TECHNOLOGY TRENDS

Why Today's Switches Can't Handle Tomorrow's Traffic

Bart W. Stuck

All of today's communications switches are designed to handle specific types of traffic. For example, central office and PBX equipment was designed to handle 64-kbps voice and low-speed circuit-switched data. The channelized T1 multiplexer is meant for packing multiple 64-kbps bit streams onto a single, higher-speed point-to-point transmission facility. Packet switches were designed for packet data traffic at 64 kbps and below. Each of these technologies was designed long before PCs and LANs entered the communications scene.

Megabit LAN speeds are becoming common for on-site data communications, and off-site calling is shifting from predictable, voice circuit-switched loads (one quarter erlang or nine CCS) to unpredictable peak packet loads. As transmission capacities expand with the deployment of fiber, new switching technologies are the inevitable consequence. Switching fabrics that have been optimized for one particular type of circuit traffic, such as today's common-control circuit switches, simply cannot handle the bursty, high peak-to-average packet transmission ratios generated by the new applications and carried by the fiber transmission links.

Although applications and transmission techniques are changing rapidly, networks are still made up of more terminals than switches. Consequently, it still makes economic sense to keep the terminals' interface to the network relatively simple. The challenge, then, is to develop a single packet switching fabric and associated software for the network nodes that can be cost effective for both small and large numbers of terminals, and for a wide range of bandwidth and bit-rate requirements (see Figure 1).

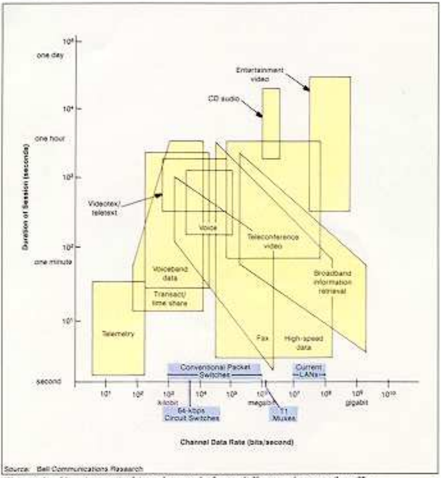


Figure 1. Varying switching demands from different classes of traffic.

Perhaps before the turn of the century, these new switching systems, capable of mindboggling speeds (e.g., terabits per second—a terabit is a trillion bits), will begin to appear, first in private and later in public networks. The switches being contemplated for tomorrow's networks will have to meet performance requirements that are fundamentally different than those we have today. They will have to economically support on-demand connections for voice, data and image, and they will serve tens of thousands of users who are sending packets from systems that operate at speeds from 10 to 1,000 Mbps. Although these switching systems are initially being developed to accommodate new levels of data and image communications, voice communications will undoubtedly follow.

This article highlights current computing and transmission trends that point to the need for new switching architectures, explains the inability of today's equipment to address the changing characteristics of network traffic and briefly introduces the new switching fabrics that are being developed.

Changing Applications

The growing dependence on computers in business operations, coupled with the impact of technological advances—such as high-capacity optical fiber, faster digital signal processors and increased-

Bart Stuck is president of Business Strategies, a consulting firm located in Westport, CT, and specializing in network computing and telecommunications.

TECHNOLOGY TRENDS

Future Electronic and Optical Switch Designs

The vendors' electronic switching design efforts are exemplified by today's so-called fast packet switch technologies, some of which are now entering the marketplace. Fast-packet switching traces its origin to switching theory work by Clos and Benes in the 1950s, to IBM- and Department of Defense-sponsored high-speed computer/memory interconnection work in the 1960s and 1970s and to high-speed switching work by Turner and Decina in the 1980s. All these efforts culminated in the development of Banyan switching, named because its operation can be compared to exploring the branches of a Banyan tree.

One strength of the Banyan switch is its use of distributed, concurrent control and information transport; one drawback is the possibility of sections within the switch experiencing contention and becoming bottlenecks under normal operating conditions, let alone under peak loading. This was ameliorated by adding a front-end Batcher sorting network (named for its inventor) to order the inputs into the Banyan switch, hence the Batcher-Banyan switch fabric. AT&T received a patent in 1985 on aspects of the Batcher-Banyan technology (J. S. Turner, Fast Packet Switching System, US Patent 4,494,230, June 15, 1985).

Ironically, StrataCom's IPX product, the first switch marketed as a so-called fast packet switch, bears no relationship to Batcher-Banyan designs. The IPX sends and receives frames of bits over links, and reads and writes these out of shared memory under the control of a single processor.

