AFRL-IF-RS-TR-2001-84
Final Technical Report
April 2001

ADVANCED COMPUTING SYSTEMS

USC/Information Sciences Institute

Sponsored by
Defense Advanced Research Projects Agency

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

20010607 013

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-IF-RS-TR-2001-84 has been reviewed and is approved for publication.

Raymond A. Liuzzi

APPROVED: RAYMOND A. LIUZZI
Project Engineer

James A. Collins

FOR THE DIRECTOR: JAMES A. COLLINS, Acting Chief
Information Technology Division
Information Directorate

If your address has changed or if you wish to be removed from the Air Force Research Laboratory Rome Research Site mailing list, or if the addressee is no longer employed by your organization, please notify AFRL/IFTD, 525 Brooks Road, Rome, NY 13441-4505. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

ADVANCED COMPUTING SYSTEMS

Herbert Schorr

Contractor: USC/Information Sciences Institute
Contract Number: F30602-91-C-0146
Effective Date of Contract: 19 September 1991
Contract Expiration Date: 31 July 1999
Program Code Number: 62301E
Short Title of Work: Advanced Computing Systems

Period of Work Covered: Sep 91 – Jul 99

Principal Investigator: Herbert Schorr
Phone: (310) 822-1511
AFRL Project Engineer: Raymond A. Liuzzi
Phone: (315) 330-3577

Approved for public release; distribution unlimited.

This research was supported by the Defense Advanced Research Projects Agency of the Department of Defense and was monitored by Raymond A. Liuzzi, AFRL/IFTD, 525 Brooks Rd, Rome, NY.

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 2001	Final Sep 91 - Jul 99

4. TITLE AND SUBTITLE

ADVANCED COMPUTING SYSTEMS

5. FUNDING NUMBERS

C - F30602-91-C-0146
PE - 62301E
PR - H137
TA - 00
WU - 01

6. AUTHOR(S)

Herbert Schorr

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

USC/Information Sciences Institute
4676 Admiralty Way
Marina Del Rey, CA 90292-6601

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Advanced Research Projects Agency
3701 North Fairfax Drive
Arlington, VA 22203-1714

AFRL/IFTD
525 Brooks Road
Rome, NY 13441-4505

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AFRL-IF-RS-TR-2001-84

11. SUPPLEMENTARY NOTES

AFRL Project Engineer: Raymond A. Liuzzi, IFTD, 315-330-3577

12a. DISTRIBUTION AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Advanced Computing Systems effort made provision for a wide range of research activities that could be conducted under the overall contract. The major tasks undertaken during the contract period included:
ACS Acquisition: The overall objective of this effort was to perform research and acquire advanced prototype computing systems in support of the President's initiative in high performance computing (HPCC).
ACS Performance Evaluation: The main objective of this effort was the use of a new methodology to evaluate the performance of parallel computers by generalizing techniques used for uniprocessor performance benchmarks.
ResNets: The primary objectives in this task included integration of maturing technologies into DoD testbeds and advanced facilities and the design, implementation and support of DARPA's wide area research nodes.
NetTech: The primary objective of this task was the integration of technologies developed in the DARPA Communicator and Spoken Language Systems programs into advanced environments and testbeds.
Each of these tasks was successfully completed and over the lifetime of this effort a number of high performance and advanced computing systems were implemented, studied and transitioned.

14. SUBJECT TERMS

Computers, Architecture, Hardware, Software, Performance Evaluation, Metrics

15. NUMBER OF PAGES 28

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

TABLE OF CONTENTS

Final Report

OVERVIEW

The Information Systems Office (ISO) and the Information Technology Office (ITO) have a two fold mission: facilitate the transfer of advanced technology R&D to deployment and operations in the Defense Information Infrastructure (DII); be the primary vehicle for leading edge information technology services within the DARPA community. Over the lifetime of this contract vehicle, a number of high performance and advanced computing systems were implemented, studied and transitioned to mainstream use.

The Advanced Computing Systems effort made provision for a wide range of research activities that could be conducted under the overall contract. The major tasks undertaken during the contract period include:

- **ACS Acquisition**

  The overall objective of this effort was to perform research and acquire advanced prototype computing systems in support of the President's initiative in high performance computing (HPCC).

- **ACS Performance Evaluation**

  The main objective of this effort was the use of a new methodology to evaluate the performance of parallel computers by generalizing techniques used for uniprocessor performance benchmarks.

- **ResNets**

  The primary objectives in this task included integration of maturing technologies into DoD testbeds and advanced facilities and the design, implementation and support of DARPA's wide area research nodes.

- **NetTech**

  The primary objective of this task was the integration of technologies developed in the DARPA Communicator and Spoken Language Systems programs into advanced environments and testbeds.

1. ADVANCED COMPUTING SYSTEMS ACQUISITION

Summary

The overall objective of this effort was to perform research and acquire advanced prototype computing systems in support of the President's initiative in high performance computing (HPCC). Main tasks were to include development of HPC support systems and technologies, development of techniques to enhance utilization of HPCC systems, prototype software for the above two tasks and experimentally support advanced computation needs in National and Defense critical applications.

1.1. HPCC Enterprise/TIE Facility and Systems

The HPCC/Enterprise facility at DARPA was the primary system developed by the contract to support the goals of the HPCC. The HPC facility was an advanced space that integrated a number of new technologies in an effort to make High Performance Computing more accessible to the DoD community. The facility was transitioned to DARPA's JPO office in 1998 to facilitate concept and component integration into mainstream DoD facilities. Key components prototyped, developed or integrated by USC/ISI in this space were:

1.1.1. Dynamically Reconfigurable HPC Environment

The facility itself was based on a modular systems support concept. Each module or "octmod" included power, cooling, lighting and communications. The octmods could be moved within the facility and serve as anchors for walls etc. Over the lifetime of the contract, the facility was reconfigured a number of times to support emerging requirements. These reconfigurations were accomplished with a minimum of disruption and time.

