

From: David M. Ihnat  
Douglas H. Price

Date: 13 July 1988

## An Overview of the Chicago Unix Electronic Mail Environment

## ABSTRACT

Recent decisions by a large corporate supplier of cooperative electronic mail services has precipitated a potential crisis in the reliable and timely delivery of electronic mail in the Chicago area. This paper seeks to describe the existing UNIX operating system electronic mail environment, the evolving corporate network provision climate, and its ramifications for users of electronic mail in the Chicago area. Some possible solutions are outlined as a departure point for further discussion.

## 1. INTRODUCTION

Recent decisions by a large corporate provider of cooperative electronic mail services has precipitated a potential crisis in the reliable and timely delivery of electronic mail in the Chicago area. This paper seeks to describe the existing UNIX<sup>1</sup> electronic mail environment, the evolving corporate network provision climate, and its ramifications for users of electronic mail in the Chicago area. Some possible solutions are outlined as a departure point for further discussion.

## 2. WHAT IS EMAIL?

Electronic Mail, (hereafter referred to as *email*), is a facility that permits individuals (and automatic processes, in some systems) to send messages to one another that are stored until the receiver is ready to read them. These messages normally contain textual data, such as memoranda or letters. Some systems may also permit the inclusion of binary data, digitized voice, images, document files, and other specialized content.

<sup>1</sup> UNIX is a registered trademark of AT&T.

Copyright (C) 1988 by David M. Ihnat and Douglas H. Price,  
all rights reserved.

This version of the document is provided as a benefit to the members of /usr/group/chicago chapter, and may be reproduced only in its entirety, including this copyright notice.

The user can create and send such messages to a single or multiple receivers.

The receiver of such a message may retrieve these messages, and reply to and/or delete these message at will. The community of individuals using email may have some assurance that the receiver will eventually see their messages. This avoids the need for "telephone tag", notes lost on scraps of paper, and many of the missed opportunities involved in both inter- and intra-office communication. Users may log their communications activity, sorting and filing their messages and the responses in a compact, efficient manner. This also permits the computer backup mechanisms to protect their correspondence along with their work data.

Email is also speedy. While company mail or the U.S. Mail service might take days to deliver paper correspondence, email can deliver it in hours, if not seconds.

In the computing and pure science research communities, network email has become a virtual necessity due to the geographically wide spread of cooperating individuals engaged on large projects. In fact, any large project requiring the exchange of large amounts of text is invariably enhanced by the available of email. This can become particularly evident after email is generally available. The amount of physical paper that must be handled, mailed, sorted, duplicated and exchanged, along with the associated delays this engenders, is sharply reduced. All in all, general availability of email services enhances productivity, and reduces a number of communications problems that many organizations (especially large ones) experience.

### **3. THE EXISTING NETWORK**

Many individuals and companies have become dependent on email for both personal and professional reasons. It has become as necessary to many as a pocket pager, cellular telephone, telephone answering machine, or a private secretary. It has the advantage of supplying a reliable method for the relatively quick delivery of important messages and moderate amounts of machine readable data to users both in the Chicago area, and around the world.

For many businesses, it is an important source of contact with their user communities, and permits a certain amount of low-volume, non-proprietary contact with other vendors, whether they be competitors or suppliers.

### **3.1 HISTORY**

While not absolutely required for its functioning, UNIX email service has grown up in a cooperative, egalitarian environment. This environment has consisted of a large number of major corporations (including AT&T), UNIX system vendors of every size, government entities, and universities, along with a smaller number of personal computer sites, tacitly supporting the forwarding of email from site to site, without regard to the source and destination of the message.

Large, well-networked (and/or well-heeled) organizations provided the long haul links, freely allowing foreign as well as internal data. This service is provided on a cooperative basis, in which it is assumed that the price for passing other people's data is that they will pass yours. Smaller sites provided the local connectivity and shorter haul links in a local area. Many of the large organizations provided gateways, that bridge email from one network to another. Several major networks are concatenated in this way into a single super email network that spans the planet, encompassing tens of thousands of machines and hundreds of thousands of users.

#### **3.1.1 Users**

The distribution and power of UNIX-based email service was strongly shaped by its original user communities. The basically informal growth of the resulting network is its greatest strength, enforcing a required flexibility that has permitted it to grow and change with the changing needs and capabilities of its user communities. Mail generated by even the earliest UNIX system may still be processed by modern UNIX email systems. Simplicity of message structure permits the creation of small and verifiable interfaces to foreign email systems.

##### **3.1.1.1 Business**

The first organization to make use of UNIX-based email was, of course, AT&T Bell Laboratories. The first documented use of UUCP inside of AT&T occurred around 1978.

Originally a curiosity, the capability expanded with the distribution of the research tool, UNIX Version 6. At this time, AT&T was a fully regulated monopoly. Since AT&T was not permitted to sell software, it gave copies of the UNIX system to educational institutions for a nominal charge. This gave rise to the next big population of UNIX email users, the universities.

### 3.1.1.2 Educational

The next group of users to make use of UNIX and its email capabilities were the educational institutions that received copies of the UNIX operating system as a teaching and research tool. Email permitted easy communication with the AT&T UNIX developers in Murray Hill, New Jersey, and served as a departure point for research efforts. These research efforts, and the associated expansion of the growing email network were greatly assisted by the next set of users of email.