To support conferencing and broadcasting without increasing switch fabric contention, a great deal of design work has been devoted to putting Banyan "copy networks" in front of the Batcher-Banyan switches. The copy networks make copies of a master transmission for distribution to multiple receivers.

Asynchronous transfer mode (ATM) specifies a set of interfaces for multiple bit rate packet switching. A Batcher-Banyan switch fabric could use ATM techniques. Most people expect fast packet implementations of ATM to enter the market by the mid-1990s. AT&T, NTI, Siemens, Alcatel, Ericsson, NEC, and others in the common carrier equipment market, as well as IBM, DEC, Unisys/Timeplex, N.E.T./Adaptive, Motorola/Codex and others in the customer premises equipment market, are all pursuing some form of electronic switching fabric designs that feature multiple processors and distributed control.

Future Photonic Switches

The types of "photonic switching" (synonymous with "optical switching") include so-called knockout switching, star coupler wavelength division multiplexing, and variations of each.

In knockout switching, contention is resolved by the lower priority message trying again (after being "knocked out" of order by a higher-priority message). Knockout switching is appropriate for small numbers of inputs and outputs, but currently has trouble scaling to thousands of ports with a diverse traffic mix.

But the two switching fabrics being most actively pursued are star coupler and photonic optical space division switching. Both of these fabrics can cost-effectively handle a small number of ports and scale up to thousands of ports. In one approach, optical space switches might be employed to replace electronic switches.

Another approach is to use star couplers and wavelength diversity switching. Commerical product development activity at Gould and British Telecom/DuPont is exploring each of these approaches as well as hybrids. Currently, products based on optical switching technologies appear to be more than five years away from commercial realization.

capacity semiconductor memories—are the leading stimuli to the development of new switching technology. The rapid rate of acceptance of PCs, LANs and FDDI are some of the most visible examples.

The increasing importance to business of bandwidth-intensive applications was illustrated at a briefing IBM conducted for leading industry analysts in October 1989 at Research Triangle Park, NC. IBM presented the following list of LAN-related applications:

- **Engineering image and graphics processing.**
- **Photorealistic animation.**
- **Business document image processing.**
- **Production of technical documents.**
- **Teleradiology and clinical monitoring.**

These applications will intensify the requirement for gateways that can interconnect local area and wide area networks at bit rates of hundreds of megabits per second. IBM is not the only vendor to recognize the impact of these applications. Products that address these applications are available from providers such as Artel, IN-NET, Fibronics, Proteon, QPSX and Ultra Network.

Other business computing applications that drive the need for new switch technology are less dramatic but no less important. Two general types of data transfer activity—the massive after-hours batch database file transfers and the on-line business-day aggregation of many smaller transaction messages—are expected to increase the need for flexible, high-capacity communication.

For example, remote data center operations raise the volume and frequency of large file transfers. This in turn increases the demand

TECHNOLOGY TRENDS

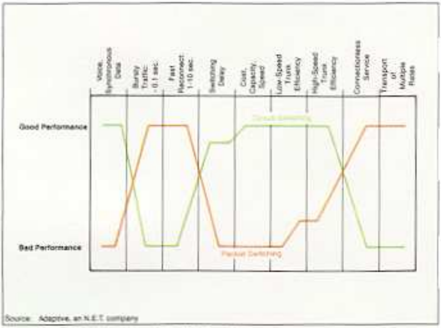


Figure 2. Technology characteristic template.

for higher-speed channel extenders, the specialized parallel interfaces that connect computers to off-site disk drives for data storage.

The current generation of channel extenders runs at a peak rate of three megabytes per second. Thus a data center with 100 disk drives could require a peak rate of [3 × 8 bits in a byte × 100 disk drives] 24,000 Mbps. Channel extender products from companies like IBM, DataSwitch, AT&T Paradyne, Computer Network Technology and others are likely to evolve to support a fiber optic channel interface, HSCI (high-speed channel interface), which IBM developed in conjunction with the Department of Energy's Los Alamos laboratory.

Other business computing needs that will require increased communication capacity include online transaction processing and the related batch file transfers of electronic checks and point-of-sale receipt images. Companies like DEC, IBM and Unisys each have special collections of processors and related products (front ends, cluster controllers, LANs, etc.) that target these business needs.

In short, demand is increasing for business computing applications that require increased bandwidth. These applications are, in turn, changing the nature of network traffic in such a way that current circuit switches will soon be overpowered (see Figure 2—from N.E.T.). The deployment of backbone fiber in public and private networks will also hasten the demise of the circuit switch.

