Applied Math 254: Computer Networks

Notes for Class 3: Definitions and Fundamental Properties

John M. McQuillan and David C. Walden

March 1975

1. DEFINITIONS

Some brief definitions are needed to isolate the kind of computer network we will be primarily considering:

Nodes. The nodes of the network are real-time computers, with limited storage and processing resources, which perform the basic packet-switching functions.

Hosts. The Hosts of the network are the computers, connected to nodes, which are the providers and users of the network services.

Lines. The lines of the network are some type of communications circuit of relatively high bandwidth and reasonably low error rate.

Connectivity. We assume a general, distributed topology in which each node can have multiple paths to other nodes, but not necessarily to all other nodes. Simple networks such as stars or rings are degenerate cases of the general topology we consider.

Message. The unit of data exchanged between source Host and destination Host.

Packet. The unit of data exchanged between adjacent nodes.

Acknowledgment. A piece of control information returned to a source to indicate successful receipt of a packet or message. A packet acknowledgment may be returned from an adjacent node to indicate successful receipt of a packet; a message acknowledgment may be returned from the destination to the source to indicate successful receipt of a message.

Store and Forward Subnetwork. The node stores a copy of a packet when it receives one, forwards it to an adjacent node, and discards its copy only on receipt of an acknowledgment from the adjacent node, a total storage interval of much less than a second.

Packet Switching. The nodes forward packets from many sources to many destinations along the same line, multiplexing the use of the line at a high rate.

Routing Algorithm. The procedure which the nodes use to determine which of the several possible paths through the network will be taken by a packet

Node-Node Transmission Procedures. The set of procedures governing the flow of packets between adjacent

nodes.

Source-Destination Transmission Procedures. The set of procedures governing the flow of messages between source node and destination node.

Host-Node Transmission Procedures. The set of procedures governing the flow of information between a Host and the node to which that Host is directly connected.

Host-Host Transmission Procedures. The set of procedures governing the flow of information between the source Host and the destination Host.

Within the class of network under consideration, there are already several operational networks and many network designs. The ARPA Network (Heart 70) is made up of over fifty node computers called IMPs consisting of about six nodes and about two Hosts per node. The Societe Internationale de Telecommunication Aeronautique (SITA) Network (Brant 72) connects centers in eight or so cities mostly in Europe. The European Informatics Network (EIN) (Barber 74), also known as Cost-11, is currently in a design stage and will be a network interconnecting about six computers in several Common Market countries. Some other packet-switching network designs include: Autodin II (Rosner 73), NPL (Davies 67), PCI (Auerbach 74), RCP (Despres 72), and Telenet (Auerbach 74).

Some of the more obvious differences among these networks can be cited briefly. The ARPA Network splits messages into packets up to 1000 bits long; the other networks have 2000-bit packets and no multipacket messages. Hosts connect to a single node in the ARPA Network and SITA; multiple connections are possible in Cyclades and EIN. Dynamic routing is used in the ARPA Network and EIN; a different adaptive method is used in SITA; fixed routing is presently used in Cyclades. The ARPA Network delivers messages to the destination Host in the same sequence as it accepts them from the source Host; Cyclades does not; in EIN it is optional. Clearly, many of the design choices made in these networks are in conflict with each other. The resolution of these conflicts is essential if balanced, high performance networks are to be planned and built, particularly since many future designs will be intended for larger, less experimental, and more complex networks.

As the ARPA Network is discussed at length throughout the remainder of this paper, we next summarize how the IMP in the ARPA Network performs its functions as a message switching center and interface between Host computers. Figure 1 shows a diagram of message flow in the ARPA Network and illustrates some of the terminology. The Host sends the

IMP a message with up to 8063 data bits. The source IMP breaks this up into packets with up to 1008 data bits. When the packet is successfully received at each IMP, an acknowledge or ack is sent back to the previous IMP. When the message arrives at the destination IMP it is reassembled, that is, the packets are combined into a message again. The message is sent to the destination Host and when it has been accepted, a Ready For Next Message which we abbreviate as RFNM is sent back to the source Host. It is also a packet and it is acknowledged. Several points are worth noting. First, acks are not actually separate transmissions, but are piggy-backed in packets to cut down on overhead. Next, packets on the phone line are checksummed in the modem interface hardware and the IMP employs a positive acknowledgement retransmission scheme. That is, if a packet is in error, it is not acknowledged. Then it is retransmitted until an acknowledge is received. Further, an IMP may send the several packets of a message out on different phone lines. For these reasons, the packets of a message may arrive at the destination IMP out of order and must be reassembled into the correct order for transmission into the Host.

