NOVEMBER 1992

NSL
Technical Note TN-10

MECCA: A Message-Enabled Communication and Information System

Anita Borg

digital **Network Systems Laboratory** 250 University Avenue Palo Alto, California 94301 USA

The Network Systems Laboratory (NSL), begun in August 1989, is a research laboratory devoted to components and tools for building and managing real-world networks. Current activity is primarily focused on TCP/IP internetworks and issues of heterogeneous networks. NSL also offers training and consulting for other groups within Digital.

NSL is also a focal point for operating Digital's internal IP research network (CRAnet) and the DECWRL gateway. Ongoing experience with practical network operations provides an important basis for research on network management. NSL's involvement with network operations also provides a test bed for new tools and network components.

We publish the results of our work in a variety of journals, conferences, research reports, and technical notes. This document is a technical note. We use this form for rapid distribution of technical material. Usually this represents research in progress. Research reports are normally accounts of completed research and may include material from earlier technical notes.

Research reports and technical notes may be ordered from us. You may mail your order to:

Technical Report Distribution
Digital Equipment Corporation
Network Systems Laboratory - WRL-1
250 University Avenue
Palo Alto, California 94301 USA

Reports and notes may also be ordered by electronic mail. Use one of the following addresses:

Digital E-net:	`DECWRL::NSL-TECHREPORTS`
Internet:	`NSL-Techreports@pa.dec.com`
UUCP:	`decwrl!nsl-techreports`

To obtain more details on ordering by electronic mail, send a message to one of these addresses with the word ''`help`'' in the Subject line; you will receive detailed instructions.

MECCA: A Message Enabled Communication and Information System

Anita Borg

November, 1992

Abstract

This note describes the design principles, functionality and prototype of the MECCA communication and information system. MECCA provides automatic administration of a membership-based electronic mail community.

digital **Network Systems Laboratory** 250 University Avenue Palo Alto, California 94301 USA

MECCA

Table of Contents

MECCA

Introduction

MECCA is a system currently under development at Digital's Network Systems Laboratory in Palo Alto, California. A beta-test with 50 members currently exists. Two larger installations, with 1500 and 250 members, will be in place by late February 1993.

The goals of the MECCA design are simplicity, generality and flexibility. MECCA provides a flexible set of tools for managing electronic mail (email) communication and accessing information via email. It gets information easily to the the right people based on who, where or what they are or on their interests.

The difference between MECCA and all other email distribution systems is that it provides automatic administration of a membership-based email community. Membership is defined as the existence of an entry in a database. The concept of membership implies the existence of policies for acceptance as a member and the provision of security against intrusion by non-members. We separate the notion of policy from that of mechanism by providing mechanisms for instituting policies rather than defining what those policies should be.

This note describes the functionality of a full MECCA system as well as the prototype currently under beta test. Section 1 is an overview of the design and functionality provided by an instance of MECCA. Section 2 describes the prototype system. Section 3 gives a more detailed description of the system's functionality.

1. Summary of Core Functionality

The system is designed to run on a single central machine that is accessible by email from all users of the system. Messages to be sent through the system are handled as they come in. Updates to the data base occur nightly (or at some regular time). The users of the system need *no special hardware or software*. They need only have access to a system which can get email to and from the central machine. The users of a single MECCA system can all be running different software on different types of hardware.

1.1. Core Services

The heart of a MECCA system is a data base of profiles of its users maintained on the central machine. Each profile is provided, maintained, and can be modified, via email, by the user it describes or by the system administrator. Using this data base, MECCA provides five related services. Currently, requests for different services are distinguished by addresses to which they are sent.

- **Administrative Services:** Handles requests to add users, change profile entries, suggest changes to the system or administrative policies, file bug reports, and ask for help or general information about the current data base contents.
- **Direct Mail Service:** Allows messages to be directed to subsets of the members based on information about the potential recipients contained in their profiles. A line in the sender's message specifies the attributes of users to whom she wishes the message sent.

- **List Creation Service:** Requests information about the users of the system. Like direct mail, the person sending the message specifies the attributes of users to be located. She also specifies which information about those users she would like to receive. This can be used to locate individual users, create mailing lists, perform statistical studies and so forth. A user can specify different levels of privacy for individual pieces of information in her profile allowing some information to be used only for message routing and not for list creation.

- **Publication Services:** Published messages contain descriptive information about their contents. Members specify in their profiles what subjects are currently of interest. Messages are forwarded only to interested members. This is essentially a subscription service. The subjects are arbitrary and can be used to establish threads of discussions that a user may choose to follow.

