FERMILAB-Conf-91/90

Future Prospects for Networking in the United States

W. Lidinsky

Fermi National Accelerator Laboratory
P.O. Box 500, Batavia, Illinois 60510

March 1991

* Presented at the *Computing and High Energy Physics Conference*, KEK, Tsukuba, Japan, March 10-15, 1991.

Future Prospects for Networking In The United States

William Lidinsky
Fermi National Accelerator Laboratory

INTRODUCTION

When considering networks on the basis of geographical scope (i.e., local to wide area networks), the focus is often on the lower three layers of the International Standards Organization Open Systems Interconnection Reference Model (OSI/RM) [ISO]. By contrast, the upper three layers of the Model deal not so much with transport as with network services that are to a first approximation independent of geographical scope. Using this theme, the paper first discusses United States trends in local, metropolitan, and wide area networks. Following this is a discription of computer services needed and advances in services offered by networks; an area of increasing emphasis in the United States. Finally, consideration is given to the impact that such trends may have on high energy physics and U S HEPnet.

The time horizon considered here is limited to mid-decade or a little beyond because studies of technology forecasting beyond five to six years have been shown to be innacurate. This paper concentrates on computer networks. Through the mid-1990s, there appears to be no significant technical or cost advantage for integrating voice and data.

LAYERS 1 TO 3: LANs TO WANs

Local Area Networks

IEEE802.3[1] and IEEE 802.5[2] are the dominant local area networks (LANs) in the United States today. Within enterprises such as large corporations, universities, national laboratories, single sites are usually populated with catanets of LANs interconnected by bridges of which almost all are either of the transparent spanning tree or source routing class. These catanets are usually arranged as a hierarchy of LANs of either two or three levels with the apex LAN referred to by the term "backbone". Figure 1 shows a typical arrangement today.

One problem with these catanets is related to their use for distributed computing that requires remote procedure calls, inter-object communication, paging, and the like. In these cases the delay incurred in crossing bridges and accessing destination LAN segments becomes too long for reasonable performance, thus requiring careful placement of clients and servers. This problem will continue to exist until the latter half of this decade when single LANs will emerge that are capable of supporting these applications without regard to placement.

Another LAN, the Fiber Distributed Data Interface (FDDI), has emerged after almost a decade of work by the American National Standards Institute's X3T9.5 committee. FDDI chip sets are available from a few vendors now. Several additional vendors will announce chip sets this year. Table 1 provides information about FDDI and contrasts it with 802.3. Currently there are on the market 15 FDDI bridge

---

[1] IEEE802.3, although different from Ethernet™, is often referred to as Ethernet from which it was derived. 802.3 is the dominant CSMA/CD local area network today. The original Ethernet has and will continue to diminish in use.

[2] IEEE 802.5 is often referred to as the IBM™ Token Ring.

and media converters, 6 routers, 12 interfaces, 8 concentrators for a total of 41 vendors offering FDDI products. An independent FDDI certification organization, the Advanced Networking Test Center, is in the business of certifying FDDI products for interoperability. A number of additional products will become available within the next few months.

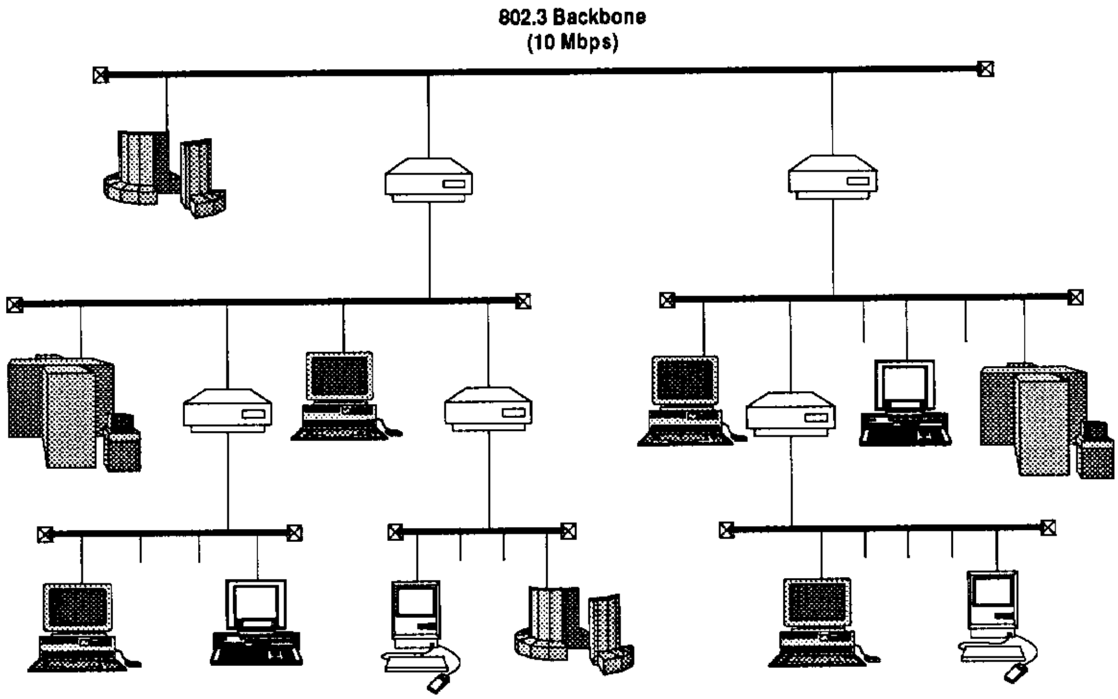

Figure 1: Typical LAN Catanet Today

	802.3 "New Std. Ethernet"	FDDI
Chip Set Cost in 1991	$50	$300
Interface Cost in 1991 (ISA, EISA)	$500	$6,000
Aggregate Bit Rate	10 Mbps	100 Mbps
Distance per segment	500 meters	2000 meters
Number of Stations per segment	20 - 100 depending on use	500