The Fiber Factor

As shown in Figure 3, advances in transmission technology precede changes in switching fabrics. Fiber optic transmission technology, in particular, has been evolving steadily since the early 1970s.

A single-mode optical fiber with a 62.5-micron core and 125-micron cladding can transmit more than 10 terabits per second over tens of kilometers with error rates of less than one bit per billion. To put that into perspective, the entire peak daily load from the U.S. interexchange networks of AT&T, MCI, US Sprint, et al., as well as the local operating telephone companies, is currently estimated to require one terabit per second of digital transmission capacity.

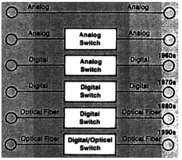
Current Bellcore projections suggest that the price of fiber optic cable may drop below that of copper within the next several years. Once fiber is installed, advances in electronics allow users to increase their transmission capacity by swapping out transceiver electronics, rather than installing additional fiber cable.

In fact, cable TV operators are already considering replacing depreciated coaxial cable transmission backbones with fiber and coax splitters with optical star couplers. As shown in Figure 4, every output port of a star coupler receives a copy of the signal received on any input port. The star coupler is the optical analog of the electronic computer bus (e.g., PC bus, VME bus, Multibus, Unibus); the principal difference is that the star coupler can handle tens of gigabits per second compared with tens to hundreds of megabits per second for the electronic bus.

> SONET permits different subrates to be extracted directly from the basic bit stream

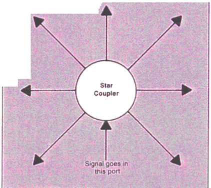
Star couplers are designed for broadcasting, but can be modified for point-to-point switching as well. Firms such as Gould and British Telecom/DuPont offer gigabit-per-second star couplers today. AT&T (with its StarLAN product line) and DCA (with its 10Net product line) also offer star couplers for interconnecting local area networks at speeds of several megabits per second, and higher bit rates will follow.

In general, fiber optic transmission standards share the frustrating multiplicity of the LAN and premises wiring standards that evolved in the 1980s—different types of fiber connectors, different cladding and core diameters, and different types of LEDs and lasers are all "standard." The CCITT is currently discussing two different interfaces, under the heading of Broadband Integrated Services Digital Network (BISDN). One is proposed to operate at 150 Mbps and another at 600 Mbps. Although a



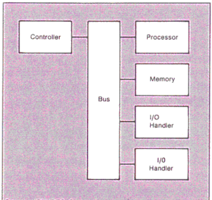
Source: AT&T Bell Laboratories

Figure 3. Switching technology follows transmission technology.



Source: BT/DuPont, Gould

Figure 4. Tomorrow's switches may be based on the optical star coupler.



Source: IBM, DEC, Compaq

Figure 5. All of today's switches use the Von Neumann architecture.

universal information outlet (*a la* the telephone system or the Xerox advertisements of the early 1980s) is desired by customers, vendors don't seem to think it's practical.

On a brighter note, however, more consensus in standards activity has been achieved with the CCITT's Synchronous Optical Network (SONET) transmission standards, which were completed in 1988. The relatively rapid development of SONET standards (about three years) illustrates their importance to the public carriers, who are the standards' principal sponsors and who have already deployed significant fiber installations.

SONET will be the basis for the next wave of carrier transmission standards and consists of a base frame supporting a serial bit stream rate of from 51.840 megabits per second (OC-1) up to 255 times this rate, or 13.2192 gigabits per seconds. Note that this rate does not even begin to tax the transmission capabilities of a single optical fiber.

The importance of SONET is that it permits a variety of different subrates to be extracted directly from the basic OC-1 bit stream—including DS0 (64 kbps), DS1 (1.536 Mbps), DS3 (28 DS1s) and the European CEPT and T3 transmission rates—without having to demultiplex them in stages. This capability allows digital cross connect functions (for drop and insert, grooming and consolidation) to be built directly into SONET equipment, thereby reducing the number of circuit packs and increasing overall reliability. Common circuit packs will reduce inventory costs and simplify operations support such as board swapping and, perhaps most important, will simplify restoration and protection. SONET-based equipment is already available, or will be within the next year, from a number of vendors, including TranSwitch, Unisys Network Systems, Artel, T3 Technologies, Optilink, Telco Systems, DSC and Network Equipment Technologies through its affiliate, Adaptive Corporation.