1.1.2. Integrated A/V Environment

All communication feeds in the HPC included A/V. The video feeds could be directed via a mixing/switching system to a number of locations, including rear screen 3D capable projection systems, tape systems, flat panel displays and traditional monitors. Audio similarly would be routed via a mixing/switching system to multiple destinations. This system enabled workstations displays from diverse locations within the facility to be mapped to any number of display devices throughout the facility.

1.1.3. CAVE System

An immersive 3D environment (CAVE) was acquired and prototyped in the HPC. Through the use of timed polarized glasses and three or more projection systems a user can immerse himself in a virtual three dimensional environment. The CAVE system and the supporting computational systems was used to prototype a number of potential DoD applications from Submarine Engineering Spaces Design to 3D Battlespace Visualization.

1.1.4. Video Teleconferencing Systems

A number of video teleconferencing systems were proto-typed in the HPC environment including multicast based mbone tools, ISDN based point to point solutions and ATM based full frame rate solutions.

1.1.5. Voice Activated and GUI based Facilities Control

The facility included computer control of all key HPC systems (lights, cooling, video, audio and security). Both iconic and voice activated systems were developed and prototyped for the HPC environment. All workstations had the ability to control HPC systems (given appropriate access level), allowing one person to dynamically control the facility from any position within the facility, or remotely (if desired) via dial-up.

1.2. Experiment and Demonstration Support

A number of demonstrations and DoD experiments in High Performance Computing and Networking were supported in the HPC over the lifetime of the project. Numerous PI meetings, reviews etc. were hosted in the facility. Demonstrations of note included:

- Support of the Synthetic Theatre of War
- Wide Area Multi-Cave Collaboration during SuperComputing 95
- Telecommunications Bill Signing in 1996
- MAGIC Testbed Final Review/Demonstration

2. ADVANCED COMPUTING SYSTEMS PERFORMANCE EVALUATION

Summary

The main objective of this effort was the use of a new methodology to evaluate the performance of parallel computers by generalizing techniques used for uniprocessor performance benchmarks. The approach explored is the use of "Micro-benchmarks" to measure specific hardware and software characteristics of a machine (such as floating point multiply speed, memory reference cost, etc.) A suite of Micro-Benchmarks can be used to measure the specific features that affect performance of specific tasks (such as numerically intensive Fortran code).

2.1. Data Analysis of Existing Machines

A number of parallel machines were used for this analysis: the CM-5 from Thinking Machines Corporation (a message passing model), the KSR-1 from Kendall Square Research (a shared memory model), the Stanford DASH (shared memory model), Silicon Graphics

Challenge and Onyx systems, Sun SparcCenter 2000 and Convex SP/X. Performance was analyzed on individual nodes. The analysis revealed a undocumented subcache structure in the KSR-1. A full technical report was submitted in Data Item A003 3/1/92-9/30/92.

2.2. Micro Benchmarks Model

The results of these modeling experiments were included in a paper presented at SuperComputing 93. This paper was included in Data Item A003 10/1/92-9/30/93 so will not be included here. The model has proven an effective means of predicting parallel processor performance for specific tasks.

3. RESNETS

Summary

The primary objectives in this task included integration of maturing technologies into DoD testbeds and advanced facilities and the design, implementation and support of DARPA's wide area research nodes.

3.1. ATDnet

Experimentation and support of the DARPA ATDnet ring node was a main component of the work performed under this task. ATDnet provided the high bandwidth network access feed to the HPC facility, enabling many of the demonstrations and experiments that contributed to the success of the facility. During the lifetime of this task, ATDnet transitioned from a point to point PVC based ATM environment to an advanced signal based system capable of emulating local area networks. A number of experiments were carried out on ATDnet. See http://www.atd.net for further details. Work continues with ATDnet under the ResNets II contract.

3.2. Web Accessible SNMP Based Network Management Tools

A web accessible SNMP based network management toolset was prototyped under this task. A report on this prototype work was given at the Fall NGI PI Meeting. Selected slides from this meeting are included here:

Current Tools

SITES REACHABLE FROM ISIE

SITE	IP ADDRESS	HOSTNAME	TIME	ROUND TRIP TIME
DARPA	134.207.34.20	darpa.wabit.atd.net	Tue Aug 25 15:00.00 GMT 1998	49.838 ms
NASA	[illegible]	[illegible]	Tue Aug 25 [illegible]	[illegible]

CAIRN Sites:

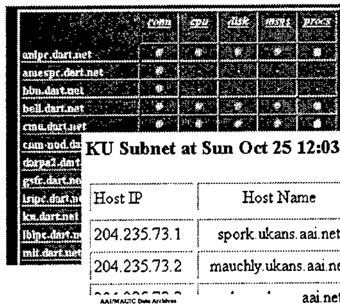
Fri Aug 21 12:15:01 GMT 1998 to Tue Aug 25 03:00:54 GMT 1998

SITE	IP ADDRESS	HOSTNAME	UPTIME PERCENTAGE
ANL	140.173.6.70	anlpc.cairn.net	91.69%
AMES	140.173.1.9	amessun.cairn.net	99.69%
			.31%

ATDNET links:

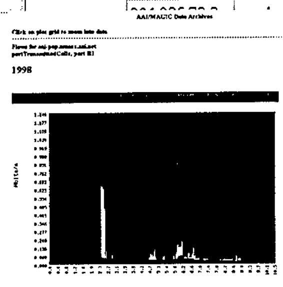
Results for Mon Aug 17 14:12:01 GMT 1998 to Tue Aug 25 04:30:05 GMT 1998

LINK	LINK TYPE (ATM capacity Mb/s)	AVG TRANSMIT RATE (Mb/s)	AVG RECEIVE RATE (Mb/s)	PEAK TRANSMIT RATE (Mb/s)	PEAK RECEIVE RATE (Mb/s)
DARPA-DISA	OC3 (149.7)	0 547	0 111	8 401	6 764
DARPA-NASA	OC3 (149.7)	0.164	0 141	7 307	7.396
DARPA-BASS	OC3 (149 7)	0.115	0 080	1 631	3.441
[illegible]	OC3	[illegible]	[illegible]	[illegible]	[illegible]



KU Subnet at Sun Oct 25 12:03:18 CST 1998

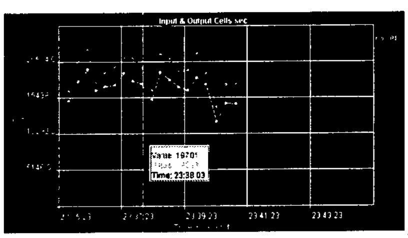
Host IP	Host Name	Round-Trip Time
204.235.73.1	spork.ukans.aai.net	4 ms
204.235.73.2	mauchly.ukans.aai.net	1 ms
[illegible]	[illegible] aai.net	2 ms
[illegible]	[illegible].aai.net	1 ms
[illegible]	[illegible] ai.net	1 ms

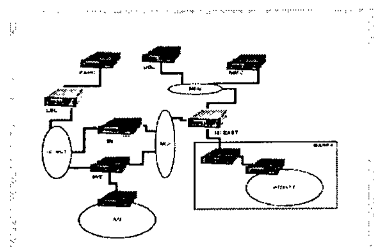


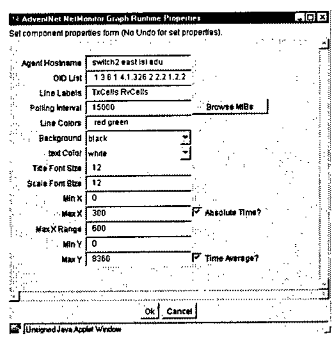
Candidate Tools

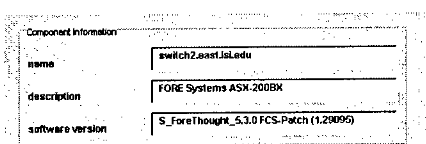
AdventNet

ISIE-MCI LINK TRAFFIC

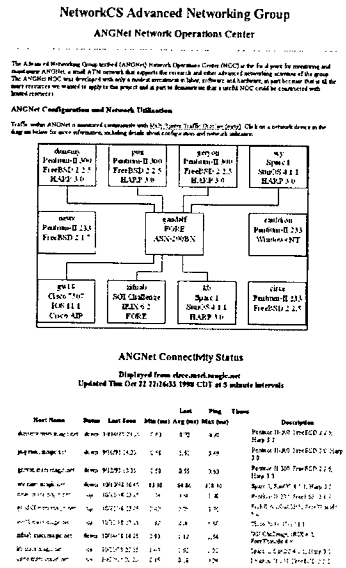




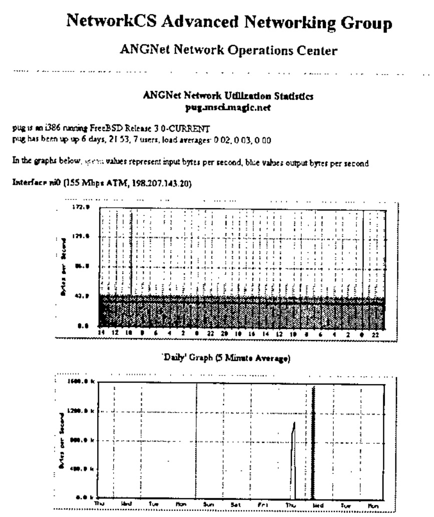




Candidate Tools (cont.)

