

Addressing and Related Issues in Electronic Mail Systems

John C. Klensin*
Lorry Scura+

Introduction

The worldwide "electronic mail network" is actually a collection of networks that do not have identical conventions for addresses or the way in which these are written. In several respects, the situation is not unlike that in the obtaining and use of food composition data. With the latter, one must first find the location of the data, then find out how to obtain it, then somehow translate its conventions, definitions, and coding and nomenclature methods into one's own. The good news about electronic mail addressing is that much more standardization-- both about the terminology used in description and the ways of accomplishing the translations-- than the situation we have with food composition data. Indeed, in many cases, a common-appearing address syntax is used, and transitions of electronic mail from one environment to another are completely transparent to the users. Nonetheless, some knowledge of alternate addressing models is useful, if only to facilitate intelligent asking for help when problems arise or unusual situations are encountered.

Different Host Systems, Different Styles

In addition to actual addressing issues, "external" electronic mail has been developed on most systems as an extension of whatever "internal" electronic mail or interpersonal note facilities were available on those systems already. Some of these extensions have not been very neat because the addressing forms needed in "wide area" heterogeneous networks have been different from those previously developed for use in local systems. Again, in parallel with issues in the design of software for dietary analysis, there has been a tension between trying to make the interface look (to the user) similar to other, familiar, software in the local environment and trying to make network interface software

look the same across all hosts on the network. The former usually wins out: the major advantages of the latter are convenience for the few of us who use the same networks from many different machines and the even fewer of us who try to write general addressing or electronic mail use guidelines or tutorials.

Two examples may illustrate the problems encountered by system designers, and passed on to users, when general network addressing is added to existing local electronic mail facilities.

- Some systems have historically been designed on the assumption that some few characters (typically about eight) are sufficient for representing the names of users and host machines. For a variety of reasons (a few discussed below), naming systems on the long-distance networks tend to require longer names, and more structure, than can be accommodated in two eight-character fields.
- Some systems "know", based on local conventions, what characters can appear in host names and people's names or other addresses. Often these assumptions exclude characters such as period ("."), at-sign ("@"), semicolon (";"), or slash ("/"), which are heavily used in various national and international network addressing systems.

The combination of these two problems results in some odd-appearing local syntax and other arrangements: the need to send mail to intermediate (e.g., eight-character or restricted syntax) addresses with internal instructions about the "real" address or to use quoted strings or special processing instructions to permit processing of locally-unacceptable addresses. Fortunately, over time, enough host operating system work seems to be going on that the number and

complexity of these "work arounds" that are required is decreasing. For example, many VAX/VMS systems use a syntax that is illustrated in some of the slides for our other paper. There, Internet or BITNET addresses must be prefaced by, e.g., 'IN%' and followed by another double-quote, to avoid conflicts with local parsing conventions. This is typical of what is done to work around local systems and their conventions.

That said, these arrangements, whether odd typing conventions or intermediate addressing is required, are just a nuisance: they need to be understood and worked out with local support personnel and then observed as a ritual without spending much further energy worrying about just what they mean.

The Networks and Their Connections

Within the US, there are five major network arrangements that can be used for electronic mail. These are interconnected, although some of the connections are easier to use than others. Hosts and other arrangements that provide connections and translations between networks are known as "gateways". Most of the rest of the world that has network connectivity is linked to one or more of these, sometimes under different names and often with less complicated arrangements. In no particular order, they are:

BITNET: An academic and research network that primarily connects to academic institutions although some government components and private research organizations are also members. "Native" BITNET addresses are of the form `username@hostname`, with the user and host names each being limited to eight characters. Many BITNET hosts are converting, or have converted, to Internet (see below) naming conventions, hence "UCBCMSA" (sometimes, by convention, called "UCBCMSA.BITNET") and "CSMA.BERKELEY.EDU" are the same machine and can be addressed either way. From the Internet, BITNET users are addressed either in `username@domain` form, where "domain" is the Internet name for the host, or using the form `user%host.bitnet@gateway`. When no better gateway is known, `cunyvm.cuny.edu` is typically used.

BITNET users at most hosts send mail to Internet addresses simply by specifying the Internet address; the exceptions are becoming less and less useful as hosts. As Internet connections become more widespread, it is possible that BITNET service will be phased out entirely during the next few years. BITNET itself is run by an organization called CREN, which is US-based. A European network called EARN, one in

Canada called NETNorth, and a few smaller ones in other parts of the world are administratively separate from, but technologically part of, BITNET.

Internet: This is actually a collection of networks running the same protocols and sufficiently interconnected so, to the user, they appear to be a single network. Its components include the so-called "NSFNet regional networks" (that terminology may be obsolete by the time you read this) such as BARRNet, CERFNet, NYSERNet, JvNCNet, NEARNet, and so on. It evolved from the old ARPANET and is expected to evolve gradually into a very high performance "National Research and Education Network" (NREN). It is practical to connect much smaller machines (e.g., individual workstations or personal computers) to the Internet than has historically been the case with BITNET, so the number of hosts has become very large, with thousands of them at some universities. Although rules and arrangements differ from one region to the next, the Internet serves a superset of the institutions that would be accepted for BITNET membership. Internet addressing takes the form `user@domain`, where either the user name or the domain name may be quite long. Some Internet sites actually use full personal names, such that `John_C_Klensin@MIT.EDU` would be a plausible address. Domain names use a hierarchical structured model to identify a particular host. For example, in the address `Lorry_SY@nutmeg.hnrc.tufts.edu`, "nutmeg" is a host at the HNRC, which is part of Tufts, which is an EDUCational institution. Internet addresses outside the US typically use the two-letter country abbreviation as the last component, e.g., `tubvm.cs.tu-berlin.de` identifies a machine associated with the Technical University of Berlin in Germany. For academic and research network purposes, the Internet may be considered the "spine" through which most other connections flow.