Table 1: 802.3/FDDI Comparisons

FDDI is being used in backbone applications at a number of universities and laboratories (see Figure 2). Because of (1) its high interface cost, and (2) the present inability of directly attached computers to take advantage of its speed, it has not yet found its way into general LAN use. Figure 3 shows the projected cost of FDDI interfaces - a curve that follows closely the history of 803.2. This cost information coupled with the knowledge that every known computer vendor either has or is developing an FDDI interface leads to the conclusion that by mid-decade 802.3 and 802.5 will begin to tail off and many computers, whether they be mainframes or workstations, will interconnect via FDDI LANs. FDDI will be the "Ethernet" of the middle of this decade.

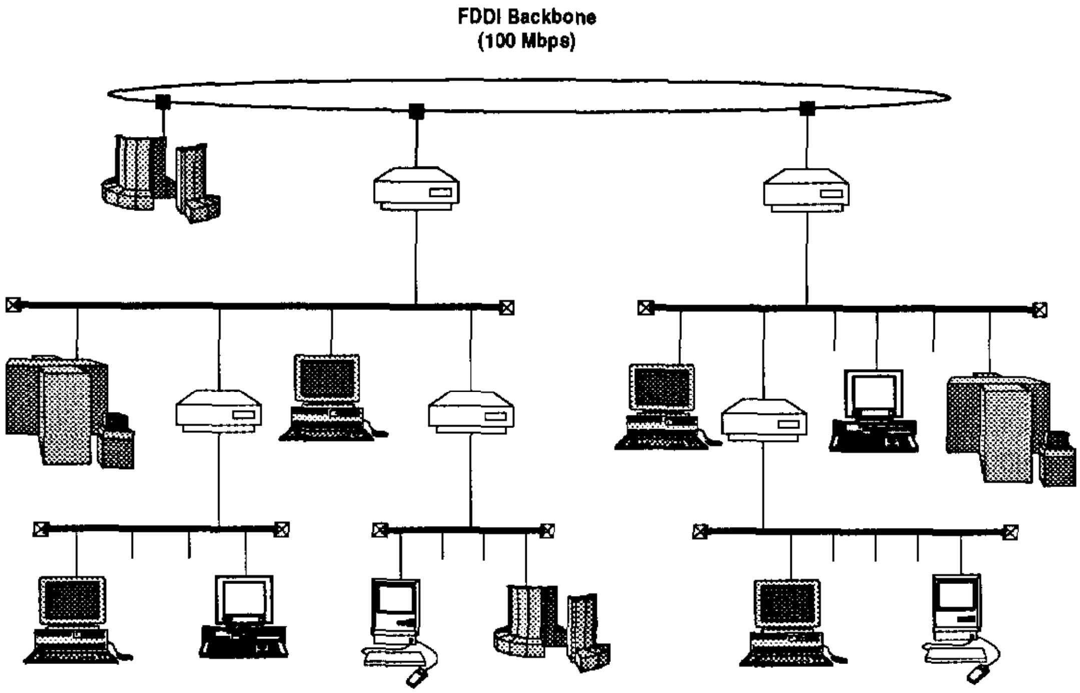

Figure 2: A LAN Catanet With FDDI Backbone

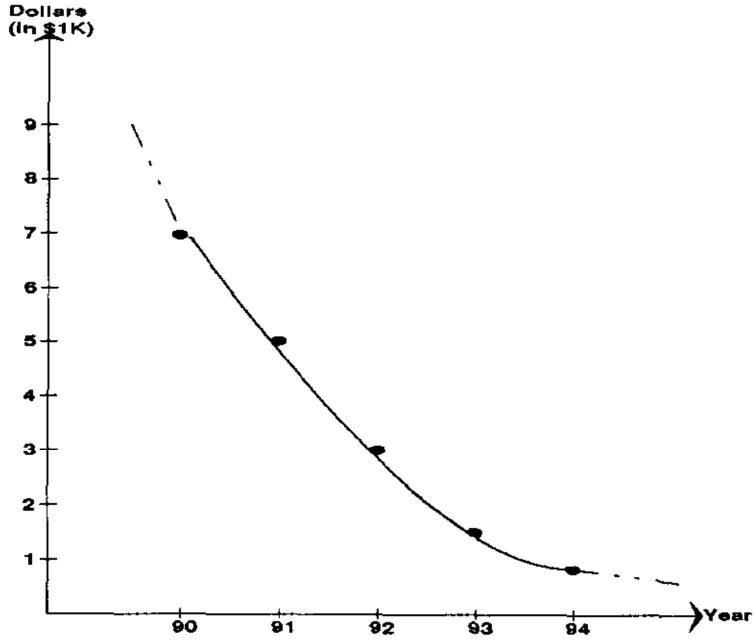

Figure 3: FDDI Interface Cost

Alternatives exist to FDDI backbones. A few vendors offer switching systems that can interconnect many LANs including FDDI and have aggregate bit rates in the range of a Gbps. As FDDI becomes the LAN of choice, these switches will provide the backbone function.

