BOLT BERANEK AND NEWMAN INC

CONSULTING · DEVELOPMENT · RESEARCH

Report No. 1837 July 1969

INTERFACE MESSAGE PROCESSORS FOR THE ARPA COMPUTER NETWORK

QUARTERLY TECHNICAL REPORT NO. 2
1 April 1969 to 30 June 1969

interesting and rather comprehensive overview of IMP operation.

— priority, — [illegible] messages

Submitted to:

Advanced Research Projects Agency
Washington, D.C. 20301
Attn: Dr. L.G. Roberts

Report No. 1837 July 1969

INTERFACE MESSAGE PROCESSORS FOR THE ARPA COMPUTER NETWORK

QUARTERLY TECHNICAL REPORT NO. 2
1 April 1969 to 30 June 1969

SUBMITTED TO:

Advanced Research Projects Agency
Washington, D.C. 20301
Attn: Dr. L.G. Roberts

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-69-0179.

TABLE OF CONTENTS

1. INTRODUCTION

This Quarterly Technical Report No. 2 describes several aspects of our progress on the ARPA Computer Network during the second quarter of 1969. During this quarter, the specifications for interconnecting a Host and an IMP were documented in BBN Report No. 1822, which was distributed to all the participating Hosts. The report, which was made available in loose leaf form, is divided into four sections: an introductory section; a section describing the physical apparatus and the requirements for its installation; a section devoted to software protocol and message formatting; and a section detailing the standard Host/IMP interface, the electrical signals on the Host cable, and the requirements for the design of the Host's special hardware interface.

Early in this quarter, we completed the hardware design and received a prototype IMP. This unit contains one IMP/Modem interface and one IMP/Host interface, as well as a normal complement of the other special features (real-time clock, watchdog timer, etc.). The prototype was checked out and has since been operating reliably.

The software design has been substantially completed and the implementation of the initial operational IMP program is well underway. In this report we describe the organization of this program and present our preliminary plans for obtaining measurements.

2. OPERATIONAL PROGRAM DESIGN

An initial operational IMP program has been designed and is being implemented for use in the ARPA network. The overall function and organization of this program is based upon the IMP design and the network protocol, as described in BBN Report No. 1763. The current program design is decribed below.

The system software has been specially designed (for the modified Honeywell D.D.P.-516 computer) to maximize packet throughput, to provide maximum autonomy for each IMP, and to maximize the available buffer space by using as little core memory for the program as possible.

The principal function of the program is the processing of packets. This processing includes segmentation of Host messages into packets for transmission, building of headers, receiving and transmitting of store and forward packets, reassembling of packets received at a destination into messages for transmission to the Host, and generating of RFNM's and acknowledgements.

Another important function of the IMP program is network monitoring, which helps to maintain reliable communication . Network monitoring includes detecting adjacent lines that have gone dead and detecting and discarding duplicate packets. In addition, information received from its neighbors at half second intervals is used to identify unreachable destinations (dead or inaccessible sites) and to inform the IMP of the minimum expected delay to each destination.

Our preliminary plans for providing measurement facilities incorporate the taking of synchronized network snapshots, the accumulating of histograms, the tracing of packets through the network, and the recording of actual packet arrival times on an input channel. These measurement facilities are discussed in Section 3.

+

The program currently occupies approximately 6000 registers of core, which is half of the available memory.* The other half of memory will be used for buffer storage (except that a portion of this buffer space will be used for the operational DDT program early in the network operation).

