 4%	3 4%

	59
	70%
		25
		30%
			p<.40
			chi-sq=1.9
			d.f.=2

(I)

Contract with first outside organization.

	Indust	Univ	
None	30 47%	19 68%	49 53%
Informal	7 11%	4 14%	11 12%
Consult.	7 11%	3 11%	10 11%
Joint	14 11%	2 11%	16 11%
Other	6 9%	0 0%	6 7%

	64
	70%
		28
		30%
			p<.23
			chi-sq=6.8
			d.f.=5

(J)

Contract with second outside organization.

	Indust	Univ	
None	24 50%	15 56%	39 52%
Informal	5 10%	9 33%	14 19%
Consult.	7 15%	1 4%	8 11%
Joint	6 13%	1 4%	7 9%
Other	6 13%	1 4%	7 9%

	48
	64%
		27
		36%
			p<.05
			chi-sq=11.3
			d.f.=5

(K)

Contract with third outside organization.

	Indust	Univ	
None	15 46%	14 64%	29 53%
Informal	4 12%	5 23%	9 16%
Consult.	3 9%	1 5%	4 7%
Joint	7 21%	1 5%	8 15%
Other	4 12%	1 5%	5 9%

	33
	60%
		22
		40%
			p<.24
			chi-sq=6.7
			d.f.=5

(L)

Contract with fourth outside organization.

	Indust	Univ	
None	10 71%	3 50%	13 65%
Informal	1 7%	3 50%	4 20%
Consult.	0 0%	0 0%	0 0%
Joint	1 7%	0 0%	1 5%
Other	2 14%	0 0%	2 10%

	14
	70%
		6
		30%
			p<.14
			chi-sq=5.4
			d.f.=5

(M)

Change in contract with first outside organization.

	Indust	Univ	
None	41 77%	21 96%	62 83%
Explicit	3 6%	0 0%	13 4%
Protect	3 6%	0 0%	13 4%
Less	6 11%	1 5%	7 9%

	53
	70%
		22
		30%
			p<.28
			chi-sq=3.9
			d.f.=3

(N)

Change in contract with second outside organization.

	Indust	Univ	
None	35 90%	20 95%	55 92%
Explicit	0 0%	0 0%	0 0%
Protect	1 3%	0 0%	1 2%
Less	3 8%	1 5%	4 7%

	39
	65%
		21
		35%
			p<.68
			chi-sq=0.8
			d.f.=2

(O)

Change in contract with third outside organization.

	Indust	Univ	
None	27 46%	17 64%	44 53%
Explicit	0 0%	0 0%	0 0%
Protect	1 3%	0 0%	1 2%
Less	1 3%	1 6%	2 4%

	29
	62%
		18
		38%
			p<.69
			chi-sq=0.7
			d.f.=2

(P)

Change in contract with fourth outside organization.

	Indust	Univ	
None	12 92%	7 88%	19 91%
Explicit	0 0%	0 0%	0 0%
Protect	1 8%	0 0%	1 5%
Less	0 0%	1 13%	1 5%

	13
	62%
		8
		38%
			p<.32
			chi-sq=2.3
			d.f.=2

(Q)

Inaccessible locations inhibit ION use.

	Indust	Univ	
Yes	44	12	56
	67%	41%	59%
No	22	17	39
	33%	59%	41%
	66	29	p<.04
	70%	30%	chi-sq=4.3
			d.f.=1

(R)

Proprietary concerns inhibit ION use.

	Indust	Univ	
Yes	41	23	64
	62%	79%	67%
No	25	6	31
	38%	21%	33%
	66	29	p<.16
	70%	30%	chi-sq=2.0
			d.f.=1

(S)

Change in cross-boundary activities attributed to ION.

	Indust	Univ	
None	7	2	9
	13%	8%	11%
Some	23	8	31
	41%	31%	38%
Quite	9	6	15
	16%	23%	18%
Very	17	10	27
	30%	39%	33%
	56	26	p<.64
	68%	32%	chi-sq=1.7
			d.f.=3

(T)

Change in communications attributed to ION.