Finally, beginning in the last half of the '90s, networks will emerge that have the very low delay required for distributed processing and offer aggregate bit rates of in excess of 100 Gbps. Such a network (shown in Figure 4) will be capable of supporting thousands of workstations and mainframes and span tens of kilometers. These networks, which already exist in laboratories will eliminate the need for catanets of more conventional LANs and at the same time solve the delay problems associated with these catanets [Lidi].

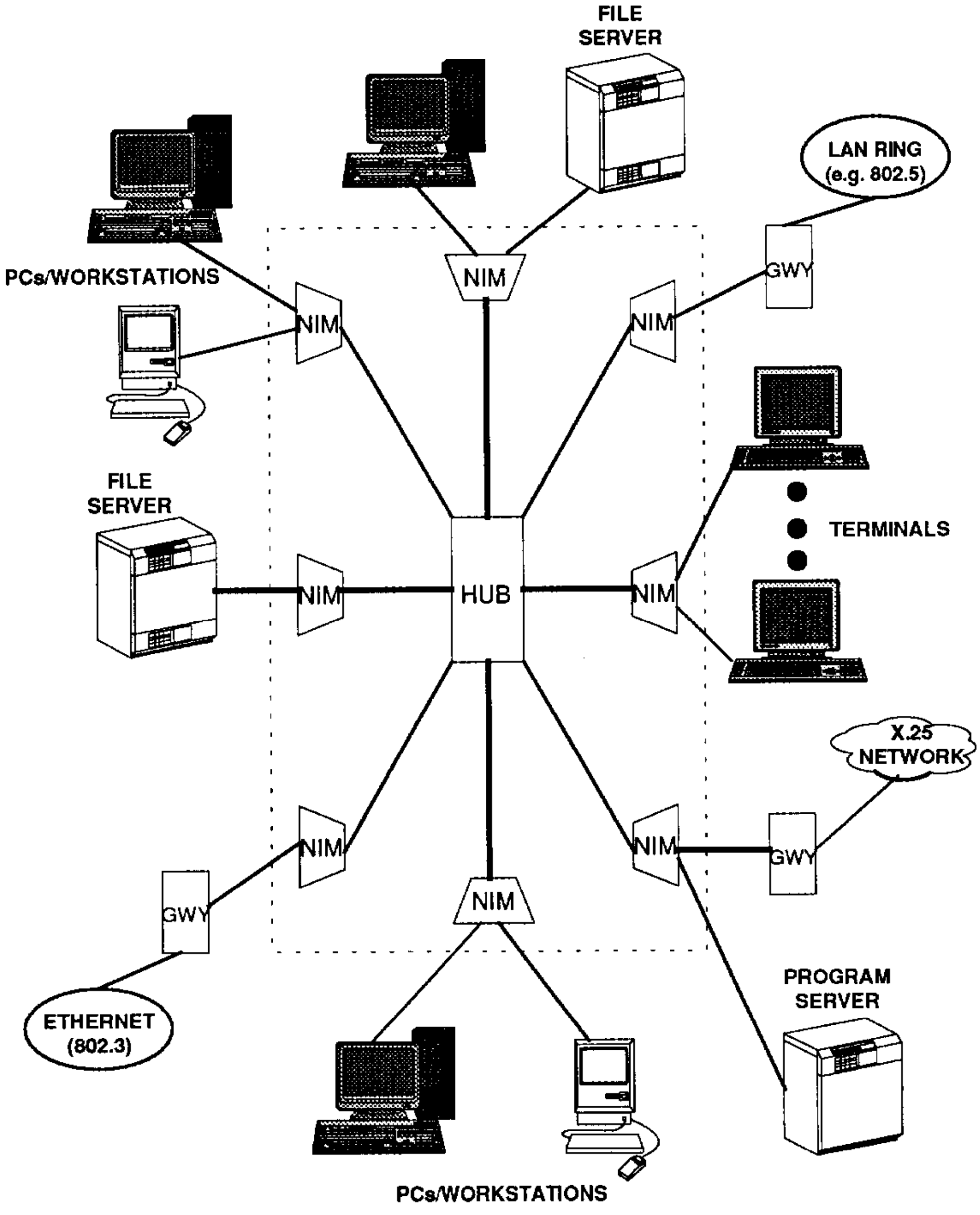

Figure 4: Large LANs for Campus Environments Beginning Last Half of Decade

Metropolitan Area Networks

The goal of Metropolitan Area Networks (MANs) is local area network performance over a metropolitan area with the network service provided by a common carrier. MAN technology centers not around packet switching but around cell switching. Two similar protocols are being standardized and energetically developed by the common carriers and their equipment suppliers: ATM and IEEE802.6. ATM is targeted at wide area networks while IEEE 802.6 is destined for MANs. These two protocols both have the same payload and headers sizes (48 bytes and 5 bytes respectively). The cell sizes are kept small intentionally in order to support time-critical transmission of voice and video as well as data. Cells carrying time critical information are given priority achieving what is called isochronous service. Some local exchange carriers intend to offer a service called SMDS (Switched Multimegabit Data Service) based upon IEEE 802.6 switching and T1 and T3 access by the middle of this decade. Figure 5 shows a MAN topology based upon SMDS. In this figure the busses run at T3 or higher rates and carry multiplexed 802.6 cells for up to several miles. The hub routes and otherwise processes cells among the busses.