### 3.1.1.3 Government

In 1969, the Defense Advanced Research Project Agency (DARPA) in conjunction with the Department of Defense, military contractors and contracted universities formed the ARPANET. This research network served a two-fold purpose:

- 1) It provided a testbed for research in computer networks in a wide area environment.
- 2) It permitted the various organizations, and their individual researchers involved in this research to cooperate and converse electronically.

A major portion of the long-haul, backbone email network still functions via the ARPANET. There are legal and practical limitations on this traffic, as will be seen later in this paper. Also, a large impediment to the continued expansion of this particular network are the requirements of high speed, dedicated links and a special communications processor per location (IMP).

An important product of this work was a heterogeneous machine environment mail standard, that eventually became the basis of the electronic mail supernet (more on this later).

### 3.1.1.4 The Second Wave Business Users

The obvious success of the ARPANET government, university, and military contractor network prompted many students leaving the universities to push for these capabilities in other vendor's products. Also, about this time, AT&T had released the UNIX system for sale by value added resellers (VARs). These VARs manufactured a variety of host computers, all running the UNIX system, and all having the UUCP-based email capability available for use by a purchaser. This unified user environment had the effect of enforcing a uniform set of expectations on users.

## 3.2 THE CHICAGO AREA EMAIL USERS

At this moment, approximately 85 machines are registered in the USENET Maps<sup>1</sup> for the Chicago area (see below). Most of the machines on the list belong to "small" businesses, where small is defined as having less than 3 UNIX machines exchanging email with the rest of the net. Others belong to larger corporations, and are *corporate gateways* that hide a much larger number of machines behind them inside the walls of the corporate office. Appendix I contains the names of the machines currently represented in the maps, and their owning organizations.

### 3.2.1 USENET and Email

USENET (also referred to as Netnews) is a network bulletin board system that has paralleled the growth of the greater email network. Dozens of machines in the Chicago area make use of it. A great number of organizations receive it, including business sites, educational sites and UNIX public access bulletin board machines.

At the time of this writing, there are over 400 subject categories under active discussion constituting approximately 3 megabytes a day of traffic. Why is this

---

<sup>1</sup> The USENET maps represent a collection of machine identity and connectivity information used to create routing tables for electronic mail, among other uses. This routing function will be discussed later.

significant? Well, a large number of email messages are generated as responses to articles posted on USENET. The traffic flow of USENET follows the major regional links into and out of the Chicago area. The amount of individual email traffic unrelated to USENET passing into and out of the Chicago area is almost trivial compared to the continuous load that USENET itself represents.

## **4. THE STRUCTURE**

While many computer systems have some form of email, most have not until recently had network capability, and have been limited to a single machine. The email network we are dealing with, on the other hand, has been traditionally related the UNIX operating system and its derivatives, but now spans many operating systems, including VAX-VMS<sup>1</sup>, MS-DOS<sup>2</sup> and other systems. This basic network mail capability, originally intended to only span alike systems, now has been *gateways* between operating system types forming a large super email network (referred to as an *internet*)

### **4.1 THE EMAIL INTERNET**

As noted earlier, electronic mail has been gatewayed between dissimilar operating systems. Software translates from a standard email form into one compatible with the native operating system mail tools. The critical component that makes this possible are network mail standards.

#### **4.1.1 Email Standard Structure**

##### **4.1.1.1 ARPANET Standards**

The ARPANET standard RFC822<sup>3</sup> specifies a physical structure for a email message on the ARPA Internet. A message

<sup>1</sup>VAX and VMS are trademarks of Digital Equipment Corporation.

<sup>2</sup>MS-DOS is a trademark of Microsoft Corporation.

<sup>3</sup>I "RFC 822: Standards for the Format of ARPA Internet Text Messages", D. H. Croker, Department of Electrical

consists of a *header* followed by a blank line, followed by a text *body*. The header must be encoded in ASCII<sup>1</sup>. This simple structure, always beginning with a line of the format:

```
>From <user-address> <date-string>
```

is in general easy to parse. This is also the form that the native UNIX system mailers expect. RFC822 compliant and other existing mailers may have other capabilities that are embodied in additional header lines. These may include such things as a subject line, the date the email was created, message identifiers, trace messages etc.

#### **4.1.1.2 ARPANET and UUCP MAILNET Standard Addressing**

There are two forms of network addressing that are of interest to us. The older is the UUCP MAILNET address (referred to as a *bangist-style route*), and the more recent is ARPANET RFC882<sup>2</sup>.

##### **4.1.1.2.1 Routes**

Bangist routing is the oldest form of UNIX system network email, and is supported directly by the UNIX system mail software that comes with virtually every UNIX system release. It consists of an explicit list of machine names, forming a route, separated by exclamation points ('!', referred to as *bangs*), terminating in the login name of the recipient on the target machine. For example:

```
mach-a!mach-b!dest-mach!user
```

Engineering, University of Delaware, Newark, DE 19711.,  
August 1982.

<sup>1</sup> The American Standard Code for Information Interchange. This is the native character set of the UNIX operating system and MS-DOS microcomputers.

Individual machine names are normally characterized by a flat namespace of no more than 7 characters in significance<sup>1</sup>.

.P

The advantages of this scheme are its implementation simplicity and ease of use in small networks. From a network administrative point of view, knowledge of the network structure is localized to each individual machine in the chain.