UUCP and FIDONET: These two names refer to protocol arrangements that have been used to create networks connecting UNIX (tm) machines and things that imitate them and IBM PC-like desktop computers and things that imitate them, respectively. They, especially the latter, have been excellent solutions for low-cost, minimum fuss connections for individuals and small companies. Both networks have become extensive, and provide the major or only connections to some countries. Both user Internet-style addressing, and gateways to other networks that use this style of addressing are typically invisible to the end-user.

Commercial providers: There are several commercial providers of electronic mail services. Despite using "network" terminology, most of them operate with a single tightly-connected cluster of hosts and, except for connections to other networks, essentially accept and deliver mail on the same machine. The best-known among these are probably ATTMAIL, CompuServe, EasyLink, MCIMail, and various services offered by BT (e.g., DIALCOM, CGNET) and US Sprint (e.g., SprintMail) (these are all trademarks of their respective owners). All of these have connections to the Internet, usually using standard Internet addressing with some prefix that designates the network. In general, these commercial providers will sell electronic mail services (and other computer, teleconferencing, and bulletin board services) to anyone who is willing to pay the bills; several of the services are remarkably inexpensive for relatively light use. The much-advertised GENIE and Prodigy do not offer external mail connections at this time. Many of these systems are beginning to use so-called X.400 addressing: insofar as the Internet's domain names represent a single hierarchy, X.400 is multi-hierarchical, with separate hierarchies for various organizational arrangements and types of identification. There is, as yet, no Standard way to write an X.400 address; the syntax differs from one system to another.

Getting Attachment Help

The best advice is to find an expert and ask. If the advice you get seems to contradict good sense or what appears above, get a second opinion. Most academic institutions in the US are connected to either BITNET, the Internet, or both, although these networks are often thought of as the property of computer and physical scientists until someone else asks. Individual practitioners are likely to find one of the commercial services attractive, especially if there are few computer resources in-house. Most of those services offer a mix of databases, fax and telex arrangements, and user support services in addition to their electronic mail and comparison shopping is sensible. UUCP and Fidonet connections can be very attractive and inexpensive once established, but can be difficult to set up initially unless one has done so before.

Electronic Mail and the Nutrient Composition Community

In discussing why one would want to use these facilities, we pointed out that topic-specific electronic mailing lists may be much more useful for some purposes than person-to-person electronic mail (or

postal mail, telephone, or fax). Electronic mailing lists are especially useful for asking a question of, or discussing issues among, a group of people. INFOODS has created such a list for the nutrient composition community, paralleling the European WHO list for Nutritional Epidemiology. The instructions for subscription to the two lists (one may, of course, subscribe to either or both) once you have electronic mail access operating are:

Nutritional Epidemiology List

Send a mail message whose text body contains the line
SUB NUTEPI your full name

The mail message should be addressed to:

From BITNET or EARN: LISTSERV@db0tui11

From other networks, use the Internet address form:
LISTSERV@tubvm.cs.tu-berlin.de

Once you are on the list, administrative requests (e.g., subscription and subscription-cancellations) should be sent to the [LISTSERV](mailto:LISTSERV@db0tui11), as shown above. Material to be posted to other list subscribers should be sent to

NUTEPI@db0tui11 or
NUTEPI@tubvm.cs.tu-berlin.de respectively.

Food Composition Discussion List

Send a mail message requesting a subscription to:
food-comp-request@infoods.mit.edu

This mail is read by a person, so the message need not be in a specific form. The address is the same from BITNET, EARN, the Internet, etc.

Once you are on the list, administrative requests (e.g., subscription and subscription-cancellations) should be sent to the "-request" address, as shown above. Material to be posted to other list subscribers should be sent to
food-comp@infoods.mit.edu

References and Useful Addresses

The following information was accurate as of the time of the conference. Some of this information changes very rapidly and, in particular, it has been claimed that any description of networks and connectivity is always obsolete by the time it can be published.

Frey, Donnalyn and Rick Adams. !%@:: *A Directory of Electronic Mail Addresses*. O'Reilly and Associates, Petaluma, CA: 1989.

Quarterman, John S. *The Matrix, Computer Networks and Conferencing Systems Worldwide*. Digital Press, 1990.

The vendors of commercial electronic mail systems are usually listed in local "white pages" telephone books. The following information may be useful for those that often are not:

CompuServe starter kits are available from most computer dealers. Telephone enrollment and more information is available at 800-

MCIMail registers new users and provides additional information by telephone at 800-444-6245 (202-833-8484 in the Washington, DC area).

SprintNet (SprintMail, Telemail) customer service and registration information is available at 800-336-0437.

*John C. Klensin
INFOODS Secretariat, Room N52-457, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139.

*Lorry Scura
USDA Human Nutrition Research Center, Tufts University, 711 Washington St., Boston, MA 02111.