DESIGN ISSUES OF PROTOCOLS FOR COMPUTER MAIL

Jose J. Garcia-Luna-Aceves
Franklin F. Kuo
Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, Hawaii 96822
Design Issues of Protocols for Computer Mail

Jose J. Garcia-Luna-Aceves
Franklin F. Luu

Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, Hawaii 96822

Office of Naval Research
Code 437
Arlington, VA 22217

Approved for public release; distribution unlimited.

In this paper we present major design considerations for the development of international computer mail protocols. A simplified functional model for computer mail systems is introduced to serve as the basis of the discussion of the protocols. The general framework of the Reference Model for Open System Architecture proposed by the ISO is followed in the description of the computer mail protocol and design issues involved at each layer of the computer mail protocol are examined. The computer mail protocol is aimed at non-interactive communication between message system users.
DESIGN ISSUES OF PROTOCOLS FOR COMPUTER MAIL

by

Jose J. Garcia-Luna-Aceves
Franklin F. Kuo
Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, Hawaii 96822

ABSTRACT

In this paper we present major design considerations for the development of international computer mail protocols. A simplified functional model for computer mail systems is introduced to serve as the basis of the discussion of the protocols. The general framework of the Reference Model for Open System Architecture proposed by the ISO is followed in the description of the computer mail protocol and design issues involved at each layer of the computer mail protocol are examined. The computer mail protocol is aimed at non-interactive communication between message system users.
INTRODUCTION

Many computer-based message systems already exist [16] and in the near future many more system will be in use by large organizations to automate the flow of messages among their members. International standards are needed to allow the open interworking of message system users in different organizations, systems and countries. Over the last few years, various studies have been made on the development of communication protocols for computer based message systems [13], [15], [17] and certain international agreements may soon be reached. In this paper we present major design considerations for the development of such standards. Our approach is to partition computer mail functions into layers following the general framework of the Reference Model of Open Systems Architecture proposed by the ISO [9]. This layering of functions permits a clean specification of the standards needed.

Section I specifies a functional model for computer mail systems that serves as the basis for the specification of the communication standards. This model is compatible with the one proposed by the North American System Environment Subgroup of IFIP W.G. 6.5 [19] but makes a distinction between communications within systems and among systems. Section II describes the requirements of protocols for computer mail and the layering of the computer mail protocol (CMP) we propose. Section III describes the session layers of the CMP. Section IV specifies the presentation layer of the CMP. Section V specifies the management layer of the CMP. Section VI discusses the message processing layer of the CMP. Section VIII summarizes highlights of this proposal and points out those areas in need for further specification.
I. A FUNCTIONAL MODEL FOR COMPUTER MAIL SYSTEMS

A. Components

We model a computer mail system by partitioning the system into functional components, each dedicated to a specific set of computer mail functions. There are three different types of functional entities in the proposed model [17]: MAILBOX, MAILER and GATEWAY MAILER.

MAILBOX (MBX) is the entity responsible for the processing, storage and retrieval of user messages. A mailbox serves as the interface between its message system user and delivery services. This component consists of:

1. Message processing modules to compose, edit, retrieve and archive user messages.
2. Communications software to transfer user messages to and from the entity dedicated to message delivery (the local mailer).
3. User files where user messages are permanently stored and (optionally) a personal directory with users' system mailbox addresses is maintained.
4. The message workspace where undelivered messages and messages being composed are maintained.

MAILER (MLR) is the component process responsible for the delivery of messages to and from a specific set of mailboxes and the identification of the system mailbox addresses of the recipients of the messages. The mailer is formed by:

1. Communication software modules dedicated to the communications of the mailer with its local mailboxes and other mailers.
2. The message buffer where messages to and from local mailboxes are temporarily stored.
3. The mailer directory database which maintains addressing information and time-stamped records of message deliveries.

A mailing network is formed by the union of logically connected mailers and corresponds to a public or private message system. Thus, mailing networks
(and their corresponding mailers) could be managed and owned by one or more organizations.

