Network Working Group V. Cerf (ARPA)
Request for Comments: 771 J. Postel (ISI)
September 1980

MAIL TRANSITION PLAN

PREFACE

This is a draft memo and comments are requested.

INTRODUCTION

The principal aim of the mail service transition plan is to provide orderly support for computer mail service during the period of transition from the old ARPANET protocols to the new Internet protocols.

This plan covers only the transition from the current text computer mail in the ARPANET environment to text computer mail in an Internet environment. This plan does not address a second transition from text only mail to multimedia mail [10,11].

The goal is to provide equivalent or better service in the new Internet environment as was available in the ARPANET environment. During the interim period, when both protocol environments are in use, the goal is to minimize the impact on users and existing software, yet to permit the maximum mail exchange connectivity.

It is assumed that the user is familiar with both the ARPANET and Internet protocol environments [1-8]. The Internet protocols are designed to be used in a diverse collection of networks including the ARPANET, Packet Radio nets, Satellite nets, and local nets (e.g., Ethernets, Ring nets); while the ARPANET protocol are, of course, limited to the ARPANET.

The Internet protocol environment specifies TCP as the host-to-host transport protocol. The ARPANET protocol environment specifies NCP as the host-to-host transport protocol. Both TCP and NCP provide connection type process-to-process communication. The problem in the transition is to bridge these two different interprocess communication systems.

The objective of this plan is to specify the means by which the ARPANET computer mail services may be extended into the Internet system without disruptive changes for the users during the transition.

MODEL OF MAIL SERVICE

The model of the computer mail system taken here separates the mail composition and reading functions from the mail transport functions. In the following, the discussion will be hoplessly TOPS20-oriented. We appologize to users of other systems, but we feel it is better to discuss examples we know than to attempt to be abstract.

In the ARPANET mail service, composition and reading is done with user programs such as HERMES, MSG, MM, etc., while mail transmission is done by system programs such as MAILER (sending) and FTPSRV (receiving).

One element of the ARPANET mail service is the assumption that every source of mail can have a direct interprocess communication connection (via the NCPs) to every destination for mail. (There are some cases where special handling and forwarding of mail violates this assumption.)

Mailbox names are of the form "MAILBOX@HOST", and it is assumed that MAILBOX is a destination mailbox on that host.

The messages are actually transmitted according to the provisions of the File Transfer Protocol. Mail may be transimitted via either the control connection (MAIL command), or via a data connection (MLFL command). In either case, the argument specifies only the mailbox since the destination host is assumed to be the host receiving the transmission.

For example: messages sent from Postel at USC-ISIF to Cerf at USC-ISIA would arrive at ISIA with the argument "Cerf" but no indication of the host.

COMPOUND AND ALTERNATE NAMES

Mailboxes are of the form "mailbox@host" where mailbox is usually a name like "Postel" and host is a host identifier like "USC-ISIF". In some cases it will be useful to allow the host to be a compound name such as:

USC-ISIA
ARPANET-ISIA
SATNET-NDRE
PPSN-RSRE
HOST1.SRINET
LSCNET/MAILROOM

or even the name of an organization:

BBN
ARPA
MIT
SRI

The only restriction is that "@" not appear in either the "mailbox" or the "host" strings in the destination address.

To actually send the message the mailer program must convert the host string into the physical address to which to transmit the message. This name-to-address conversion is typically done by looking the name up in a table and finding the physical address in another field of that table entry. This means that all the compound and organization names (and any other alternate names or synonyms) must also be in the host table.

HIDDEN HOSTS

Sometimes the mailbox part of the destination address is a compound name and is used to mark a set of mailboxes which are not really on the host at all, but rather on another host which is connected to this host in a non-standard way.

It is important to users of computer mail that replies to messages may be easily composed with automatic assistance from the mail processing programs. To preserve this capability it is important that a host understand the mailbox part of every address in every message it sends if the host part of the address is itself.

That is, for every message, in every header field, in every address "m@h", host h must understand all values of m. Thus when a host prepares a message it should check all the addresses that appear in the header and for any address whose host part is this host the mailbox part should be verified.