good convey

A. Program Organization

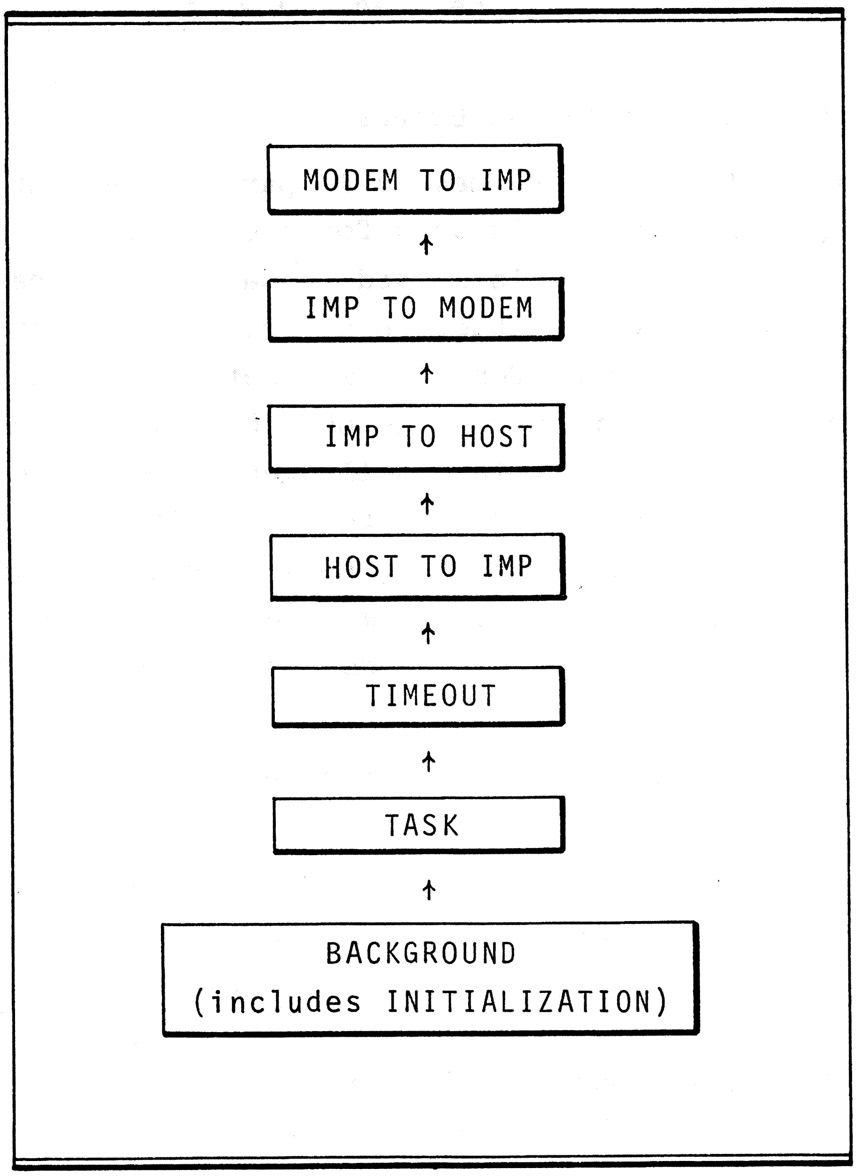
The program is composed of five main pieces: a *task routine* that performs the major portion of the IMP packet processing; the Modem routines (IMP to Modem and Modem to IMP) that handle interrupts and resetting of buffers for the Modem channels; the Host routines (IMP to Host and Host to IMP) that not only handle interrupts and resetting of buffers for the Host channels but also build packet headers during input and construct RFNMs that are returned to the source Host during output; a *timeout routine* that maintains a software clock, activates certain deferred tasks, and attends to infrequent events; and a *background routine* that handles all jobs of sufficiently low priority that the exact time of their execution is relatively unimportant. The background routine also includes an initialization routine that is used during startup.

The transfer of program control from one routine to another is completely governed by the occurrence of priority interrupts. After the initialization routine is executed, program control remains in the background routine, which is cycled through repeatedly until an interrupt occurs causing control to be transferred to another routine. Each routine has associated with it a priority that allows it to interrupt any other routine situated below it in the hierarchy. The hierarchy of program control is shown in Fig. 1 below.

---

*The IMP cabinet has space for an additional 4000 words of core.

*7-1 a half
of buffers for packets left!



ascending priority

Fig. 1. Program Control.

When all interrupts have finally been serviced, control is again returned to the background routine, which resumes at the instruction that was about to be executed when the first interrupt occurred.

B. Buffer Allocation and Queues

A large fraction of core memory is partitioned into fixed size buffers, each of which is used for storing a single packet. A buffer in which no packet is stored is said to be free, and all free buffers are chained together on a free buffer list. A buffer is removed from the free buffer list and set up for input each time a buffer is needed to store the next packet from a Host or Modem channel or to store an internally generated packet. Buffers that are not located on the free buffer list either are set up to store an incoming packet, are having their contents transmitted on one of the output channels, are being processed by the IMP program, or have been dispatched to one or more system queues for subsequent processing.

There are seven principal system queues:

Task Queue

The task queue is organized into a Host task queue and two Modem task queues. Packets received on Host channels are placed on the Host task queue. All received acknowledgements, dead Host and routing information, I Heard You and Hello packets are placed on a fast Modem task queue. All other packets from the Modems are placed on a regular Modem task queue. The fast Modem task queue will be serviced first. The IMP program currently services the Host task queue before the regular Modem task queue, but this latter arrangement is not yet considered final.

Output Queues

A separate output queue is constructed for each Modem channel and each Host channel. All packets waiting their turn for transmission on a given channel are chained together on the queue for that channel. Each Modem output queue is divided into an acknowledgement queue, a priority queue, a RFNM queue, and a regular message queue, which are serviced in the indicated order. The Host output queue is divided into a control message queue, a priority queue, and a regular message queue, which are also serviced in the indicated order.

priority order in queues?

