Annette L. DeSchon

INTERMAIL, An Experimental Mail Forwarding System
**INTERMAIL, An Experimental Mail Forwarding System**

**Annette L. DeSchon**

**USC/Information Sciences Institute**
4676 Admiralty Way
Marina del Rey, CA 90292-6695

**Defense Advanced Research Projects Agency**
1400 Wilson Blvd.
Arlington, VA 22209

**MDA903 81 C 0335**

**September, 1985**

**18**

**Unclassified**

**This document is approved for public release; distribution is unlimited.**

**The INTERMAIL mail forwarding system connects the DARPA Internet mail system to a number of commercial mail systems, including MCI Mail, Telemail, and the IEEE CompMail system. It is implemented on a DEC TOPS-20 system using the SAIL programming language.**
Annette L. DeSchon

INTERMAIL, An Experimental Mail Forwarding System

This research is supported by the Defense Advanced Research Projects Agency under Contract No. MDA903-81-C-0035. Views and conclusions contained in this report are the author’s and should not be interpreted as representing the official opinion or policy of DARPA, the U.S. Government, or any person or agency connected with them.
1. Introduction

The evolution of large electronic mail systems testifies to the increasing importance of electronic mail as a means of communication and coordination throughout the scientific research community. These systems include the DARPA Internet mail system, the GTE Telemail system, the MCI Mail system, and the IEEE Compmail system. Until now these systems have operated autonomously, and no convenient mechanism has existed to allow users of one system to send electronic mail to users on another system. INTERMAIL is an experimental mail forwarding system which allows users to send electronic mail across such mail system boundaries. The use of INTERMAIL is transparent, in that users on each system are able to use their usual mail programs to prepare, to send, and to receive messages. No modifications to any of the mail programs on any of the systems are required.

This paper describes the implementation of the INTERMAIL system and discusses some of the problems associated with the interconnection of mail systems. The following sections provide

- background information on the various mail systems using INTERMAIL
- an overview of the current INTERMAIL implementation
- alternative approaches to addressing in a mail forwarding system
- a discussion of the INTERMAIL user interface
- a review of the open issues in the area of mail forwarding.

2. Background

The DARPA Internet mail system is made up of a number of hosts, connected by mail delivery programs that utilize the Simple Mail Transfer Protocol (SMTP). Each of these hosts contains various user mailboxes. For historical reasons, a DARPA Internet mail system mailbox is often referred to as an "ARPANET mailbox". It should be noted that these mailboxes may be located on any host which is a part of the Internet. An ARPANET mailing address consists of two parts, a mailbox name and a host name, separated by the "@" character. For example, a user named "John Smith" at the Internet host "HOST-XYZ" might have the ARPANET mailing address "Smith@HOST-XYZ". Consistent with this convention, the ARPANET mailing address for the INTERMAIL system is "INTERMAIL@USC.ISIF".

A user mailbox is usually a file that contains the messages that a user has received. On a TOPS-20 system such as USC.ISIF, a person's mailbox is a file named "MAIL.TXT" in his login directory. These messages are appended to the mailbox by the mail distribution component of the mail system. The user interface component of a mail system, or mail program, allows the user to compose and to send messages. It also enables the user to perform mailbox management functions such as reading new messages and deleting old messages. When a message is being composed, the user interface program prompts the user for the header information and the text of the message. Many mail programs contain a built-in text editor. Once the user has indicated that the message is complete by entering a "SEND" command, the mail program turns the message over to the mail distribution component of the system.

Figure 1 illustrates the path of a message sent by a user on one Internet host to a user on a remote Internet host. The user sends mail using the mail program. The message is picked up by the local
mailer program to be delivered to the mailer program on a remote host. The remote mailer program appends the message to the mailbox of the addressee. The user mail program can be used to read, to file, and to delete messages in a user mailbox.

Figure 1

A message consists of a header and text. The header section of a message consists of various fields. Some of these fields, such as the "To" field and the "Cc" field, are used by the mail system to determine the destination(s) of the message. Others, such as the "Date" field, the "Subject" field, and the "From" field are provided for the information of the recipient. In many mail systems the "Date" field and the "From" field are provided automatically by the mail system, as the message is sent.

The following is an example of a typical ARPANET message.

Date: 12 Feb 1985 11:17 PST
Subject: SNDMSG Problem
From: Anna-Lena Neches <ALNECHES@USC-ISID>
To: DESCHON@USC-ISIF.ARPA, Postel@USC-ISIF.ARPA
Cc: ALNeches@USC-ISID.ARPA

Hi Annette,

There was a problem with our HOSTS3.BIN-file which was installed last night. I have temporarily rolled back the latest version of the host tables and reinstalled the version from yesterday......