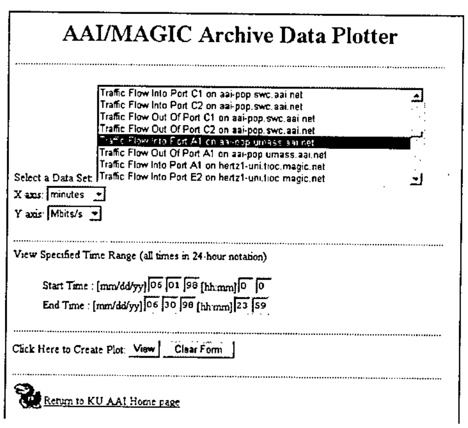


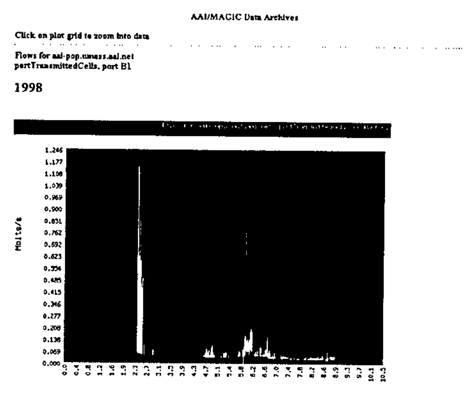
MRTG

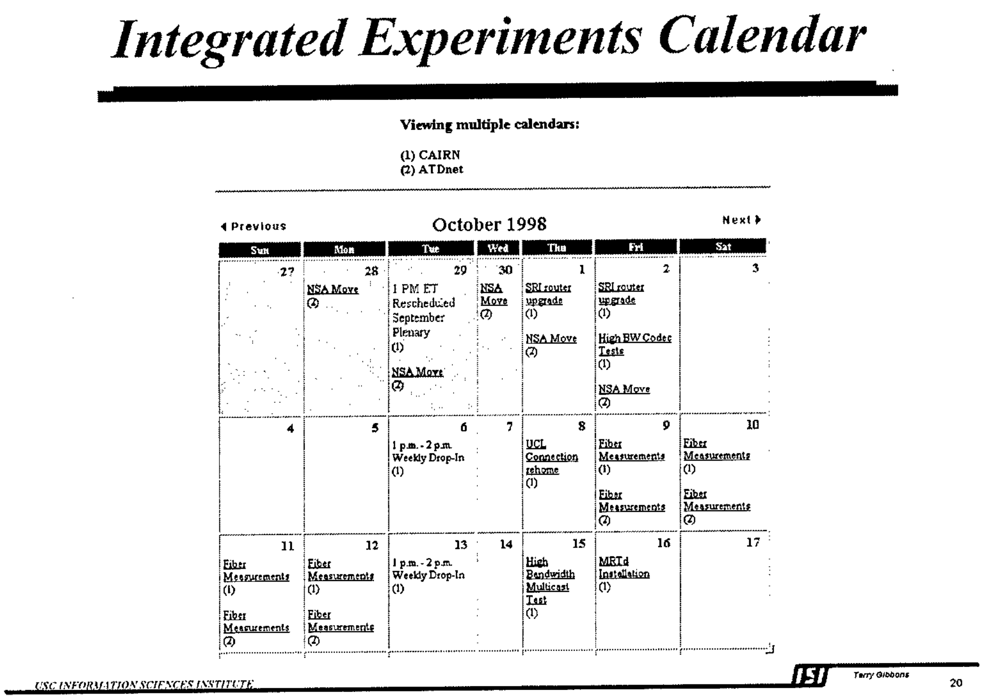


Candidate Tools (cont.)

KU Plotter / LBNL Netlogger







This toolset was transitioned to the CAIRN and SuperNet testbeds and is still being developed under the ResNets II contract. The current toolset can be viewed at http://www.cairn.net and http://www.ngi-supernet.org . There are plans to expand use of this toolset to the Mid-Atlantic Crossroads Gigapop in 2001.

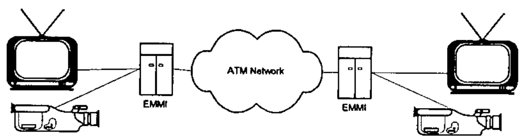
3.3. DNSSEC implementation in CAIRN

Recognizing the importance of secure reliable DNS information to the stability of the Internet, the preliminary DNS Security work from TIS was prototyped on FreeBSD routers and transitioned into DARPA's CAIRN research testbed. This work has lead to extensive involvement in the IETF's DNSSec working group. Work on DNSSec continues under DARPA's ACAIRN contract. Further details can be seen at http://www.cairn.net/DNSSEC/index.html.

3.4. Network Based Collaboration Tools

A variety of tests were performed on ATM (Lucent EMMI , Fore Systems AVA) and IP based (Mbone suite) collaboration tools to evaluate the strengths and weaknesses of emerging VTC technologies. A summary report to DARPA was given in January 1998. Selected slides are included here:

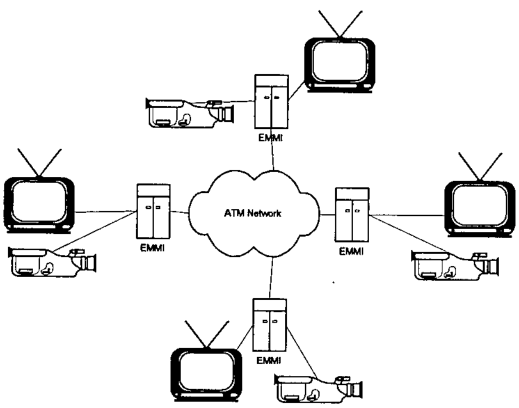
ATM Based Collaboration



Phase I

• Point to Point ATM based

• Available Today

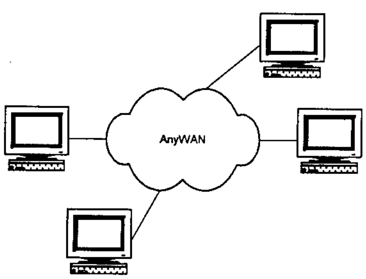


Phase II

• Four Way Multiplexed ATM based

• Available in 3 to 6 months

IP Multicast Based Collaboration



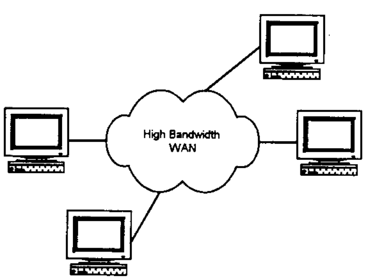
Phase I

• Moderate Quality

• Some Application Sharing

• Low Bandwidth Requirements

• Available Today



Phase II

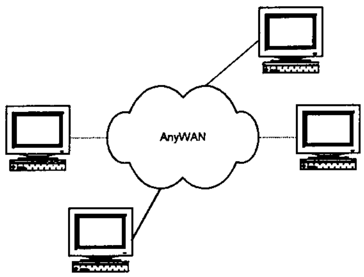
• High Quality Audio and Video

• High Bandwidth

• Better Application Sharing

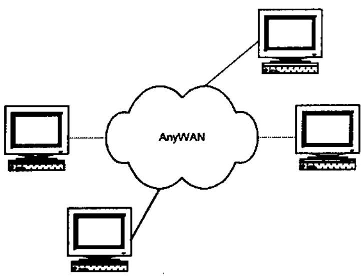
• Available 3 to 6 months

IP Multicast Based Collaboration (cont.)