	Indust	Univ	
None	0	0	0
	0%	0%	0%
Some	16	4	20
	24%	14%	21%
Quite	16	11	27
	24%	38%	28%
Very	34	14	48
	52%	48%	51%
	66	29	p<.14
	70%	30%	chi-sq=2.4
			d.f.=2

(U)

Type of first organization.

	Indust	Univ	
Univ	52 62%	20 54%	72 60%
Indust	21 25%	9 24%	30 25%
Gov't	6 7%	8 22%	14 12%
BBoard	4 5%	0 0%	4 3%
Other	1 1%	0 0%	1 1%

	84
	70%
		37
		30%
			p<.13
			chi-sq = 7.1
			d.f. = 4

(V)

Type of second organization.

	Indust	Univ	
Univ	40 65%	16 44%	56 57%
Indust	10 16%	16 44%	26 27%
Gov't	9 15%	2 6%	11 11%
BBoard	3 5%	2 6%	5 5%
Other	0 0%	0 0%	0 0%

	62
	63%
		36
		37%
			p<.02
			chi-sq = 10.1
			d.f. = 3

(W)

Type of third organization.

	Indust	Univ	
Univ	26 63%	13 46%	39 57%
Indust	8 20%	12 43%	20 29%
Gov't	2 5%	1 4%	3 4%
BBoard	4 10%	1 4%	5 7%
Other	1 2%	1 4%	2 3%

	41
	60%
		28
		40%
			p<.29
			chi-sq = 5.0
			d.f. = 4

(X)

Type of fourth organization.

	Indust	Univ	
Univ	6 43%	7 70%	13 54%
Indust	5 36%	2 20%	7 29%
Gov't	1 7%	0 0%	1 4%
BBoard	2 14%	1 10%	3 13%
Other	0 0%	0 0%	0 0%

	14
	58%
		10
		42%
			p<.55
			chi-sq = 2.1
			d.f. = 3

(Y)

Total number of projects with outside input.

	Indust	Univ	
0	19 21%	4 11%	23 18%
1	27 30%	11 30%	38 30%
2	20 22%	8 22%	28 22%
3	13 15%	7 19%	20 16%
4	1 1%	2 5%	3 2%
>5	9 10%	5 14%	14 11%

	89
	70%
		37
		30%
			p<.52
			chi-sq=4.2
			d.f.=5

(Z)

Number of years of ION use.

	Indust	Univ	
0	5 6%	2 5%	7 6%
1	25 28%	4 11%	29 23%
2	21 24%	6 16%	27 21%
3	10 11%	6 16%	16 13%
>4	28 32%	19 51%	47 37%

	89
	70%
		37
		30%
			p<.12
			chi-sq=7.3
			d.f.=4

(AA)

Frequency of computer use.

	Indust	Univ	
Daily	55 71%	27 73%	82 72%
Constant	22 29%	10 27%	32 28%
Other	0 0%	0 0%	0 0%

	77
	68%
		37
		33%
			p<1
			chi-sq=0.0
			d.f.=1

(BB)

Area of research/development.

	Indust	Univ	
Software	61 69%	19 51%	80 64%
Hardware	6 7%	1 3%	7 6%
Theory	6 7%	6 16%	12 10%
Systems	9 10%	10 27%	19 15%
AI	6 7%	1 3%	7 6%

	88
	70%
		37
		30%
			p<.04
			chi-sq=10.1
			d.f.=4

(GG)

Job title.

	Indust	Univ	
Manager	22 25%	2 5%	24 19%
Faculty	0 0%	22 60%	22 18%
Rsch Staff	22 25%	7 19%	29 23%
Tech Staff	37 42%	1 3%	38 30%
Scientist	5 6%	2 5%	7 6%
Other	2 2%	3 8%	2 3%

	88
	70%
		37
		30%
			p<.00
			chi-sq = 75.6
			d.f. = 6

(II)

Resources exchanged (in addition to information).

	Indust	Univ	
No	44 49%	21 57%	65 52%
Yes	45 51%	16 43%	61 48%

	89
	70%
		37
		30%
			p<.58
			chi-sq = 0.3
			d.f. = 1

(HH)

Number of outside organizations reported.

	Indust	Univ	
1	27 30%	1 3%	28 22%
2	21 24%	8 22%	29 23%
3	28 32%	18 49%	46 37%
4	13 15%	10 27%	23 18%

	89
	70%
		37
		30%
			p<..00
			chi-sq = 13.3
			d.f. = 3

(JJ)

Total number of outside organizations (non-ION included).