The primary disadvantage is that the user must discover and in future remember a usable path to the target machine. While tools have been provided to make this scheme more palatable, it is nevertheless fraught with occasional routing breakdowns. On the other hand, it makes minimal demand on central administration in moderately sized networks, and requires no special coordination between cooperating machines. Machine names are also **relative** in that more than one machine on the same network can have the same name, it is the route to the machine which determines the correct destination. This can lead to occasional confusion.

#### 4.1.1.2.2 Domains

The ARPANET domain naming scheme is, on the other hand, is an absolute, hierarchical namespace. It is characterized by domains of naming authority, with a single central name authority called the ARPANET Network Information Center (NIC), currently maintained by SRI International under contract to DARPA. The NIC delegates and maintains a full list of naming authorities. These authorities, in turn have the authority to issue names, or further divide their allocated namespace into *subdomains*, each in turn able to subdivide still further. Here are some examples of existing domain host names:

ihnp4.ATT.COM  
A.CS.UIUC.EDU

---

<sup>1</sup> This is due to a limitation that AT&T incorporated in UUCP. Names of greater than 7 characters may cause file transfers to go to the wrong destination, or fail altogether.

sun.Sun.Com  
tis.llnl.gov  
packard.garage.nj.att.com  
vfrot.CHI.IL.US

It can be seen that a great amount of flexibility is available in naming, and that there is no possibility of a name collision, given the hierarchical nature of the name.

On the negative side, the names can get rather long, and read from a "funny" direction. Also, in order for the scheme to work, a hierarchy of naming authorities must exist and cooperate with one another. The hierarchical pattern of the namespace and authority does push the decision making out towards the leaves of the tree, but the authority structure is still required.

There are additional benefits, however. Since there is an authority, and that professionally managed authority collects and distributes information about all of the domains to each of the subordinate authorities, definitive understandings about the "best" route to a particular destination can be assumed in the construction of the email network. Even if the local host has no direct knowledge of the existence of the target host, the domain name gives the local host a deterministic method for either contacting the target directly, or routing the message in the correct direction. This is done via *domain name servers* and *domain forwarders*.

## **4.2 EMAIL SOFTWARE CAPABILITIES**

The following discussion will concern the easiest and most popular way of linking UNIX systems together.

## **4.3 UUCP**

The operating systems we have discussed deliver network email via many networking media, but for the moment, we will look at the use of the UUCP (Unix to Unix Copy) utility. This program is provided with virtually every copy of the UNIX operating system in the field, and represents a common denominator of networking.

### **4.3.1 UUCP Capabilities**

UUCP is basically a file transfer mechanism. It permits the transfer of data from one UUCP equipped machine to another. It also permits remote execution of restricted list of commands. This is the capability important to email<sup>1</sup>.

#### **4.3.1.1 Local Area Networks**

UUCP is capable of transferring email via local area networks between UNIX hosts in an individual location. This is often the user's first introduction to network email. Networks commonly used include broadcast networks such as Ethernet and direct-connect RS-232 links between machines.

#### **4.3.1.2 Dialup Circuits**

UUCP can use a number of network media, but all versions of this program in the field can make use of asynchronous modems and telephone lines. Common network speeds used include 1200 baud (100 bytes per second) and 2400 baud (200 bytes per second). Some 300 baud links exist. Modems with speeds higher than 2400 baud are becoming popular.

It is this dialup capability that excites our interest, since this represents a cost-effective tool for the exchange of email between organizations, rather than just internally.

## **4.4 PROPERTIES**

---

<sup>1</sup> Recently, two packages, UULINK from VORTEX Technologies, and a public domain package called UUPC had become available. These packages permit MS-DOS compatible computers to engage in UUCP conversations with true UNIX hosts, and include the basic email software to permit these machines to participate in the email internet. This development has the potential to sharply increase the number of UNIX-style email participant machines.

It is important to understand what capabilities email has, and also what it cannot do.

#### **4.4.1 Permissible Content**

The majority of email systems in UNIX machines have certain limitations. Email systems based on other operating systems may have other limitations. Since the email internet consists of a number of email networks concatenated together, one must assume the limitations of all of them, if we are to pass mail through a number of machines.

- A. Some email systems limit single messages to less than 65,000 bytes, other systems are limited to 32,000 bytes. This limitation may either be due to restrictions introduced by system administrators, or may be due to signed integer problems on 16 bit machines. In any case, it is often thought to be antisocial to ship large amounts of data via the email network.
- B. Most email systems are currently limited to textual data only. Some email systems permit binary data in email, but the standards are currently in flux. In any case, there is no guarantee that every mail system in the route will be able to handle binary data.
- C. Some email systems possess extra facilities, such as delivery receipts, route probes, etc. There is no guarantee that any facility, aside from the delivery of the message itself is a provided function of the net. Sites along the route may strip the message header of the information necessary to activate the function on the destination machine, as well.

#### **4.4.2 Security and Authentication**

It is important to note that there are no standards for the protection of messages that pass through the net. They are sent in clear text, and any administrator of a machine along the route has the ability (though it is frowned upon)

to look at all messages flowing through the system. Some end-user email programs have encryption capabilities, however. Most are constructed to permit an encrypted message to pass transparently through foreign mail systems, so no difficulty should be encountered using such an email program.

Messages may also be forged. Since the UNIX system (and the email network in general) does not have any real authentication system for messages, it is possible for a normal user anywhere in the network to create and inject a message into the system, with no reasonable possibility of discovering the originator.