Phase III

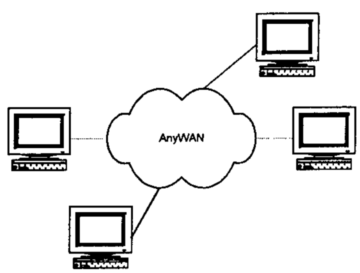
- Quality scales to available network resources (multilayered video, etc.)
- Good Application Integration with A/V
- Integration of Speech Recognition
- Available 6 to 12 months



Phase IV

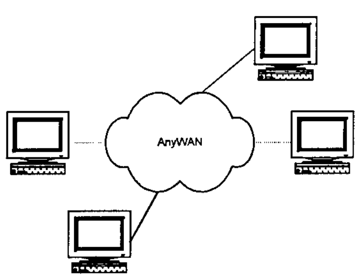
- All of Phase III
- Integrate Reliable Multicast Transport
- Integrate Secure Transport
- Integrate Stereo Audio
- Available 12 to 24 months

IP Multicast Based Collaboration (cont.)



Phase V

- All of Phase IV
- Integrate Multiple Videos in Shared Collaborative Space
- Available 24 to 36 months



Phase VI

- All of Phase V
- Develop "Realistic" Presence Audio/Video
- Available 36 to 48 months

ATM vs Multicast Collaboration

ATM

- Requires ATM WAN
- Not workstation based
- Not multipoint without multiplexer
- Only offers voice and video
- Security via secondary equipment (Fastlane, etc.)
- Very high quality

Multicast

- Runs over any IP WAN
- Workstation based
- Inherently multipoint
- List of collaboration applications is growing
- Security eventually at IP layer
- Quality rapidly approaching ATM

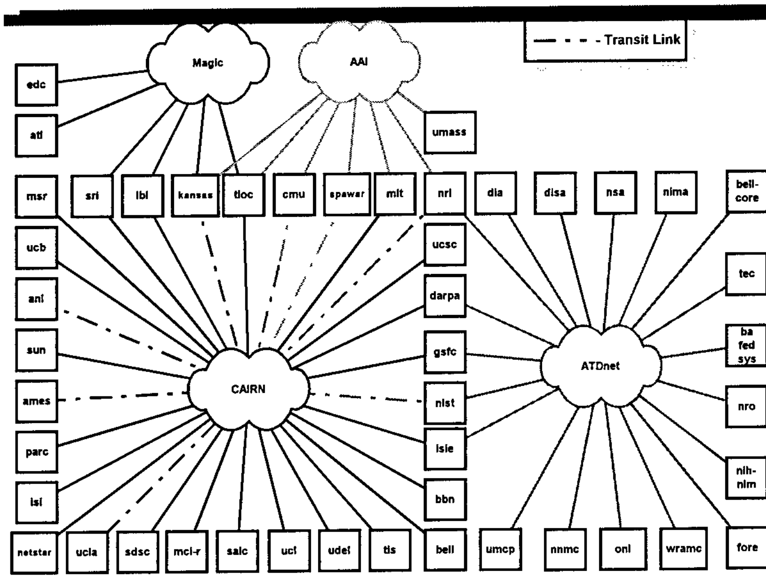
Multicast based collaboration tools were integrated into ATDnet monthly meetings, facilitating the transfer of this technology to DoD organizations via DISA and DIA ATDnet participants. Work continues under the ResNets II contract to integrate multicast based tools into SuperNet and specific SuperNet experiments such as Matisse.

3.5. AMAC Integration

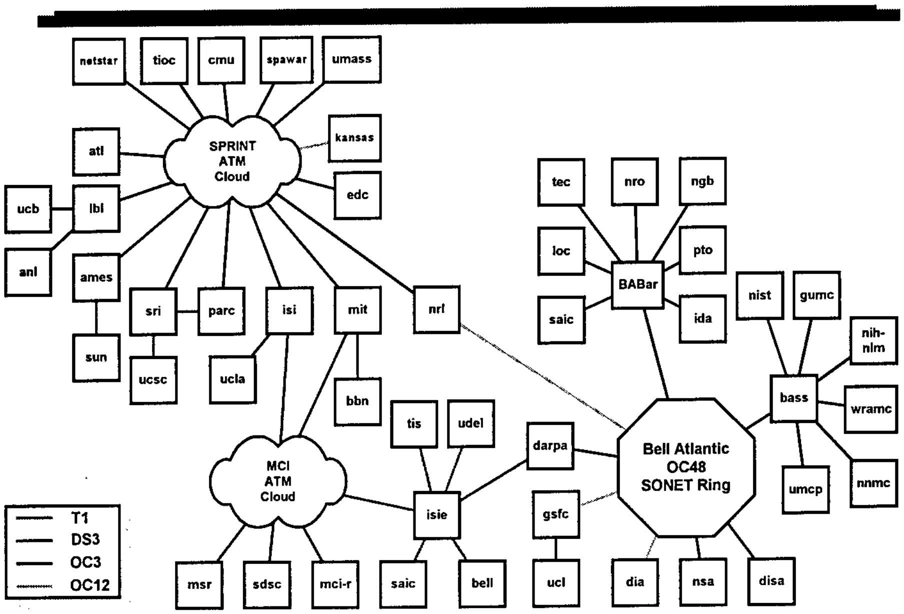
The four primary DARPA funded research networks were integrated into a single layer 3 network in this fore-runner to SuperNet. AMAC (ATDnet, Magic, AAI and CAIRN) used a combination of PNNI signaling, emulated LAN technology and FreeBSD based routers to allow any researcher on AMAC to speak to another AMAC site at Layer 3. This integration served as a road map for the integration work to be performed under ResNets and ResNets II in the creation of SuperNet from the NTON, HSCC, ATDnet, BossNet and OnRamp networks.

