A White Paper On

Internet E-Mail Services

Version 1.3

Prepared for

The Electronic Mail Association PRMD Operators' Committee

Prepared by

Mohsen Banan

Tele-Voice

17005 SE 31st Place

Bellevue, WA 98008

September 16, 1993

Contents

Executive Summary

Electronic Messaging is a technology which has seen a rapid increase in usage, and hence importance, in almost every company. A recent poll showed that over 95% of the Fortune 100 has installed some kind of electronic mail. Broadly speaking there are two types of end-user systems installed within corporations today. First there are the vendor- supplied systems. Some examples of these are IBMs PROFS, Digital Equipment Corporation's ALL-IN-1, and Lotus's cc:Mail. The second type is usually found on Unix systems and engineering workstations. It is often either free software or else it comes bundled in with the operating system.

Companies have spent a lot of time and money integrating their internal e-mail systems. However, the integration has often taken a separate path for the two types of environments. For the vendor supplied systems, the path was to either integrate them via a proprietary gateway provider or to use protocols defined by the CCITT X.400 Series of Recommendations. This internal integration was then often followed by implementing an external connection to a public service provider (often via X.400 protocols) to enable communications with trading partners and other companies.

The Unix and engineering environments, on the other hand, usually had network (TCP/IP) and e-mail (Simple Mail Transfer Protocol – SMTP) interconnection protocols bundled in with their operating systems. This allowed them a level of interconnection just by installing a physical LAN connection. Externally, this environment interconnected via a network known as "The Internet." This environment has an estimated 20 million e-mail users and has defined protocols which allow for file transfer, virtual terminal and dial-up connections, bulletin boards and other services, in addition to electronic mail.

Although growth of the Internet over the past ten years has been very rapid and steady from the perspective of messaging in the business world, the presence of the Internet in the messaging scene has been felt rather suddenly and abruptly. Recently, integration of Internet e-mail into corporate messaging systems has become a more urgent and significant requirement.

While there is some excellent work going on to incorporate X.400 in the Internet environment, there has been a limited amount of successful integration between these two groups within PRMDs. One reason is that expertise which overlaps these two groups is, sadly, lacking. It is not at all uncommon to have an intra-company meeting with representatives from each environment present, but only limited communication resulting because of the lack of overlapping knowledge.

This paper is intended to bridge this gap. While there have been many excellent papers, books, lectures etc. (many of which are either cited or included here) on the subjects of both the Internet and X.400, there has been not been a paper on one environment with the intended audience of the other environment. This is a paper about the Internet e-mail and related services with the intended audience of middle managers who are familiar with X.400 and vendor supplied e-mail products.

Companies can no longer afford to have two separate e-mail systems. They need to have some level of integration with every e-mail user within their own environment. On the other hand, nobody is willing to sacrifice functionality, service or, most important of all, information flow.

This paper offers an overview of the Internet e-mail services, a more detailed look at end-user interfaces, the message transfer services and addressing on the Internet. It also discusses related services that users of the Internet expect as part of their overall e-mail service. These sections cover such things as mailing lists and directory services. Finally, this paper explores related topics that often are understood only by experienced Internet users: Acceptable Use Policy, Proper Etiquette and the Usenet conferencing system.

The intent of this paper is to provide information which will allow readers who are unfamiliar with the Internet e-mail environment to have a base level of understanding, explained in terms which are understandable to the inexperienced reader, of what the people who use the Internet within their company are really using.

Chapter 1

Introduction

Users of X.400-based e-mail systems may be baffled when examining the resources available via the Internet, the global network of networks connecting academia and industry. The Internet was not a well-planned enterprise, but instead arose in more of an "organic" way. In all, the Internet's interconnected and overlapping services may appear to the X.400 user as a jungle of incomprehensible and dangerous-looking vines.

This paper is intended as a machete to hack through the Internet jungle.

The paper takes a snapshot of electronic mail and related services that exist in Internet today. The Internet e-mail services refer to all that is necessary to accomplish electronic mail communication among Internet hosts and others.

The description given in this document is in the context of the X.400 and X.500 models. It is assumed that the reader is familiar with the concepts and terminology of X.400.

Internet e-mail services as they exist throughout Internet today are not based on a pre-designed and homogeneous system. Internet e-mail carries a great deal of historic baggage and compromises with itself. As such it has proven to be difficult to understand and manage for the uninitiated. Yet, e-mail is regarded by many as their most valuable application. For these reasons a document such as this can be of value to many.

This document addresses many aspects of Internet e-mail for a wide spectrum of readers. Those who are not familiar with Internet e-mail will find it useful to read this entire document. After reading this document,

```
SMTP, DNS, MX, UUCP, MMDF, Sendmail, MH, ELM, PEM, MIME, ...
```

will no longer be a random sequence of letters and each component's place in the over all picture will be clear.

The level of detail provided here is not intended to be sufficient for a system administrator or an end user to solve his problems. This document is primarily intended to be used as a road map, where a general overview of all the significant components is provided and inter-relationships of these components are described. Reference to relevant material where additional detail can be obtained is always provided.

1.1 Purpose and Scope

PRMD Operators Committee has determined that an educational paper discussing the Internet and its electronic mail facilities would be of value to the members of the Electronic Mail Association as a whole. This paper is prepared for that purpose.

Many corporations have not fully integrated their corporate e-mail networks and their engineering research and development e-mail networks. Often these research and development networks are linked to the Internet.

Integration of Internet e-mail into corporate messaging systems is often a requirement. This integration can be more efficient if the characteristics, services and the culture surrounding Internet e-mail is better understood by PRMD Operators. This paper specifically addresses this need.

Furthermore, the maturity and widespread availability of Internet e-mail software makes SMTP a possible choice for implementing corporate messaging backbones in certain environments. This paper contains information that can contribute to better decisions about such issues.

1.2 Intended Audience

PRMD Operators' management staff, PRMD Operators' technical staff, messaging system architects and those wishing to understand Internet e-mail services can all benefit from this paper. The structure of this paper is intended to allow for all of these audiences.

The primary audience of this paper is the middle level management staff operating a PRMD. It is assumed that the reader has some knowledge of X.400 Message Handling Systems and of the practical aspects of deploying mail user agents (UAs) and gateways. Where necessary, the paper provides technical information (although generally not in great detail) and therefore should prove valuable to technical PRMD Operator staff as well as to the managers responsible for realizing transfer of application data.

It is intended that the PRMD Operators Committee presents this paper to the Electronic Mail Association. Members representing public mail carriers, private companies, government entities, academic institutions, and those in the standards communities will find the issues discussed in this paper of immediate importance to them.

1.3 Disclaimer

Because of the very dynamic nature of the Internet, the information contained in this paper may not be accurate by the time you read this. The author has made every possible effort to ensure accuracy of information contained in the paper at the time of the writing.

However, neither Electronic Mail Association nor Tele-Voice provide any warranty that the information included in this paper is fit for a specific purpose. In particular, not all the information in the chapter on Internet e-mail software (see section **??** [Internet E-Mail Software], page **??**), has been independently verified and use of no particular software is recommended in this paper.

Much of the information in this paper may be regarded as a collection of existing papers available through out Internet which express the collective knowledge of the Internet community about various aspects of Internet e-mail services.

About structure of this paper

The structure of this paper is as follows.

Chapter 1, (see section **??** [Introduction], page **??**), defines the purpose and scope of this paper.

Chapter 2, (see section **??** [Overview of Internet E-Mail Services], page **??**, provides an overview of the services described in the paper.

Chapter 3, (see section **??** [Internet E-Mail Related Protocols], page **??**, introduces the underlying protocols that are the foundation of Internet e-mail.

Chapter 4, (see section **??** [Internet E-Mail Software], page **??**), provides a description of the popular User Agents that are being used in various environments.

Chapter 5, (see section **??** [Proper use of Internet E-Mail Services], page **??**) deals with the *proper* use of interpersonal messaging in Internet. This is expressed from the Internet communities' perspective. It is important to note that *proper* use as expressed in this chapter should not be viewed as universal; it simply is an attempt to express the Internet Community's perspective. Proper etiquette for e-mail communication is amongst the topics covered in this chapter.

Chapter 6, (see section **??** [Internet E-Mail Directories], page **??**), describes e-mail directory services available throughout Internet. If you know some person's name and domain chances are that with a bit of investigative work you can find their e-mail address. The information in this chapter can help you in that search.

In addition to a comprehensive bibliography, an electronic references list is also included. Electronic references such as, [E1], identify documents that are electronicly available through the Internet.

Chapter 2

Overview of Internet E-Mail Services

2.1 The Internet

The Internet is a collection of networks spanning the globe, with a large number of research, educational and commercial networks connected together into one global network. It has gateways linking organizations throughout the world and yet it has no central, worldwide, technical control point. This loosely-organized, international collaboration of autonomous, interconnected networks, supports host-to-host communication through voluntary adherence to open protocols and procedures defined by Internet Standards.

At present, the organizations that connect through the Internet include corporations, universities and colleges, research laboratories – government and private, libraries, specialized scientific centers, state agencies, K-12 (Kindergarten-12th Grade) organizations and individuals. Each one of these organizations has its own network which is in turn connected to the Internet. These smaller networks all use a common suite of networking protocols known as TCP/IP. This commonality of protocols, coupled with a commonality of network functionality and interoperability, creates what to a user seems to be a single, integrated, worldwide, virtual network.

The Internet serves as both a live test-bed for on-going networking research and a daily communications tool for millions of users. It is a major tool in academic and industrial research in computer technology, physics, and astronomy, and increasingly in biological, social, and other sciences.[1] RFC-1296,[2] [?], provides statistics on the

[1] See "The Matrix", [?], for additional information.

[2] RFCs (Request For Comment) are a series of Internet technical documents. See section **??**

growth of the Internet by examining the number of Internet hosts and domains over a 10-year period. As of July 1993, 1.8 million IP hosts were on the Internet.[3] Estimates of number of users connected to Internet vary, depending upon the definition of connectivity used. It is estimated that some 20 million users around the world are linked through the Internet.

The Internet is growing at a phenomenal rate, with no deceleration in sight. Every month thousands of new users are added. New networks are added literally almost every day. During the period from 1981 to 1992 the number of hosts has nearly doubled every year. The current growth is running at some 85% per year. Figures 2.1 (see section **??** [Internet Growth (1981-1991)], page **??**) [**?**], illustrate the exponential rate at which the number of Internet hosts has been increasing through the past decade.