Sent Queue

A separate sent queue for each Modem channel contains packets that have already been transmitted on that line but for which no acknowledgement has yet been received.

Retransmission Queue

If a received acknowledgement frees a buffer that is not first on the sent queue, all the skipped buffers are placed on the retransmission queue.

Reassembly Queue

The reassembly queue contains those packets that are being reassembled into messages for the Host.

Conversion Queues

The conversion queues consist of an input queue of fully reassembled messages, which are intended for the Host and are awaiting the application of the Host specified unique transformation, and an output queue of packets, which have just been received from the Host and are awaiting transformation.

C. Program Routines

Modem to IMP

The Modem to IMP routine is executed whenever a packet arrives on a Modem channel. This routine sets up a free buffer for storing subsequent input on that channel and places the buffer containing a correctly received packet on a Modem task queue. If no free buffers are available, the first buffer is removed from its place on the regular Modem task queue. Should no such buffer be available, the last received packet is discarded by setting up its buffer for input.

IMP to Modem

The IMP to Modem routine is executed each time a packet transmission is completed. Buffers containing a RFNM, a priority, or a regular message are placed on the sent queue, and the next buffer on the output queue is set up for transmission. If no packets are present on the output queue, this routine will be restarted when another buffer is placed there.

IMP to Host

The IMP to Host routine is executed whenever the transmission of a packet to the Host is completed. This routine removes the next buffer from the output queue to the Host and sets the buffer up for transmission to the Host. In addition, it builds a RFNM packet and places the RFNM on the Host task queue. If the output queue is empty, a next buffer is not set up and the routine is restarted when the queue is no longer empty.

Host to IMP

The Host to IMP routine is executed whenever the Host finishes its input or the Host input buffer is filled. The routine builds a header into the buffer, which is then placed on the end of the Host task queue. A free buffer is then set up for additional input from the Host. This routine will discard the filled Host buffer, if the Host Error flip flop went on during the transmission; in addition, all succeeding packets will be discarded until an end-of-message signal is detected. This routine also checks the message type for validity, verifies that the destination is not dead and that the link table is not full or the link blocked.

Timeout

The timeout routine is started every 25.6 ms by a clock interrupt. Timeout is used to increment a 16-bit software clock and to insure the occurrence of events that either are expected to be infrequent, or must occur at regular intervals, or have been marked for execution by the timeout routine.

The timeout routine is divided into a fast, medium, and slow timeout routine. The slow timeout routine is run every 25 clock interrupts. The medium timeout routine is run every fifth interrupt, except when the slow routine runs. In all other cases the fast routine is run.

The software clock is updated during each timeout routine. During the medium timeout routine, buffers on the retransmission queue are routed onto an output queue; buffers that are on the sent queue and that have timed out are routed onto an output queue for retransmission, as are all buffers on an output queue for a Modem channel that has just gone dead. During the slow timeout routine, the following tasks are performed.

- One dead line, if any, will be located and marked.
- A flag will be set if the Host ready line is off.
- A new routing table will be built and placed on each output queue to the Modems.

- Counters that might normally overflow are reset.
- A flag is set for each link that has timed out.
- A flag is set for each message that has timed out in reassembly.

Task

The task routine removes the next packet from the top of the task queue for processing and then either discards it or places it at the end of one of the system queues. If the first task is an acknowledgement, its four header words are compared against the header entries on the sent queue. If no match occurs, the buffer is released to the free buffer list and the next task is obtained. If a header match is obtained with the first entry, the buffer is removed from the sent queue and both buffers are freed. If a match occurs at other than the first entry, all the skipped entries are placed on the retransmission queue.

Buffers containing received routing and dead Host information are copied and then released. If a buffer contains an I Heard You packet, a flag will have been reset and the buffer will simply be released. If it is an Hello packet, a flag is set to send an I Heard You packet and the buffer is released.

When a packet reaches its destination the task routine checks the message number and determines whether sufficient buffer storage is available. The buffer is then either placed on the reassembly queue, on the output queue to the Host, or else is discarded.

If sufficient buffer storage is available, store and forward packets are routed onto an output queue to a Modem. If buffer storage is insufficient, the packets are discarded. Discarded packets are not acknowledged and will eventually be retransmitted.

Background

The background routine performs all the low priority processing. It handles all packets that are generated at the IMP teletype or destined for it, all completed trace and statistics buffers, all buffers on the conversion queues, as well as all packets generated by or intended for the operational DDT program. The background routine cannot interrupt any other routine, and, therefore, operates only when no other processing remains to be done.

The initialization routine, which is included in the background, builds and initializes all tables, builds the initial queue pointers, builds the free buffers, and creates a free buffer list. It also creates a list of free trace buffers and a list of free reassembly slots and then starts up all the output routines.

