

NETWORK WORKING GROUP
RFC #475
NIC #14919

ABHAY K. BHUSHAN
MIT-DMCG
MARCH 6, 1973

14919
Postal

FTP AND NETWORK MAIL SYSTEM

Rev'd 15 MAR 73

This paper describes my understanding of the results of the Network Mail System meeting SRI-ARC on February 23, 1973, and the implications for FTP (File Transfer Protocol). There was general agreement at the meeting that network mail function should be within FTP.

FTP currently provides two commands for handling mail. The MAIL command allows a user to send mail via the TELNET connection (the server collects the mail and determines its end by searching for the character sequence "CRLF.CRLF"). The MLFL (mail file) command allows a user to send mail via the data connection (requires a user-FTP to handle the command but transfer is more efficient as server need not search for a special character sequence). These commands are being used to provide network mailing facilities. Local mail and SNDMSG programs have been modified at many sites to include network mailing (e.g, USER\$HOST at BBN-TENEX and MAIL host user at MIT-DMCG).

The network mail system should provide a facility whereby users can conveniently send messages to other network users who have "mailboxes" at one or more hosts. It is not required that the messages or mail be delivered in real-time. The network mail system is not an interactive inter-console communication facility, but it may be possible for some sites to deliver "urgent" mail to users in real-time (e.g., print mail at user console if user is currently logged-in). The mail system also does not provide a general inter-process communication facility, though it may be possible to deliver messages to programs which have mailbox addresses. Inter-process and inter-entity communication facilities are very desirable but are beyond the scope of the network mail system.

The concepts of "mailbox" and "mailbox addresses" are central to this discussion of network mail system. A mailbox is a place where the mail is stored before a user picks it up. It may be a file in the user's directory or it may be a bin for hard-copy. The mailbox address is the address required by the sender in order to send the mail to its destination mailbox. For users who have an "on-line" network mailbox, the mailbox address contains the Host address and the user's mailbox identification at that Host. The mailbox identification is that which is required by an FTP-server in order that it may put the mail in the desired mailbox. The terms mailbox address and address will be used to refer to the on-line network mailbox address.

NETWORK MAIL SYSTEM FUNCTIONS

The network mail system should provide the following six functions:

1. CREATING: This refers to the manner in which the user creates or composes his message. The FTP servers do not explicitly provide any message editing capability (server's editing conventions may be applicable in the case of MAIL command). Editing conventions such as those for character delete and line cancel vary widely over the network. The user is most familiar with his local Host conventions and these should be used for network mail editing. The user also has access to local editing systems which can be used for composing message files. The message file may then be transmitted via the MAIL or NLFL commands (NLFL being preferable). The present FTP approach of assuming the creation of messages to be sender's responsibility seems adequate. TIP users if they desire editing facilities should use intermediate Hosts for creating and sending messages.

2. LOCATING: How sender determines receiver's address. FTP assumes that the sender knows the receiver's correct address. There is no published or "on-line" list of mailbox addresses. There is, however, a list of network participants maintained (on-line) and published by the Network Information Center (NIC) at SRI-ARC. The network users have been assigned a unique "NIC Ident" and Host site by the NIC. It was therefore specified in FTP that FTP-servers maintain a table that maps NIC Idents to mail-box identifications. The NIC will maintain on-line and publish the local mailbox address information for network participants. It would be possible for users to look up a published list, or query the NIC on-line to locate destination addresses. The NIC will also provide an on-line facility (similar to FTP) that can be used by programs for retrieving the address information. This latter approach of the NIC's maintaining addresses has several advantages. The user can obtain a number of addresses for a group, and use these to transmit mail. The FTP servers need not maintain NIC Ident Tables, and the NIC can provide a good facility for locating addresses from last names, NIC idents, or even sketchy information. It may still be desirable that FTP servers accept NIC idents, last names, and other standard forms as mailbox identifiers.

3. SENDING: How message is sent to the destination mailbox. The messages may be sent directly to the destination mailbox (via TELNET or Data connections) or via an intermediate Host such as the NIC. FTP does not explicitly provide for mail forwarding by intermediate Hosts but FTP servers may be able to recognize addresses as not being local, and forward mail. In the event mail is to be forwarded, a desirable facility is to have the intermediate site return an acknowledgment (by request) upon delivery of mail or if delivery fails within a specified time. The current FTP specifications recommend that FTP-servers accept multiple addresses but do not require this.

4. STORING: Where mail is stored before reading and if information is available for later reference or retrieval. The FTP does not require that sender store mail or keep duplicate copies. It is the receiver's responsibility to store the information for reading, reference, or retrieval. The receiver need not store the mail as a data file but can directly print it out on a user console or line printer. FTP does not specify the procedures for storage handling by intermediate sites. If intermediate site is used for forwarding the mail then it should be the responsibility of that site to store mail until it is delivered to its final destination. If the mail is undeliverable then the intermediate site should return the undelivered information to the sender. A similar situation arises when sending of mail is deferred by the sending site (destination host may be down). The sending site then acts as an intermediate forwarder insofar as the user is concerned.