As expected, the traffic on the Internet has been growing as well. Figure 2.2 illustrates the growth of traffic on the Internet, as measured by the number of packets of information sent through the network.

Not only has the population of Internet been growing at a phenomenal rate but the diversity of services available through Internet have been growing very rapidly and steadily as well. Internet-enabled interpersonal communication services in general, and electronic mail in particular are amongst the most popular features of Internet. Other "basic" functions provided are access to remote computational and informational facilities, and file transfer. Such informational facilities include distributed file systems, distributed bulletin board services (e.g., Usenet), a variety of distributed information systems and a very extensive network of file archives.

These services allow users to make available information to others, increase productivity through "cooperative work tools" or offer a way to share resources and coordinate work among users. This document primarily focuses on use of electronic mail in Internet.

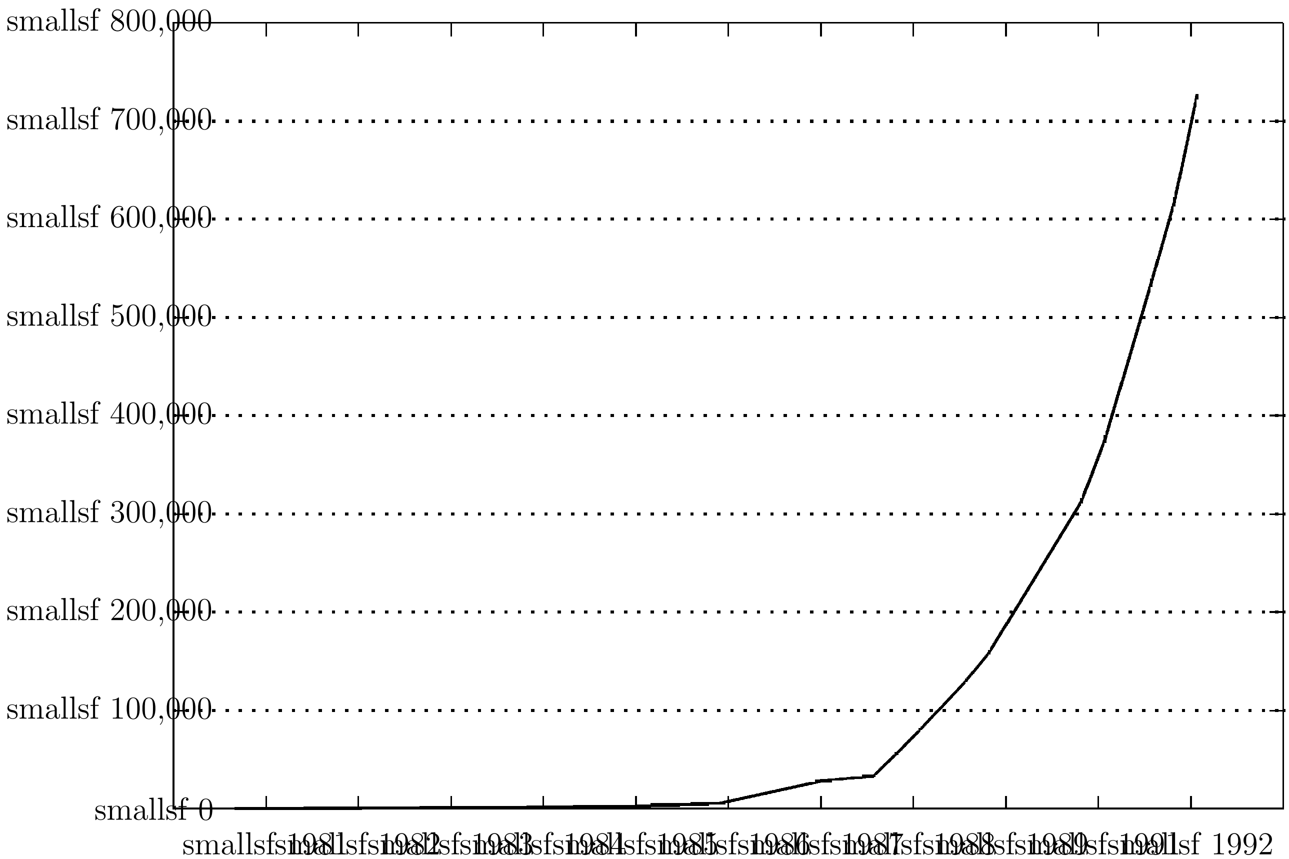
The next sections provides an overview of Internet e-mail and related services.

2.2 Internet E-Mail Services

Mail is the most common service available on Internet. With capabilities far exceeding traditional mail services or telephone messaging systems, its importance has grown dramatically over the past decade. Its unhindered and continuous operation is critical to many organizations worldwide for exchange of information. This is to the extent

[About RFCs], page **??**, for more details.

[3] One easily computed measure of Internet connectivity is the count of host address records within the Domain Name System. See section **??** [Domain Name System (DNS)], page **??**.



otesize SOURCE: RFC-1296

Figure 2.1: Number of Internet Hosts (Linear)

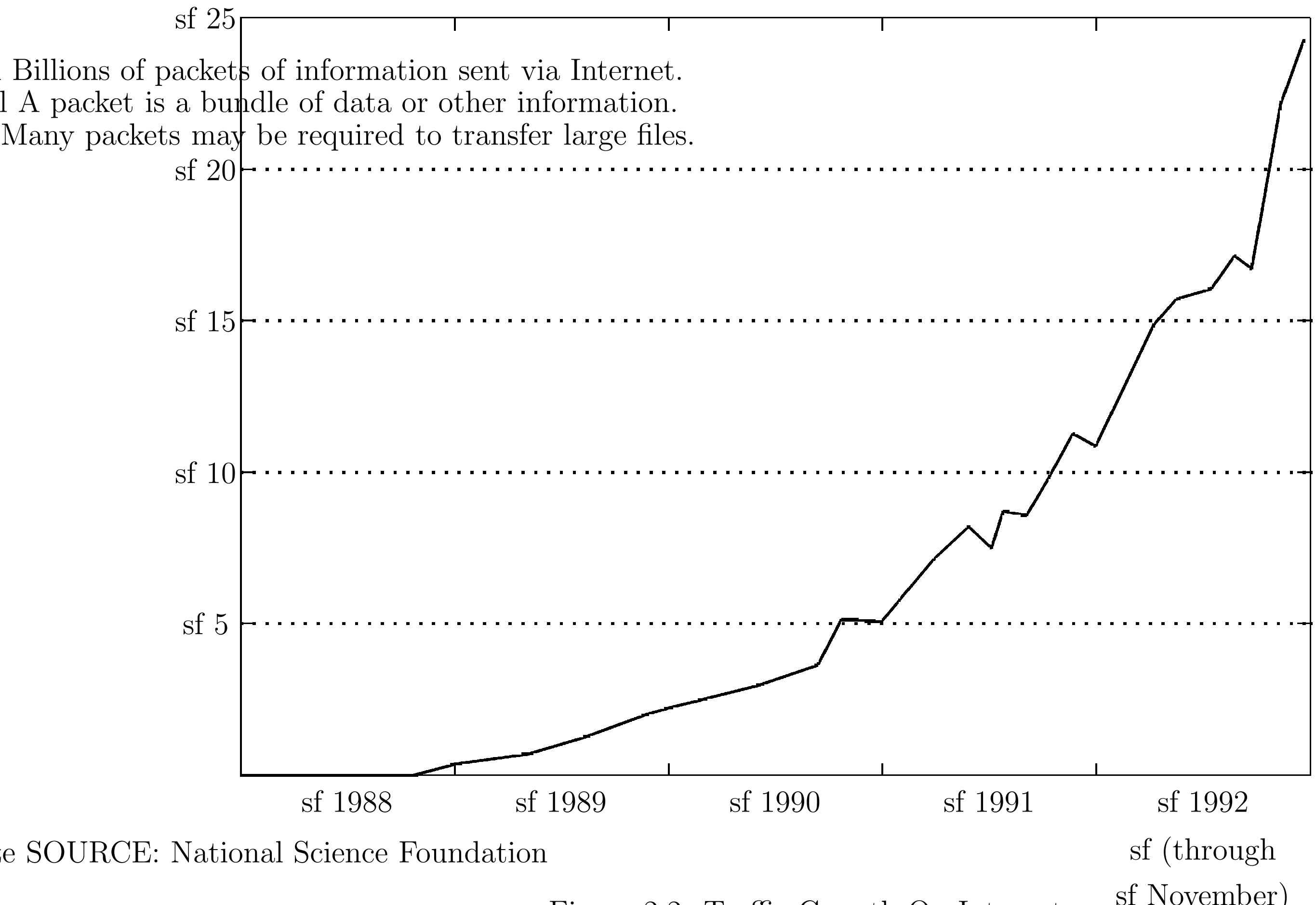


Figure 2.2: Traffic Growth On Internet.

that a one day interruption in service is viewed with grave concern and all efforts are made to restore service as soon as possible.

Over the years, use of Internet e-mail has grown especially rapidly in the academic and research community: colleges, universities, government and private research institutions and scientific centers. Topics discussed and information exchanged through this medium range from politics and history to critical experimental data and theoretical developments in the sciences and engineering. Often such dialogues are carried out through *mailing lists* – that is, long-lasting *distribution lists* involving people who want to hold extended discussions on the same subjects. There are also computer conferencing systems available which allow large groups of people to post messages to all members of the group. Such systems provide for sophisticated user interfaces and separation of messages into categories by topic. They are used for detailed threads of discussions within continuous topics.

Table 2.1 presents an overview of the services available, and some general comments about them.

The common Envelope and Header fields, From, To, Cc, Bcc, Date, Subject, ... are all supported. These fields identify the sender of the message, the receiver(s), message date and the subject (as specified by the sender) of the message. Non-Delivery Reports are supported. Delivery reports and Notifications of receipt of e-mail are not supported.

Delivery can be to a number of users at the same time, as specified in a *distribution list*. Forwarding of mail to other address(es) is also supported.

Security issues such as privacy and authentication are included as enhancements to the basic service through Privacy Enhanced E-Mail (PEM). Incorporation of complex body parts is through the Multipurpose Internet Mail Extensions (MIME).

Frequently the issues of "size" and "message routing" are of interest to users of Internet e-mail. As with many evolutionary realities, support for "size" is not uniform throughout the whole network. However facilities that provide for segmentation and re-assembly do exist. At a minimum, message sizes up to 64 K bytes can be reliably routed, though in many cases this figure is 1 M Byte or more.