- **Retrieval Service:** Allows the retrieval of all published messages and some direct mail messages from an archive.

In addition, by modifying the contents of her profile, a user may:

- Specify that some mail is to be summarized on a daily or weekly basis and other mail is to be sent in full.

- Turn on or off the receipt of either direct or published mail (*e.g.* while on vacation).

- Request that some or all fields in the profile that is used to route direct mail may not be returned as the value of a list creation request.

- Have mail delivered either as it arrives or in batched collections (full copies of all messages on a subject concatenated into a single message).

1.2. Flexibility

MECCA is designed to be extremely flexible. As mentioned above, mechanism and policy are separated. For example different policies for membership and security can be implemented for each instance of MECCA. In addition, the data base used to represent membership can be built as a part of a MECCA installation, or the email portion of MECCA can reference an external data base. Finally, the user interface can be changed. The current interface, a query language described below, is only one possibility.

1.3. Extensibility

The system is designed to be flexibly extensible. The mechanism used to extend services is the data base profile. A profile in the data base serves two purposes. It gives the person it represents the right to send messages through the system and it allows that person to selectively receive messages going through the system. In fact, a profile can represent a person or it can represent a program. The program it represents can provide an arbitrary message based service to the members of the data base. It is logically irrelevant whether the program executes locally or on some distant machine. A wide range of services can be provided this way.

For example, the archiver will be implemented as an extended service rather than as part of the core of MECCA. It will be a separate program represented by a profile. The profile will specify that the archiver subscribes to all published mail and any mail sent explicitly to it. Any number of different archive services could be made available.

Existing external services can be accessed via MECCA. For example, one might provide access to an on-line news service, representing that service with a profile. The profile would admit wire service-generated messages into the system. Each individual member could describe in their profile the news areas that are of interest.

Another possible kind of extension is one in which a profile represents a program which performs some operation on messages going through the system. For example, a profile could represent a program which collects statistics about some or all of the messages going through the system and them periodically generates reports that are available to members.

2. The Prototype Implementation

The current working prototype implementation includes all administrative services, direct mail, published mail. List creation, archives, and summarization are to be added soon. It is under beta test with approximately 50 users. The first two installations, with 250 and 1500 users, will take place in late February, 1993.

The prototype system runs on top of Ultrix and implements its own simple data base rather than using a commercial data base system. The choice of query language, as well as the membership and security policies are those described in the next section and are specific to the particular instance of MECCA.

3. A more detailed look at functionality

This section takes a closer look at each function. We use sample internet style address of the form service@company.com to represent the five addresses to which the request types are sent.

3.1. Administrative Functions

A message sent to admin@company.com could initiate any of the following services depending on the content of the `Subject:` line in the message.

Subject:	Service performed:
--------	------------------
subscribe	Returns a description of MECCA and an application for membership.
adduser	Expects a filled out application for membership in the body of the message. If correct, i.e., the applicant meets the membership requirements, the user is added and an introductory message is returned.
profile-send	Returns a copy of the user’s current profile.
profile-change	Replaces the user’s current profile with a profile

	contained in the message unless it fails to parse. The actual update which will affect the results of queries takes place over night.
profile-delete	Removes the users profile, completely deleting her from the data base.
summarize-profiles	Returns a list of all keywords used in the data base or a list of the values for specific keywords if a Keywords: <list of keywords> line appears in the message.
add-address	Adds (overnight) a new legitimate source address to the user's profile. Message must contain valid True-Name: and Password: fields. The source address added is taken from the From: field of the message header.
delete-address	Same as above, but the address is deleted.
bug	Message is forwarded to the implementation team.
suggestion	Message is forwarded to the administrator.
help	Returns the short form of the MECCA documentation.
help-long	Returns the long form of the MECCA documentation.
help.ps or help-long.ps	As above, but in postscript.

3.2. Profiles

At the heart of MECCA is a data base of member profiles. This section describes the type of data base currently provided for a MECCA installation. The goals of this design is to provide maximal flexibility for the user to describe herself and her interests.

A profile is a series of entries of the form

```
<keyword>: <value string>
```

There are a few required keywords and a few keywords with constrained values, but, in general both keywords and values are completely arbitrary. That is, a user can make up keywords and values at will.

To change her profile, the user edits a copy of her current profile and sends it to admin@company.com with the string `profile-change` in the `Subject:` line.

A current copy can be requested by sending a message to admin@company.com with `profile-send` in the subject line.