GATEWAY MAILER (GMR) is the entity responsible for internetwork communication. Each mailing network has associated with it a gateway mailer, which represents half of the gateway between a mailing network and any other mailing network. A computer mail system consists of a set of interconnected mailing networks. The gateway mailer is formed by:

1. Communication software modules to communicate the local mailing network with other mailing networks.
2. The message buffer, where messages to and from mailers in the local mailing network are temporarily stored.
3. The gateway directory database, which maintains internetwork addressing information and time-stamped records of message deliveries.

B. System Operation

In our architecture delivery and addressing are functions transparent to the message system users. We define a user-level naming format called the NOLS address [7], [8], which consists of four fields that contain information about the recipient of a message, as is shown in Table 1. NOLS addresses are used by message system users to describe recipients. At the time the sender of a message enters his message he also enters a NOLS address with what he knows about the recipient's name and/or title, organization, geographical location and (perhaps) message system. The sender is not concerned with the recipient's system mailbox address or how the message is delivered, and the system assists the sender in identifying the system mailbox address from the user-level description (NOLS address) of the recipient.

Figure 1 shows in a simplified form the steps followed to deliver a message. Message delivery is carried out in two phases: The NEXUS establishment phase and the message delivery phase. A NEXUS is an end-to-end virtual
connection established between the sender's and the recipient's mailboxes. A NEXUS address is the specification of a NEXUS in the system as is shown in Table 1. During the NEXUS establishment phase the system maps the NOLS address (user-oriented standard) into a NEXUS address (system-oriented standard) and establishes the end-to-end virtual connection. The message delivery phase consists of the dispatch of the message through the NEXUS.

In an internetworking environment, with tens of thousands of users, several organizations and various countries involved, it would not be feasible or desirable to maintain the system mailbox addresses of all message system users in the database of each mailer or a centralized entity. Consequently, in our model each mailer maintains only the mailbox addresses of the users served by that mailer, together with a set of mailer addresses (pointers) that the mailer associates with user groups and organizations. The NOLS address entered by a sender serves as an identification query to the system for obtaining the correct system mailbox address. As is shown in Fig. 1, the procedure followed to resolve such a query is a store-and-forward process in which the NOLS address is forwarded among mailers and gateway mailers according to the network (internetwork) pointers maintained in their directory databases. When a mailer (gateway mailer) receives a query, the mailer (gateway mailer) examines the fields of the NOLS address and based on its directory database it decides whether to forward the query, or to reply with a positive or negative acknowledgement. Once the sender's mailer obtains a positive acknowledgement, the NEXUS between sender's and recipient's mailboxes is created (Fig. 1) and the message can be delivered. Since users may enter NOLS addresses lacking key information, various gateway mailers and/or mailers may have to be queried to establish the NEXUS.

An "electronic envelope" is added to the user message or identification
query being transmitted that specifies the control information needed for
correct routing, structuring of information, protection of users' privacy,
billing and accounting.

II. THE COMPUTER MAIL PROTOCOL (CMP)

Today, office automation systems are becoming more and more sophisticated
with the merging of techniques for the communication, processing, storage and
retrieval of information in office environments. However, the evolution of
these sytems has been such that a highly segmented and specialized market
exists [16]. Under such conditions, customers (individuals and organizations)
manage and own message systems that may differ from each other in many
different ways. In order to allow the open interworking between message
system users who access different systems, international standards are
needed. These standards must specify:

1. The form in which the dialogue is established between communicating
   parties;
2. The structure of the information exchanged between processes;
3. The form in which the computer mail services are managed; and
4. The message processing procedures implied in the messages
   transmitted.

This set of standards constitute the computer mail protocol (CMP). The
CMP must be such that:

1. It is simple enough to be implementable;
2. It permits the communication with existing systems;
3. It supports all the services needed;
4. It is extensible, so that new features can be added to it; and
5. It is independent of implementation considerations.