THE TRANSITION PLAN

The basic ground rules for the transition are:

1. ARPANET mailbox names must continue to work correctly.

2. No changes should be required to mail editor software which parses message headers to compose replies and the like. Specifically, non-ARPANET mailbox designators must be accommodated without change to the parsing and checking mechanisms of mail processing programs.

3. Automatic forwarding of messages between NCP and TCP environments without user (or operator) intervention.

For the communication of messages between NCP and TCP hosts a mail relay service will be provided on a few hosts that implement both TCP and NCP. These will be "well known" in the same sense that sockets or ports for contacting Telnet or FTP servers are well known.

To make use of these relay servers changes will be made to the mailer programs. The mailer program will be responsible for determining if the destination address of the message is directly reachable via the interprocess communication system it has available (TCP or NCP or both), or if the mail must be relayed. If the mail must be relayed, the mailer must choose a relay server and transmit the message to it.

The basis for the decision the mailer must make is an expanded host name table. There will be a table which translates host names to physical addresses. The physical addresses in this table will be the 32-bit Internet addresses. (This makes sense for even NCP-only hosts, since after 1 January 1981 even they must use 96-bit leader format which requires 24-bit ARPANET physical addresses). Each entry in this table will also have some flag bits.

The flag bits will include information to indicate if the host in this entry is (1) a NCP host with "old tables", (2) a NCP host with "new tables", (3) a TCP host, or (4) some other kind of host. All TCP hosts are assumed to have "new tables". "Old tables" are those without these flag bits, while "new tables" do have these flags.

A separate table may be useful to list the addresses of the hosts with relay servers.

When a message is sent to a relay server, the control information (in the argument of the mail transfer command) must be augmented to include the destination host identifier.

The relay server may accept messages to be relayed without knowing that destination mailbox is actually reachable. This means that it may later discover that the destination mailbox does not exist (or some other condition prevents mail delivery). To be able to report the error to the originating user, the mailbox (mailbox@host) of the originating user must be included in the argument of the mail transfer command. If the argument does not contain the address of the originating user no error response is attempted. The error report, which is itself a message, does not carry an originator address in the command argument to avoid the possibility of a endless chain of error reports (however, an originator address does appear the header).

Since the originating host will act as if the mail was successfully delivered when it is accepted by the relay server, it deletes any back up copies of the message it was keeping in case of errors. For this reason, the relay server must include the complete message in any error report it sends to the originator. The relay server should parse the addresses in the argument before accepting a message. If it does not understand how deliver locally, or both relay and reply (if the originating address is present) to the message, it should not accept it.

There are enough differences in the transmission procedure that the relay server will use a distinct mail transfer protocol, separate from the file transfer protocol.

MAIL TRANSFER PROTOCOL

The mail trasfer protocol to be used by the relay server and all TCP hosts is documented in reference [9].

CONNECTIVITY

There are nine cases of mail exchange, the three by three matrix of (1) old-table NCP hosts, (2) new-table NCP hosts, (3) TCP hosts. There are also two transfer mechanisms: file transfer and mail transfer. The diagonal is easy, each type of host can exchange mail with other hosts of its type. The other cases are more subtle.

An old-table NCP host is assumed to have a table with 32-bit physical addresses, but no flag bits. It has NCP and file transfer. It does not have the separate mail transfer protocol.

An new-table NCP host is assumed to have a table with 32-bit physical addresses, and the flag bits. It has NCP and file transfer. It also has the new separate mail transfer.

An TCP host is assumed to have a table with 32-bit physical addresses, and the flag bits. It has the new separate mail transfer. It probably has a file transfer, but does not use it for mail.

1. Old-table NCP to Old-table NCP

   This transfer is direct and uses the old mechanisms -- NCP and file transfer.

2. New-table NCP to Old-table NCP

   This transfer is direct and uses the old mechanisms -- NCP and file transfer.

3. TCP to Old-table NCP

   This transfer must use a relay server. The first transfer (from the TCP host to the relay server) is via TCP and the mail transfer protocol. The second transfer (from the relay server to the old-table NCP) is via NCP and file transfer protocol.

4. Old-table NCP to New-table NCP

   This transfer is direct and uses the old mechanisms -- NCP and file transfer.