Routing of messages in a network of 1.8 million hosts is non-trivial. A great deal of the success of Internet E-Mail can be attributed to Domain Name System (DNS). DNS is a directory service that maps host names to host addresses and also accommodates routing of e-mail.

Service	*Comments*
From:, To:, Cc:, Bcc:, Date:, Subject:	All Supported. See section **??** [Internet E-Mail Format], page **??**.
Non-Delivery Report	Supported.
Distribution Lists	Supported.
Delivery Reports	Not Supported.
Receipt Notifications	Not Supported.
Security (Authentication, Confidentiality, Integrity, ...)	PEM. See section **??** [Privacy Enhanced E-Mail (PEM)], page **??**.
Complex Body Parts	MIME. See section **??** [Multipurpose Internet Mail Extensions (MIME)], page **??**.
Addressing	Simple and It Works. See section **??** [Addressing and Routing], page **??**.
Message Size Limitations	64K Bytes always permitted, most cases 1Meg or more is allowed.
Message Transfer Service (MTA) Limitations	SMTP (Simple Mail Transfer Protocol). See section **??** [SMTP], page **??**. SMTP is limited to 7 bit ASCII. Extension specifications for SMTP exist.
Probe	Not supported.
Trace Information	At Every MTA.
Message Routing	DNS holds it all together. See section **??** [Domain Name System (DNS)], page **??**.
Mail Address Directory	Many not One (Finger/WAIS/Quipu/...). See section **??** [Finding E-Mail Addresses], page **??**.
Mail Gateways	Lots.
Conferencing Systems	Usenet. See section **??** [Usenet], page **??**.

Table 2.1: Overview of Internet E-Mail Services

Chapter 3

Internet E-Mail Related Protocols

The architecture and technical specifications of the Internet are the result of numerous research and development activities conducted over a period of over two decades, performed by the network research and development community, by service and equipment vendors, and by government agencies around the world.

This chapter focuses on Internet e-mail related specifications. First we provide an overview of how Internet specifications are published. Next the Internet standardization process is described. Then an overview of e-mail specifications and their inter-relationship is presented. Each specification is then briefly discussed.

3.1 About RFCs

Each distinct version of an Internet specification is published in the form of a "Request for Comment", or RFC. The RFCs form a series of publications of networking technical documents, begun in 1969 as part of the original Defence Advanced Research Projects Agency (DARPA) wide-area networking (ARPANET) project. From a historical perspective it is interesting to note that the earliest e-mail related RFC dates back to 1971.

RFCs cover a wide range of topics, from early discussion of new research concepts in computer communications to status memos about the Internet. RFCs are continuously updated. However, they are never revised or reissued with the same number.

Not all RFCs are Internet standards. RFCs do not become Internet standards until their correctness and implementability, reliability, deployability, manageability, and practicality has been proven. There is an RFC which says what RFCs are current Internet standards.

3.2 Internet Standardization Process

In general, an Internet Standard is a specification that is stable and well-understood, is technically competent, has multiple, independent, and interoperable implementations with operational experience, enjoys significant public support, and is recognizably useful in some or all parts of the Internet.

Not all specifications of protocols or services for the Internet should or will become Internet Standards. Such non-standards track specifications are not subject to the rules for Internet standardization; generally, they will be published directly as RFCs at the discretion of the RFC editor. These RFCs will be marked as "Experimental" or "Informational". Such non-standards-track documents do not necessarily reach the level of Standard.

Organization

The Internet Activities Board (IAB) is the primary coordinating committee for Internet design, engineering, and management.[1] The IAB has delegated to its Internet Engineering Task Force (IETF) the primary responsibility for the development and review of potential Internet Standards from all sources. The IETF forms Working Groups to pursue specific technical issues, frequently resulting in the development of one or more specifications that are proposed for adoption as Internet Standards.

Final decisions on Internet standardization are made by the Internet Engineering Steering Group (IESG). Any member of the Internet community with the time and interest is urged to attend IETF meetings and to participate actively in one or more IETF Working Groups. Working Groups accomplish most of the tasks via e-mail; IETF attendance is not a requirement for contributing meaningfully to a Working Group. Participation is by individual technical contributors, rather than formal representatives of organizations. The process works because the IETF Working Groups display a spirit of cooperation as well as a high degree of technical maturity; most IETF members agree that the greatest benefit for all members of the Internet community results from cooperative development of technically superior protocols and services.

The status of specifications on the Internet standards track is summarized periodically in a summary RFC entitled "IAB Official Protocol Standards". This RFC

[1] RFC-1310, [?], describes the Internet Standards Process. A summary of that document is reproduced here.

shows the level of maturity and other helpful information for each Internet protocol or service specification. The IAB views the RFC publication process to be sufficiently important to warrant including the RFC Editor in the IAB membership.

The "IAB Official Protocol Standards" RFC is the authoritative statement of the status of any particular Internet specification, and it is the "Publication of Record" with respect to Internet standardization.

Process

Protocols which are to become standards in the Internet go through a series of states or maturity levels (proposed standard, draft standard, and standard) involving increasing amounts of scrutiny and testing.

Advancement of a protocol to proposed standard is an important step since it marks a protocol as a candidate for eventual standardization (it puts the protocol "on the standards track"). Advancement to draft standard is a major step which warns the community that, unless major objections are raised or flaws are discovered, the protocol is likely to be advanced to standard in six months.

Protocol State

In addition to being designated as "standard", protocols can be assigned other "states". These states are *informational*, *experimental*, and *historic*.

An informational protocol is one developed by standard organizations other than the IAB, or by vendors or others outside the purview of the IAB. Such protocols may in some cases also be recommended for use in the Internet by the IAB.

Experimental protocols are typically those developed as part of an ongoing research project not related to an operational service offering. While they may be proposed as a service protocol at a later stage, and thus be put on the standards track, they are in general not intended for operational use.

Finally, historic protocols are those that are unlikely to ever become standards in the Internet either because they have been superseded by later developments or due to lack of interest.

Protocol Status

In addition to a state (like "Proposed Standard"), a protocol is also assigned a status, or requirement level. These can be "Required", "Recommended", "Elective", "Limited Use", and "Not Recommended".

STATUS STATE	Required	Recommended	Elective	Limited Use	Not Recommended
Standard	X	XXX	XXX		
Draft Std.	X	X	XXX		
Proposed Std.		X	XXX		
Informational		X	XXX	XX	X
Experimental			X	XXX	XX
Historic				X	XXX

SOURCE: RFC-1360

Table 3.1: Protocol Categorization

At any given time a protocol occupies a cell of Table 3.1. Protocols are likely to be in cells in about the following proportions (indicated by the relative number of Xs). A new protocol is most likely to start in the (proposed standard, elective) cell, or the (experimental, not recommended) cell.

3.3 Overview of Internet E-Mail RFCs

A listing of those RFCs pertinent to e-mail appears in Figure 3.1.[2] These cover the following topics:

- Internet E-Mail Format (RFCs 822, 1049)
  See section **??** [Internet E-Mail Format], page **??**.
- Addressing and Routing (RFCs 920, 974, 1034, 1035)
  See section **??** [Addressing and Routing], page **??**.
- Transfer protocols (RFCs 821, 976)
  See section **??** [Transfer Protocols], page **??**.
- Enhancements to the basic service (RFCs 1341, 1342, 1421, 1422, 1423, 1424)
  See section **??** [Enhancements to the Basic Service], page **??**.
- Interoperability with X.400 (RFC 1327)
  See section **??** [Interoperability with X.400], page **??**.

[2] Instructions on how to obtain RFCs are included as appendix to this paper. See section **??** [Obtaining Request For Comments], page **??**

In Figure 3.1, the official Internet standards are marked by "*".

We shall present an overview of these topics in the following sections.

3.4 Internet E-Mail Format

The format of mail used in Internet is defined by RFC-822, [?]. The format of headers is quite simple. The entire header is encoded in 7 bit ASCII as lines of text. Furthermore, RFC-822 only specifies a format for messages and does not specify any delivery mechanisms. The simplicity and flexibility of the format of RFC-822 messages are its main strengths, as it can easily be implemented on all systems. It is also its weakness: support of non-ASCII character set is missing; the structure of the header is very limited and body of messages are unstructured. The entire scope of Internet e-mail can be defined as RFC-822 compliant.

Extensions to RFC-822 exist that address its shortcomings. Many of these extensions are described in this paper.

In 1988, a Content-Type header field was added to Internet mail messages. RFC 1049, [?], recognized the need for using the mail system for sending around information with a greater degree of structure, while remaining within the constraints imposed by the limited character set.

3.5 Addressing and Routing

Once we have an agreed upon system of formatting an electronic message, we will have to institute an addressing system, between sites that transmit and receive e-mail. This is done through *Domains*.

3.5.1 E-Mail Addresses

E-Mail in the Internet is addressed to a user or other named facility at a host using a two part address of the form: `<user or facility>@<host>`, where host is identified by a domain name (see Domains, below).

For the purposes of E-Mail a host may be a true internet host with an assigned set of IP addresses, a non-internet host searched through a gateway with an assigned set of IP addresses or a provider host covering all users on an organization (e.g. dec.com, apple.com, ssw.com, hp.com, etc.) serviced by a directory host with an assigned set of IP addresses.