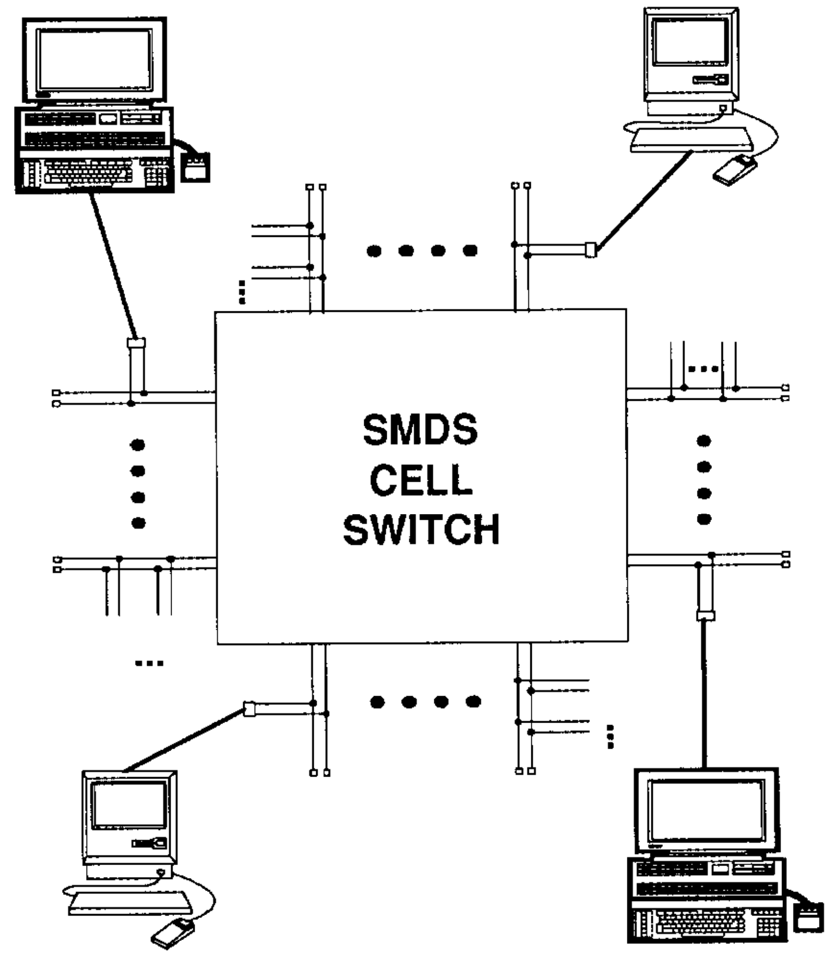

Figure 5: MAN Topology for SMDS Based Upon IEEE 802.6

Because significant implementation problems remain unsolved, the deployment of SMDS by mid decade is problematical. These problems include the ability to process millions of cells per second and privacy issues eminating from the fact that 802.6 has what is fundamentally a bus topology.

Wide Area Networks

In the United States today, wide area computer networking is provided from three sources: (1) common commercial carriers such as AT&T, MCI, SPRINT; (2) value-added network service providers such as Telenet and Tymenet; and (3) networks owned and operated by specific organizations such as the National Science Foundation (NSFnet), the Department of Energy (ESnet), and the many networks of private corporations. To date, common carriers have had little success in providing computer networks, although they have provided dedicated leased lines. Value-added providers have had somewhat more success especially with smaller organizations not able to support their own networks. However, organization-specific computer networks predominate - especially for large organizations.

WANs and Commercial Network Vendors

For the next few years commercial networking vendors, be they value-added or those offered by the traditional voice carriers, will not improve their relative position in computer networking in the US. Because of the increased availability and attractive costs of sophisticated routers and network management tools, these vendors offer little advantage over organization-specific networks. However, one possible offering that may change this is Frame Relay.

The traditional voice common carriers have focused on X.25, ISDN, and lately Frame Relay and Asynchronous Transfer Mode (ATM). X.25 is not widely deployed in the US and has demonstrated inadequate performance due to high overhead internal network implementations. ISDN will be widely deployed for voice; and it will be attractively priced. The commercial carriers hope that ISDN will also be used for data. However, ISDN is connection-oriented, making it a cost-ineffective match for the bursty unpredictable nature of data. Also, by then even Primary Rate ISDN will have inadequate bit rates for many data applications. It will be attractive, however, for local access to wide area networks from places such as a home or small office. ISDN, although it is not likely to play a part in wide area computer networking for medium to large organizations, is clearly the evolutionary basis for voice networks. As such it may still find some data uses beyond local access.