5. New-table NCP to New-table NCP

   This transfer is done with the NCP and the mail transfer protocol, that is, using the old interprocess communication system and the new mail transmission scheme.

6. TCP to New-table NCP

   This transfer must use a relay server. The first transfer (from the TCP host to the relay server) is via TCP and the mail transfer protocol. The second transfer (from the relay server to the new-table NCP) is via NCP and mail transfer protocol.

7. Old-table NCP to TCP

   This transfer must use a special relay server. The first transfer (from the old-table NCP to the relay server) is via NCP and the file transfer protocol. The second transfer (from the relay server to the TCP host) is via TCP and mail transfer protocol. This relay server must be special because the messages coming from the old-table NCP host will not have the destination host information in the command argument. This relay server must have a list of registered TCP user mailboxes and their associated TCP host identifiers. Since such a registry could be potentially large and frequently changing (and will grow as more TCP hosts come into existence) it will be necessary to limit the mailboxes on the registry.

8. New-table NCP to TCP

   This transfer must use a relay server. The first transfer (from the new-table NCP to the relay server) is via NCP and the mail transfer protocol. The second transfer (from the relay server to the TCP host) is via TCP and mail transfer protocol.

9. TCP to TCP

   This transfer is direct and uses the new mechanisms -- TCP and the mail transfer protocol.

In general, whenever possible the new procedures are to be used.

MULTIPLE RECIPIENTS

A substantial portion of the mail sent is addressed to multiple recipients. It would substantially cut the transmission and processing costs if such multiple recipient mail were transfered using the multiple recipient technique available for use in both the old file transfer protocol [12] and new mail transfer protocol [9].

The relay servers will attempt to use a multiple recipient commands whenever applicable on transmitting messages, and will accept such commands when revceiving messages.

COMPOSITION AND READING PROGRAMS

The impact on the mail composition and reading programs is minimal. If these programs use a table to recognize, complete, or verify host identifiers, then they must be modified to use the new table.

To assist the user in replying to messages it will be important that all addresses in the header fields (TO:, CC:, etc.) be complete with both the mailbox and host parts. In some cases this has not previously been necessary since the addresses without host parts could be assumed to be local to the originating host, and the sending host was recorded by the receiving host. When the messages were sent directly the originating host was the sending host, but when messages are relayed the originating host will not be the host sending the mail to the destination host.

REFERENCES

[1] Cerf, V., "The Catenet Model for Internetworking," IEN 48, DARPA/IPTO, July 1978.

[2] Postel, J., "Internet Protocol," RFC 760, USC/Information Sciences Institute, NTIS ADA079730, January 1980.

[3] Postel, J., "Transmission Control Protocol," RFC 761, USC/Information Sciences Institute, NTIS ADA082609, January 1980.

[4] Postel, J., "Telnet Protocol Specification," RFC 764, USC/Information Sciences Institute, June 1980.

[4] Postel, J., "File Transfer Protocol," RFC 765, USC/Information Sciences Institute, June 1980.

[5] Postel, J., "Assigned Numbers," USC/Information Sciences Institute, RFC 762, January 1980.

[6] Postel, J., "Internet Protocol Handbook," USC/Information Sciences Institute, RFC 766, July 1980.

[7] Feinler, E. and, J. Postel, "ARPANET Protocol Handbook," NIC 7104, Network Information Center, SRI International, January 1978.

[8] Crocker, D., J. Vittal, K. Pogran, and, D. Henderson, "Standards for the Format of ARPA Network Text Messages," RFC 733 7104, Network Information Center, SRI International, November 1977.

[9] Sluizer, S. and, J. Postel, "Mail Transfer Protocol," USC/Information Sciences Institute, RFC rrr, September 1980.

[10] Postel, J., "Internet Message Protocol," USC/Information Sciences Institute, RFC 759, August 1980.

[11] Postel, J., "A Structured Format for Transmission of Multi-Media Documents," USC/Information Sciences Institute, RFC 767, August 1980.

[12] Harrenstien, K., "FTP Extension: XRSQ/XRCP," SRI International, RFC 743, December 1977.