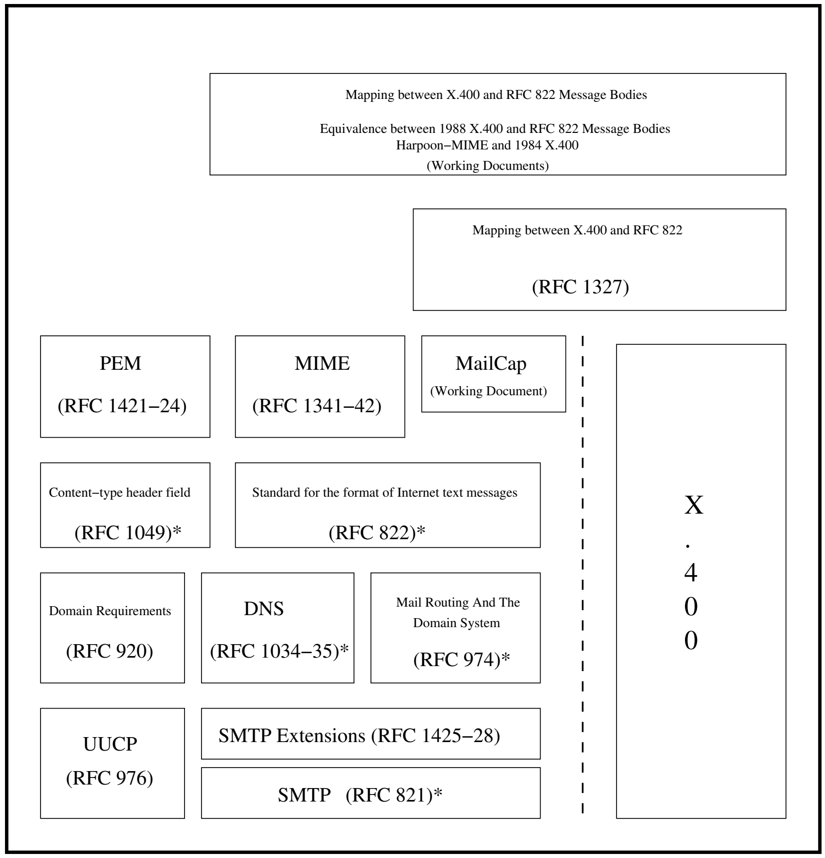


Figure 3.1: Mail Related RFCs

3.5.2 Domains

For the purpose of addressing, domains[3] are a way of uniquely specifying a destination. Much like a postal address, a domain specifies a set of progressively more restrictive "domains" of the potential address space.

Domain requirements are considered in RFC 920, [?]. The Domain Name System is talked about in RFCs 1034-35, [?], [?], and issues related to mail routing and the domain system are considered in RFC 974, [?].

As an example consider the following destination specification:

`mohsen@rostam.neda.com`

This is an example of an e-mail address. It is a method of specifying a given person on a specific domain.

The top level domain (com in the above example) has many possible values. In the United States, "`com`", "`mil`", "`edu`", and "`gov`" are fairly standard. Elsewhere, the top level domain tends to be a country code, the second level tends to be a province or state, OR a classification like "`edu`" or "`ac`" for academic (such as `ac.jp`, `go.jp`, `ac.uk`, `edu.au`, etc.) and the third an organization.

3.5.3 Domain Name System (DNS)

In order to transmit a message, an internet mailer needs to transform a host name into an IP address. In the early days of the Internet, this was achieved by a simple host name to IP address mapping tool on each host that was distributed at regular intervals from a control location. This tool continues to exist as a fallback mechanism on many hosts, but copies are no longer distributed. They became too large and too dynamic and another approach is now employed. This apporach is the Domain Name System (DNS).

The DNS is a set of data bases organized in a hierarchical fashion. Each root domain (e.g. com, mil, edu, gov, uk, ca, etc.) maintains a root database (in detail they are replicated at a number of hosts) at a set of known, IP addresses. The root database contains records that identify the IP addresses of all other root databases and all sub-domains in its domain (e.g. ssw.com, dec.com, nsf.gov, etc.) The databases for each sub-domain contains the IP addresses of their parents and all of their own sub-domains. These processes can continue to any depth. In detail, a sub-domain

[3]Internet domains are in many ways different from X.400 domains.

```
host name 0 - mx1 - host name 1 - mxx - host name x - Ax - IP1
            \ mx2 - host name 2 - A2 - IP2a, IP2b
            \ mx3 - host name 3 - A3 - IP3
```

Figure 3.2: Example DNS Entry

may maintain its own database, or have its entries entered (by agreement) in its parent's DNS database.

From an e-mail perspective a DNS database contains two types of records. "A" (address) records that map a host name or partial (wild-carded) host name to an IP address and "MX" (mail exchange) records that map a host name, or partial (wild-carded) host name to another host name. There may be multiple "A" or "MX" records in a data base for a host. Multiple "A" records allow an unordered set of IP addresses to be associated with a single host name. Multiple "MX" records allow an ordered sequence of hosts to be associated with a host name (MX records contain a weight that is used to order them).

When a host name is resolved on the DNS, the result is a sequence of IP addresses derived by traversing the MX and A record tree (See Figure 3.2).

MX records have two roles in the e-mail Internet. First to identify an Internet gateway host providing access to non-internet hosts, and second to identify internet "firewalls". From within an organization the IP addresses associated with the highest priority MX record would be accessible. From without an organization, the IP addresses associated with the highest priority MX record can be blocked and thus delivery will be affected via the host identified by a lower priority MX record and its associated IP addresses.

RFCs 1034-5 include detailed information on the Domain Name System (DNS). Issues related to domain requirements are talked about in RFC 920. Mail routing and the Domain system itself are the subject of RFC 974.

(DNS) is a distributed, replicated name service whose primary purpose are to:

- map host names into corresponding Internet addresses,
- map Internet addresses into hostnames.
- locate forwarders for electronic mail transfer.

Over a million computers implement the Internet's Domain Name System, making it the world's most distributed database. DNS client software is known as a resolver, and many resolver implementations exist for most popular computer systems.

At its heart, DNS consists of a naming taxonomy that partitions a namespace of virtual hostnames (such as `uw-isdl.ee.washington.edu`) and corresponding host IP addresses across a hierarchal collection of *DNS servers*. Each server holds some portion of the partitioned database of hostname/address mappings and is capable of responding to user queries on the subset of attributes and values it contains.

In operation, the resolution of a domain name consists of sending a series of queries to a subset of DNS servers, each one responsible for some portion of the corresponding fully qualified domain name.

DNS is also used to distribute information about host hardware, operating system configurations and electronic mail exchanger addresses, and it is possible to query the system for wild-card matches (for example, it is possible to ask for all records of a particular type matching a particular string).

DNS has been an operational success, having expanded continuously since its inception to now cover over 1 million machine names. Despite this success, there are some problems with both the basic architecture and the specific implementations now in service on the Internet.

Maintenance of the system is distributed, with the required information usually entered into flat text files (usually by hand) at the site of each authoritative sub-domain server. This can lead to inconsistencies and errors in the database that can only be corrected through human intervention. There is no internal consistency checking of this information by the system itself (for example, to verify that registered hosts actually exist on the net).

Another problem can arise during operation. If the authoritative server for a particular sub-domain remains offline for an extended period, then users will find that they cannot perform hostname to address conversion. In this case, users can find themselves unable to access a host, even though that particular host is available.

Mail Exchange (MX) Records

A non-SMTP/Internet site that wishes to register on the Internet will need to get a "nearby" Internet site to set up an MX record for them. An MX record is essentially a domain-server database record that (effectively) registers your domain name on the Internet, and indicates that the Internet site knows how to forward mail to you. Usually, the forwarding is done via some non-SMTP/Internet route, such as UUCP (Unix to Unix Copy Program). You can get an MX record for one site, or a "wildcard" MX record so that you can have your own sub-domains.

3.6 Transfer Protocols

With the message format and the addressing system in place, one has to decide on protocols for transmitting a message from one site to another. For sites on the Internet, this transfer protocol is known as SMTP (Simple Mail Transfer Protocol). For other sites not on the Internet, message transmission takes place using some privately arranged mechanism. Many private mechanisms are used in practice; one of the more popular ones is UUCP.

3.6.1 SMTP

RFC 821 [**?**] and RFC 1123, [**?**], specify this transfer protocol. This is the communications protocol used most commonly over TCP/IP links. SMTP usually operates directly between the source and destination machines, so intermediate machines don't get involved (except for gateways, see below). SMTP is usually implemented as part of the MTA. Extensions to SMTP are specified in RFCs 1425 [**?**], 1426 [**?**], and 1427 [**?**].

3.6.2 UUCP

Details of this protocol are provided in RFC 976, [**?**]. UUCP is an old protocol. It was very heavily used when UNIX systems communicated only over RS232 serial lines, usually over modems. UUCP is implemented as a suite of programs developed back in the early 70's to provide this communications link. All that UUCP does is transfer files from one system to another. There is an additional mechanism where one system can direct the destination system to run a file through a specific program. Electronic mail in UUCP is simply requesting the destination machine to run ‘‘mail” on a data file.

UUCP communicates by means of ‘‘protocols”, the most common being ‘‘g”, a method for transmission of data over telephone lines and ensuring that the data is not corrupted. There are several other protocols, none universally available, and most oriented towards communication media other than telephone voice lines (such as dialup X.25, PAD X.25, or LAN connects).

UUCP operates over fixed system-to-system links, so sending mail from one system to another often has to traverse other intermediate systems.

The UUCP network is that set of machines that talk with each other via UUCP. Sending mail through this network requires that the sender know the network topology of UUCP links and specify a path from one machine to the next. The UUCP Mapping Project hides much of this complexity.

3.7 Enhancements to the Basic Service

3.7.1 Privacy Enhanced E-Mail (PEM)

Initial specification of Internet e-mail did not address security issues. In particular, security mechanisms to provide Data Confidentiality, Authenticity, Integrity and Non-Repudiation were missing.

Security extensions to Internet e-mail that address these shortcomings have been specified in the form of a set of related RFCs. (RFCs 1421 [?], 1422 [?], 1423 [?], and 1424 [?]).

These RFCs defines message encipherment and authentication procedures, in order to provide privacy enhancement services for electronic mail transfer in the Internet.

Those familiar with X.509 will detect a great deal of similarities between PEM mechanisms and procedures for Digital Signatures defined in X.509.

The procedures defined in these RFCs are intended to be compatible with a wide range of key management approaches, including both symmetric (secret-key) and asymmetric (public-key) approaches for encryption of data encrypting keys. Use of symmetric cryptography for message text encryption and/or integrity check computation is anticipated.

Privacy enhancement services (confidentiality, authentication, and message integrity assurance) are offered through the use of end-to- end cryptography between originator and recipient User Agent processes, with no special processing requirements imposed on the Message Transfer System at endpoints or at intermediate relay sites. This approach allows privacy enhancement facilities to be incorporated on a site-by-site or user-by-user basis without impact on other Internet entities. Interoperability among heterogeneous components and mail transport facilities is supported.

3.7.2 Multipurpose Internet Mail Extensions (MIME)

Recently, the Internet Engineering Task Force (IETF) has developed a document titled "Multipurpose Internet Mail Extensions" or MIME, RFC-1341, [?] and 1342 [?].

MIME is the official proposed standard format for multimedia Internet mail encapsulated inside standard Internet RFC 822 messages. Facilities include sending multiple objects in a single message, character sets other than US-ASCII, multi-font text messages, non-textual material such as images and audio fragments, and other extensions.