ATM is viewed by the common carriers and their equipment suppliers as the on-demand replacement for time division multiplexing (TDM). It relays 48 byte plus 5 byte header cells. It is being standardized by CCITT with the eventual goal being the support of voice, data and video at up to 45Mbps. ATM is targeted for initial deployment in the mid-1990s or beyond. It has, among other problems, severe cell processing rate requirements that cannot be economically met today.

Frame Relay [Mars] can be thought of as the high performance successor to X.25. It is the one offering from the common carriers that holds promise in the next few years. It is especially attractive for LAN-interconnect applications. Frame Relay operates with a variable length frame that can be the same length as that of the LANs or End-User-Systems (EUS) that it interconnects. It relays frames at Layer 2 using LAPD (I.122) which is very similar to LAPB, the Layer 2 protocol of X.25. Frame Relay is specified at up to 2 Mbps. It assumes an intrinsically more reliable network than was assumed when X.25 was defined (a good assumption now that digital switching and optical fiber links are widely deployed), and that higher layer protocols in the EUS will deal with errors, flow control, and the like. By eliminating the overhead associated with such functions, Frame Relay promises to have low latency. It will be offered in 1991 in the US as a Permanent Virtual Circuit (PVC) service with on-demand Switched Virtual Circuit (SVC) available beginning in 1992. SVCs will be setup using Q.931 as does ISDN. The LAPD header contains a Data Link Control Identifier (DLCI) for the purpose of establishing and transmitting over multiple virtual circuits on one physical interface . The common carrier network supporting Frame Relay performs the routing.

Organization-Specific WANs

Organization-specific networks are dominated by three protocol architectures: SNA, DECnet, and TCP/IP. The first two are proprietary network designs while the last has been widely (if not always officially) standardized. IBM's SNA is very popular in the commercial world and is probably the most widely deployed. It is, however, of the least interest to us because it is closely tied to IBM computing systems. DECnet and TCP/IP are the network protocols most extensively used within the scientific and education communities. Organization-specific networks are, in general, both dissimilar and separate from each other. Moreover, even instances of the same network type (e.g., TCP/IP) are often interconnected in a tenuous fashion. These networks are generally comprised of links leased from the common carriers and router nodes dedicated to the specific network. The Energy Sciences Network, ESnet, shown in Figure 6, is typical. [ES]

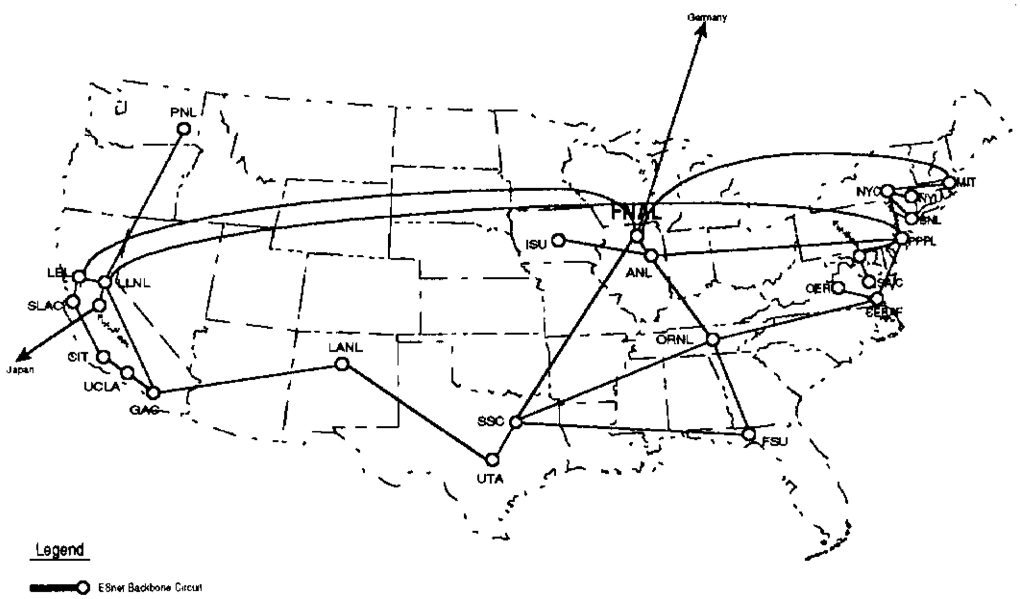

Figure 6: A Typical Wide Area Network - ESnet

Increasingly both in the US and beyond it's borders, internetworks (e.g., the INTERNET) are providing connectivity (but not interoperability) between networks. Growing numbers of organization-specific networks have attached to internetworks (through protocol-mapping gateways if necessary) achieving connectivity with various levels of interoperability. Moreover both internetworks and organization-specific networks are becoming multi-protocol due to the commercial availability of routers capable of concurrently supporting multiple protocols. ESnet, for example, now supports both DECnet and TCP/IP. In addition, there has been an increasing emphasis on standards. The federal government has an initiative (GOSIP) for moving agencies and their contractors toward ISO standards. Many vendors of private networks and network components are also moving toward these standards. A result is that TCP/IP will begin to be replaced by ISO's TP4/CLNP.