#### **4.4.3 Error Reporting**

One area that is of serious concern when dealing with network email is not only the ability to deliver the mail, but also to tell when it isn't being delivered. This may happen due to a variety of causes, such as a particular node being temporarily or permanently down; an inaccurate path being specified for a USENET mailing; a non-existent destination user; or an inappropriately maintained or corrupted database for a domain nameserver. In any event, the problem of how to notify the sender that a failure has occurred is not uniformly addressed. In most implementations of UUCP, either the entire message, or at least the header, is returned via the inverse of the path by which it arrived. Often this may take the form of a warning that contact with some intermediate node, or the terminal destination, has not been achieved for some period of time, but that attempts will continue for a some interval. Eventually, the request will be terminated. Similarly, failures due to unknown intermediate nodes in a UUCP path, or a terminal machine for UUCP or domain systems, will usually result in a single failure return along the delivery path. Finally, an unknown user will be reported by the destination node itself.

#### **4.4.4 Postmaster**

All network addressing, whether a bangist-style UUCP address, or a domain-style node/user combination, ultimately requires a destination user, or login ID. In

cases where such an individual destination is unknown, a convention was developed (first on ARPANET, and adopted by BSD UNIX) that a dummy user called *postmaster* would be available to receive mail of a general nature--i.e., queries for contact information, broadcasts of general interest to network managers, etc. Unfortunately, this convention has never been adopted for UNIX systems as provided by AT&T; thus, mail directed to 'postmaster' on an AT&T-derived destination node may well be 'bounced', or fail. For this reason, if attempting to contact a node where an individual recipient is unknown, often the best approach is to direct the message to both 'postmaster' and 'root'. (The greatest disadvantage of mailing to root, of course, is that this login often receives an amazing amount of 'junk' mail concerning system functions; often so much that many administrators become lax in screening mail to root.)

#### **4.5 COMMERCIAL EMAIL SERVICES**

The value of inter- and intra-machine email has not been lost on the commercial marketplace. Several services have existed for some time that exclusively provide email services to business and individuals--MCIMAIL, SPRINT MAIL, and ATTMAIL are prime examples. In addition, general-purpose public computer services such as Compuserve have provided email capabilities as an integral part of their offerings. An important item to note is that, while these services originally evolved with no provision for gateway networking, the increasing importance of this type of interconnectivity is now leading them to examine methods of joining the rest of the networks in this manner.

Finally, while most connections to date have consisted of donated/negotiated reciprocal links between cooperating machines, a commercial UUCP node--**uunet** --has, in the last year or so, been successfully providing this service on a commercial basis for a fee.

### **5. GETTING CONNECTED**

In order to connect into the email internet, Two things are required: The physical ability to send/receive email and/or data files; and the visibility to guarantee that senders can find the node. The easiest way to obtain the first is using UUCP.

## **5.1 CONNECTIONS TO NEIGHBORS**

Simply enough, to be able to physically receive and send UUCP traffic, all that is needed is one cooperating neighbor which has links to enough other machines, or at least another well-connected machine. This typically consists of exchanging telephone numbers, node names, physical link characteristics (i.e., 1200/2400 baud) and a login/password combination. However, such a link must be negotiated--often, this simply consists of the agreement that in return for forwarding mail to or through your node, you will be willing to forward mail routed to or through their node. Also, both nodes owe each other the security of treating the connection information as confidential. Finally, once the link becomes a known, functioning path, there is an implied obligation to maintain its functional status--mail 'constipation' can be the result of failing to do so. Of course, any issue may be negotiated with the neighboring site, such as being a polled-only site (phone calls initiated in only one direction), or not advertising the existence of the link to the net-at-large; but this is on a case-by-case basis.

## **5.2 THE USENET MAP ENTRY**

Provided that the new node would like the world to know of its existence and connectivity, the next step is to submit a USENET map entry to the UUCP Map Project. This is a project of the USENIX Association, wherein voluntary submissions are collected and redistributed such that any interested site may incorporate the information in a local routing database. The current instructions and an example form for the map entry are appended to this document.

## 6. THE CRISIS

Now, the problem. Notice that, as stated previously, the costs associated with providing inter-machine communications have been borne on a voluntary, non-charging basis. Unfortunately, the cost of providing these services has grown--both in terms of telephone charges, and administrative overhead engendered by the growing volume of mail due to Netnews and the proliferation of small, private machines ('fuzzball' sites). Many corporations, universities, and individuals are now closely examining these formerly accepted expenses. One of the companies which have decided to reduce and/or eliminate formerly provided participation in such email internet cooperation is AT&T. While certainly not the only company which has decided to take such steps, due to its former heavy support and provision of backbone sites and links, the AT&T decision has the potential to greatly disrupt the network connectivity across the country, as well as specifically in the Chicago region.

### 6.1 THE AT&T DECISION: A SUMMARY

To briefly summarize, the amount of UUCP mail traffic flowing through AT&T machines which is simply pass-through mail (that is, not having an AT&T machine as either the source or destination), has been determined to be excessive. This decision was reached after examining the issue in light of company resource usage, and corporate goals. To reduce this usage in an orderly manner, in a short while all such pass-through mail will result in a warning being sent to the originator, noting the impending cutoff. Finally, all such pass-through mail will be bounced (rejected) back to the message originator. The tentative date for the email cutoff is Sept. 1, 1988. Netnews traffic will continue to be redistributed.