MIME defines structure for Internet message bodies through enhancements to the Content-Type field.

In MIME, the term "content-type" is used to refer to an information object contained in the body of a message. In contrast, X.400 uses the term "body part type".

RFC 1341 covers only message bodies, not message headers; to see how to represent non-ASCII characters in message headers, see Internet RFC 1342: K Moore, "Representation of non-ASCII text in Internet message headers" (June 1992).

MIME, the Multi-purpose Internet Mail Extensions, is a freely available specification that offers a way to interchange multi-media e-mail among many different computer systems. MIME supports not only several pre-defined types of non-textual message contents, such as audio, GIF files and PostScript images, but also permits you to define your own types of message parts.

3.7.2.1 Pre-Defined MIME Types

Each part of a multimedia message identifies what type of information is carried in the message part. For example, a message part containing audio data might have either type `audio/basic` or type `audio/x-next`. Both are `audio` types; the *subtypes* are `basic` and `x-next`.

An entire MIME message—as opposed to an individual part of a multipart message—can also have a type. For example, a message might have the type `text/plain`, and consist entirely of plain text. A MIME message containing parts of different types has the umbrella type `multipart/mixed`.

Here are brief summaries of the pre-defined types and subtypes in MIME version 1.0.

application/octet-stream is for "other" kinds of data, either uninterpreted binary data or information to be processed by a mail-based application.

application/postscript is for PostScript programs (typically documents represented in PostScript form).

audio/basic is for audio data encoded using 8-bit ISDN (Pulse Code Modulation). A sample rate of 8000/second and a single channel is assumed.

image/gif is for GIF (graphical interchange format) image data.

image/jpeg is for JPEG (joint picture experts group) image data.

message/external-body is for specifying large bodies by reference to an external data source.

message/partial is for partial messages, to permit the fragmented transmission of bodies that are thought to be too large to be passed through mail transport facilities.

message/rfc822 is an encapsulated RFC 822 conformant message which may have its own type.

multipart/alternative represents the same data in multiple formats.

multipart/digest is for multipart entities in which each part is an encapsulated message.

multipart/mixed is the primary way to represent a MIME message containing parts of various different types.

multipart/parallel is for data intended to be viewed simultaneously.

text/plain indicates plain (unformatted) text.

video/mpeg is for MPEG (motion picture experts group) video data. Video requires the capability to display moving images, typically including specialized hardware and software.

There may be other registered types and subtypes down the road. MIME also allows arbitrary subtypes whose names are prefixed with "x-", but anything else is reserved for registered types.

RFC 1341, [**?**], contains more detailed explanations of required or optional attributes to be used with particular types.

3.8 Interoperability with X.400

Integration of X.400 and X.500 services into the Internet are in progress. Coexistence of X.400 and SMTP are expected.

Since the introduction of X.400(84), there has been work ongoing for defining mappings between MHS and RFC 822. The most recent work in this area is RFC 1327, [?], which focuses primarily on translation of envelope and headers.

Two recent Internet Draft documents:

1. Mapping between X.400 and RFC-822 Message Bodies
2. Equivalences between 1988 X.400 and RFC-822 Message Bodies

focus on mapping of message bodies.

> **Note:** Internet Drafts are working documents.They may be updated, replaced or obsoleted by other documents at any time. Those interested in more information on these papers should contact Internet Engineering Task Force MIME-MHS Interworking Working Group.

These documents have been specifically designed to provide optimal behavior for three different scenarios:

1. Allow a MIME user and an MHS user to exchange an arbitrary binary content;
2. Allow MIME content-types to "tunnel" through an MHS relay (that is, two MIME users can exchange content-types without loss through an MHS relay); and,
3. Allow MHS body parts to "tunnel" through a MIME relay (that is, two MHS users can exchange body parts without loss through a MIME relay).

Chapter 4

Internet E-Mail Software

Much of Internet e-mail services are delivered through “freely”[1] available software. In order to have a better understanding of Internet e-mail it is important to have some understanding of the software that delivers the services.

The information contained in this chapter is intended to be used as reference information. This chapter is primarily targeted for consumption by technical staff.

Those not interested in such details may benefit from simply “glancing” over this chapter.

E-Mail softeware can best be discussed based on two dimensions.

1. The platforms that the software runs on (e.g., UNIX, DOS, MAC)

2. The services that the software implements (e.g., User Agent, MTA).

Two electronic papers,

- UNIX E-Mail Software, [E1]

- UUCP Mail, News and Gateways Software for PCs and MACs, [E2]

provide a comprehensive list of freely available Internet e-mail software.

These electronic papers introduce many types of E-Mail software which include:

- User Agents

- E-Mail Directories

- E-Mail Sorters

- Message Transfer Agents

[1]See Richard Stallman for one definition of “Free Software”, [**?**].

- E-Mail Gateways
- Mailing List maintenance software
- Programming libraries for processing Internet e-mail messages.

In this paper, we focus on freely available UNIX Mail User Agent software.

User Agent Software

Depending on the level of sophistication of the user and the extent of messaging they do, there are a variety of choices.

Unix Mail User Agents can be categorized into three classes based on the environments they run under.

1. Character Based.
2. X Windows.
3. Emacs.

As an example of the type of software that is used in the Internet, the remainder of this section focuses on the author's choice of messaging environment.

4.1 Emacs Based User Agent Software

GNU Emacs is an advanced, self-documenting, customizable, extensible, real-time display editor. GNU Emacs is so customizable and extensible that calling it just an editor is a misnomer. Emacs is an **Editor Centered User Environment.**

Most of editing commands in Emacs are written in Emacs Lisp (elisp). Elisp provides for practically unlimited extension of GNU Emacs.

GNU Emacs runs on 100s of hardware/software platforms. Emacs runs on almost all flavors of UNIX. VMS, TOPS-20, ... are among other operating systems that Emacs has been ported to. GNU Emacs is freely available.

GNU Emacs provides a level of service adequate for providing most basic office services on many hardware/software platforms in a unified fashion.

To better introduce EMACS, the following section has been reproduced from "The new Hacker's Dictionary", [**?**].

> EMACS */ee'maks/* [from Editing MACroS] n. The ne plus ultra of hacker editors, a program editor with an entire LISP system inside it. It was originally written by Richard Stallman in **TECO** under **ITS** at the MIT AI lab, but the most widely used versions now run under UNIX. It includes facilities to run compilation subprocesses and send and receive mail; many hackers spend up to 80% of their **tube time** inside it.
>
> Some versions running under window managers iconify as an overflowing kitchen sink, perhaps to suggest the one feature the editor does not (yet) include. ...

Amongst the many features that GNU emacs offers are a set of very rich and integrated Electronic Mail facilities. The remainder of this chapter enumerates some of the more popular ones.

4.1.1 Rmail

At A Glance

```
Package Name: Rmail
Latest Version:  Emacs 18.58
Developer/Author: Free Software Foundation
Availability: FTP prep.ai.mit.edu
Sources: UNIX
Platforms: All platforms that emacs supports
Support: gnu.emacs
Folder Format: BABYL
Notable Features: Comes standards with emacs distribution.
                  Implemented entirely in elisp.
```

Reviews:

- **From: "Emacs Documentation"**

  Rmail is an Emacs subsystem for reading and disposing of mail. Rmail stores mail messages in Rmail files in BABYL format (originally used under the ITS operating system), although it can incorporate new mail from MMDF and Unix format files, or mixed-format files. Reading the messages in an Rmail file is done in a special major mode, Rmail mode, which redefines most letters to run commands for managing mail.

Rmail can do the standard things such as displaying, deleting, filing, or replying to messages. Replying uses another Emacs subsystem, Mail mode. Messages can be saved in either BABYL or Unix format. Rmail maintains per-message attributes and user-defined labels. Rmail can burst message digests.

4.1.2 MH-E

At A Glance

```
Package Name: MH-E
Latest Version:
Developer/Author: Stephen Gildea <gildea@bbn.com>
Availability: FTP prep.ai.mit.edu
Sources: Freely available
Platforms: All platforms that run EMACS
Support: Usenet newsgroups (gnu.emacs and comp.mail.mh)
Folder Format: mbox
Notable Features: Fully integrates MH into Emacs.
```

Reviews:

- **From: "MH Documentation"**

  MH-E is an interface to MH from within GNU Emacs. MH must be installed before MH-E can be used. MH must have been compiled with the MHE compiler flag. MH-E is distributed with both GNU Emacs and MH.

4.1.3 VM

At A Glance

```
Package Name: View Mail (VM)
Latest Version: VM 4.41
Developer/Author: Kyle Jones <kyle@uunet.uu.net>
Availability: Via anonymous ftp
   ab20.larc.nasa.gov:pub/vm/vm-4.41.tar.Z
   ftp.uu.net:pub/vm-4.41.tar.Z
   archive.cis.ohio-state.edu:
      pub/gnu/emacs/elisp-archive/packages/vm-4.41.tar.Z
Sources: Openly Available
Platforms: Every where Emacs Runs.
Support: Usenet newsgroup gnu.emacs.vm.info and by sending mail to
         info-vm-request@uunet.uu.net.
Folder Format: mbox
Notable Features: Entirely implemented in emacs-lisp.
```

Reviews:

- **From: "VM Documentation"**

  VM (View Mail) is a GNU Emacs subsystem that allows UNIX mail to be read and disposed of within Emacs. Commands exist to do the normal things expected of a mail user agent, such as generating replies, saving messages to folders, deleting messages and so on. There are other more advanced commands that do tasks like bursting and creating digests, message forwarding, and organizing message presentation according to various criteria.

4.1.4 SuperCite

At A Glance

```
Package Name: SuperCite
Latest Version: SC 3.0
Developer/Author: Barry A. Warsaw <bwarsaw@cen.com>
Availability:
Sources: Yes.
Platforms: Everywhere EMACS runs.
Support: mailing list supercite-request@warsaw.nlm.nih.gov
Folder Format: mbox
Notable Features: This is not a user agent. It is a user agent
                  add-on to accomplish sophisticated citing and
                  attributing of the original messages.
```

Reviews:

- **From: "UNIX EMail Software - a Survey"**

  Supercite is a GNU Emacs package written entirely in elisp which interfaces to common mail and news reading subsystems, and provides sophisticated citing and attributing of the original messages. Supercite has a very specific and limited role in the process of composing replies to both Netnews and Electronic Mail.