These trends clearly suggest that by mid-decade proprietary network architectures will largely be a thing of the past. This will have important benefits. Not only will interconnectivety and interoperability be improved but the golden handcuffs that bind computing to networking acquisitions will be gone. Replacing it will be a multiprotocol, multi-vendor *open interworking paradigm* with internetworking playing a key role.

WAN Links

Most of the organization-specific WANs in the US today lease 56Kbps or T1 links from a common carrier, with the more wealthy networks using at least some T3 links. Both reductions in cost and also computing needs are inducing migration from 56K to T1 and from T1 to T3. An interesting cost issue currently exists. In the past, due to cost, shorter links tended to be T3 while the longer links were T1. Presently, the cost of T3 links has dropped to the point where a T3 link costs no more than two or three T1 links, independent of distance. This is probably partially due to the fact that the common carriers in the US installed over 2 million miles of fiber during the 1980s and currently have excess capacity. However, tariffs and competition also play a part. As a consequence of these movements in cost, organization-specific WANs will move heavly to T3 links by mid-decade.

WAN Routers

Routers, as already mentioned, have begun to support multiple protocols. Hence, today we have networks supporting mixtures of proprietary and standard protocols on the same network. As new or standard protocols evolve and become popular or required, router vendors will find it in their best interest to support them as well. This then becomes an extremely attractive manner in which to evolve a network. Begin with a single protocol, add protocols as necessary, migrate users to the newer protocols when need for the old ones diminishes, and then delete the old protocols; a nice and tidy migration process. Supporting multiple protocols on the same router has some costs, but they are usually reasonable because links and much of the routing hardware is shared.

There are many unsolved issues associated with the move in the US toward the *open interworking paradigm.* Technical concerns include naming and addressing, routing protocols, congestion control, network features and the tools to build them, management, load balancing, accounting, and security. Management issues include evolution strategies, administration, control of development, and funding. These issues are being addressed and solved.

LAYERS 5 TO 7: NETWORK SERVICES

The connectivity provided by the lower three to four layers of the ISO/RM are not the only networking issues. Of growing importance in our heterogeneous computing environment is the interoperability among networked computing systems that are dissimilar not only with respect to computing hardware but also with respect to network-related software running on that hardware. These issues are the province of the upper three layers of the OSI/RM. When one gets to these layers, the bit rate so much discussed at the lower layers is seldom mentioned. In its place is response time - the time it takes for information to move from an application process on one EUS to an application on another possibly dissimilar EUS.

Layer 5 (Session) is the first layer that is explicitly aware of End-User-System processes and objects. It job is not to network computers but to network processes. It concerns itself with process synchronization issues such as checkpointing and token management. Remote procedure calls (RPCs) and inter-object communication (IOC) are yet other concerns of Layer 5. [Tane]

Layer 6 (Presentation) concerns itself with functions requested often enough to warrent finding general solutions. This layer deals with the syntax and semantics of transmitted and received information. Examples include character code translation, pixel image format translation, encryption, data compression, and the translation of files, networked file systems, and abstract date structures. NFS

(Networked File System) is an example of a distributed file system that will operate across a variety of platforms at Layer 6.

Layer 7 (Application) deals with a variety of commonly needed protocols. Networked virtual terminals provide a meta terminal model to efficiently deal with the 100s of incompatible terminal types in the world. An extention of this to bit mapped terminals is the network-based X-windowing standard. Other examples include remote paging, remote swapping and electronic mail. The advent of Graphical User Interfaces (GUIs) presented Layer 7 with yet another area of concern.

Some Trends

There are, in the U S some trends in Layer 5 - 7 and related issues. X-windows and more specifically X11R4 (X version 11, release 4) has become the defacto network windowing standard for workstations and personal computers. With the advent of 80486 and 68040 microprocessors, the performance differences between high-end personal computers and low-end workstations has dissappeared. UNIX is displacing proprietary operating systems because UNIX computing platforms are much lower in cost and the operating system and applications that run on it are easily portable. There are two competing UNIX standards in the U S with no clear winner: OSF UNIX from the Open Software Foundation and UNIX SVR4 from UNIX International. Each has their own graphical user interface: Motif from OSF and Open Look from UNIX International. Network File System (NSF) mentioned above is a network file system that runs and can interoperate across a number of dissimilar operating systems including several varieties of UNIX, MSDOS, MSDOS/MSWindows3.0, VMS, MacOS, and others.

Electronic mail is an area where standards are making substantial headway. CCITT's X.400 email standard and X.500 directory standard are being implemented by many vendors. By mid-decade email systems employing these standards will be widespread and email connectivity should be much better than it is now.

The above information indicates that useable amounts of Layer 5 - 7 software is becomming increasingly generic, running on a many different computers under several "high runner" operating systems. Selecting computers and software from among this network software will substantially increase the likelihood of both networkability and interoperability with the golden handcuffs described earlier.