However, there's no guarantee that this will continue to necessarily be the case--they're not required to do so, after all. In addition, the opinion has been expressed by at least two other companies that, if AT&T won't handle pass-through for them, they won't handle passthrough for AT&T. Taken to its logical conclusion, this can result in

a crippling of the connectivity required to keep the network viable.

## **6.2 THE CONCLUSION**

This isn't the end of the email network--and that's the purpose of this paper, and the meeting at which it is being summarized. In September, there will be a one-day symposium, hosted jointly by the Chicago chapters of /usr/group and the IEEE Computer Society, to address exactly this issue. There are a number of potential remedies ranging from wider distribution of the load, to utilization of commercial services such as *uunet*, to wider implementation of the currently-experimental creation of regional domains and domain parks. It isn't the purpose here to attempt to present all the possible solutions, or even to probe the problem in great depth. Rather, we hope to have provided a solid basis of understanding such that the attendees may both understand the issues and topics being presented, and actively participate in a detailed problem definition and resolution through the September seminar.

--dmi        David M. Ihnat  
--dhp        Douglas H. Price

## Appendix I

### Chicago Area Email Users

simon	
111a	AT&T Communications
abtcser	Abbott Labs.
aocgl	Alpha Omega Consulting Group, LTD
asiux1	Ameritech Applied Technologies, Inc.
aicchi	Analysts International Corp., Chicago Branch
apcichi	Apollo Computer Inc., Chicago Regional & Branch Offices
ncsel	Apple Computer Inc.
acich	Automated Concepts, Inc. (Chicago)
stox	Autorad, Inc.
bartek1	Bartek Associates
beattres	Beatrice Company
buadm	Bradley University, Admissions Department
cyber	Bradley University, Computer Center
buea3b2	Bradley University, E.E. Department
bluefl	Bradley University, English Department
butelcom	Bradley University, Telecommunications
bradley	Bradley University, Text Processing
bu-jrn	Bradley University, Text Processing
crt	CRT
csdev	Challenger Software
chinet	Chinet - Public Access Unix
poot	Chinet - Public Access Unix
cxhq	Computer X Inc.
cxintg	Computer X Inc.
datamen	Datamension Corporation
dcrig1	Defense Contract Administration Services Region Chicago
edgesa	Edge Systems, Inc.
extel	Extel, Incorporated
vpnet	GPS Microdata
htc-chi	Hull Trading Co.
icom	Icom Systems, Inc.
hcfeams	Illinois Bell ARSB Support
ilunix	Illinois Bell ARSB Support
gndctl	Institute For Learning and Teaching of Mathematics (Univ. of IL @ Chicago)
invest	Investment Futures, Inc.
jpusal	J.P.U.S.A.
rowell	John Rowell & Associates
katar	Katar Innovations
laidbak	Lachman Associates, Inc. (LAI)
luccpud	Loyola University of Chicago
wlbngr	Loyola University of Chicago
ddsw1	Macro Computer Solutions, Inc.
mcs1	Macro Computer Solutions, Inc.

illusion	Magic Numbers Software
n8ino	Mainstream
masschgo	Masscomp
ilmss	Media Software & Systems, Inc.
amtfocus	Motorola, Inc., General Systems Group, Advanced Manufacturing Technology
mcdchg	Motorola, Microcomputer Division
nucsrl	Northwestern University
obdient	Obedient Software Corporation
sam	Ogburn Stouffer Center, University of Chicago
will	Ogburn Stouffer Center, University of Chicago
plxchi	PLEXUS Computers, Inc.
vfrot	Private Machine
homebru	Private Machine
pyrchi	Pyramid Technology Corp. - Central Region
richpl	Rich Inc.
riccb	Rockwell Telecommunications, Inc.
ohare	Salse Force Companies, Inc.
limerick	Private machine
indep1	Software Architects
indep2	Software Architects
sunbird	Sun Microsystems Inc.
tarkus	Private machine
tellabs5	Tellabs, Inc.
spl	The Software Public Library, Inc.
spl1	The Software Public Library, Inc.
sol1	The Solution
uniq	Uniq Digital Technologies, Inc.
crsp	University of Chicago
sphinx	University of Chicago Computation Center
gargoyle	University of Chicago, Computer Science Department
anubis	University of Chicago, Computer Science Department
bourbaki	University of Chicago, Mathematics Department
oddjob	University of Chicago; Deparment of Astronomy and Astrophysics
paideia	University of Chicago; Education Department
uicbert	University of Illinois at Chicago; Dept of E.E./C.S.
chaos	University of Oklahoma High Energy Physics
gorgon	Unorganized
fineart	WFMT-FM Chicago's Finearts Radio
vijit	Wang Laboratories, Inc.
wheaton	Wheaton College
harper	William Rainey Harper College
harpervm	William Rainey Harper College
jolnet	jolnet, Public Access UNIX, Joliet, IL

## **Appendix II**

### **UUCP Map Registry**

#### Appendix II

The UUCP map is posted to newsgroup comp.mail.maps.

From rn, the map can be easily unpacked with a command such as

```
43-46w | (cd ~uucp/uumap ; sh)
```

or you can use John Quarterman's script to automatically unpack the files.

All files intended as pathalias input begin with "d." and "u.", thus

```
pathalias Path.* uumap/[du].*
```

is a useful command to run. (You supply Path.\* with local additions.)

The map is also available on a demand basis at a number of hosts who have volunteered to make their copy available to the general public; details of this access are posted separately in file "network".