To specify the CMP we follow the Reference Model for Open System
Architecture proposed by the ISO [9]. The CMP extends the higher three layers of the ISO model [10], namely: The Application, Presentation and Session layers. As is shown in Figure 2, the CMP is partitioned into four layers: The Message Processing layer and the Management layer, which together constitute the ISO application layer, the Presentation layer, and the Session layer. The CMP relies on the existence of a transport service that permits the establishment of logical connections between processes in different host computers to transfer arbitrary messages in the proper sequence, independently of the communication subnetwork(s) being used [3], [6]. Each command of the CMP is formed by a set of three nested headers that constitute the "electronic envelope" of the information to be processed by mailboxes, mailers and gateway mailers. Such an envelope provides all control information.

III. SESSION LAYER

The session layer of the CMP supports the dialogue between mailboxes, mailers and gateway mailers, and between message system users and their mailboxes. It adds computer mail services to the bare transport service used to communicate processes in different hosts. The functions at this layer are:

1. The establishment and termination of sessions (virtual connections) between presentation level entities.
2. The exchange of session identification information.
3. The establishment of session mode capabilities, control of information transfer and negotiation of workstation profiles.
4. Error recovery and control.

In general, different modes of sessions may be established between users and mailboxes (full duplex, half duplex), and between mailboxes, mailers and gateway mailers (one-to-one, one-to-many). Because of the non-interactive
form of communication supported in computer mail, a datagram-oriented form of communication can be used to support the dialogue between functional components. With this approach sessions are opened and closed with the same command used to send the presentation-level information through the session. Thus, the dispatch of presentation-level commands is done by means of datagrams. The reply to a datagram specifies what portion of the command has been already received, so that an error recovery procedure can be implemented and failures at the transport layer masked. The session-layer header must also specify:

1. A unique identifier for the commands;
2. The next destination (or route) of the command;
3. The mailers (gateway mailers) that have handled the command; and
4. The type of command (i.e., datagram or reply to datagram).

Replies to datagrams must specify in addition:

1. The identifier of the command being acknowledged.
2. The mailers (gateway mailers) that handled the command.

This information is used to route the datagram through various mailers and gateway mailers and ensure loop-free deliveries.

IV. PRESENTATION LAYER

Even though today's computer mail systems only support text messages, the merging of new data communications techniques, storage technologies, and processor and terminal technologies will allow future computer mail systems to become multimedia communication tools which integrate text, digitized voice, facsimile, graphics and (perhaps) video within the same message.
structure. If standards are not set in advance, the usefulness of such communication tools will be diminished as users with different terminal equipment will be unable to communicate with each other.

The presentation layer of the CMP is necessary to allow the open interworking between heterogeneous office workstations (e.g. teletypes, graphic terminals, word processors) and to permit the exchange of information in standard formats between mailers, gateway mailers and mailboxes. For such purposes, this layer defines a set of structuring rules to specify formats that integrate alphanumeric text, digitized voice, graphics and video pictures in the messages exchanged in the system.

The presentation layer of our CMP relies on the concept of virtual workstation which is an extension of the virtual terminal concept [1], [6]. The virtual workstation consists of a hypothetical office workstation with standard functional characteristics known by the management and message-processing-level processes.

The model of a virtual workstation consists of a data structure which is specified by its dimensions, the data types it contains and the operations defined on it [6]. The access to the data structure is asymmetrical, always initiated by the sender's process. Each workstation is assigned a workstation class and profiles based on the functional characteristics of its input/output devices and the available forms to represent information. A mailbox maps the virtual workstation characteristics defined in the CMP into the physical characteristics of the local workstation based on the workstation class and profiles.

The presentation layer is carried out in three phases: The negotiation phase, the data transfer phase and the termination phase. These three phases assume the existence of a session between the communicating parties. The
negotiation phase allows the presentation-level processes to agree on a particular structuring technique. Once an agreement is reached, the data exchange takes place until no more information is to be transferred, at which time the transaction is terminated. The negotiation of profiles is asymmetric, with the sender's process always being the master of the negotiation. This asymmetry is suitable for Computer Mail and guarantees that the negotiation procedure is loop-free.

Profiles can be negotiated only when the NEXUS is being established. That is, when the identification query is transmitted. Once the NEXUS is established, no further negotiation is possible. The recipient's mailbox stores the message as received. When the message is retrieved, those portions that the mailbox cannot output are deleted and the user is notified of the event.