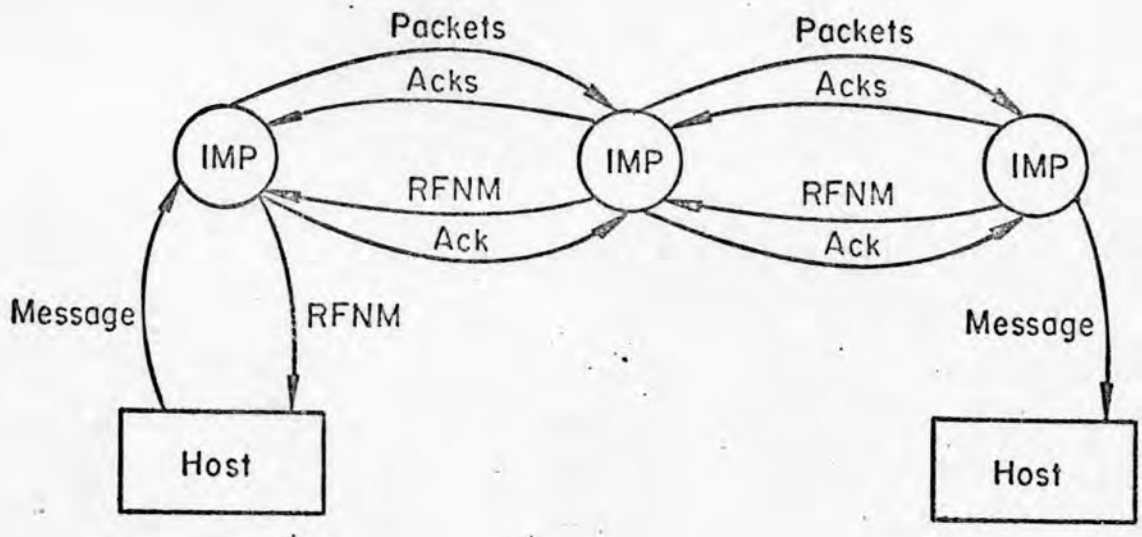

Figure 1 Message Flow in the ARPANET

2. FUNDAMENTAL ISSUES

In this section we define what we believe are fundamental properties and requirements of packet-switching networks.

We begin by giving the properties central to packet-switching network design. The key assumption here is that the packet processing algorithms (acknowledgment/retransmission strategies used to control transmission over noisy circuits, routing, etc.) result in a virtual network path between the Hosts with the following characteristics:

a. Finite, fluctuating delay -- A result of the basic line bandwidth, speed of light delays, queueing in the nodes, line errors, etc.

b. Finite, fluctuating bandwidth -- A result of network overhead, line errors, use of the network by many sources, etc.

c. Finite packet error rate (duplicate or lost packets) -- A result of the acknowledgment system in any store-and-forward discipline (this is a different use of the term "error rate" than in traditional telephony).

Duplicate packets are caused when a node goes down after receiving a packet and forwarding it without having sent the acknowledgment. The previous node then generates a duplicate with its retransmission of the packet. Packets are lost when a node goes down after receiving a packet and acknowledging it before the successful transmission of the packet to the next node. An attempt to prevent lost and duplicate packets must fail as there is a tradeoff between minimizing duplicate packets and minimizing lost packets. If the nodes avoid duplication of packets whenever possible, more packets are lost. Conversely, if the nodes retransmit whenever packets may be lost, more packets are duplicated.

d. Disordering of packets -- A property of the acknowledgment and routing algorithms.

These four properties describe what we term the store-and-forward subnetwork.

There are also two basic problems to be solved by the source and destination* in the virtual path described above:

e. Finite storage -- A property of the nodes.

f. Differing source and destination bandwidths -- Largely a property of the Hosts.

A slightly different treatment of this subject can be found in (Pouzin 74).