Fiber optic standards efforts did not end with SONET. The CCITT and the U.S. Exchange Carrier Standards Association's T1 Committee also have a variety of interface and higher-level framing standards activities under way under the general umbrella of either BISDN or ATM (asynchronous transfer mode). The point is that these will be built on top of the framing techniques that have already been specified by SONET.

Although SONET facilitates the transmission services that can handle multimedia integrated voice, text and image applications, these applications cannot really come to life without new switching architectures and related software. Today's switching equipment is inadequate in several respects.

Today's Switches vs. Tomorrow's Traffic

All of today's PBX, public and private network switches are designed for a single bit rate (64 kbps for PBX and CO switches, channelized T1 for public and private muxes) and use a single shared control processor, which relies on so-called von Neumann architecture (see Figure 5). Examples include AT&T's 5ESS or System 85/75, NTI's DMS-100 or SL-1, N.E.T.'s IDNX and StrataCom's IPX. In all these systems, the bus, switch matrix, and processor are serially reusable resources—each must finish one operation before starting the next. Consequently, each can become a bottleneck.

Advances in solid state technology and in interconnection technology (i.e., interconnecting the switch's internal components) have taken the form of higher clock rates, more innovative bus designs, additional stages of switching that create more paths between inputs and outputs (e.g., the time-space-time division switching of the

Abbreviations and Acronyms Used in This Article

- ATM – Asynchronous Transfer Mode
- BISDN – Broadband Integrated Services Digital Network
- CCITT – International Telephone and Telegraph Consultative Committee
- CCS – 100 Call Seconds
- CEPT – Conference of European Postal and Telecom administrations
- DS1 – Digital Service, Level 1
- FDDI – Fiber Distributed Data Interface
- HSCI – High-Speed Channel Interface
- OSI – Open Systems Interconnection
- SONET – Synchronous Optical Network
- TCP/IP – Transmission Control Protocol/Internet Protocol

N.E.T. IDNX) and the appearance of so-called frame relay switches (e.g., the StrataCom IPX). Although these advances have improved switch performance, they cannot eliminate the fundamental one-thing-at-a-time serial processor bottleneck or the limitations of single-bit-rate switching.

To handle tomorrow's needs for multiple bit rates on demand, switches will have to break the serial processor bottleneck and adopt radically different designs. Tasks that are performed sequentially by serially reusable resources in today's switches will be distributed among multiple resources operating in parallel and concurrently. By exploiting parallelism and concurrency, the basic processing rate of each element stays the same, but the total system throughput rate increases.

The Limits of LANs

Distributed control is not a new idea. The simplest Ethernet LANs take advantage of it, with each LAN station controlling its own access to the shared bus. Each Ethernet station listens, and if all is quiet transmits a 64-byte preamble then listens for the signal to return to see if it is garbled. If the signal comes back ungarbled, the station goes ahead with its transmission; if it is garbled, at least one other station is trying to transmit, and both stations stop and retry. As long as there aren't too many stations trying to access the bus simultaneously—a problem switch designers refer to as contention—Ethernet works fine.

Speeding up the clock rate—say from 10 to 100 Mbps—will not commensurately increase the Ethernet's throughput, since the 64-byte preamble transmitted for 51.2 microseconds at 10 Mbps must become a 640-byte preamble transmitted for 51.2 microseconds at 100 Mbps. More transmission capacity is being used to control station access, making less available for information transfer.

On the other hand, the design of token-ring LANs does permit performance to improve commensurately with clock rate increases. In a token-ring LAN, control is also distributed, but the maximum signal propagation delay—that between the two adjacent stations on the ring that are furthest apart—is much less than that between the two stations on the ends of an Ethernet bus. FDDI and the 802.6 standard DQDB are examples of token-passing techniques that can successfully increase their performance with increases in the clock rate.

But the single-processor bottleneck reappears as soon as any type of LAN server is added. Because a server must serially process each station's request, contention will occur when the number of stations or the number of their requests exceeds the server's ability to handle them. Adding more servers will not solve the basic problem, since the device that allocates requests among the multiple servers is now the potential bottleneck.

The central switch design problem isn't really quantitative but qualitative. New switching fabrics are crucial not just because larger files are being transferred more frequently at higher speeds, but also because these transfers will vary unpredictably in their bandwidth and bit rate requirements. Tomorrow's switches will have to handle concurrently the multiple bit rates and bandwidth requirements of voice, data, video and image traffic.