The presentation-level headers of the commands define the virtual workstation class and profiles being negotiated. The header of the replies specify the class and profiles supported by the recipient's mailbox. Information is structured in a self-identifying form by means of a set of data types. A data type is made up of a tag and a value component. The tag component specifies the data type, its size, its lexical level address [12] in the message and its attributes. The value component contains management-level information. A cell is the basic unit of reference consisting of c bits. A data type is a string of adjacent cells that encodes a piece of information. Table 2 shows the eight data types defined in the CMP. The integer data type encodes control information in a compact form. String data types are used to integrate multimedia information in the same message. These data types have a tag component describing the characteristics of the integer or string, and a value component with the application level
The structure data type is a heterogeneous collection of data types used to structure messages (e.g., in pages, chapters, sections, etc.) and organize messages (e.g., group various messages in a message bag). A management-level command is encoded as a structure. The array data type is used to organize homogeneous data types. The pointer data type refers to another data type in the same message or a previous message. It is used to transmit multidestination messages with low redundancy.

The "type" subfield of the tag component specifies the data type. The lexical level address provides for the identification of the data type within the message. It consists of the lexical level of the data type and the index (sequential number) of the data type within the level. The attributes specified in the tags depend on the techniques used to encode text, voice, graphics and video, and may have to be negotiated. The encryption subfield indicates whether or not the value component of the data type is encrypted and what technique is used.

IV. MANAGEMENT LAYER

The management layer of the CMP manages the delivery of user messages and the distribution and maintenance of identification information. Table 3 gives the seven commands defined in the CMP to deliver user messages, dispatch identification queries and update directory databases. The management-level header contains the information required to:

1. Specify the command to be executed.

2. Identify the sender and the recipient, and date and time of delivery. This information is used to enforce closed user groups (and user privacy in general), and to establish "time stamps" of every message delivery and identification query processed.
Time stamps of transactions permit the system to trace the message when problems arise. In addition, they constitute an accurate source of information for billing and accounting purposes.

The information component of a management-level command contains the information to be processed by mailboxes (in message deliveries) or mailers and gateway mailers (in identification queries and updates). Table 4 gives the structure of the commands defined in the management layer.

The establishment of a NEXUS between sender's and recipient's mailboxes is carried out at this layer in the form described in Sec. I. As is shown in Table 3, NOLS addresses are distributed by IDENTIFY commands, and DELIVER commands are used to dispatch user messages. A NEXUS is an end-to-end virtual connection at the management layer that relies upon sessions established at the session layer. Various NEXUS's may have to be established for the delivery of a multidestination messages and various sessions may have to be established to support a NEXUS if the message has to be routed through various mailers and gateway mailers.

Network and system level updating of the pointers maintained in the directory databases of mailers and gateway mailers is done by means of the UPDATE command (Table 3), which is dispatched when forwarding mailers or gateway mailers detect erroneously delivered commands. Information about local users in the mailers' directory databases is independently maintained by each mailer. End-to-end error recovery from erroneous deliveries is supported by the provision of the Transaction-ID of the headers and the forwarding of the commands (when appropriate) to the proper mailer or gateway mailer.
VI. MESSAGE PROCESSING LAYER

The message processing layer of the CMP comprise the functions of editing, filing, sorting and reviewing messages at the user MAILBOX. The message processing layer typically includes a set of commands to process textual information. Various such message processing applications have been implemented [11], [18]. Since message processing functions differ from system to system, it is not known at this time what standards are needed at this layer or if standards are even needed at all. However, we can identify a generic set of message processing functions that fit general needs. These functions are:

1. User session establishment and termination;
2. Information retrieval;
3. Message editing and composition;
4. Message delivery, forwarding and reply;
5. Filing and archiving of information;
6. User help, documentation and training;
7. Customization of user interface;
8. Billing and statistics; and

VII. CONCLUSIONS

In this paper we have presented major design considerations for the development of international standards for computer mail. We followed the Reference Model proposed by the ISO to describe our proposed protocol. The layering of the computer mail protocol into four layers permits a clean specification of the standards needed. The message processing layer deals with the functions of composing, filing and sorting of the messages. The management layer deals with the management of message deliveries and the
identification of system addresses. The presentation layer integrates multimedia information in the same message and structures application-level commands. The session layer incorporates computer mail functions to the lower level packet transport services and permits a dialogue between computer mail processes.