Regards,

Anna-Lena
The commercial systems (MCI Mail, Telemail, and IEEE CompMail) all have a similar model of operation. Each system has a user mail program that is used to send mail. There is a mailbox in which mail is collected, whether the user is logged on or not. When the user logs on, he is informed of any new unread messages. At this point, he may use the user mail program to read messages, to delete messages, to file messages for future reference, or to send a reply. The most significant difference between the commercial systems and the internet mail system is that the commercial systems generally assume that the user mailboxes are all on the same machine.

3. The INTERMAIL System

The INTERMAIL system currently runs on USC-ISIF, a DEC TOPS-20 system that is a host on the DARPA Internet. Figure 2 illustrates the relationship between INTERMAIL and the mail systems involved. INTERMAIL communicates with each mail system using the same mail programs that a human user would use. Since it uses standard user-interface programs on each system, it does not have to be involved in the inner workings of any mail system.
Figure 3
Figure 3 illustrates the path that a message takes between a user on a host in the Internet and a user on a commercial system. INTERMAIL consists of several program modules. TELEMAILER, COMPMAILER, MCIMAILER, and the Internet mailer are used to transfer messages between USC-ISIF and other systems. The IM program examines the forwarding information contained each message to determine the destination system. The scheduling of these modules is coordinated by a batch job that runs several times each day. INTERMAIL has a login account and a mailbox on each of the commercial mail systems. A user on a commercial system sends mail destined for a user on a host on the Internet to the "Intermail" mailbox on his local system. INTERMAIL periodically picks up this mail, determines the destination ARPANET address and turns it over to the local Internet mailer for delivery to the appropriate Internet host. In the other direction, a user on an Internet host sends mail destined for a user on a commercial mail system to "INTERMAIL@USC-ISIF". INTERMAIL periodically reads messages from its mailbox on USC-ISIF, determines the destination mail system, and sends the message using TELEMAILER, COMPMAILER, or MCIMAILER.

The programs that make up INTERMAIL include one program that parses the forwarding information on each message, and several mail interface programs that send and receive mail for a particular mail system. The IM program examines the forwarding information contained each message that enters the system, and reformats it for processing by the appropriate mail interface program. If any message contains bad forwarding information or an address that is not accepted by the destination mail system, the program that encounters the problem will copy the message into a "BAD-MSG" file. These messages which contain errors are processed later by a human postmaster, who hopefully can determine what went wrong, and contact the originator of the message. The five mail interface programs are MM, SEND-MSGS, TELEMAILER, MCIMAILER, and COMPMAILER. MM is a TOPS-20 mail program that is used to interface to the "INTERMAIL" ARPANET mailbox. The MM program is used to move mail out of the INTERMAIL mailbox into separate files. The SEND-MSGS program picks up the mail which is destined for an ARPANET mailbox and sends it using the MM program. The TELEMAILER program exchanges messages with the TeleNet mail system (Telemail) using a permanent leased line. The COMPMAILER program exchanges messages with the IEEE CompMail system using a permanent leased line. The MCIMAILER program exchanges messages with MCI Mail using a dial-out modem. We chose to implement TELEMAILER, COMPMAILER, and MCIMAILER in the SAIL programming language because of the built-in text handling functions.

Figure 4 illustrates the communication between the INTERMAIL program modules, through the use of files that are written and read from a common directory. "INTERMAIL Mailbox" is an ordinary MAIL.TXT file in the <INTERMAIL> directory on USC-ISIF. The other files shown contain messages (one per file) at various stages in the forwarding process. A batch job that logs into the <INTERMAIL> directory on USC-ISIF is used to run these programs in the following order:

1. MM reads messages from INTERMAIL's ARPANET mailbox into "IN1" files.

2. IM converts "IN1" files to "(--UNSENT-TELEMAIL--)", "(--UNSENT-MCIMAIL--)", "(--UNSENT-COMPMAIL--)", or "(--UNSENT-ARPAMAIL--)" files.

3. TELEMAILER sends messages from "(--UNSENT-TELEMAIL--)" files to Telemail and reads messages from the Telemail "INTERMAIL" account into "IN1" files.

4. MCIMAILER sends messages from "(--UNSENT-MCIMAIL--)" files to MCI Mail and reads messages from the MCI Mail "INTERMAIL" account into "IN1" files.
5. COMPMAILER sends messages from "(--UNSENT-COMPMAIL--)" files to CompMail and reads messages from the CompMail "INTERMAIL" account into "IN1" files.