Design Issues for New Switches

One of the only principles of conventional switch design to remain unshaken is the engineering decision to maintain switching functions in network nodes rather than in terminal devices. The need for switches will persist until unlimited bandwidth eliminates the need to aggregate network traffic and until network traffic can route itself. In the meantime, today's designers of tomorrow's switching fabrics must face two important issues: the choice of packet over circuit switching to achieve the desired concurrency and parallelism and the need for new switching protocols and related software.

It now appears that packet switch-

Speeding up Old Switch Designs

Analog crossbar switching is almost as old as voice telephony, but switch designers have recently resurrected it. Embedding crosspoint arrays in silicon can dramatically increase the speed of the old design. One approach is discussed in a recently published paper by G. J. Scott and C. J. Anderson of IBM. (see "A GaAs 16×16 MESFET Crosspoint Switch at 1700 Mbps," Proceedings, 14th European Conference on Optical Communications September 11-15, 1988, Brighton, England, pp. 280-283.)

The authors describe a crosspoint electronic switching fabric connecting any of 16 inputs to any of 16 outputs, with the connection done using solid state gallium arsenide transistors. Each port can handle up to 1.7 gigabits per second, resulting in a total switch throughput of (16 × 1.7) 27.2 gigabits per second.

This exceeds the throughput of today's common control switches. For example, AT&T's 4ESS has 100,000 ports, each capable of 64 kbps, resulting in a total switch capacity of 6.4 gigabits per second. Although Scott and Anderson's design looks promising, it will be difficult to build a solid state crosspoint switch large enough to handle tens of thousands of inputs and outputs. Noise, timing, power dissipation, and interconnection technology (of internal switch components) all limit this approach to relatively small switches.

Even if these scaling problems were successfully addressed, the fundamental problem of a single controller remains—in this case it is required to arbitrate contention for paths through the switch□

ing will eclipse circuit switching because it can more readily handle multiple processors and fluctuating traffic requirements for a broad range of bit rates. In contrast, circuit switches are designed to rely on a single processor and to handle traffic at a single bit rate (64 kbps).

Conference bridging and broadcasting for voice, data and image applications are also better supported by packet switches than by circuit switches, since the latter are point-to-point rather than point-to-multipoint designs. Hybrid switch fabrics that will support concurrent circuit and packet switching have begun to appear (e.g., Netrix), but circuit switching will likely be used in the hybrids for a narrow range of transmission rates for specific applications such as voice. Companies like AT&T and NTI that have adopted so-called evergreen policies of not abandoning the installed customer base are thereby forced to adopt hybrid approaches, at least for the next 10 years.

Second, software development notoriously lags behind hardware technology. For example, IBM announced PS/2 in 1987 and Unix workstations in 1990. In both cases, the hardware is years ahead of fully featured, stable software.

Or consider the popular data networking protocol suites—SNA, DECnet, TCP/IP and OSI. These protocols have taken substantially longer to develop than the transmission and switching technologies they use. These protocols have also been developed specifically for handling information transfers over 64-kbps or lower bit-rate lines, far below SONET speeds. The industry's limited field experience with high-speed fiber optic switch fabric and transmission systems seems to virtually guarantees further delays in the development of appropriate switching, control and information transfer protocols and related software.

New Switching Technologies

Two different approaches, each with several variations, are being taken to the design of this single fabric switch (see "Future Electronic and Photonic Switch Designs," p. 100). In one of these approaches, the photonic signals carried by fiber optic transmission links are converted for switching to electronic signals. An electronic switch fabric is used that relies on electrical signal processing and control performed by solid state integrated circuitry.

In the second approach, the photonic signals that are carried by the fiber-optic network are not converted to electrical signals but are optically switched and multiplexed. Inside the switch, optical components—such as lasers, LEDs, detectors and star couplers—handle both signaling and information transport. Optical switch fabrics use electronic components to control the switch, and they also employ conventional communication techniques such as frequency division multiplexing, digital signal processing and equalization.

Conclusion

Both approaches are being pursued in the laboratory today by AT&T, NTI, Siemens, Alcatel, Ericsson, NEC, and others in the common carrier switching equipment market, as well as by IBM, DEC, Unisys/Timeplex, N.E.T./Adaptive, Motorola/Codex and others in the customer premises switching equipment market. Beyond electronically controlled switching of optical signals, these companies are also investigating photonic switching fabrics controlled via optical logic.

All these new switching technologies employ concurrent or parallel control for arbitration and information transfer, and hence they avoid the common control bottleneck that is intrinsic to current generation switching systems. Hybrid opto-electronic fabrics are closest to delivery, and should appear within the next one to three years, with optical fabrics following after another decade.