5. RECORDING: Should the mail be catalogued and recorded for later reference and retrieval. FTP currently does not provide an explicit mechanism for the receiver to record mail. If an intermediate site (the NIC) is used for mail distribution then a function of such a site could be to record mail, if so requested. NIC is ideal for recording mail, but other sites may also wish to record mail. If the mail is recorded, then it is not necessary to send the entire contents of the mail. Instead only a citation for the document can be sent and the receiver can retrieve the mail only if he wants to. This is particularly useful for large documents such as NWG/RFC which are distributed to a group. The citation may contain author, title, retrieval pathname, and perhaps an abstract.

6. READING: How the mail is finally presented to and read by the user. FTP currently assumes that mail reading is entirely the receiving site's function. However, there are ways in which the sender can aid the receiver in providing improved mail reading facilities. For example, the receiving system, if it knows a message to be urgent can deliver it immediately at a user console. Long messages may be put in separate files with notification in user's regular mail. Alternately, mail could be a citation that the reading program can retrieve upon user request. Selective handling of different classes of mail is important for an improved network mail system.

MODELS FOR MAIL SYSTEM USE

The user of a mail system can use intermediate site for locating addresses, recording and/or distributing mail, and for creating and reading mail. We therefore have the following models for mail system use:

1. The user connects directly to the destination FTP server and sends mail using the MAIL command. Local editing functions are limited to character delete and line cancel (assuming user is in line-at-time mode) and server conventions may also apply. The user only needs a user-TELNET program at his site but needs to know the destination address. This model is specially applicable to TIP and other mini-Host users who do not have a user-FTP or user-Mail programs.

2. The user composes the mail using a local editor (or mail system) and then requests his user-FTP or mail program to send the mail directly to the destination via the FTP MAIL or MLFL commands. The user needs to know the destination address. The mail can be deferred by the sending program if the destination Host is down. TIP users can use this model by using the facilities of a "home-base" Host.

3. The user uses an intermediate site such as NIC (other sites may provide forwarding services too) for mail distribution. The user need not know the destination addresses but can use NIC idents for individuals and groups of individuals. The mail can be recorded on request and its sending can be deferred (the destination Host may be down, or it may be more economical to defer mail). The message to be mailed may be created at the local site using local editing facilities, or it may be created directly at the intermediate site.

4. The user may send a citation of the mail instead of the complete mail item. The citation refers to an existing document which can be retrieved on-line (such as the NIC number of a NIC journal communication).

MAILING TO TIP USERS

The TIP does not currently provide an FTP server or mailbox facilities. While it is possible to send mail to TIP terminals (such as line printers) it seems undesirable to do so because of the possibility of losing mail, the lack of privacy, and the fact that user may be several (or several hundred) miles away from the location of the TIP. The TIP users normally have a "home-base" computer where they do their computing work most of the time. The TIP user problem is best solved by requiring that TIP users rent mailboxes at their "home-base" Host. Such a Host can provide good mail reading and query facilities. A TIP user can request his "home" Host to send him notification of mail on a TIP terminal. If RDML command (NWG/RFC 458) is accepted in FTP, TIP users could use such a command. More important, if the user has a number of mailboxes on different Hosts, the RDML (or RDMF) command can be used to read his mail at all the sites where he has mailboxes.

ACCESS CONTROL IN MAIL SYSTEM

It has been suggested that FTP specification should require that mail function (for receiving mail) should be "free", i.e., FTP servers should not require the user to "login" (send the USER, PASS, and ACCT commands). In the absence of the access control commands the FTP server should charge the cost of receiving mail to an overhead or browsing account. It should be noted that this "free" mail function using default "USER" account may not allow non mail-related commands without reinitializing. This requirement will improve communication among the network users.

Some systems, such as Multics, have mechanisms for access control in the receipt of mail. That is a user can specify who is eligible to send him mail (normally users give the access " *.*.*", i.e., any one can send mail). The access control commands would be required to gain privileged access. The USER command does not seem the best way to identify the sender of mail. Consider users logged in as GUEST, ICCO, NETWORK, MIT-DMCG, and NETWORK-USER. A separate FROM command seems desirable. Such a command can be used to identify the sender as well as to send acknowledgments and replies. The receiving site can tag the mail as: FROM AKB at MIT-DMCG, logged in as GUEST. The receiver can then send reply to the mailbox address AKB at Host 70 (SENDMSG AKB@DMCG or MAIL DMCG AKB).

NETWORK INFORMATION CENTER FUNCTIONS

The NIC is a very special facility for handling mail. It provides facilities for recording and distributing mail to individuals and groups of individuals, and for locating users' addresses. The NIC will also undertake to provide distribution of unrecorded mail. Currently the NIC requires that users log into the NIC and use NLS to create and distribute mail. Using NLS for creating mail has been a frustrating experience for many who are used to different editing systems. Recently there has been a problem that NIC is overloaded at most times of the day and even if one can get a "network terminal" and log in, the interaction is quite slow. As NIC (or NLS) is designed for character-at-a-time interaction with remote echo, the use is inefficient. Using NIC is particularly unbearable when the user falls behind in his echo by as much as an entire line.