6. IM converts "IN1" files to "(--UNSENT-TELEMAIL--)", "(--UNSENT-MCIMAIL--)", "(--UNSENT-COMPMAIL--)", or "(--UNSENT-ARPAMAIL--)" files.

7. SEND-MSGs sends messages from "(--UNSENT-ARPAMAIL--)" files, destined for ARPANET mailboxes, using the MM program.

The commercial mail systems that are connected by the INTERMAIL system were all designed to be used by human users, as opposed to other computer programs. Therefore, these "user friendly" interface programs tend to use English sentences, which may or may not be phrased exactly the same way from day to day. In addition, line noise is sometimes a problem when the connection to the remote system is via a dial-out modem. The most challenging aspect of these implementations has proved to be recovering from a situation in which a prompt or an error message has been misunderstood.

4. Addressing

The addressing information used by Intermail is included at the beginning of the text of the message. This approach was chosen largely in order to facilitate the use of the existing commercial systems. Each of the mail systems connected by INTERMAIL has its own addressing conventions:

- An ARPANET mailing address consists of a mailbox name and a host name, separated by the "@" character, for example, "Smith@USC-ISIF". Mailbox names are set up to be unique within each Internet host.

- An MCI Mail address consists of a unique seven digit user id, for example, "123-4567". The corresponding user name, for example "JSmith", may be used instead but user names are not guaranteed to be unique, so the sender may be asked to choose the correct user id from a list of possible matches.

- A CompMail address consists of a host number and a user id separated by the ":" character, for example, "64:CMP0123". In addition, a user name, for example "J.Smith", may be used so long as it is locally defined for the sender's login directory.

- A Telemail mailing address consists of a user name and one or more organization names separated by the "/" character, for example, "JSMITH/ACCTG/XYZCORP". An abbreviated form, for example "JSMITH", may be used so long as it is locally defined.

The INTERMAIL system has an account on each of these mail systems. Its ARPANET mailing address is "Intermail@USC-ISIF". On MCI Mail it is "107-8239" or "Intermail". On CompMail it is "64:CMP0617" or "Intermail". On Telemail it is "INTERMAIL/USCISI". Therefore on each of these systems, the messages that are to be forwarded to a user on an Internet host are addressed to "Intermail" in the header section of the message.

The INTERMAIL system requires forwarding information consisting of the name of the destination mail system and the name of the addressee, as he is known in the destination system. In setting up forwarding in the CompMail-to-ARPANET direction, we encountered a minor problem with the
character set interpretation. The "@" character, which is used in every ARPANET mailing address, is also used as the default line-delete character on the CompMail system and therefore must be reset before any messages containing an ARPANET mailing address can be composed.

Another approach to addressing is to embed the address of the recipient on the remote mail system in the "mailbox name" part of the ARPANET mailing address. Such an address might be "JSmith%MCMail@HOST" or "J.Smith%CompMail@HOST". We call this the relay approach. Certain modifications to the mail programs on the host running INTERMAIL, would make this approach possible in the ARPANET to commercial system direction. It would, however, be impossible in the other direction, since there is no way to modify the addressing scheme used in a commercial system.

One advantage to the relay approach is that it is possible to include users from different mail systems on the same ARPANET mailing list. Another advantage is that the "reply" command in the user interface program can be used to send a message back to the originator of the message being replied to, even when the message came from a different mail system. The disadvantage is that if this approach were adopted, the addressing conventions would no longer be the same in both directions. In addition, it would be impractical to make modifications to all of the TOPS-20 mail programs.

5. The User Interface to INTERMAIL

Our goal was to design a user interface that is flexible enough to accept the addressing information in a form familiar to the sender. In order to insure that the header information from the source system can be mapped into the header information of the destination system, INTERMAIL interprets only the header fields that are common to all four mail systems. These fields are the "To" field, the "Cc" field, the "Subject" field, and the "From" field. The header fields that can be delivered to the end recipient unchanged, are said to have end-to-end significance. In this system the fields that are in this category are the "Subject" field and the "From" field. The syntax for a list of addressees for the "To" field and the "Cc" field vary from system to system. In the Internet mail system, the addressees are entered in a list, separated by commas. Continuation lines begin with at least one space.