Example Profile:

```
Name: Anita Borg
Email-Address: borg@pa.dec.com
Incoming-Email-Addresses: borg borg@pa.dec.com
        borg@nsl.pa.dec.com
Accepting-Mail: All
Geographical-Area: SFBayArea
Work-Address:   Network Systems Lab
                Digital Equipment Corp
                181 Lytton Ave.
                Palo Alto, CA 94301
Work-City: PaloAlto
Work-State-Province: CA
Work-Country: USA
Work-Telephone: 415-688-1367
Technical-Interests: data base architecture email
        operating-systems performance memory
Technical-Expertise: operating-systems performance memory
Current-Work: operating-systems performance memory database
Type-of-Work: design program  research
Type-of-employer: industry

Conferences-Attended: asplos sosp
Employer-Name: Digital
Memberships-Professional: ACM IEEE

Highest-Degree: Doctorate PhD
Highest-Degree-Year: 1981
Highest-Degree-Area: CS
Highest-Degree-School: New York University

Home-City: PaloAlto
Home-State-Province: CA
Home-Country: USA

Available-For:  review-for-conference program-committees
        review-for-journal speaker
```

In the sample profile, the required keywords are

- `Name`
- `Email-Address` where mail is to be sent
- `Incoming-Email-Addresses` legal incoming addresses from this person
- `Accepting-Mail` gross control of the type of mail this user wants possible values are `all, direct, publish, none`

All other keywords and values are arbitrary.

3.3. Keywords and Values in Use

For the freedom to choose arbitrary keywords and values to be useful, multiple users must agree to the same string to mean the same thing. To that end, a user may request a list of all keywords currently in use or a list of all values associated with a a keyword.

The list of the currently used keywords and a list of values currently associated with keywords are compiled once a day.

To get a list of the currently keywords, send a message to admin@company.com with `summarize-profiles` in the `Subject:` line.

To instead get a list of values for a set of keywords, include at the beginning of the message a `Keywords:` line containing the keywords.

In the returned message, the frequency with which a keyword or value appears in the data base is indicated with asterisks. One asterisk indicates that the value or keyword is used in 10 or more profiles. Two asterisks indicate that the value or keyword is used in 100 or more profiles.

3.4. The Query Language

The current interface to the data base of members is a simple query language. This interface is appropriate for the technical users our early installations, but is not an inherent part of the system. Alternative interfaces can be layered on top of the query language either as part of the core system or as message translators. Queries are used in messages and in the profiles as described in later sections.

The legal primitive queries are:

```
(exists? <field name>)
(empty? <field name>)
(contains? <word> <field name>)
```

They result in case insensitive word matches.

Compound queries may be formed using and, or and not:

```
(and query-1 query-2 ...)
(or query-1 query-2 ...)
(not query)
```

In all cases, a <word> is any string that does not contain white space (blanks and tabs). Punctuation is ignored. The legal values for <field name> depend on the use of the query and are described in the paragraphs below. Queries are used for four purposes in this system.

3.5. Direct Mail

Direct mail allows the sender to specify the subset of the data base to which a message will be targeted. Direct mail is sent to direct-mail@company.com.

The message must include a line of the form

```
Directed-To: <query>
```

This must appear as part of or immediately following the message header. There are no constraints on the form of the rest of the message.

The message, including the query, is sent to any member whose profile matches the query and who is currently accepting direct mail. Direct mail is always sent in full and cannot currently be summarized.

Direct mail is by default not archived, as this would defeat the purpose of restricting its audience. By including the primitive query, `(contains? archiver name)`, it is possible to cause the mail to be archived.

3.6. List Building

The list building facility allows the user to collect information about data base members whose profiles match a particular query. List building requests are sent to make-list@company.com.

A list building request contains a query in the `Subject:` field. The query is similar to that used for direct mail, but may not contain references to the `-(private)` versions of profile keywords. Only public information may be queried in and returned as the result of a list building request. As usual, continuation lines must begin with white space. The message may contain a `Reply-with:` field containing a comma separated list of profile keywords. This specifies which fields the requester would like returned. The body of the message may contain descriptive information.

The query is matched against all profiles which specify that they are not hidden, i.e, contain `Hidden: No`. Information is returned from each matching profile. If the `Reply-with:` field is present, only the fields specified are returned, otherwise all public fields are returned for every matching profile. The body of the request message is put in the `Context:` field of the reply. This can be used as a reminder of the purpose of the request.

