Acnet Maximum Message Size

System enhancement

Thu, Oct 29, 1998

It has been decided that all front ends must be able to support Acnet messages sizes up to 16384 bytes. The main part of this effort for the IRM/Local Station systems was to increase the maximum datagram size supported via IP. In principle, it was not difficult, but heavy use of such a facility is likely to negatively impact system real time performance. It is hoped that any such use of large datagrams will be limited. This note explains the changes that were made to accomplish this support.

General system resources

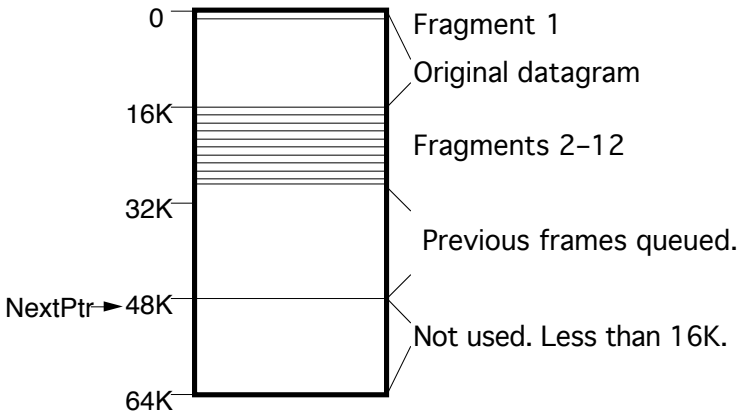
The system supports IP without major resources. The Linac stations that operate with 68020 CPUs have only 1 MB of dynamic memory plus 1 MB of non-volatile memory. The amount of dynamic memory that is available for general allocation is 384K. Part of it is used for resident system requirements. For modern systems, the amount of available dynamic memory is quite limited. The non-volatile memory is almost entirely dedicated for various system configuration tables.

Network transmission

The network transmit frame buffer area is currently 64K bytes in size. Frames are constructed in this memory in a circular fashion. It is assumed that the construction of frames to be transmitted does not result in a backlog even approaching 64K bytes. The general attitude of the software is to keep as much network communications active as is possible. When a frame has been built, it is queued to the network hardware. The software does not wait for its completion before building another frame. Multiple frames may be queued waiting for the network hardware to catch up. When a network transmit interrupt is received, and if another frame has already been queued to the network, then the next frame is passed to the hardware interface right away, without waiting for task level activity to proceed. The software expects to support a 15 Hz synchronous control system front end, so it is important to keep things moving.

The circular buffer area used for frame transmission, for the case of transmitting a 16K-byte datagram, must first be used to construct the full datagram, which must in contiguous memory within this 64K area. If there is not 16K bytes available from the end of the last frame transmitted, then it must be formed beginning at the start of the area. Thus, the worst case might be if there is just under 16K bytes available, when a 16K datagram must be formed. Then it will be determined that the datagram will have to be fragmented. (If it does not have to be fragmented, it is transmitted directly from this area.) The first fragment is sent from this area, and then the subsequent datagrams are formed in sequence in the region following the 16-Byte datagram originally formed. If there are 12 fragments, say, the first one is sent from the beginning of the entire datagram area. The rest of the fragments are built beyond the complete datagram, with the frame and IP headers copied and the IP checksum recalculated for each.

Graphically, the picture of this worst case might look as follows:



As soon as the fragments are built and queued to then network hardware, only a bit more than 32K remain for building additional frames. As soon as the first fragment has been sent, however, more than 48K is available for building more frames. This is the worst case scenario. Usually there will be even more free room available.

Network reception

For reception of very large messages, one must use IP, with its support for fragmentation and reassembly. On token ring, we use 2K-byte fragmentation, whereas on ethernet we use 1500-byte fragmentation. When a fragment is received, memory is allocated to hold it until all fragments that comprise a complete datagram have been received. When the last fragment is received, which is not required to be in order, the fragments are combined into a completed datagram in another allocated block of memory. With a 16K maximum datagram size, this block may have to be as large, which means that contiguous memory of that size must be available. As the fragments are copied in order to build the complete datagram, each fragment block is freed, of course. But it was necessary to have allocated a total of 32K bytes for the worst case datagram received. Building a 16K-byte datagram from 12 fragments requires about 17 ms for a 68020, or 5 ms for a 68040 (IRM). A 16K ping reception and response requires 54 ms on a 68020, or 13 ms on an IRM. These are significant loads for a 15 Hz front end.

Status

The new support software has been installed in a test node. It has passed preliminary testing, including supporting ping requests of 16K bytes as well as Acnet RETDAT requests of nearly 16K bytes. Some time will be needed for more testing and upgrading of the operational IRMs and local stations.