An alternative to direct use of NIC is to use the NIC via FTP and programs at the user's site. The user can create journal documents using his own local editing system and then transfer it to NIC via FTP. The user may have to specify such information as author, title, where the acknowledgment should be sent, and journal number if the item is to be recorded. It should also be possible for users to send sequential files to NIC and have them restructured into NLS form without having to do an "input sequential" (a suggestion is to "NLS" the file if its name is suffixed with a .NLS). Alternately

it should be possible for user's to retrieve journal documents and other sequential files without having to do a previous "output sequential".

The NIC currently delivers mail via hardcopy and/or on-line. On-line currently means that user must log into NIC to see if he has a message and read it by "print branch". The messages are not seen by the destination users for several days and many users get their hard copy before they have had a chance to examine their on-line NIC mail. If the NIC were to deliver mail via FTP to network users, then the mail turn-around time will be greatly speeded and the users will not have to log into the NIC. Large documents need not be mailed to the user in their entirety but only a citation need be sent. The NIC will have to collect the information on the mailbox addresses of Network participants for delivering mail, especially since it appears that many FTP servers are not "respecting" NIC idents. It is recognized that a user may have more than one valid mailbox address, but the NIC needs to have only one (the most used) of these addresses.

The NIC identification subsystem (currently accessible via NLS only) contains information on users (such as affiliation, US Mail address, telephone numbers, etc.) and groups (members, etc.). The on-line mailbox address information can be added here. The NIC will undertake to provide a facility whereby the identification subsystem can be queried by programs, allowing mailing programs to retrieve the addresses automatically. This facility will be separate from FTP.

FTP MODIFICATIONS

The FTP currently does not provide explicit facilities for recording mail, communicating sender's address, sending program readable citations, specifying author and title for documents, requesting acknowledgments, and indicating message type (urgent, ordinary, and long). To overcome these deficiencies, we can take any of the following approaches:

1. Kludge the desired features in the pathname syntax of the MAIL and MLFL commands, justifying the kludge on the grounds that most of the functions are to be used only by the NIC.
2. Add new commands for the desired functions and alter the MAIL and MLFL commands somewhat to recognize the existence of the new commands.
3. Define a new mail command which incorporates the missing functions (in the process defining new commands for the desired functions). The MAIL and MLFL commands can be used in their

present form but may be gradually phased out.

The first approach seems undesirable to me as many of the missing functions can be used by other sites as well. In addition it will be easier to write programs to deal with commands rather than a complex syntax. The second and the third approaches are not very different from each other. The third approach seems preferable as it will allow existing mail programs to function in their present form. Using the third approach consider the following new FTP commands:

1. MLTO (mail to): The argument is one or more mailbox identifiers separated by "," (commas). It is suggested that if there is no argument, the mail should be sent to some responsible user or printed on a printer. This command starts the sequence of optional FTP mail related commands described below. The sequence ends with the TEXT, FILE, or CITA (citation) commands.
2. FROM: The argument is the address of the sender or senders. It is in a standard form that can be interpreted by programs as well as human users. The information is to be used for identifying the sender(s), for sending replies, and for sending acknowledgments if the receiver is an intermediate forwarding site.
3. MTYP (mail type): This identifies the type of mail as U (urgent), O (ordinary), and L (long). The receiving system can take the appropriate actions from this knowledge. The default assumption is ordinary mail.
4. RECO (record the mail): The argument if present is the identifying information for recording (such as NIC Journal number). If no argument is present the server will assign the recording information and send an appropriate reply (real-time or deferred).
5. AUTH (author): Identifies the author of the document in a form acceptable to the server (NIC ident may be required by NIC).
6. TITL (title): Identifies the title of the document. The argument is an ASCII string ending with the sequence "CRLF.CRLF".
7. ACKN (acknowledge): Relevant for intermediate forwarding sites. Asks the server to send acknowledgment on delivery or if delivery fails within a specified time.

8. TEXT: No arguments. Starts the transfer of mail over TELNET connection in an identical manner as MAIL.
 - . FILE: No arguments. Starts transfer of mail over the data connection in an identical manner as MLFL.
10. CITA (citation): Argument is the pathname of a retrievable file.

We also need to define new reply codes for handling mail. Some sites have expressed the need for replies such as "send only X bytes of mail". Other replies could specifically request additional commands such as USER/PASS/ACCT for privileged mailing, FROM/ACKN for mail forwarding, and AUTH/TITL for recorded mail. Another suggestion that may be given consideration is allowing TYPE/BYTE other than A/8 for FILE command. Mailing large files between like machines such as PDP-10s is more efficient in I/36. The RDML and RDMF commands proposed by Bressler and Thomas (NWG/RFC 458) also merit consideration as they would aid the handling of mail for users who have mailboxes at different Hosts.