--

--

PDIS and The Information Highway

Jonathan M. Smith
CIS Dept., University of Pennsylvania, Philadelphia, PA
`jms@aurora.cis.upenn.edu`

ABSTRACT

Advanced networking technologies offer opportunities for new and better PDIS.

1.0 The Information Hypeway and Reality

> ''*For a successful technology, reality must take precedence over public relations, for Nature cannot be fooled.*'' [4]

The so-called ''Information Highway'' attracts a great deal of attention, and yet a precise definition eludes us. This tends to be the result of definitions of convenience and overuse of the roadway analogy. There are two trendlines that are real. The first, and the one that will have the most impact, is the expansion of the Internet and commercial analogues. The impact stems directly from the scale and services of these networks, and the fraction of the population which can access them. The other trendline is the technological changes in higher performance networking elements. This trendline presents problems and opportunities for PDIS researchers, unlike the largely operational difficulties of expanding access. Since the backbone technologies will extend their reach to the home, the problems of scale are only postponed.

We now focus on Asynchronous Transfer Mode (ATM).

2.0 ATM as Information Asphalt

In the past half-decade three major developments in the high-speed networking [11] occurred: *(1) deployment of fiber optic technology; (2) availability of fast packet switches; and (3) network architectures which support paced cell traffic.*

Fiber-optic technologies such as Synchronous Optical Network (SONET) [7], which has commonly deployed rates of OC-3 (155 Mbps), OC-12 (622 Mbps), and OC-48 (2.4 Gbps) are now available, providing a high-bandwidth transmission infrastructure. Such infrastructures have been deployed and tested, as shown in the AURORA Gigabit Testbed's ATM network topology (partial), **Figure 2**, after an illustration of its geography in **Figure 1**.

The highest-performance networks are based on the idea of packet or cell-switching to share links between topologies and to share bandwidth between applications. The concept of Asynchronous Time-Division Multiplexing (ATDM), used in modern cell-switched architectures, makes link-sharing *programmable*.

Switched communications networks, as exemplified by the Public-Switched Telephone Network [10] use a digital Hierarchy based on Synchronous Time Division Multiplexing (STDM) internally. Higher-rate channels are aggregated from multiple lower-speed channels in a process called ''multiplexing''; the inverse process disintegrates the higher-speed channel into its lower-speed components. In a synchronous hierarchy, a clock is used for channel identification.

Synchronous multiplexing can thus be implemented entirely in hardware, yielding a multiplexing scheme which is simple, fast, and relatively inflexible with respect to bandwidth dynamics.

The computer communications community has attempted to address the dynamics of typical computer communications traffic through the idea of packet-switching [2] which is well-suited to the burstiness inherent in computer communications. Most data communications networks are optimized for this bursty behavior, especially local-area technologies [9].

With the Asynchronous Time Division Multiplexing (ATDM) [5] technique a hardware clock ''tick'' is replaced with a virtual ''channel'' identifier (VCI). Thus, logical sub-channels are assigned with a tag ('circuit ID'). This tagging allows burstiness to be accommodated by allocating more of the tags to a busy queue. In addition, this tagging can clearly be software controlled. The fact that this multiplexing behavior can be software controlled means that the problem has been transformed from a simple hardware system to a much more complex, but much more flexible, combination of hardware and software. The control of tagging is the key to resource allocation, and brings with it all the problems and opportunities of distributed control, resource allocation versus dynamics, real-time software, reservations, and policy versus mechanism. ATM is a specific implementation of the ATDM technique, using 48 byte ''cells'' of data with 5 byte headers containing the VCI.

As of about five years ago [3], the main cell-switching challenges were believed to be: *(1) switching at high speed (per-cell routing); (2) delivering of data to applications in their own data units at high speed; and (3) balancing ATM flexibility with distributed resource* *management to limit queueing and congestion.*

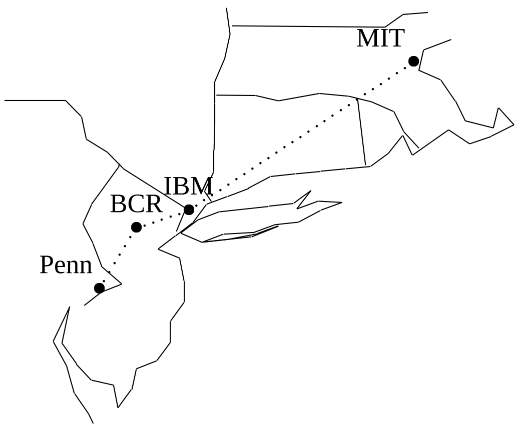

Figure 1: AURORA Geography

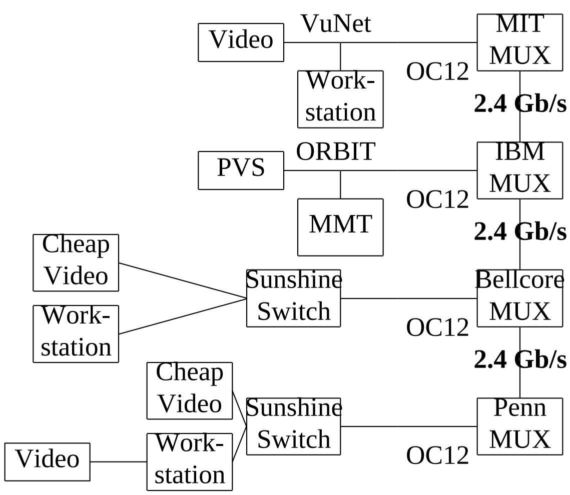

Figure 2: Partial AURORA Logical Topology

The first two research questions have been answered [1]. For example, the Sunshine [6] experimental switch prototype used for the ATM network infrastructure of the AURORA testbed is capable of 155 Mbps per port, with a scalable architecture capable of supporting large numbers of ports. An ATM host interface for computer workstations designed by C. Brendan S. Traw at Penn [13] has been shown to deliver application-to-application throughput of over 130 Mbps [12]. The resource management problems of queueing and congestion remain whenever resource sharing induced overloads (e.g., of an ATM switch outport port) can occur.