  Supercite is only useful in conjunction with mail/news reading subsystems such as VM, GNUS, RMAIL, etc. (hereafter referred to collectively as *readers*). Supercite is typically called through a hook, defined by the reader, when the initial reply buffer is set up. Thereafter, supercite's many commands and formatting styles are available in that reply buffer until the reply is sent, at which time supercite is re-initialized and ready for the next reply.

4.1.5 Bbdb

At A Glance

```
Package Name: BBDB
Latest Version: 1.48
Developer/Author: Jamie Zawinski <jwz@@lucid.com>
Availability: Through anonymous FTP
  archive.cis.ohio-state.edu:
       /pub/gnu/emacs/elisp-archive/packages/bbdb.tar.Z
Sources: Openly Available
Platforms: Every where Emacs Runs
Support: Mailing list info-bbdb-request@@lucid.com
Folder Format:
Notable Features: BBDB is not a user agent. It is an e-mail integrated
                  rolodex-like database.
```

Reviews:

- **From: "BBDB Manual"**

  BBDB is a rolodex-like database program for GNU Emacs. BBDB stands for *Insidious Big Brother Database*, and is not, repeat, *not* an obscure reference to the Buck Rogers TV series.

  It provides the following features:

  - Tight integration with mail and news readers, with little or no interaction by the user.
  - easy (or automatic) display of the record corresponding to the sender of the current message.
  - automatic creation of records based on the contents of the current message.
  - automatic addition of data to arbitrary fields of the record corresponding to the sender of the current message.
  - Listing all records which match a regular expression;
  - Listing all records which match a regular expression in a particular field (`company` or `notes,` for example).

4.2 UNIX Based User Agent Software

The following are some of the more popular UNIX based user agents.

4.2.1 MH

At A Glance

```
Package Name: MH
Latest Version: 6.8
Developer/Author: Many. Includes RAND Corp.,
                  University of California Irvine.
Maintainer:  John Romine <Bug-MH@ics.uci.edu>
Availability: FTP from,
   anonymous@ics.uci.edu[128.195.1.1]:pub/mh/mh-6.8.tar.Z
Sources: Yes
Platforms: UNIX
Support: comp.mail.mh includes a Frequently Asked Questions (FAQ).
Folder Format: individual files
Features: Very rich in features. Very customizable. Very extendible.
          Higher layer X11 and Emacs user interfaces to MH are
          also available.
```

Reviews:

- **From: "UNIX EMail Software - a Survey"**

  The big difference between MH and most other *mail user agents* is that you can use MH from a UNIX shell prompt. In MH, each command is a separate program, and the shell is used as an interpreter. So, all the power of UNIX shells (pipes, redirection, history, aliases, and so on) works with MH–you don't have to learn a new interface. Other mail agents have their own command interpreter for their individual mail commands (although the MUSH mail agent simulates a UNIX shell). Mail messages are stored in individual files.

A new version of MH that supports multi-part multi-media mail is also available.

4.2.2 Elm

At A Glance

```
Package Name: ELM
Latest Version: Elm 2.3 PL11
Developer/Author: H.P.
         Usenet Trust
         Syd Weinstein <syd@dsi.com>
Availability: comp.sources.unix vol. 22, issues 60-85
              Patches in vol. 24 issues 25-35
Sources: Yes
Platforms: UNIX, PC
Support: comp.mail.elm
Folder Format: mbox
Notable Features: Screen oriented. Good for novices.
```

Reviews:

- **From: "UNIX EMail Software - a Survey"**

  Elm is designed to run with `sendmail` or `/bin/rmail` (according to what's on your system) and is a full replacement of programs like `/bin/mail` and `mailx`. The system is more than just a single program, however, and includes programs like `frm` to list a 'table of contents' of your mail, `printmail` to quickly paginate mail files (to allow 'clean' printouts), and `autoreply`, a system wide daemon that can auto-answer mail for people while they're on vacation without having multiple copies spawned on the system.

  The most significant difference between Elm and most other mail systems is that Elm is screen-oriented. Upon further use, however, users will find that Elm is also quite a bit easier to use, and quite a bit more intelligent about sending mail and so on.

Elm is particularly good for novices. The only drawback that I've heard is that elm is a bit less user configurable than, say, MUSH.

4.2.3 Pine

At A Glance

```
Package Name: Pine
Latest Version: Pine 3.07
Developer/Author: Mike Siebel, Mark Crispin,
                  and Laurence Lundblade <pine@cac.washington.edu>
Availability: Through anonymous FTP from:
              ftp.cac.washington.edu:
              pub/mail/pine3.07.tar.Z

Sources: Openly available.
Platforms: UNIX
Support: Pine mailing list: send a request to
         pine-info-request@cac.washington.edu
Folder Format: mbox
Notable Features: Screen oriented. Good for novices.
```

Reviews:

- **From: "UNIX EMail Software - a Survey"**

  Pine is a mailer developed by the University of Washington Office of Computing and Communications. It has been designed for ease-of-use and with the novice computer user in mind. It is based on Internet mail protocols (e.g. RFC-822, SMTP, IMAP, and MIME) and currently runs on a variety of UNIX platforms.

  The guiding principles for achieving ease-of-use in Pine were: careful limitation of features, one-character mnemonic commands, always-present command menus, immediate user feedback, and high tolerance for user mistakes. It is intended that Pine can be learned by exploration rather than reading manuals.

A stand-alone version of Pico, Pine's message composition editor, is also included. It is a very simple and easy to use text editor with text justification and a spelling checker.

Features:

- Mail index showing a message summary which includes the status, sender, size, date and subject of messages.
- View and process mail with the following commands: forward, reply, save, export, print, delete, capture address and search.
- Address book for saving long complex addresses and personal distribution lists under a nickname.
- Multiple folders and folder management screen for filing messages.
- Message composer with easy-to-use editor and spelling checker. The message composer also assists entering and formatting addresses and provides direct access to the address book.
- Online help specific to each screen and context.
- Supports access to remote mail repositories via the IMAP2 protocol defined in RFC-1176.
- Soon to support multi-part mail conforming to proposed MIME Internet standard, allowing sending of sounds, graphics such as GIF and TIFF files, and binary files such as spreadsheets.

Pine, including source code, is freely available via anonymous FTP from ftp.cac.washington.edu on the Internet. Other provisions for distribution have not been made. From the Internet, you may try out Pine and leave comments by telneting to "demo.cac.washington.edu" and logging in as "pinedemo". To join the Pine mailing list for announcements send a request to "pine-info-request@cac.washington.edu".

Pine is very portable and runs on a variety of UNIX machines including DECstations, NeXTs, VAX's and Suns. Pine was originally based on Elm, but it has evolved much since, ("Pine Is No-longer Elm").

For further information send e-mail to pine@cac.washington.edu. Pine is the work of Mike Siebel, Mark Crispin, and Laurence Lundblade at the University of Washington.

Chapter 5

Proper use of Internet E-Mail Services

This chapter includes various types of information which enable users of Internet e-mail services to better enjoy and participate in the Internet community.

It is important to note that *proper* use as expressed in this chapter should not be viewed as universal; it simply is an attempt to express the Internet Community's perspective.

5.1 Acceptable Use Policy

When a user obtains connectivity to Internet through a network service provider, he/she must comply to the acceptable use policy of the network service provider. Often the acceptable use policy of the immediate network provider refers to other network's acceptable use policy. It is expected of network users to comply with policies of all service providers when using the network. In addition to the network service provider, the user can set and maintain an acceptable use policy of its own for its user community. Examples of acceptable use policies are included in this paper as an appendix. See section **??** [Examples of Acceptable Use Policy], page **??**.

In particular, transfer of commercial data on Internet is subject to some restrictions. These are discussed in the next section.

Commercial Traffic Over the Internet

The Internet's U.S. backbone, called NSFnet is funded by the National Science Foundation, NSF, which is a government agency. The NSF has an ''acceptable use'' policy that prohibits the transmission of commercial data over its backbone. However, many network service providers offer services for commercial data that bypass the NSFnet.

Uunet Technologies Inc., which operates Alternet, Performance Systems International Inc., which operates Psinet, General Atomics Inc., which operates Cerfnet, and other commercial network providers have set up an association called the Commercial Information Exchange (CIX). CIX provides an alternative commercial backbone. At present members of CIX carry each other's traffic on their network at no charge.

Advanced Network and Services Inc. (ANS), a nonprofit organization which manages NSFnet, has set up a subsidiary to use the spare capacity of NSFnet to carry commercial traffic.

Commercialization of Internet has started and worldwide commercial traffic on Internet is expected to grow.

5.2 Etiquette

Much of the information contained in this section has been derived from various articles posted on Usenet and available through out Internet.

This section is a guide to using Internet e-mail politely, effectively and efficiently. Communication by computer is new to many and there are certain aspects that can make it a frustrating experience. This document should help you avoid the worst traps.

This section summarize the Internet communities practices and are not recommendations on the part of EMA.

- **Use Descriptive Subject Headings**

  Always use the subject field. Messages with blank subject fields are less useful to the recipient. The subject line of a message is there to enable a person with a limited amount of time to decide whether or not to read your message immediately. Some sites truncate the length of the subject line to 40 characters so keep your subjects short and to the point.

- **Limit Line Length and Avoid Control Characters**

  Try to keep your text in a generic format. Many (if not most) of the people reading e-mail do so from 80 column terminals or from workstations with 80 column terminal windows. Try to keep your lines of text to less than 80 characters for optimal readability. If people quote part of your message in a follow up, short lines will probably show up better, too.

Also realize that there are many, many different forms of terminals in use. If you enter special control characters in your message, it may result in your message being unreadable on some terminal types; a character sequence that causes reverse video on your screen may result in a keyboard lock and graphics mode on someone else's terminal. You should also try to avoid the use of tabs, since they may also be interpreted differently on terminals other than your own.

- **Be Brief**

  Never say in ten words what you can say in fewer. Say it succinctly and it will have a greater impact. Remember that the longer you make your message, the fewer people will bother to read it.

- **Be Careful with Humor and Sarcasm**

  Without the voice inflections and body language of personal communications, it is easy for a remark meant to be funny to be misinterpreted. Subtle humor tends to get lost, so take steps to make sure that people realize you are trying to be funny. No matter how broad the humor or satire, it is safer to remind people that you are being funny.

  The Internet community has developed some symbols/icons for expressing emotions in electronic messages. This mechanism is often referred to as *smiley* or *emoticon*.