Performance

One question that needs to be addressed relates to the network performance requirements for high requirement Layer 5-7 functions such as paging, swapping, windowed graphical image transfer, etc. Figure 7 quantifies this performance [Lidi]. For example, pages range in size from about 512 bytes to 2048 bytes or about 5K - 20K bits. To achieve paging over a network, the page response time needs to be of the order of 1 millisecond. To do this, bit rates of at least 10 Mbps are necessary. Remote procedure calls and inter-object messaging need response times that are an order of magnitude faster than paging but are also an order of magnitude smaller. Thus 10 Mbps is OK here also. Bit-mapped window transfers that are essentially computer graphic stills may be satisfied by 10 Mbps bit rates but may need 100 Mbps bit rates if the images are complex.

Keep in mind, however, that response time is not just $\frac{\text{work unit size}}{\text{bit rate}}$ but includes network access time and also network software processing time and I/O time in the EUSs as well. These latter times are often greater than the time spent in the network. Consider also that many protocols use multicast and broadcast for functions such as remote startup of EUSs, and server-server synchronization in client-server systems. Examples include the 802.1D spanning tree protocol, APR, and Bootp. The trend toward transparent distributed computing implies increased amounts of such traffic.

From Figure 7 one can conclude that 802.3 LAN and FDDI segments (10 Mbps and 100 Mbps respectively can satisfy page, swap, and rpc needs and some window needs if the networks are lightly enough loaded and there are few enough EUS per segment so that access time is kept to perhaps 10% of the needed response time. 802.3 and FDDI, because they are common medium networks, also deal with broadcast and multicast easily. The same cannot be said for wide area networks that have irregular mesh topologies.

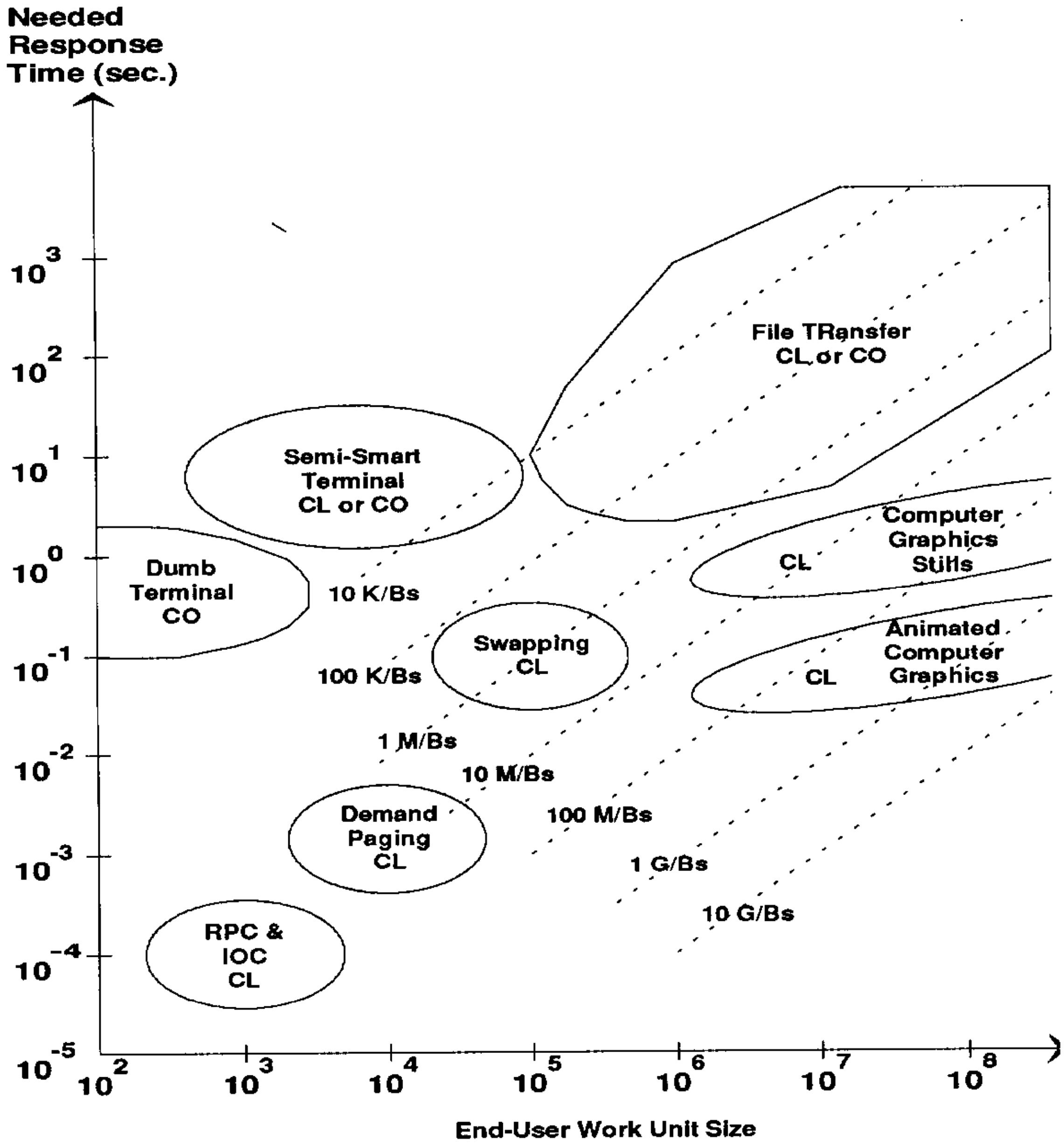

Figure 7: Needed Response Time for Certain Layer 5-7 Network Functions

Conclusions

In the mid-1990s networking in the United States will consist of LAN catanets interconnected by both organization-specific WANs and internets. WANs will also be interconnected by internets. MANs will not exist yet. Local area catanets will consist of a mixture of 802.3, 802.5, FDDI, and 1Gbps backbone hubs. EUS interfaces will be a mixture of 802.3, 802.5, and FDDI with FDDI growing most rapidly.