3. Effect on PDIS

A result of ATM is the ability to provide network configurations with both selectable Quality of Service (QoS) characteristics and the high bandwidth necessary for transporting and reproducing sensory data at remote sites. The implication is that very complex systems, incorporating sensory data from the real world, can be combined with models, simulators and control paradigms from the computational world. Such systems might be applied to many tasks in which integrating distributed information and specialized resources can speed the process of turning concepts into real-world artifacts. For example, a two-hand robotic system with a visual input system might be used to explore, visually and kinesthetically, a remote situation.

Common medium-speed Local-Area Networks, such as the Ethernet, are based on shared-bus media. Thus the degree of sharing, which is not known *a priori* can give rise to unpredictable network delays, which are unacceptable. Many other non-shared media either are too costly for general-purpose use or lack the bandwidth needed for the delivery of sensory data.

New network architectures based on high-speed packet switches have been developed. While a key design feature of these networks is the combination of their high (Gbps) bandwidth and the programmability of their multiplexing behavior, they offer, in addition, bounded delay in switching components and bounded packet interarrival times. This allows real-time systems to be built from real-time subsystems connected by network fabrics which provide time-constrained services.

This has enormous impact on the design of Parallel and Distributed Information Systems. Consider the following examples:

1. *Failure Detection* of links through timeouts becomes more reliable.
2. *Data Types* such as sensory data can become first-class elements of PDIS.
3. *Required QoS* can be *programmed*; for example, interarrival delays and bandwidths. Thus, biases in the communication desiderata for control portions of transactions and data movement portions could be exploited in the network.

An additional implication of high throughput networks is that many bottlenecks in systems relocate, in particular to elements such as disks. Storage systems based on such networks [8] may offer an architectural solution.

4. What needs to be done?

The control algorithms for the large-scale deployment of ATM are under development, in both theoretical and practical settings. Applications such as PDIS should make their needs and traffic characteristics known to

-- --

the ATM community, specifically those (such as the ATM Forum) involved in the design and standardization of signaling protocols for ATM switches and devices. This ensures that design assumptions reflect the needs of the PDIS community.

The Information Highway, at least the ATM portion, offers the potential for new and improved types of information systems built from distributed elements. This potential comes from performance, programmability, and distributed control algorithms for network behavior. This makes the network a first-class element in information systems design rather than a problem which must be compensated for.

5. References

[1] U.S. Congress, Office of Technology Assessment, ‘‘Advanced Network Technology - Background Paper,’’ OTA-BP-TCT-101, U.S. Government Printing Office, Washington, DC (June 1993).

[2] V. G. Cerf and R. E. Kahn, ‘‘A protocol for packet network intercommunication,’’ *IEEE Transactions on Communications* **COM-22**, pp. 637-648 (May 1974).

[3] David D. Clark, Bruce S. Davie, David J. Farber, Inder S. Gopal, Bharath K. Kadaba, W. David Sincoskie, Jonathan M. Smith, and David L. Tennenhouse, ‘‘The AURORA Gigabit Testbed,’’ *Computer Networks and ISDN Systems* **25**(6), pp. 599-621, North-Holland (January 1993).

[4] Richard P. Feynman, *What do YOU Care What Other People Think?*, Norton (1988), p. 237.

[5] A. G. Fraser, ‘‘Early Experiments with Asynchronous Time Division Networks,’’ *IEEE Network* **7**(1), pp. 12-26 (January 1993).

[6] J. Giacopelli, J. Hickey, W. Marcus, W. D. Sincoskie, and M. Littlewood, ‘‘Sunshine: A High-Performance Self-Routing Broadband Packet Switch Architecture,’’ *IEEE Journal on Selected Areas in Communications* **9**(8), pp. 1289-1298 (October, 1991).

[7] Anna Hac and Hasan B. Mutlu, ‘‘Synchronous Optical Network and Broadband ISDN Protocols,’’ *IEEE Computer* **22**(11), pp. 26-34 (November 1989).

[8] John H. Hartman and John K. Ousterhout, ‘‘The Zebra Striped Network File System,’’ in *Proceedings, 14th ACM Symp. OS Principles* (December, 1993), pp. 29-43.

[9] R.M. Metcalfe and D.R. Boggs, ‘‘ETHERNET: Distributed Packet Switching for Local Computer Networks,’’ *Communications of the ACM*, pp. 395-404 (July 1976).

[10] MTSs, *Engineering and Operations in The Bell System*, AT&T Bell Laboratories, Murray Hill, NJ (1983). ISBN #0-932764-04-5

[11] Craig Partridge, *Gigabit Networking*, Addison-Wesley, Reading, MA (1993). ISBN 0-201-56333-9

[12] Jonathan M. Smith and C. Brendan S. Traw, ‘‘Giving Applications Access to Gb/s Networking,’’ *IEEE Network* **7**(4), pp. 44-52, Special Issue: End-System Support for High-Speed Networks (Breaking Through the Network I/O Bottleneck) (July 1993).

[13] C. Brendan S. Traw and Jonathan M. Smith, ‘‘Hardware/Software Organization of a High-Performance ATM Host Interface,’’ *IEEE Journal on Selected Areas in Communications (Special Issue on High Speed Computer/Network Interfaces)* **11**(2), pp. 240-253 (February 1993).