  To better introduce this concept, the following section has been reproduced from "The new Hacker's Dictionary", [**?**].

> emoticon /*ee-moh'ti-kon*/ n. An ASCII glyph used to indicate an emotional state in e-mail or news. Hundreds have been proposed, but only a few are in common use. These include:
>
> `:-)` 'smiley face' (for humor, laughter, friendliness, occasionally sarcasm)
> `:-(` 'frowny face' (for sadness, anger, or upset)
> `;-)` 'half-smiley' (**ha ha only serious**); also known as *semi-smiley* or *winkey face*.
> `:-/` 'wry face'
>
> (These may become more comprehensible if you tilt your head sideways, to the left.)
>
> The first 2 listed are by far the most frequently encountered. Hyphenless forms of them are common on CompuServe, GEnie, and BIX; see also **bixie**. On **USENET**, *smiley* is often used as a generic term synonymous with **emoticon**, as well as specifically for the happy-face emoticon.

It appears that the emoticon was invented by one Scott Fahlman on the CMU **bboard** systems around 1980. He later wrote: ‘‘I wish I had saved the original post, or at least recorded the date for posterity, but I had no idea that I was starting something that would soon pollute all the world's communication channels.” [GLS – Guy L. Steele Jr. – confirms that he remembers this original posting].

Note for the **newbie**: Overuse of the smiley is a mark of loserhood! More than one per paragraph is a fairly sure sign that you have gone over the line.

- **Summarize What You are Responding To**

  When you are responding to a message, summarize the parts of the message to which you are responding. This allows readers to appreciate your comments rather than trying to remember what the original message said.

  Summarization is best done by including appropriate quotes from the original message. Do not include the entire message if it is not relevant to the points you are making. Summarize only the major points you are discussing.

- **Cite Appropriate References**

  If you are using facts to support a cause, state where they came from. Don't take someone else's ideas and use them as your own.

- **Don't Overdo Signatures**

  Signatures are nice, and many people can have a signature added to their messages automatically. Don't overdo it. Signatures can tell the world something about you, but keep them short. A signature that is longer than the message itself is considered to be in bad taste. Every signature should include at least your return address relative to a major, known site on the network and a proper domain-format address.

5.3 Mailing Lists

The concept of an Internet Mailing List is very similar to X.400's Distribution List.

A *mailing list* can be thought of as an alias pointing to groups of users, which allow mail to be sent to the whole group at once. Mailing lists are often set up to carry certain topics.

Internet mailing lists often follow a convention for having a separate alias for administrative aspects of mailing list management. This convention consists of adding a `-request` to the local part of the mailing list address.

As an example, consider `mobile-ip@parc.xerox.com` mailing lists which discusses extensions to the Internet Protocol to accomodate mobility of hosts. In order to subscribe to this mailing list, one simply can send a message to `mobile-ip-request@parc.xerox.com` asking to be added to that mailing list.

It is good practice to keep a record of mailing lists to which you subscribe. Having that information comes very handy when you need to change your e-mail address.

5.4 Vacation

When for any reason you know that for an extended period of time you will be unavailable to respond to your incoming mail, it is good practice to let people know about your un-availability. Example of such situations are when you are on vacation or on extended business trips.

To address this requirement, the `vacation` program exists.

5.5 Citation

When responding to a message it is often good practice to include the original message to which you are responding. Often it is not necessary to include the entire message; summarizing the specific points to which you are responding are sufficient. It is important to properly attribute the original message. Many user agents facilitate citing and attributing of the original messages. For example, `supercite` (see section **??** [SuperCite], page **??**, for details) provides sophisticated citing and attributing of the original messages.

5.6 Signatures

It is customary to identify yourself at the end of the message and tell people something about yourself.

It is important to remember that the main purpose of a signature is to help people locate you. Long signatures are considered to be in bad taste. Keeping your signatures to less than four lines is generally good practice.

5.7 Junk Mail

Since commercial use of Internet has been quite limited, the problem of *Junk Mail* as we know it in paper mail has not been significant. Generally speaking unwanted e-mail is uncommon in Internet.

User agent feature/add-ons that sort out and prioritize incoming messages based on the Originator, Recipient, Subject and other information are used by some busy e-mail users.

5.8 Privacy

Security mechanisms were not designed into Internet e-mail. Traditional Internet e-mail systems lack security features for data confidentiality, authenticity and non-repudiation.

Security extensions to Internet e-mail that address these short comings exist. See section **??** [Privacy Enhanced E-Mail (PEM)], page **??**.

5.9 Usenet

Much of the information contained in this section has been derived from various articles posted on Usenet and available through out Internet. See, [E3], [**?**] for more complete information about Usenet.

Defining Usenet is difficult. It is almost impossible to generalize over all Usenet sites in any non-trivial way.

Usenet can be defined as the set of people who exchange articles tagged with one or more universally-recognized labels, called "newsgroups" (or "groups" for short). Usenet encompasses government agencies, large universities, high schools, businesses of all sizes, home computers of all descriptions, etc.

Gateways between mailing lists and newsgroups are very common. In this ‘‘shadowy” world of news-mail gateways, the line between Usenet and not-Usenet becomes very hard to draw. Furthermore, Usenet software and e-mail software often share many common features and can sometimes be used interchangeably.

Control through out Usenet is loose. No person or group has authority over Usenet as a whole. Every administrator controls his own site. No one has any real control over any site but his own. To help hold Usenet together, various articles are periodically posted in newsgroups in the "news" hierarchy. These articles are provided as a public service by various volunteers. Among the periodic postings are lists of active newsgroups, both "standard" (for lack of a better term) and "alternative."

Propagation of Usenet traffic is primarily through UUCP and NNTP.

In the old days, UUCP over long-distance dialup lines was the dominant means of article transmission.

In March 1986 a package was released implementing news transmission, posting, and reading using the Network News Transfer Protocol (NNTP) (as specified in RFC 977, [**?**]). This protocol allows hosts to exchange articles via TCP/IP connections rather than using the traditional uucp. It also permits users to read and post news from machines which cannot or choose not to install the USENET news software.

Chapter 6

Internet E-Mail Directories

The two basic e-mail Directory questions are:

1. I use e-mail network *MYNET*. How can I get to people on e-mail network *YOURNET*?

2. I know someone's name, and I think he/she might have an electronic mail address somewhere. How can I find it?

This chapter focuses on the above two questions.

6.1 Inter-Network Mail Addressing

Details of Inter-Network Mail Addressing, i.e., ways of getting from *MYNET* to *YOURNET*, can be found in a white paper by J.J. Chew entitled "Inter-Network Mail Guide", [E6]. Here we present the basic methodology, along with some examples.

E-mail Networks, in their simplest form consist of two things:

- A list of e-mail Network Identifiers; for instance `applelink, bitnet, compuserve, mci, sprint` ....

- A sequence of connection blocks describing how to get from one network to another. Each block looks something like:

#FROM:	mynet
#TO:	yournet
#RECIPIENT:	youraddress
#CONTACT:	contactaddress
#INSTRUCTIONS:	I send to 'youraddress@thegateway'

Consider the following list of EMA accounts on various e-mail networks.

EasyLink, Envoy 100, GEnie, MCI, Sprint:	**EMA**
AT&T Mail:	**!EMA**
CompuServe:	**70007,2377**
Dialcom:	**52:PRD003**
IBM:	**IBMMAIL(USEMA001)**

Given the above information about EMA addresses, we apply the information in the "Inter-Network Mail Guide" and obtain the following:

- From Internet to CompuServe

#FROM	internet
#TO	compuserve
#RECIPIENT	70007,2377
#INSTRUCT	send to '70007.2377@compuserve.com'

- From CompuServe to Internet

#FROM	compuserve
#TO	internet
#RECIPIENT	mohsen@neda.com
#INSTRUCT	send to %INTERNET:mohsen@neda.com

6.2 Finding E-Mail Addresses

There is no single e-mail directory for the Internet. However, if you know some person's name and domain chances are that with a bit of investigative work you can find their e-mail address. The information in this section can help you in the search.

6.2.1 Usenet-addresses server

If the person that you are looking for has posted anything to Usenet his/her name may be in the USENET address database on the machine `pit-manager.mit.edu`.

To get his/her e-mail address send mail to:

```
To: mail-server@pit-manager.mit.edu
Subject: send usenet-addresses/THE.NAME.GOES.HERE
```

6.2.2 WHOIS Service

The WHOIS database contains information on registered network users, hosts, mailboxes, and organized groups of people to name a few. The database information for each of these categories is gathered via registration with the Network Information Center (NIC) at SRI International. The official specification of the WHOIS service, as well as the policy for network user registration is provided in RFC 954 [?]. Another commonly used name for this service is "NICNAME".

The following example shows how a typical user can access the whois service.

```
whois help
nicname -h
```

to get a help message. The WHOIS and NICNAME programs will check the database for the given names. For example,

```
nicname "name"
whois "name"
whois -h "host" "names"  (e.g., whois -h stanford.edu ''names'')
```

The following hosts are known to run the whois server:

- nic.ddn.mil
- rs.internic.net
- stanford.edu

6.2.3 Finger

The official specification of the FINGER service is provided in RFC 742, [?]. This RFC provides an interface to the finger programs at remote sites.

"finger @domain" provides a list of all online users at the site named "domain". "finger user@domain" provides about user named "user". Security ramifications of running a finger server should be considered before making this service available.

For example,[1] the White House has a secure finger server.

```
rostam-4>date
Mon Jun  7 19:38:47 PDT 1993
rostam-5>finger clinton@whitehouse.gov
[whitehouse.gov]

    Finger service for arbitrary addresses on whitehouse.gov is not
supported. If you wish to send electronic mail, valid addresses are
"PRESIDENT@WHITEHOUSE.GOV", and "VICE.PRESIDENT@WHITEHOUSE.GOV".
```

It is possible to provide access to an organization's phone book through finger. One such example is the MIT finger server.

MIT Finger

The MIT finger facility is a prototype for an MIT-wide online directory service. To use it, simply type "finger lastName@mit.edu". To get instructions for usage, type "finger help@mit.edu".

6.2.4 UUCP Maps

UUCP maps can often be consulted for finding the administrative contact of UUCP connected domains. The following example illustrates the type of information that can be obtained using the UUCP maps.