In the future, more general session and presentation layers may evolve, as the development of Teletex protocols indicates [2], and it is important to identify standards for the transmission of multimedia information. There have been various studies related to the presentation layer [13],[17] and the management layer [17] of the CMP. However, there have been few attempts to provide solutions to the addressing problems that arise in large internet systems [8], [15]. This problem will constitute a major design problem for system interconnection. The specification of message processing functions is still an open problem, mainly because users are just starting to understand office automation and because such functions are directly related to the purposes of the communication between users.
REFERENCES


<table>
<thead>
<tr>
<th>FORMAT TYPE</th>
<th>COMPONENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOLS address</td>
<td>&lt;N-field&gt; &lt;O-field&gt; &lt;L-field&gt; &lt;S-field&gt;</td>
</tr>
<tr>
<td>N-field -- Contains information about the recipient(s) of the message such as his Name and/or title.</td>
<td></td>
</tr>
<tr>
<td>O-field -- Contains information about the Organization(s), group(s) or system server(s).</td>
<td></td>
</tr>
<tr>
<td>L-field -- Contains information about the geographical Location of the organization(s) or group(s) referred in the O-field.</td>
<td></td>
</tr>
<tr>
<td>S-field -- Contains information about the System which offers the computer mail services to the recipient(s).</td>
<td></td>
</tr>
<tr>
<td>Proposed 3-layer system mailbox address</td>
<td>[&lt;internetwork layer&gt; &lt;intranetwork layer&gt; &lt;local layer&gt;]</td>
</tr>
<tr>
<td>Internetwork layer--Contains the name of the network specified according to the system-wide formats.</td>
<td></td>
</tr>
<tr>
<td>Intranetwork layer--Contains the name of the mailer in the mailing network specified in the internetwork layer. The mailer's name is specified according to formats particular to the network.</td>
<td></td>
</tr>
<tr>
<td>Local layer--Contains the name of the mailbox in the mailer specified in the intranetwork layer. The name of the mailbox is specified according to formats particular to the mailer.</td>
<td></td>
</tr>
<tr>
<td>NEXUS address</td>
<td>([sender's system] [mailbox address] , [recipient's system] [mailbox address])</td>
</tr>
</tbody>
</table>
Table 2 - DATA TYPE FORMATS

<table>
<thead>
<tr>
<th>DATE TYPE</th>
<th>FORMAT</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTEGER</td>
<td>TAG = &lt;X&gt; VALUE = &lt;integer base-2&gt;</td>
<td>If size field is equal to 0, the value component is defined to be empty.</td>
</tr>
<tr>
<td>TEXT STRING</td>
<td>TAG = &lt;X&gt; &lt;character attributes&gt; VALUE = &lt;character string&gt;</td>
<td>Character attributes describe: character set (e.g., ASCII), character appearance (e.g., type face, boldness, color, slanting, size underlining) and inter-spacing.</td>
</tr>
<tr>
<td>AUDIO STRING</td>
<td>TAG = &lt;X&gt; &lt;vocoding&gt; VALUE = &lt;bit string&gt;</td>
<td>Vocoding describes the parameters used in the encoding technique to digitize voice [4].</td>
</tr>
<tr>
<td>GRAPHIC STRING</td>
<td>TAG = &lt;X&gt; VALUE = &lt;output commands&gt;</td>
<td>Output commands specify the creation, manipulation and management of segments and units in segments [5], [16].</td>
</tr>
<tr>
<td>VIDEO STRING</td>
<td>TAG = &lt;X&gt; &lt;coding&gt; &lt;attributes&gt; VALUE = &lt;bit string&gt;</td>
<td>Attributes describe the dimensions of the space where information is to be shown and the characteristics of the device [16], [17].</td>
</tr>
<tr>
<td>STRUCTURE</td>
<td>TAG = &lt;X&gt; &lt;element count&gt; VALUE = &lt;data types&gt;</td>
<td></td>
</tr>
<tr>
<td>ARRAY</td>
<td>TAG = &lt;type&gt;&lt;lexical address&gt;&lt;dimensions&gt;&lt;size dimension 1&gt;&lt;size dimension n&gt;&lt;nested tag&gt;&lt;framing attributes&gt; VALUE = &lt;value components of data types&gt;</td>
<td>Nested tag is the tag component of the homogeneous data types in the value component. If nested tag = text string, then framing attributes describe the page format where information is to be presented. Otherwise field is empty.</td>
</tr>
<tr>
<td>POINTER</td>
<td>TAG = &lt;type&gt; &lt;lexical address&gt; &lt;crypto&gt; &lt;nested tag&gt; VALUE = &lt;referenced lexical address&gt; &lt;command identifier&gt;</td>
<td>The contents of the value component is located relative to other data type. Nested tag is used for error checking, this field and crypto should be equal to the tag fields of the referenced data type.</td>
</tr>
</tbody>
</table>