To: Smith@HOSTA, Jones@HOSTB, Lewis@HOSTC, Scott@HOSTD, Williams@HOST1, York@HOST2
Cc: DeSchon@USC-ISIF, Postel@USC-ISIF

In MCI Mail, the addressees are entered on separate lines. If a postal address or a telex address is used, the address requires more than one line.

To: AJones
To: John Smith (ADDRESS)
       1234 Maple Street
       Los Angeles, Ca. 90025
To: 123-4567
Cc: JWashington

In the CompMail system, addresses are entered in a list separated by either commas or spaces. There is no separate "Cc" field. In addition, the "&" character is used to signify that a line is being continued.
To: J.Smith A.Jones F.Lewis CC J.Washington & L.Ford S.Chan M.Williams

In the Telemail system addresses are entered in a list separated by commas.

To: JSMITH/NASA, AJONES/XYZCORP, FLEWIS/USCISI
Cc: JWASHINGTON/XYZCORP

It is therefore necessary that INTERMAIL accept addresses in the format that is used on the destination mail system. Some allowances for the practices employed on the source mail system are made when there is no conflict. For example, when a message destined for CompMail contains an explicit "Cc" field, or uses the Internet conventions for a continuation line, INTERMAIL reformats the addressees into a form that is acceptable to the CompMail user interface program.

Two formats can be used to specify the addressing information in the forwarding section: "Simple Forwarding" and "Source Route Forwarding". INTERMAIL originally supported only Source Route Forwarding. It is more general than Simple Forwarding, in that it provides the mechanism for multi-hop source routing. Simple Forwarding was later designed for easier use and it will handle most cases.

5.1. Simple Forwarding

In Simple Forwarding, the forwarding section contains the name of the destination mail system and the address(es) of the recipient(s) of the message, in the form

```
Forward: <destination-mail-system>
To: <recipient's-address>
Cc: <someone-else's-mailing-address>
(blank line)
```

The IM program module processes the "Forward" field to determine which mailer program (SEND-MSGS, MCIMAILER, COMPMAILER, or TELEMAILER) will send the message. It then uses the "Subject" field from the header, and the addressing information from the forwarding section to make up a new header for the unsent mail file. At this point it deletes the forwarding section from the text, replacing it with the "From" field of the original header, and copies the text of the message into the unsent mail file. The header section that the recipient sees when the message is ultimately delivered is the header section that is created when the message is sent by the appropriate mailer program. For example, the following ARPANET message
Date: 18-May-84 11:52:21 PST
To: Intermail@USC-ISIF
From: Annette DeSchon <DeSchon@USC-ISIF>
Subject: Sample Message

Forward: MCIMAIL
To: 123-4567

This the text of the message.

might be delivered to the MCI Mail recipient as

Date: Fri May 18, 1984 11:58 am PST
From: Intermail / MCI ID: 107-8239
To: * John Smith / MCI ID: 123-4567
Subject: Sample Message

From: Annette DeSchon <DeSchon@USC-ISIF>

This the text of the message.

5.2. Source Route Forwarding

Source Route Forwarding was developed to handle situations in which a message will travel multiple hops before reaching its destination. As in Simple Forwarding, the forwarding information is included at the beginning of the text of the message. Instead of a single destination mail system and address list, the forwarding section contains a series of Forward Source Route (FSR) lines, which specify the path that the message will take as it is being forwarded. The first FSR is the Internet mailing address of the mail forwarding program. The second FSR is used by the IM program to create the address section of the message for the next hop that the message will take. The forwarding section must contain at least one Return Source Route (RSR) line which ultimately may be used by the end recipient to create the forwarding section of a reply. Each FSR or RSR consists of a network or mail system name and the name of a user as he is known in the destination network. The network of the first FSR must match the network of the first RSR. This is a simple example of the source routing section in a message which is to be forwarded to MCI Mail.

FSR: [ARPANET]Intermail@USC-ISIF
FSR: [MCIMAIL]<recipient's-MCI-address>
RSR: [ARPANET]<sender's-Internet-mailing-address>