The fundamental requirements for packet-switching networks are dictated by the six properties enumerated above. These requirements include:

a. Buffering -- Buffering is required because it is generally necessary to send multiple data units on a communications path before receiving an acknowledgment. Because of the finite delay of the network, it may be desirable to have buffering for multiple packets in flight between source and destination in order to increase throughput. That is, a system without adequate buffering may have unacceptably low throughput due to long delays waiting for acknowledgment between transmissions.

b. Pipelining -- The finite bandwidth of the network may necessitate the pipelining of each message flowing through the network by breaking it up into packets in order to decrease delay. The bandwidth of the circuits may be low enough so that forwarding the entire message at each node in the path results in excessive delay. By breaking the message into packets, the nodes are able to forward the first packet of the message through the network ahead of the later ones. For a message of P packets and a path of H hops, the delay is proportional to P + H - 1 instead of P * H, where the proportionality constant is the packet length divided by the transmission rate.**

c. Error Control -- The node-to-node packet processing algorithm must exercise error control, with an acknowledgment system in order to deal with the finite packet error rate of the circuits. It must also detect when a circuit becomes unusable, and when to begin to use it again. In the source-to-destination message processing algorithm, the destination may need to exercise some

---

*The question of whether the source and destination nodes or the source and destination Hosts should solve these problems is addressed in a later lecture.

**See (McQuillan 74a) for a derivation and more exact result.

controls to detect missing and duplicated messages or portions of messages, which would appear as incorrect data to the end user. Further, acknowledgments of message delivery or non-delivery may be useful, possibly to trigger retransmission. This mechanism in turn requires error control and retransmission itself, since the delivery reports can be lost or duplicated. The usual technique is to assign some unique number to identify each data unit and to time out unanswered units. The error correction mechanism is invoked infrequently, as it is needed only to recover from node or line failures.

d. Sequencing -- Since packet sequences can be received out of order, the destination must use a sequence number technique of some form to deliver messages in correct order, and packets in order within messages, despite any scrambling effect that may take place while several messages are in transit. The sequencing mechanism is frequently invoked since it is needed to recover from line errors.

e. Storage allocation -- The fact that storage in the nodes is finite means that both the packet processing and message processing algorithms must exercise control over its use. The storage may be allocated at either the sender or the receiver.

f. Flow Control -- The different source and destination data rates may necessitate implicit or explicit flow control rules to prevent the network from becoming congested when the destination is slower than the source. These rules can be tied to the sequencing mechanism, with no more messages (packets) accepted after a certain number, or tied to the storage allocation technique, with no more messages (packets) accepted until a certain amount of storage is free, or the rules can be independent of these features. In satisfying the above six requirements, the algorithm often exercises contention resolution rules to allocate resources among several users. The twin problems of any such facility are:

fairness -- resources should be used by all users fairly;

deadlock prevention -- resources must be allocated so as to avoid deadlocks.

Notice that each of these algorithms is implemented by a distributed computation. IMP-to-IMP transmission control involves cooperation between every pair of neighboring IMPs. Source-to-destination transmission control is a distributed process between the pair of IMPs exchanging a message. Finally, flow control and routing are distributed algorithms which involve all the IMPs in the network. Each IMP makes

local decisions about global functions. The process of routing messages from source to destination involves all the IMPs in the network, in order that the best path for the message be chosen and agreed upon. Such distributed computations are quite different from conventional algorithms. They are not initialized, nor do they run to completion and halt. In a real sense, the ARPA routing calculation, flow control techniques, and so on, have been in progress for 5 years now, since some part of the Network has always been running for all of that time. These algorithms are continuously active processes on a large number of different processors. In fact, the number of processors and the interconnection between them is subject to change at any moment. They must run completely without human intervention. They perform contention resolution among bidders for shared resources, and they must do so without races or deadlocks. (We have also come to believe that it is essential to have a reset mechanism to unlock "impossible" deadlocks and other conditions that may result from hardware or software failures.)

ISSUES:	LEVELS:
BUFFERING	NODE TO NODE
PIPELINING	SOURCE NODE TO DESTINATION NODE
ERROR CONTROL	HOST TO NODE
PRIORITY	TERMINAL TO NODE
STORAGE ALLOCATION	HOST TO HOST (INCLUDES TERMINAL TO HOST)
FLOW CONTROL	USER PROCESS TO HOST
FAIRNESS	
LOCKUP PREVENTION	
ADDRESSING	

Low Delay

ave. round trip delay

High Throughput

fraction of total

bandwidth obtained

Low Cost

$/month membership

$/bit or message

High Reliability

% time disconnected

% messages undelivered

Nodal Bandwidth

bits/sec/node

Line Bandwidth

bits/sec/line

Nodal Delay

sec/node

Line Delay

sec/line

Nodal Storage

bits/node

NETWORK PERFORMANCE GOALS

LOW DELAY EOR INTERACTIVE TRAFFIC

HIGH THROUGHPUT FOR LONG DATA TRANSFERS

LOW COST FOR NETWORK CONNECTIVITY AND USE

HIGH RELIABILITY FOR NETWORK CONNECTIVITY AND USE