	Indust	Univ	
1	27 30%	10 27%	37 29%
2	21 24%	6 16%	27 21%
3	16 18%	7 19%	18 18%
4	6 7%	2 5%	8 6%
>5	19 21%	12 32%	31 25%

	89
	70%
		37
		30%
			p<.70
			chi-sq = 2.2
			d.f. = 4

(KK)

Average frequency of ION use (per week).

	Indust	Univ	
0	7 8%	4 11%	11 9%
1	29 34%	13 35%	42 34%
2-5	39 46%	20 54%	59 49%
6-10	6 7%	0 0%	6 5%
>10	4 5%	0 0%	4 3%

	85
	70%
		37
		30%
			p<.25
			chi-sq=9.0
			d.f.=7

(LL)

Average number of people in outside org (via ION).

	Indust	Univ	
0	1 1%	1 3%	2 2%
1	31 37%	19 51%	40 41%
2	28 33%	11 30%	39 32%
3	6 7%	4 11%	10 8%
4	5 6%	1 3%	6 5%
5	3 4%	0 0%	3 3%
>5	10 12%	1 3%	11 9%

	84
	70%
		37
		30%
			p<.82
			chi-sq=8.4
			d.f.=14

Appendix D

Cross-Tabulations of Types and Classes of Information and Resources by Site Type

(A)

Research ideas exchanged.

	Indust	Univ	
No	32 49%	8 28%	40 42%
Yes	034 52%	21 72%	55 58%

	66
	70%
		29
		30%
			p<.09
			chi-sq=2.8
			d.f.=1

(B)

Research results exchanged.

	Indust	Univ	
No	43 66%	12 41%	55 59%
Yes	22 34%	17 59%	39 42%

	65
	70%
		29
		30%
			p<.04
			chi-sq=4.1
			d.f.=1

(C)

Joint authorship comments exchanged.

	Indust	Univ	
No	55 83%	19 66%	74 78%
Yes	11 17%	10 35%	21 22%

	66
	70%
		29
		30%
			p<.10
			chi-sq=2.8
			d.f.=1

(D)

Information for solving a particular problem.

	Indust	Univ	
No	23 35%	9 31%	32 34%
Yes	43 65%	20 69%	63 66%

	66
	70%
		29
		30%
			p<.90
			chi-sq=0.0
			d.f.=1

(E)

Information about tools and techniques.

	Indust	Univ	
No	25 38%	8 28%	33 35%
Yes	41 62%	21 72%	62 65%

	66
	70%
		29
		30%
			p<.46
			chi-sq=0.5
			d.f.=1

(F)

Administrative scheduling.

	Indust	Univ	
No	31 47%	14 48%	45 47%
Yes	35 53%	15 52%	50 53%

	66
	70%
		29
		30%
			p<1
			chi-sq=0.0
			d.f.=1

(G)

Other information types exchanged.

	Indust	Univ	
No	58 88%	26 90%	84 88%
Yes	8 12%	3 10%	11 12%

	66
	70%
		29
		30%
			p<.1
			chi-sq = 0.0
			d.f. = 1

(H)

Software exchanged.

	Indust	Univ	
No	36 55%	15 52%	51 54%
Yes	30 46%	14 48%	44 46%

	66
	70%
		29
		30%
			p<.98
			chi-sq = 0.0
			d.f. = 1

(I)

Computer resources accesssed/exchanged.

	Indust	Univ	
No	59 89%	25 86%	84 88%
Yes	7 11%	4 14%	11 12%

	66
	70%
		29
		30%
			p<.92
			chi-sq = 0.0
			d.f. = 1

(J)

Remote applications accessed.

	Indust	Univ	
No	58 88%	27 93%	85 90%
Yes	8 12%	2 7%	10 11%

	66
	70%
		29
		30%
			p<.69
			chi-sq = 0.2
			d.f. = 1

(K)

Database accessed.

	Indust	Univ	
No	62 94%	29 100%	91 96%
Yes	4 6%	0 0%	4 4%

	66
	70%
		29
		30%
			p<.42
			chi-sq = 0.6
			d.f. = 1

(L)

Other resources accessed/exchanged.