3. NETWORK MEASUREMENTS

As a vehicle for experimentation and study, the ARPA network is intended to provide a continuing supply of data to aid in understanding network performance. Two sources of this data are the direct measurement by the IMP program of its activities and the tracing of selected packets within the network.

In this section we present our preliminary plans for the design of these data gathering facilities. These plans are still subject to change, but in the next quarter we expect to finalize them and to complete our work on the measurement software. At that time we will prepare a separate document describing these facilities in detail.

A. Measurement Facilities

Each measurement facility may be activated or deactivated independently through the use of Host generated control messages. If all the facilities are activated at the same time, the system may experience a substantial measurement load, which might affect the flow of traffic and, consequently, the measurements themselves.

We have also provided for synchronization of the software clocks throughout the network. The clock in each IMP is synchronized to the clock of the IMP (that is not dead) with the lowest destination number. Each IMP transmits the reading of its 16 bit software clock to its neighboring IMPs every 1/2 second along with the routing information. The IMP uses the received time from the neighboring IMP, which is nearest the time reference, to adjust its own clock.

Periodic Snapshots

Once every half second, the IMP program will record the following information (referred to as *snapshots*) and transmit it to the measurement center.

- Length of the free buffer list
- Length of each Modem output queue
- Length of each Host output queue
- Length of the sent queue
- Length of the retransmission queue
- Length of the conversion queues
- Length of the background queue
- Number of messages in reassembly
- Number of transmit links in use
- Number of receive links in use
- Delay table (delay/HOP/Dead Hosts)
- Routing table (lines)
- Number of transmit characters converted
- Number of receive characters converted
- Received time from each neighboring IMP
- Own time

Snapshots at all IMPs will be synchronized to occur at approximately the same time.

Accumulated Histograms

Histograms are accumulated over an interval of ten seconds and then transmitted to the measurement center. Certain histograms are actual counts of number of occurrences of a particular event, while other histograms actually approximate a distribution. Specifically, the following information is recorded.

a. Per Modem

- Total number of regular and priority message words on each output line.
- Number of output Hello packets.
- Number of output RFNM packets.
- Number of output acknowledgement packets.
- Number of transmitted packets with:
  - 1 word of text
  - 2–3 words of text
  - 4–7 words of text
  - 8–15 words of text
  - 16–31 words of text
  - 32–63 words of text.
- Number of input I Heard You packets.
- Number of input packets in error.
- Number of good packets on input line.
- Number of good packets discarded because of no free buffers.
- Number of regular Modem task buffers liberated on input.
- Number of acknowledgements on input line.
- Number of retransmitted packets.

b. Per Host

- Number of messages on the input line with:
  - 1 word of text
  - 2–3 words of text
  - 4–7 words of text
  - 8–15 words of text
  - 16–31 words of text
  - 32–63 words of text.
- Number of messages on the input line with:
  - 2 packets
  - 3 packets
  - 4 packets
  - 5 packets
  - 6 packets
  - 7 packets
  - 8 packets.

- Number of messages on the output line with:
  - 1 word of text
  - 2–3 words of text
  - 4–7 words of text
  - 8–15 words of text
  - 16–31 words of text
  - 32–63 words of text
- Number of messages on the output line with:
  - 2 packets
  - 3 packets
  - 4 packets
  - 5 packets
  - 6 packets
  - 7 packets
  - 8 packets.
- Number of words to Host.
- Number of words from Host.
- Number of control messages from Host.
- Number of control messages to Host.
- Number of input messages to each destination.

c. Other

- Number of trace messages generated.
- Number of packets handled for each destination.

Interarrival Times

Upon command from the measurement center the IMP will record, for use in modeling the interarrival time distribution on the lines, the times at which the input interrupts on a given line occurred. A table containing approximately 60 IMP words will be used for this purpose. If more than 60 interrupts occur within 1/2 second, the measurements will be temporarily suspended for the remainder of the half-second interval and then automatically resumed.

B. Tracing

In the leader of each Host message is a trace bit which, if set to a one will cause that particular message to be traced through the network. Each IMP that handles the packets of that message will record the following information and transmit it to the network measurement center.

- Packet header.
- Time packet arrives at IMP.
- Time packet (buffer) is placed on the output queue.
- Time packet (buffer) is set up for transmission.
- Time the acknowledgement is returned or the time the buffer is placed on the retransmission queue. (For Modem Channels).
- Time the transmission is completed (For Host Channels).
- Output channel.
- Acknowledged or retransmitted.

To enable its tracing mechanism, an IMP must receive a command from the measurement center. This command will include the destination to which the trace packets should be sent and a link number for the IMP to use. If the tracing mechansim is not enabled, the IMP will ignore the trace bit. Once enabled, though, a second control message is needed to disable the mechanism.