The files are organized by country, using the ISO 3166 3 letter country code for each country. Each file has a name like u.iso.r1.r2.s, where "iso" is the country code, r1, r2, etc are regions and subregions (e.g. states in the USA, provinces in Canada, etc.) and s is a sequence number (usually 1, but sometimes 2, 3, and up may be provided to keep individual files down to a reasonable size, thus, u.usa.ca is separated into two regions: [135] for southern, [246] for northern.) In a few cases where very large companies post their maps, separate files are used.

For instance, \*.b.\* is the Bellcore file.

The map contains two types of files: u.\* and d.\* files. The d.\* files are for domains registered in the UUCP Zone. The u.\* files are for UUCP hosts that do not have officially registered domains, but rather belong to the unofficial ".UUCP domain". Membership in the UUCP Zone allows organizations and individuals to register official, unique, domain names, recognized by all major academic computing networks worldwide. For more information about joining the UUCP Zone, send electronic mail to the UUCP Project at one of the addresses

```
uucp-query@stargate.com
{uiucdcs,att}!stargate!uucp-query
att!stargate!uucp-query@seismo.css.gov
```

or, if you cannot send electronic mail, telephone

+1 213 868 1134

We strongly encourage you to send email if at all possible, since it cuts down on telephone tag and is much more efficient our volunteer workforce.

This map can be used to generate mail routes with pathalias. Pathalias was posted to Usenet in January 1986 and will be posted again as needed. The map is also useful to determine the person to contact when a problem arises, and to find someone for a new site to connect to.

Please check the entry for your host (and any neighbors for whom you know the information and have the time) for correctness and completeness. Please send corrections and additional information to uucpmap@rutgers.UUCP or rutgers!uucpmap or uucpmap@rutgers.edu

This map is maintained by a group of volunteers, making up part of the UUCP Project. These people devote many hours of their own time to helping out the UUCP community by keeping this map up to date. The volunteers include:

Jeff Janock - jeff@necntc.nec.com

USA: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont

Nicholas (Nike) Horton - horton@reed.uucp

USA: Delaware, Maryland, New Jersey, Virginia, Washington DC, West Virginia

Rayan Zachariassen rayan@ai.toronto.edu

CANADA: All provinces

Bill Blue - bblue@crash.uucp

USA: Arizona, California (Southern half)

Erik Fair - nca-maps@ucbvax.berkeley.edu

USA: California (Northern half)

David Schmidt - davids@iscuva.iscs.com

USA: Alaska, Idaho, Oregon, Montana, Washington, Wyoming

Doug McCallum - dougm@ico.isc.com

USA: Arkansas, Colorado, Iowa, Kansas, Louisiana, Mississippi, Nebraska, New Mexico, Oklahoma, Texas, Utah

Piet Beertema - Europe (piet@cwi.nl)  
Europe: all countries (unless otherwise noted)

Mikel Manitius - map-request@codas.att.com  
USA: Florida

Jeff Lee - jeff@ics.gatech.edu  
USA: Georgia

Bill Welch - zaiaz32!uucpmap@zaiaz.UUCP  
USA: Alabama, South Carolina

Tim Thompson - tgt@stargate.com  
USA: Illinois, Indiana, North Carolina, North Dakota, Ohio,  
Pennsylvania, South Dakota, Tennessee, Wisconsin

Bob Leffler - bob@rel.eds.com  
USA: Michigan

Rob Robertson - rob@philiabs.philips.com  
USA: New York

Hokey - hokey@plus5.com  
USA: Missouri

David Herron - david@e.ms.uky.edu  
USA: Kentucky

Brian Richter - brianr@rosevax.rosemount.com  
USA: Minnesota

Mark Horton - mark@stargate  
ATT: all regions

Torben Nielson, Bob Cunningham - torben@uhmanoa.UUCP,  
bob@uhmanoa.UUCP  
USA: Hawaii

Haesoon Cho - nmc@sorak.kaist.ac.kr  
Korea: all regions

Tohru Asami - asami@kddspeech.kddlabs.jp  
Japan: all regions

Robert Elz, Dave Davey - map-coord@munnari.UUCP  
Australia: all regions

Larry Harrison - [larryh@pyuxe.UUCP](mailto:larryh@pyuxe.UUCP)

Bell Communicates Research (Bellcore): all sections

Paul Graham - pjg@unrvax.unr.edu  
USA: Nevada

Benny Somali - somali@indovax.uucp  
Indonesia: all regi

Mel Pleasant - pleasant@rutgers.edu  
Singapore: all regions  
New Zealand: all regions

Please note that the purpose of this map is to make routers within UUCP work. The eventual direction is to make the map smaller (through the use of domains), not larger. As such, sites with lots of local machines connected together are encouraged to create a few gateway machines and to make arrangements that these gateways can forward mail to your local users. We would prefer not to have information listing the machines on your local area networks, and certainly not your personal computers and workstations. If you need such information for local mail delivery, create a supplement in pathalias form which you do not publish, but which you combine with the published data when you run pathalias. We also do not want information about machines which are not on UUCP, that is, which are not reachable with the ! notation from the main UUCP cluster.

If you don't have pathalias, it has been posted to mod.sources most recently in January 1986. If you don't have access to a mod.sources archive, contact the mod.sources moderator (currently Rich \$alz, {rayssd,mit-eddie}!mirror!sources-request or sources-request@mirror.tmc.com)

The remainder of this file describes the format of the UUCP map data. It was written July 9, 1985 by Erik E. Fair <ucbvax!fair>, and last updated July 12, 1985 by Mark Horton <stargate!mark>.