3.7. Published Mail

The published mail facility allows the potential recipient of a message to decide whether or not to receive the message based on the content of certain fields in the message. It also allows the recipient of mail to decide whether the mail is to be sent in full, one message at a time, or is to be summarized, compiled, and sent out daily or weekly. Published mail is sent to publish@company.com.

The two profile entries

```
Subscribe-Full: <query>
Subscribe-Summary: <query>
```

are used to specify what published mail one wishes to receive. Queries may refer to the From: and Subject: fields in the message header and the optional Thread: and Keyword: fields in the message body.* See the description of the query language for precise query formats.

A published mail message may be an ordinary message on any topic that is sent to publish@company.com.

A published message may optionally contain either of two special fields which may be queried via the subscription mechanism described above.

```
Keyword: <topics>
```

allows the sender of published mail to specify what she feels are the most relevant topics covered in the message. "Keyword:" must start the line. A list of white-space separated topics follows. If the list takes more than one line, continuation lines must begin with white space.

```
Thread: <thread-name> [new]
```

is used to connect a series of messages in a named conversation called a thread. It allows individuals to find out about new threads and to subscribe to particular threads.

A user wishing to begin a new conversation called xxxxx includes in her message the line

```
Thread: xxxxx new
```

Anyone subscribing to new threads will receive the message. A user subscribes by conjoining the primitive query (contains? new thread) appropriately with her Subscribe-Full or Subscribe-Summary query. Upon receipt of the first message of a thread, she may chose to update her subscriptions with (contains? xxxxx thread) in order to continue getting related messages. It is up to users to put the right thread names in their messages when they wish them to be part of an ongoing discussion.

* We may eventually provide the ability to search the entire message body if performance is not a problem.

3.8. Combining Direct and Published Mail

To assure that a certain subset of members gets a message AND that anyone interested in the subjects also can get it, include a line

```
Directed-To: <direct-mail query>
```

in a message sent to publish@company.com. The message will be sent at most once to any member, thus avoiding the duplication of messages that could occur were a message twice, once as published mail and once as direct mail.

3.9. Security and Privacy

There are two issues addressed in this section: Security of the system as a whole and privacy of an individuals data base entry. Security for the system involves authenticating all messages to assure that only members of the data base can use the data base. Privacy is assured by providing users with a number of ways of controlling the nature of the access that legitimate members have to their profile.

3.9.1. Authentication

Our goal has been to provide some degree of security from unauthorized access without burdening users or requiring anything more of a member than that she have email access. The system is only as secure as is normal private email.

With the exception of mail requesting addition to the data base, only messages from members is accepted. Membership is validated on one of two ways:

- The message is from an email address that is recognized as being that of a member, i.e., it is contained in the the `INCOMING-EMAIL-ADDRESSES:` field of some profile in the data base.
- The message contains `True-Name:` and `Password:` fields with correct values for a member of the data base.

Clearly, forgeries are possible. To limit their effectiveness and assure that they are detected, we decree that all mail to a member is sent to her email address of record, i.e., the contents of the `outgoing-email-address:` field, rather than to the return address on a message claiming to be from her. While this can result in some inconvenience when interacting with the data base from a new address, it is well worth the added protection.

It is also important that there is at least notification when any changes are made to a user's profile. Any change to the value of the `outgoing-address:` field results in a copy of the new profile being sent to both the new and old outgoing addresses. Thus, a user will receive notification of any unauthorized change to her profile.

To reiterate:

- You may send mail through the system from anywhere as long as you include your `True-Name:` and `Password:`.
- But, any results, e.g. the answer to a list-buiding query, will be sent to your address of record.
- To receive mail elsewhere you must change your profile and wait until the next day for the change to take effect.

3.9.2. Privacy

In addition to providing assurances that only members of the data base have access to its services, it is important that the information in an individual's profile be used only as that person sees fit. The mechanism described in the above section attempts to assure that only a member may modify her profile, and that any unauthorized access will be noticed. There are also mechanisms to allow a member control over the use of individual fields in her profile.

First, if you don’t want anyone ever under any circumstances to find out something about you, don’t put it in your profile. These mechanisms are not 100% fool proof.

There are two kinds of information in a profile: public, private. A field is private if its keyword ends with the suffix `-(private)`. Otherwise it is public. Both kinds of information can be used to direct mail to you. If you don’t want mail directed to you based on some piece of information, don’t put that information in your profile. Public fields can also be used in list building queries and can be returned as the result of those queries. Therefore, if you want mail sent to you on the basis of some value, but you do not want that value returned to curious queriers, put it in a private field. Both the private and public versions of a most keywords can appear in a profile. For more details, see the sections on list-building and on keywords.