Legend: <X> = <type><lexical address><size><crypto>
Table 3 - BASIC SET OF COMMANDS FOR INTERPROCESS COMMUNICATION

<table>
<thead>
<tr>
<th>COMMAND</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>DELIVER</td>
<td>Contains the user message, together with the NOLS and NEXUS address of both sender and recipient.</td>
</tr>
<tr>
<td>IDENTIFY</td>
<td>Contains the user query, i.e., a NOLS address created by the sender. It also identifies the sender for management purposes.</td>
</tr>
<tr>
<td>ACK-ID</td>
<td>Contains the positive reply to an IDENTIFY command with the system mailbox address and NOLS address of recipient.</td>
</tr>
<tr>
<td>NACK-ID</td>
<td>Contains negative reply, possibly including a &quot;list of similar names&quot;.</td>
</tr>
<tr>
<td>ACK-DEL</td>
<td>Contains positive acknowledgement to a successful delivery.</td>
</tr>
<tr>
<td>NACK-DEL</td>
<td>Contains negative acknowledgement to DELIVER command and is possibly due to component failure, unauthorized sender, or non-existent mailbox address.</td>
</tr>
<tr>
<td>UPDATE</td>
<td>Calls for an update in the directory database and contains time-stamped information.</td>
</tr>
</tbody>
</table>

Table 4 - MANAGEMENT-LEVEL COMMAND STRUCTURE

<table>
<thead>
<tr>
<th>FIELD</th>
<th>SUBFIELD</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEADER</td>
<td>TRANSACTION-ID</td>
<td>System address of the two parties (message system user of system's process) involved in the transaction.</td>
</tr>
<tr>
<td></td>
<td>O-NOLS</td>
<td>System-generated description of the sender of the message, query or update.</td>
</tr>
<tr>
<td></td>
<td>TIME STAMP</td>
<td>Date and time of transmission of the command.</td>
</tr>
<tr>
<td></td>
<td>COMMAND TYPE</td>
<td>Specification of the command to be executed.</td>
</tr>
<tr>
<td></td>
<td>OPTION SPECIFICATION</td>
<td>Specification of the service requested for the command.</td>
</tr>
<tr>
<td>INFORMATION</td>
<td>Contains message processing-level information, that is: sender's message, identification query, updating information, or reply information.</td>
<td></td>
</tr>
</tbody>
</table>
ADDRESS IDENTIFICATION

sender's MBX

MLR

to other MLRs and/or GMR

MLR

Store and Forward Process

QUERY(NOLS)

MLR

recipient's mailbox address

DIRECTORY DATABASE

NEXUS ESTABLISHMENT

sender's MBX

MLR

recipient's MBX

recipient's system mailbox address

GMR

GMR

MESSAGE DELIVERY

sender's MBX

MLR

MESSAGE nexus

MLR

recipient's MBX

Figure 1 - SYSTEM OPERATION
Figure 2 - COMPUTER MAIL PROTOCOL