The entire map is intended to be processed by pathalias, a program that generates UUCP routes from this data. All lines beginning in '#' are comment lines to pathalias, however the UUCP Project has defined a set of these comment lines to have specific format so that a complete database could be built.

The generic form of these lines is

#<field id letter><tab><field data>

Each host has an entry in the following format. The entry should begin with the #N line, end with a blank line after the pathalias

data, and not contain any other blank lines, since there are ed, sed, and awk scripts that use expressions like /<sup>#N \$1/,/^\$/ for the purpose of separating the map out into files, each containing one site entry.</sup>

```

#N UUCP name of site
#S manufacturer machine model; operating system & version
#O organization name
#C contact person's name
#E contact person's electronic mail address
#T contact person's telephone number
#P organization's address
#L latitude / longitude
#R remarks
#U netnews neighbors
#W who last edited the entry ; date edited
#
sitename .domain
sitename remotel(FREQUENCY), remote2(FREQUENCY),
      remote3(FREQUENCY)

```

Example of a completed entry:

```

#N ucbvax
#S DEC VAX-11/750; 4.3 BSD UNIX
#O University of California at Berkeley
#C Robert W. Henry
#E ucbvax!postmaster
#T +1 415 642 1024
#P 573 Evans Hall, Berkeley, CA 94720
#L 37 52 29 N / 122 13 44 W
#R This is also UCB-VAX.BERKELEY.EDU [10.2.0.78] on the
internet
#U decvax ibmpa ucsfcgl ucbscad
#W ucbvax!fair (Erik E. Fair); Sat Jun 22 03:35:16 PDT 1985
#
ucbvax .ucbvax.Berkeley.EDU
ucbvax decvax(DAILY/4), ihnp4(DAILY/2),
      sun(POLLED)

```

#### Specific Field Descriptions

#N system name

Your system's UUCP name should go here. Either the uname(1) command from System III or System V UNIX; or the uname(1) command from Version7 UNIX will tell you what UUCP is using for the local UUCP name.

One of the goals of the UUCP Project is to keep duplicate UUCP host names from appearing because there exist mailers in the world which assume that the UUCP name space contains no duplicates (and attempts UUCP path optimization on that basis), and it's just plain confusing to have two different sites with the same name.

At present, the most severe restriction on UUCP names is that the name must be unique somewhere in the first six characters, because of a poor software design decision made by AT&T for the System V release of UNIX.

This does not mean that your site name has to be six characters or less in length. Just unique within that length.

With regard to choosing system names, HARRIS'S LAMENT:

"All the good ones are taken."

#S machine type; operating system

This is a quick description of your equipment. Machine type should be manufacturer and model, and after a semi-colon(;), the operating system name and version number (if you have it). Some examples:

```
DEC PDP-11/70; 2.9 BSD UNIX
DEC PDP-11/45; ULTRIX-11
DEC VAX-11/780; VMS 4.0
SUN 2/150; 4.2 BSD UNIX
Pyramid 90x; OSx 2.1
CoData 3300; Version 7 UniPlus+
Callan Unistar 200; System V UniPlus+
IBM PC/XT; Coherent
Intel 386; XENIX 3.0
CRDS Universe 68; UNOS
```

#O organization name

This should be the full name of your organization, squeezed to fit inside 80 columns as necessary. Don't be afraid to abbreviate where the abbreviation would be clear to the entire world (say a famous institution like MIT or CERN), but beware of duplication (In USC the C could be either California or Carolina).

#C contact person

This should be the full name (or names, separated by commas) of the person responsible for handling queries from the outside world about your machine.

#E contact person's electronic address

This should be just a machine name, and a user name, like 'ucbvax!fair'. It should not be a full path, since we will be able to generate a path to the given address from the data you're giving us. There is no problem with the machine name not being the same as the #N field (i.e. the contact 'lives' on another machine at your site).

Also, it's a good idea to give a generic address or alias (if your mail system is capable of providing aliases) like 'usenet' or 'postmaster', so that if the contact person leaves the institution or is re-assigned to other duties, he doesn't keep getting mail about the system. In a perfect world, people would send notice to the UUCP Project, but in practice, they don't, so the data does get out of date. If you give a generic address you can easily change it to point at the appropriate person.

Multiple electronic addresses should be separated by commas, and all of them should be specified in the manner described above.

#T contact person's telephone number

Format: +<country code><space><area code><space><prefix><space><number>

Example:

#T +1 415 642 1024

This is the international format for the representation of phone numbers. The country code for the United States of America (and Canada) is 1. Other country codes should be listed in your telephone book.

If you must list an extension (i.e. what to ask the receptionist for, if not the name of the contact person), list it after the main phone number with an 'x' in front of it to distinguish it from the rest of the phone number.

Example:

#T +1 415 549 3854 x37

Multiple phone numbers should be separated by commas, and all of them should be completely specified as described above to prevent confusion.

#P organization's address

This field should be one line filled with whatever else anyone would need after the contact person's name, and your organization's name (given in other fields above), to mail you something by paper mail.

#L        latitude and longitude

This should be in the following format:

#L DD MM [SS] "N"|"S" / DDD MM [SS] "E"|"W" ["city"]

Two fields, with optional third.

First number is Latitude in degrees (NN), minutes (MM), and seconds (SS), and a N or S to indicate North or South of the Equator.