Wide area networks will, in all probability, be primarily of the organization-specific kind supporting multiple protocols with significant TCP/IP presence, growing TCP/CLNP use, but still with some occurrence of proprietary protocols. WANs will be used largely to interconnect LAN catanets which will have an important impact on the offered traffic and network interfaces. There will be small numbers of high performance interfaces and the traffic offered will already have been concentrated. Both catanets and wide area networks will interconnect via internetworks such as the Internet, providing more global connectivity.

MANs will still be in the future because of packet processing. Frame multiplexing schemes (e.g., TCP/IP, OSI CLNS, OSI CONS, and Frame Relay) require bridges, gateways, and routers to process frames at rates of between 10,000 to 60,000 frames per second depending upon link bit rate and packet size. Todays bridges, switches, and routers are constrained to less than 100,000 packets per second by a variety of constraints including processor speed, internal bus bit rate, and memory access time. Cell multiplexing schemes like 802.6 and ATM require frames to be fragmented, processed at rates of 2 million to 12 million cells per second and then recombined into frames. This capability does not now exist.

An important question left unanswered is whether Frame Relay will displace organization-specific wide area networks to any significant degree. This will depend on performance, deployment, and tariff considerations. If Frame Relay performs as intended, is widely introduced at both T1 and T3 rates, and is tariffed in a manner that is competitive to the costs of organization-specific networks, then it will succeed and organization-specific networks will begin to go away. Additionally, since Frame Relay will be implemented by the common carriers, they will handle internetworking which may begin to take over from the existing internetworks such as Internet.

Traffic within the catanets will undoubtly include rpc, ioc, swapping, paging, X-window traffic, and many flavors of distributed client-server computing. Multicast traffic will be large. There will be a need to extend this sort of network computing beyond the catanets. T3 bit rates in wide area networks will allow some of this but the irregular mesh topology of the WANs will inhibit multicast and the store and forward protocols of the WAN nodes will cause unacceptable delays resulting by mid-decade in the need for "cut-through" protocols in the WAN nodes. However, such protocols will probably not be standard. It is likely that combinations of low delay switch-based replacements for the catanets combined with small diameter WANs will be the evolution path to continent-wide low delay networks suitable for distributed computing.

HEPnet

Currently HEPnet consists primarily of (1) a wide area network (ESnet) with nodes at Dept. of Energy laboratories and major universities, (2) a number of mostly dedicated "tail circuits" emanating from the nodes to the sites of collaborators often at universities, (3) catanets of LANs at the node sites, and (4) links to sites beyond the United States borders. The ESnet links are all T1 and the tail circuits ranging for the most part from 9600 bps to 1.5 Mbps. The catanets are mostly 802.3 LANs interconnected with bridges and routers, but include other LANs as well.

ESnet presently runs two protocols (TCP/IP and DECnet) at Layers 3 and 4 and may add the ISO protocol pair TP4/CLNS over the next few years. This upgrade might take the form of DECnet Phase 5 which Digital Equipment Corp. promotes as being ISO standard compatible. It will probably upgrade its links to T3 in the same time frame. Depending on how Frame Relay evolves, it may represent an alternative to ESnet.

Within the next few years the catanets often have FDDI backbones and other FDDI segments as well.

Offshore links will likely need to upgraded to T1 by mid-decade.

A topic not mentioned so far is regional networks. These are wide area networks located and funded on a regional basis in the US where a region might be a state or a part of a state, or a contiguous small group of states. The technological ease and cost-effectiveness of creating WANs has caused regional networks to proliferate. The implication of this on US HEPnet is that the current tail circuits will likely be replaced in many cases by these regional networks.

Traffic demands on HEPnet will parallel closely those described above, and perhaps, as in the past, will lead in terms of network need.

References

[ES] U.S. Dept. of Energy, Office of Scientific Research, Scientific Computing Staff and ESnet Steering Committee, *ESnet Program Plan*, November 26, 1990.

[ISO] International Standards Organization, *Open Systems Interconnection-Basic Reference Model*, IS 7498, 1984.

Lidi] Lidinsky, W.P., *Data Communications Needs*, IEEE Network, v. 4, no. 2, March 1990, pp. 28-33.

[Tane] Tanenbaum, A.S., *Computer Networks*, 2nd ed., Prentice Hall, 1988.

[Mars] Marsden, P., *Interworking IEEE802/FDDI LANs Via the ISDN Frame Relay Bearer Service*, Proc. of IEEE, v. 79, no. 2, February 1991, pp 223-229.