A report on this integration work was given at the Fall 98 NGI PI meeting. Selected slides are included here:

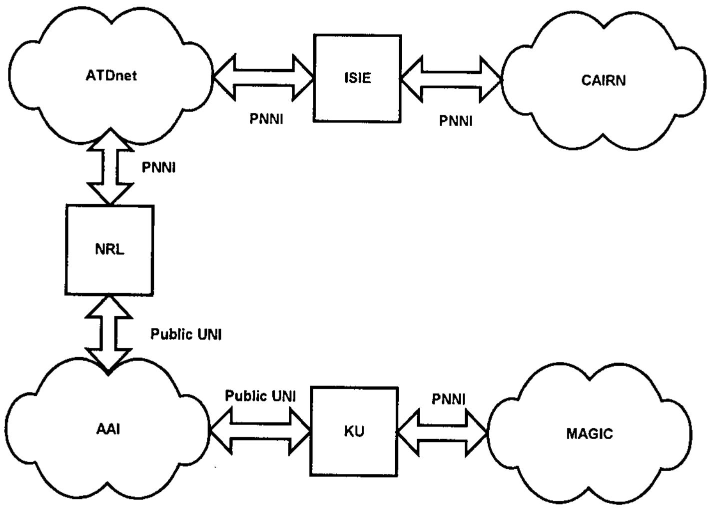
Network Sites



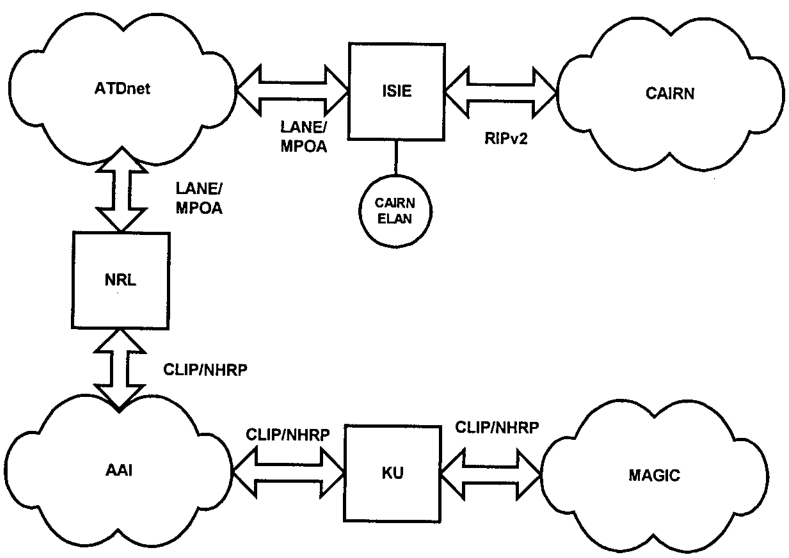
Physical Layer

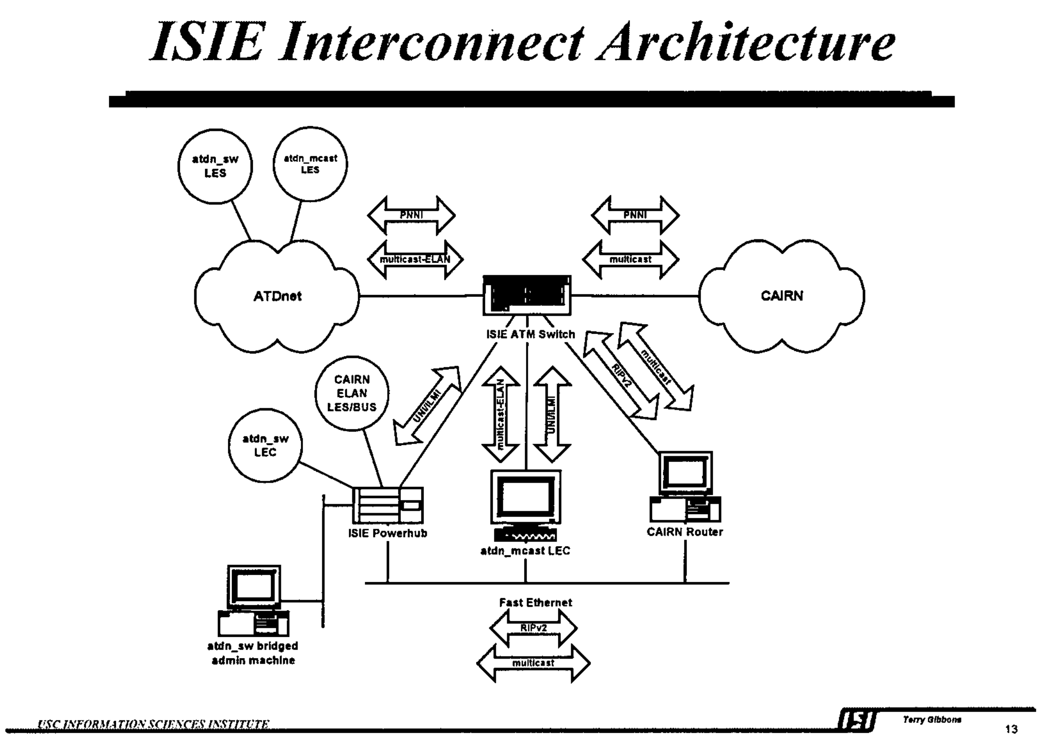


Layer 2 Integration



Layer 3 Integration





3.6. SuperNet Preliminary Integration Designs

The preliminary designs for the integration of the NTON, HSCC, BossNet, ATDnet and OnRAMP networks into DARPA's SuperNet were put together under this contract. The initial design of the DC SuperNet POP, along with the preliminary work on integration of MONET and BossNet were also done under this contract. SuperNet is active today as is the DC SPOP and LA SPOP. This work continues under the ResNets II contract. See http://www.ngi-supernet.org/ for further details.

4. NETTECH

Summary

The primary objective of this task was the integration of technologies developed in the DARPA Communicator and Spoken Language Systems programs into advanced environments and testbeds.

4.1. Wide Area Distribution of Interactive Drama Computer Learning System

ISI consulted with the developers of DARPA's Interactive Drama (ID) system. The ID system was developed under the Communicator program and was initially developed as a standalone system. ISI proposed a model suitable for ID content distribution via high bandwidth wide area networks such as DARPA's SuperNet. A number of media distribution systems and web based front end modules were investigated. A final architecture recommendation was prepared and submitted to the ID development team (Appendix A).

4.2. Web on a Disk