	Indust	Univ	
No	60 92%	29 100%	89 95%
Yes	5 8%	0 0%	5 5%

	65
	70%
		29
		30%
			p<.30
			chi-sq = 1.1
			d.f. = 1

(M)

Publicly available information exchanged.

	Indust	Univ	
No	13 20%	8 28%	21 22%
Yes	53 80%	21 72%	74 78%

	66
	70%
		29
		30%
			p<.56
			chi-sq=0.3
			d.f.=1

(N)

Internal only documents exchanged.

	Indust	Univ	
No	35 54%	12 41%	47 50%
Yes	30 46%	17 59%	47 50%

	65
	70%
		29
		30%
			p<.37
			chi-sq=0.8
			d.f.=1

(O)

Unpublished research exchanged.

	Indust	Univ	
No	25 39%	7 24%	32 34%
Yes	40 62%	22 76%	62 66%

	65
	70%
		29
		30%
			p<.26
			chi-sq=1.2
			d.f.=1

(P)

Unreased system/product information exchanged.

	Indust	Univ	
No	45 69%	17 59%	62 66%
Yes	20 31%	12 41%	32 34%

	65
	70%
		29
		30%
			p<.44
			chi-sq=0.6
			d.f.=1

(Q)

Proprietary information exchanged.

	Indust	Univ	
No	53 82%	24 83%	77 82%
Yes	12 19%	5 17%	17 18%

	66
	70%
		29
		30%
			p<1
			chi-sq=0.0
			d.f.=1

(R)

Other classes of information exchanged.

	Indust	Univ	
No	63 97%	26 90%	89 95%
Yes	1 3%	3 10%	4 5%

	65
	70%
		29
		30%
			p<.34
			chi-sq=0.9
			d.f.=1

(S)

Widely available resources accessed/made accessible.

	Indust	Univ	
No	41 63%	22 76%	63 67%
Yes	24 37%	7 24%	31 33%

	65
	70%
		29
		30%
			p<.33
			chi-sq = 1.0
			d.f. = 1

(T)

Resources that are limited accessed/made accessible.

	Indust	Univ	
No	40 62%	19 66%	47 63%
Yes	25 39%	10 35%	35 37%

	65
	70%
		29
		30%
			p<.89
			chi-sq = 0.0
			d.f. = 1

(U)

Costly resources.

	Indust	Univ	
No	57 88%	25 86%	82 87%
Yes	8 12%	4 14%	12 13%

	65
	70%
		29
		30%
			p<1
			chi-sq = 0.0
			d.f. = 1

(V)

Resources that are critical for internal operations.

	Indust	Univ	
No	59 91%	26 90%	85 90%
Yes	6 9%	3 10%	9 10%

	65
	70%
		29
		30%
			p<1
			chi-sq = 0.0
			d.f. = 1

(W)

Proprietary resources.

	Indust	Univ	
No	64 99%	28 97%	92 98%
Yes	1 2%	1 3%	2 2%

	66
	70%
		29
		30%
			p<1
			chi-sq = 0.0
			d.f. = 1

(X)

Other classes of resources accessed/made accessible.

	Indust	Univ	
No	57 88%	28 97%	85 90%
Yes	8 12%	1 3%	9 10%

	65
	70%
		29
		30%
			p<.33
			chi-sq = 1.0
			d.f. = 1

OFFICIAL DISTRIBUTION LIST

1985

Director 2 Copies
Information Processing Techniques Office
Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

Office of Naval Research 2 Copies
800 North Quincy Street
Arlington, VA 22217
Attn: Dr. R. Grafton, Code 433

Director, Code 2627 6 Copies
Naval Research Laboratory
Washington, DC 20375

Defense Technical Information Center 12 Copies
Cameron Station
Alexandria, VA 22314

National Science Foundation 2 Copies
Office of Computing Activities
1800 G. Street, N.W.
Washington, DC 20550
Attn: Program Director

Dr. E.B. Royce, Code 38 1 Copy
Head, Research Department
Naval Weapons Center
China Lake, CA 93555

Dr. G. Hopper, USNR 1 Copy
NAVDAC-OOH
Department of the Navy
Washington, DC 20374