A multi-line format FSR or RSR may be used when a header field does not fit onto a single line. This format is used to enter a postal address (for MCI Mail), or when the list of recipients is too long for one line. The multi-line format also allows the creator of the message to specify optional fields such as the "Cc" field, and the "Subject" field. In a multi-line format FSR or RSR, the first line contains only the label and the network or mail system name. The lines following, up to the next FSR or RSR, will be copied into the appropriate unsent mail file (e.g., "[..UNSENT-MCIMAIL--]"). Therefore, the syntax of these lines will be similar to the syntax for the corresponding fields in the destination mail system. This is an example of a forwarding section containing a multi-line format FSR.
As a message is forwarded, the forwarding section is modified by the IM program to reflect each hop that the message has taken. Therefore, the forwarding section of the message as it is received by a user on the destination system is not the same as it was when it was originally sent. In addition, the header section (or "envelope") that the recipient sees is the header section that is created when the message is sent by the mail interface program, rather than the original header. For example, the following ARPANET message

```
Date: 18-May-84 11:52:21 PST
To: Intermail@USC-ISIF
From: DeSchon@USC-ISIF
Subject: Sample Message

FSR: [ARPANET]Intermail@USC-ISIF
FSR: [MCIMAIL]
RSR: [ARPANET]DeSchon@USC-ISIF

This the text of the message.
```

might be delivered to the MCI Mail recipient as

```
Date: Fri May 18, 1984 11:58 am PST
From: Intermail / MCI ID: 107-8239

to: * John Smith / MCI ID: 123-4567
Subject: Sample Message

FSR: [MCIMAIL]123-4567
RSR: [ARPANET]Intermail
RSR: [ARPANET]DeSchon@USC-ISIF

This the text of the message.
```

Figure 5 illustrates a hypothetical situation in which Source Route Forwarding is used to accomplish multi-hop forwarding. So far, in practice all of our forwarding is accomplished in a single hop. This example is included to demonstrate how Source Route Forwarding would work in a situation where Simple Forwarding could not be used. In this example, each mail system uses an addressing convention different from the addressing convention used by the other mail systems pictured. (In actuality these differences can take the form of disagreement about what a host name looks like, what a user name looks like, or the syntax used to express "user at host".)

In the example pictured in Figure 5, host names in Mail System A must begin with a letter. Host names in Mail System B must begin with the characters "Il", and in Mail System C, host names must begin with the characters "$S$". Since hosts XYZ and ABC are each connected to two mail systems which use different addressing conventions, the source route from Person1 at host XXX to Person2 at host ZZZ will contain different address names than the source route which specifies the opposite
direction. When Person1 at host XXX sends the message to the forwarder at host XYZ, the forwarding section of the message consists of

FSR: [A]INTERMAILXYZ
FSR: [B]INTERMAILABC
FSR: [C]Person2$$ZZZ
RSR: [A]Person1XXX

The forwarder at host XYZ then deletes the first entry from the Forward Source Route list, and inserts its return address at the beginning of the Return Source Route list. When the message is sent from
host XYZ to the forwarder at host ABC, the forwarding section consists of

FSR: [B]Intermail@$$ABC
FSR: [C]Person2@$$ZZZ
RSR: [B]Intermail@$$XYZ
RSR: [A]Person1@XXX

Next, the forwarder at host ABC performs the same operation on the source route, and sends the message to Person2 at host ZZZ. At this point the forwarding section consists of

FSR: [C]Person2@$$ZZZ
RSR: [C]Intermail@$$ABC
RSR: [B]Intermail@$$XYZ
RSR: [A]Person1@XXX

In order to reply to a message that has been forwarded using this means, the following procedure may be used to construct the source route. Starting with the message as it is received, exchange the FSRs for RSRs, and the RSRs for FSRs. Then place the FSRs in front of the RSRs, at the beginning of the source routing section. This yields the following source route

FSR: [C]Intermail@$$ABC
FSR: [B]Intermail@$$XYZ
FSR: [A]Person1@XXX
RSR: [C]Person2@$$ZZZ

At this point Person2 would send the message to "Intermail@$$ABC" and the same steps would take place in the other direction.

6. Areas for Future Study

Thus far, the INTERMAIL system has been developed and operated as an experiment. Should we decide to establish an actual service, the issues of billing and access control will have to be addressed. In addition, more can be done in the area of automating the postmaster's job. We would like to improve the program-generated error messages and possibly develop the capability to return messages containing addressing errors automatically.

The most commonly encountered problem is that users completely omit the forwarding section of the message. This error is especially common when the user is replying to a message. This is probably because the "reply" command completely takes care of all of the addressing requirements when the reply is to a sender within the same mail system. To improve the situation on the Internet side we would like to establish a user interface program that would prompt the user for the forwarding information, and insert it into the text in the correct format. Using this approach it will also be possible to automatically construct replies to messages that have been forwarded to the Internet mail system.
7. Acknowledgements

The INTERMAIL forwarding system is based on work done by Danny Cohen, Jon Postel, Joel Goldberg, and Lee Richardson. It was originally developed to communicate with users of the MOSIS VLSI system via the Telemail system.

8. References