ISI began preliminary work on a system of dynamic web caching suitable for use by deployed DoD units (i.e. aboard ship etc.). A preliminary architecture was designed utilizing a land based master caching system that trickled updates to a shipboard counter-part via partial use of a low bandwidth connection. This work was carried forward under a follow-on proposal to DARPA. For further details, see the HICD project.

4.3. News on Demand

CMU's News on Demand news synthesis project developed under the Communicator program was deployed to a variety of locations including DARPA's HPC facility, USC/ISI, SPAWAR and other DoD agencies. This contract provided support for the deployment of the systems and underlying network connections between the CMU master servers and the deployment sites.

Appendix A

A Distributed Multimedia Streaming Architecture for Conversim

In this document we propose an overall model for a distributed Conversim architecture. Addressed are both the functions and the communication between the client and server sides of the applications. Figure 1 depicts the overall suggested architecture. The installed applications on the client workstation include the client portion of Conversim, a streaming multimedia player (TBD), and a web-browser. The server is composed of a streaming multimedia server, the server portion of Conversim, and an HTML server. Data and files will be stored on the server hard drive. Both client and server must have the TCP/IP networking protocols installed and configured.

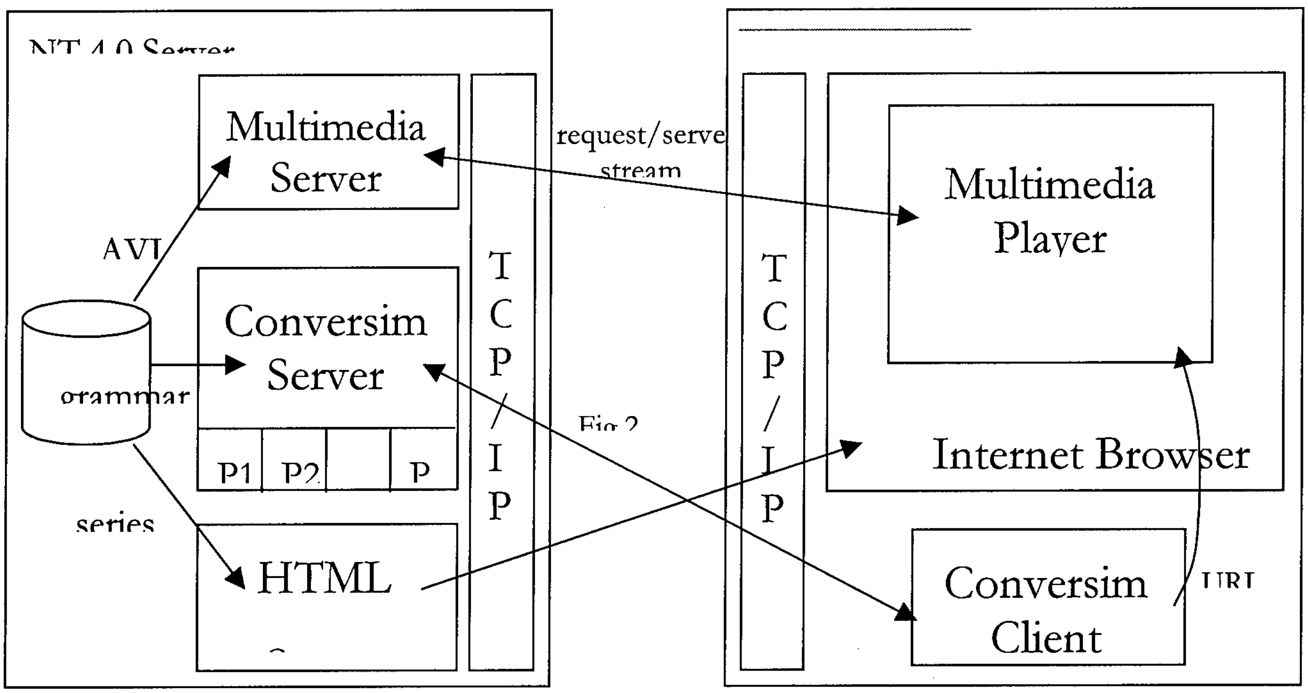


Fig 1

We will first describe the system start up procedures, then explain the client and server interaction, and lastly we discuss the Conversim server and the multimedia server in more detail.