[1] All examples in this paper are real.

```
rostam-194>uuhosts icxn.com

UUCP mail path from rostam to icxn.com:
icxn.com        icxn!\%s

UUCP mail information for host icxn.com
#Name                   icxn, .icxn.com, icxn.com
#F      uunet.uu.net
#System-CPU-OS          AST 386C; Interactive 386/ix
#Organization           InterConnections, Inc.
#Contact                Kevin J. Dunlap
#Electronic-Address     kjd@icxn.icxn.com
#Telephone              +1 206 881 5773
#Postal-Address         14711 NE 29th Bellevue, WA  98007
#Latitude-Longitude     47 36 N / 122 58 W city
#Remarks
#Written-by             kjd@icxn.icxn.com (Kevin J. Dunlap);
#USENET seaeast.wa.com
#
icxn    .icxn.com(LOCAL)
icxn=   icxn.com
icxn    uunet(DEMAND), neda(DIRECT)
```

6.2.5 Mail Addresses through WAIS

WAIS stands for Wide Area Information Servers, and is an architecture for a distributed information retrieval system. WAIS is based on the client server model of computation, and allows users of computers to share information using a common computer-to-computer protocol.

Many of the sources of information for finding a person's electronic mail address are available through WAIS. These include the Network Information Center (NIC) `whois` data base, the UUCP Maps and Usenet address server.

One advantage of WAIS is that it provides a unified interface to many types of information sources.

In addition to use of WAIS for finding e-mail addresses, many use it for the purpose of indexing their e-mail archives. WAIS is very useful for searching loosely structured text such as e-mail messages.

6.2.6 QUIPU

The Directory is one of OSI's newer standards. As such there are few implementations available. However, one implementation is openly available, QUIPU, as a part of ISO Development Environment. See [?] and [?] for more information. Several white pages pilot projects have been based on QUIPU.

The community of interworking QUIPU implementations providing a Directory Service is quite large.

6.2.7 College E-Mail Addresses

If you are trying to find e-mail addresses for undergraduate and graduate students, faculty and staff at various colleges and universities, then the "College E-Mail Addresses" electronic paper, [E4], is the best starting point. This paper is periodically updated.

An example entry in the College e-mail directory is:

```
Univ. of California/Irvine:
        CS Undergrads: bonnie.ics.uci.edu
        CS Grads/Staff/Faculty: ics.uci.edu
        EE: orion.uci.edu
        Userid usually FLLLLLLL (8 char max); sometimes FFFFFFFL.
        Faculty userids are usually their last names.
```

Appendix A

Internet Growth (1981-1991)

The following information has been reproduced from RFC-1296, [?].

```
N.1 Number of Internet Hosts
   The chart below shows the number of IP hosts on the Internet.  These
   are hosts with at least one IP address assigned.  Data was collected
   by ZONE except where noted.  The following two sections are graphs of
   the data in this chart.

        Date              Hosts
        08/81               213          Host table #152
        05/82               235          Host table #166
        08/83               562          Host table #300
        10/84             1,024          Host table #392
        10/85             1,961          Host table #485
        02/86             2,308          Host table #515
        11/86             5,089
        12/87            28,174
        07/88            33,000
        10/88            56,000
        01/89            80,000
        07/89           130,000
        10/89           159,000
        10/90           313,000
        01/91           376,000
        07/91           535,000
        10/91           617,000
        01/92           727,000
```

```
                        Number of Internet Hosts (logarithmic)


    | 1000000
    |                                                                *.*
    |                                                        ..*.*..*
    |                                                     ...
    | 100000                                          ..**
    |                                              *.*
  H |                                          ...*
  o |                                        .*
  s | 10000                                ..
  t |                                    ..
  s |                               ....*
    |                         ...*.*
1000|                   ...*..
    |                ...
    |            ...*
    | ..*....*...
 100|.
     -------------------------------------------------------------------
     8     8     8     8     8     8     8     8     8     9     9     9
     1     2     3     4     5     6     7     8     9     0     1     2
                                    Date

     "*"  = data point,  "." = estimate
```

Appendix B

Obtaining Request For Comments

The following information has been reproduced from:
  '`anonymous@ftp.nisc.sri.com:rfc/rfc-index.txt`'.

```
Many RFCs are available online; if not, this is indicated by (Not
online in the index of RFCs).  Paper copies of all RFCs are available
from SRI, either individually or on a subscription basis (for more
information contact nisc@nisc.sri.com or call 1-415-859-6387).  Online
copies are available via FTP from ftp.nisc.sri.com as rfc/rfc####.txt
or rfc/rfc####.ps (#### is the RFC number without leading zeroes).

Additionally, RFCs may be requested through electronic mail from SRI's
automated mail server by sending a message to mail-server@nisc.sri.com.
In the body of the message, indicate the RFC to be sent, e.g. "send rfcNNNN"
where NNNN is the number of the RFC.  For PostScript RFCs, specify
the extension, e.g. "send rfcNNNN.ps".  Multiple requests can be sent
in a single message by specifying each request on a separate line.
The RFC Index can be requested by typing "send rfc-index".
```

Appendix C

Examples of Acceptable Use Policy

C.1 NSFNET

```
        THE NSFNET BACKBONE SERVICES ACCEPTABLE USE POLICY

GENERAL PRINCIPLE:

(1) NSFNET Backbone services are provided to support open research and
    education in and among US research and instructional institutions,
    plus research arms of for-profit firms when engaged in open
    scholarly communication and research.  Use for other purposes is
    not acceptable.

SPECIFICALLY ACCEPTABLE USES:

(2) Communication with foreign researchers and educators in connection
    with research or instruction, as long as any network that the
    foreign user employs for such communication provides reciprocal
    access to US researchers and educators.

(3) Communication and exchange for professional development, to
    maintain currency, or to debate issues in a field or subfield of
    knowledge.

(4) Use for disciplinary-society, university-association,
```

government-advisory, or standards activities related to the user's research and instructional activities.

(5) Use in applying for or administering grants or contracts for research or instruction, but not for other fundraising or public relations activities.

(6) Any other administrative communications or activities in direct support of research and instruction.

(7) Announcements of new products or services for use in research or instruction, but not advertising of any kind.

(8) Any traffic originating from a network of another member agency of the Federal Networking Council if the traffic meets the acceptable use policy of that agency.

(9) Communication incidental to otherwise acceptable use, except for illegal or specifically unacceptable use.

UNACCEPTABLE USES:

(10) Use for for-profit activities (consulting for pay, sales or administration of campus stores, sale of tickets to sports events, and so on) or use by for-profit institutions unless covered by the General Principle or as a specifically acceptable use.

(11) Extensive use for private or personal business.

This statement applies to use of the the NSFNET Backbone only. NSF expects that connecting networks will formulate their own use policies. The NSF Division of Networking and Communications Research and Infrastructure will resolve any questions about this Policy or its interpretation.

2/92

C.2 PSI, Inc.

These statements represent a guide to the acceptable use of PSI facilities.

PSI, Inc. recognizes as acceptable all forms of data communications across its network, except where Federally subsidized facilities of transport (such as national backbones like NSFNET) may require limitations, or State/Federal/ International law prohibits. In those cases, use of the network should adhere to the general principle of advancing science and education through interexchange of information among research and educational institutions except where specific policy statements exist (see for example, next section).

In cases where data communications are addressed to recipients outside of the PSI supplied network and are carried over other regional networks or the Internet, PSI users are advised that acceptable use policies of those other networks apply and may; in fact; limit use.

If PSI is informed by Federal/State/International authorities of inappropriate or illegal use of PSI facilities and/or other networks these determinations will be binding on the Customer.

C.3 User Establishment of an Acceptable Use Policy

Note: The following is just an example and neither the author nor the EMA recommend these policies.

```
A typical company could interpret an acceptable use policy for employee
notification in the following manner:

                 Notice to company Employees regarding
                    the use of the Public Internet

General:

Various networks that participate in the Internet have different
rules or "acceptable use policies" governing the nature of traffic
over their facilities.  In general, the policies require at a minimum
that traffic:
```

```
Is legal
Is not harassing
Does not disrupt others' use of the network
Is not advertising

These requirements are consistent with applicable company policies
and company use must honor them.

Specific:

Our public E-mail carrier's connection to the Internet is to a
non-commercial network that connects to services provided by the U.S.
National Science Foundation.  Its acceptable use policy is based on
the NSF's policy, which states in part:


"NSFNET Backbone services are provided to support open research and
education in and among US research and instructional institutions, plus
research arms of for-profit firms when engaged in open scholarly
communication and research.  Use for other purposes is not acceptable."

In order to meet the NSF policy, E-mail directed to an Internet address
must meet the following criteria:

Not permitted:

1. Communicating with other commercial corporations, such as
   computer vendors, software retailers, and independent consultants,
   unless the communication is related to education, research, operation
   of the internet, or other purposes consistent with the basic policy above.

2. Communicating from one company site or person to another company site or
   person.

3. Personal, commercial, or profit-oriented activities.

Permitted:

1. Communication with U.S. government agencies
```

2. Communications with standards bodies and professional societies

3. Communication with universities, their professors, research staffs, and students, unless it is for research or consulting private to the company.

4. Arranging for university staff to visit the company for consulting purposes in connection with research.

5. Use of university and not-for-profit research centers for company computing, when there is an agreement for research between the university and the company.

6. Correspondence with individual university students and faculty in support of recruiting.

If you are uncertain about the propriety of use of the Internet for a specific correspondent, consult your local computer and network security advisor.

Electronic References

[**E1**] Chris Lewis <clewis@ferret.ocunix.on.ca>, ‘‘UNIX EMail Software - A Survey”, Posted periodically to news.answers newsgroup.

[**E2**] Kames H. Thompson <jimmy_t@verifone.com>, ‘‘UUCP Mail, News and Gateways Software for PCs and MACs”, Posted periodically to news.answers newsgroup.

[**E3**] Chuq Von Rospach, ‘‘A Primer on How to Work With the Usenet Community”, Posted periodically to news.answers newsgroup.

[**E4**] Mark Kantrowitz <mkant+@cs.cmu.edu>, ‘‘FAQ: College Email Addresses”, Posted periodically to news.answers newsgroup.

[**E5**] Jonathan Kamens <jik@mit.edu>, ‘‘FAQ: How to find people’s E-mail addresses”, Posted periodically to news.answers newsgroup.

[**E6**] John J. Chew <poslfit@gpu.utcs.utoronto.ca>, ‘‘Mail Guide”, Posted periodically to news.answers newsgroup.

Concept Index