A Slash Separator.

Second number is Longitude in degrees (DDD), minutes (MM), and seconds (SS), and a E or W to indicate East or West of the Prime Meridian in Greenwich, England.

Seconds are optional, but it is worth noting that the more accurate you are, the more accurate maps we can make of the network (including blow-ups of various high density areas, like New Jersey, or the San Francisco Bay Area).

If you give the coordinates for your city (i.e. without fudging for where you are relative to that), add the word 'city' at the end of the end of the specification, to indicate that. If you know where you are relative to a given coordinate for which you have longitude and latitude data, then the following fudge factors can be useful:

1 degree =	69.2 miles	=	111 kilometers
1 minute =	1.15 miles	=	1.86 kilometers
1 second =	102 feet	=	30.9 meters

For LONGITUDE, multiply the above numbers by the cosine of your latitude. For instance, at latitude 35 degrees, a degree of longitude is  $69.2 * 0.819 = 56.7$  miles; at latitude 40 degrees, it is  $69.2 * 0.766 = 53.0$  miles. If you don't see why the measure of longitude depends on your latitude, just think of a globe, with all those N-S meridians of longitude converging on the poles.

You don't do this cosine multiplication for LATITUDE.

Here is a short cosine table in case you don't have a trig calculator handy. (But you can always write a short program in C. The cosine function in bc(1) doesn't seem to work as documented.)

deg    cos    deg    cos    deg    cos    deg    cos    deg    cos    deg    cos

0	1.000	5	0.996	10	0.985	15	0.966	20	0.940	25	0.906
30	0.866	35	0.819	40	0.766	45	0.707	50	0.643	55	0.574
60	0.500	65	0.423	70	0.342	75	0.259	80	0.174	85	0.087

The Prime Meridian is through Greenwich, England, and longitudes run from 180 degrees West of Greenwich to 180 East. Latitudes run from 90 degrees North of the Equator to 90 degrees South.

#R        remarks

This is for one line of comment. As noted before, all lines beginning with a '#' character are comment lines, so if you need more than one line to tell us something about your site, do so between the end of the map data (the #? fields) and the pathalias data.

#U    netnews neighbors

The USENET is the network that moves netnews around, specifically, mod.announce. If you send mod.announce to any of your UUCP neighbors, list their names here, delimited by spaces. Example:

#U    ihnp4 decvax mcvax seismo

Since some places have lots of USENET neighbors, continuation lines should be just another #U and more site names.

#W        who last edited the entry and when

This field should contain an email address, a name in parentheses, followed by a semi-colon, and the output of the date program.

Example:

#W    ucbvax!fair (Erik E. Fair); Sat Jun 22 03:35:16 PDT 1985

The same rules for email address that apply in the contact's email address apply here also. (i.e. only one system name, and user name). It is intended that this field be used for automatic aging of the map entries so that we can do more automated checking and updating of the entire map. See getdate(3) from the netnews source for other acceptable date formats.

PATHALIAS DATA (or, documenting your UUCP connections & frequency of use)

The DEMAND, DAILY, etc., entries represent imaginary connect costs (see below) used by pathalias to calculate lowest cost paths. The cost breakdown is:

LOCAL        25        local area network

DEDICATED	95	high speed dedicated
DIRECT	200	local call
DEMAND	300	normal call (long distance,
anytime)		
HOURLY	500	hourly poll
EVENING	1800	time restricted call
DAILY	5000	daily poll
WEEKLY	30000	irregular poll
DEAD		a very high number - not usable path

Additionally, HIGH and LOW (used like DAILY+HIGH) are -5 and +5 respectively, for baud-rate or quality bonuses/penalties. Arithmetic expressions can be used, however, you should be aware that the results are often counter-intuitive (e.g. (DAILY\*4) means every 4 days, not 4 times a day). This is because the numbers represent "cost of connection" rather than "frequency of connection."

The numbers are intended to represent cost of transferring mail over the link, measured very roughly in elapsed time, which seems to be far more important than baud rates for this type of traffic. There is an assumed high overhead for each hop; thus, HOURLY is far more than DAILY/24.

There are a few other cost names that sometimes appear in the map. Some are synonyms for the preferred names above (e.g. POLLED is assumed to mean overnight and is taken to be the same as DAILY), some are obsolete (e.g. the letters A through F, which are letter grades for connections.) It is not acceptable to make up new names or spellings (pathalias gets very upset when people do that...).

#### LOCAL AREA NETWORKS

We do not want local area network information in the published map. If you want to put your LAN in your local Path.\* files, read about the LAN syntax in the pathalias.1 manual page.

#### WHAT TO DO WITH THIS STUFF

Once you have finished constructing your pathalias entry, mail it off to {rutgers,ucbvax}!stargate!uucpmap, which will be sent to the appropriate regional map coordinator. They maintain assigned geographic sections of the map, and the entire map is posted on a rolling basis in the USENET newsgroups mod.map.uucp over the course of a month (at the end of the month they start over).

Questions or comments about this specification should also be directed at stargate!uucpmap.

REFERENCES

1. RFC822: Standards for the Format of ARPA Internet Text Messages, D. H. Croker, Department of Electrical Engineering, University of Delaware, Newark, DE 19711., August 1982.
2. RFC882: Domain Names-Concepts and Facilities, P. Mockapetris, ARPANET Working Group Requests for Comment; No. 882, SRI International, Menlo Park, CA., Nov 1983.