Client Startup

The user starts the process off by pointing their web browser to the HTML server. The web site presents the user with a collection of series to choose from (i.e. Arabic language training series, breast cancer information series). Once the user selects a series by clicking on the html link, a Java applet is transferred from the HTML server to the client and started in the web browser. The Java applet checks that the Conversim client software is installed, a sound card and driver is installed and active, and a media player capable of playing a multimedia stream of a specific format (.avi). If either or both of the Conversim client application or the media player application (with media plug-in if required) aren't installed, the Java applet will initiate a download and install of the needed software from the server. If a sound card, driver and specific multimedia codec for the chosen series aren't installed, a message should be displayed prompting the user to install them and then re-run the series. After the Java applet finishes performing its checks and software installation, it will start the Conversim application on the client passing it the location (IP address) of the server, and the series which was chosen. The Java applet will then be closed having completed its functions, and all data processing and interaction with the user and server will be handled by the Conversim client application. The Conversim client application should perform the following steps on start up:

- initialize a TCP communication session with the server

- download the grammar, session control information, and initial questions for the series

- start the media player window pointing it to the initial introductory multimedia stream for the series

Client/Server Interaction

All Conversim client and server communication (Fig 2.) to each other is carried out through a static, pre-arranged TCP port number. The main Conversim server application continually listens on this port and when it receives a request from a new client, will spawn a new Conversim server process (Pn) to handle that client's data requests. The server process is responsible for parsing incoming client video requests (ie. video index), returning the URL pointer of the multimedia stream clip to the client, maintaining the client question/video playback history, and sending the next set of questions to prompt the user. The Conversim client passes the URL to the multimedia player already active in the browser, which then in turn initiates the streaming request, receives and plays the multimedia stream from the server. When the Conversim client shuts down, it will request that the server process and TCP communication port also close.

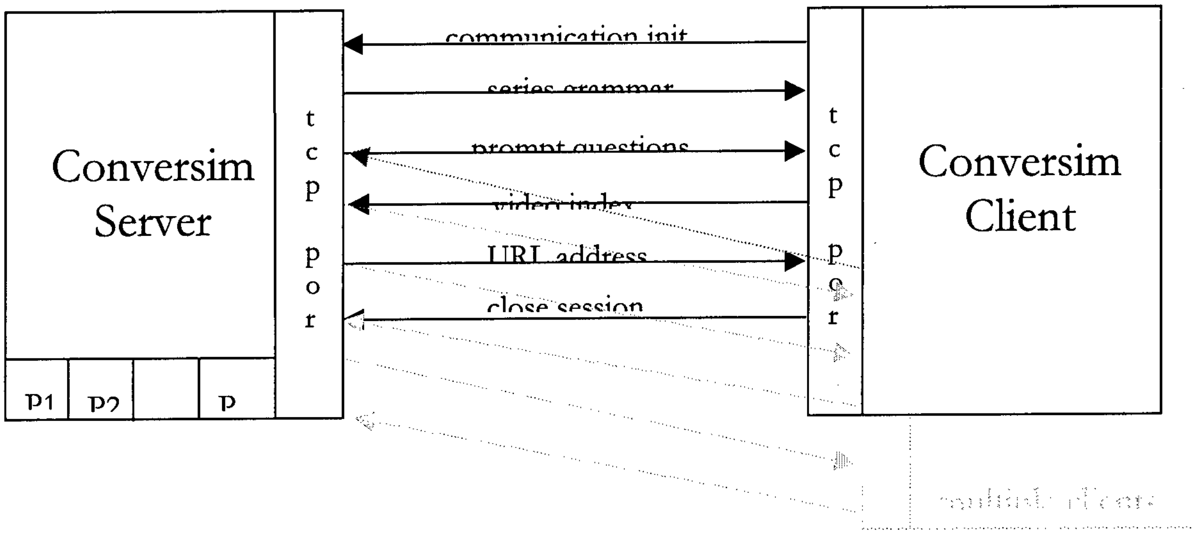


Fig 2

Conversim Server

The Conversim server continually listens (passive TCP socket) on its pre-arranged TCP port for client initialization request (active TCP socket) and following communications. When the server receives a request from a new client, it spawns off a new Conversim process to specifically server that client's requests. Thus, each server application process is able to maintain the client context and state information enabling it to generate the next logical set of user prompts. On initialization, the Conversim server process will transfer the series grammar (speech recognition signature files, video clip index), the initial set of user prompt questions, and the URL pointer for the introductory multimedia stream. Each process is also responsible for translating client video index requests into URLs.

Multimedia Player and Server Interaction

Using the information given from the Conversim client (URL to the multimedia clip), the multimedia player issues a stream request to the multimedia server. The multimedia client and server applications iterate through a setup routine to initiate the stream session. The server will then send the multimedia stream to the client.

As mentioned above, the multimedia player and server require some handshaking to take place in order to start sending and receiving each multimedia stream. This "back and forth" communication will be the main source of lag-time in starting up and playing a multimedia stream. In a later iteration of the Conversim client/server architecture, a Java applet running on the client could be continually listening for multimedia streams sent from the server. The Java applet would then forward the stream directly to the multimedia player application, and act as the player controller to stop/start playing streams.

*U.S. GOVERNMENT PRINTING OFFICE: 2001-610-055-10087

MISSION
OF
AFRL/INFORMATION DIRECTORATE (IF)

The advancement and application of Information Systems Science and Technology to meet Air Force unique requirements for Information Dominance and its transition to aerospace systems to meet Air Force needs.