HIPPI

the first standard for high-performance networking

Stephen C. Tenbrink and Donald E. Tolmie

In the late 1980s, Laboratory researchers looked at the links in their local-area network—specifically, the connections between supercomputers, file storage, and other devices inside the machine room—and saw potential bottlenecks. The network had a maximum data-transmission rate of 50 million bits per second, one of the highest in the country, but numerical simulations on the fastest supercomputers were generating increasingly large data files. The time would come when simply moving the files would take almost as much time as generating them. Visualizing the data presented another bottleneck. For example, a researcher might want to simulate the formation and merging of eddies in a turbulent fluid and then view the simulation at a rate that the eye interprets as smooth, continuous motion—about 30 frames per second. To show 30 frames each second on a high-quality, 24-bit color monitor with a 1024-by-1024 pixel display requires a data-transmission rate of about 750 million bits per second—15 times higher than what the network could support in 1987.

The solution to that problem required more than just a system for high-speed data transmission. In 1987 there was already a sufficiently high-speed communication link—the proprietary Cray HSX interface, which operated at approximately 850 million bits per second. However, the Laboratory (and other supercomputing facilities) wanted to be able to transmit data freely among computers, monitors, storage devices, and so forth. A data-transmission system for that purpose would have to meet two requirements that the HSX interface did not meet. First, the equipment for the system had to be nationally standardized so that any manufacturer could give its devices high-speed ports compatible with the new data link. Second, since connecting every device to every other device with a separate cable was impractical, the new system had to be a network in which switches would connect any device to any other as needed. Therefore Laboratory researchers decided to develop and standardize a new generation of networks that would transmit data at about a gigabit (one billion bits) per second.

National standards are developed by committees of volunteers under the auspices of standards organizations such as the American National Standards Institute (ANSI). In 1987 Laboratory personnel presented their plans for a gigabit-network national standard to an ANSI group working on high-speed data transmission. That committee had been considering a more modest standard, specifying transmission speeds of 100 to 200 million bits per second, to be used in permanent connections between computers and data storage. The committee's reception of the Los Alamos proposal was decidedly cool—the members referred to advocates of gigabit speeds as the "lunatic fringe."

Two months later, proponents of gigabit networks from both the Laboratory and industry joined the committee. Within a year that committee produced the first draft of the standards for the new gigabit network, now called HIPPI, the high-performance, parallel interface. As soon as word got out that the new standard was on the way, vendors began to build components for the gigabit network. The first commercial products appeared in 1988: a HIPPI port for the IBM 3090 and a fiber-optic extender for HIPPI networks. In the same year, the Laboratory built a prototype of a HIPPI switch. By 1989, the first commercial HIPPI switches were marketed, and HIPPI ports on computers made by IBM, Cray Research, and the Thinking Machines Corporation were being developed.

The ANSI committee developed the HIPPI standards rapidly by basing them as much as possible on off-the-shelf technology. In November 1991, three years after the initial draft was delivered to ANSI, HIPPI became the first national standard for gigabit-per-second data transmission. (More precisely, the HIPPI standard gives specifications for transmitting data at two speeds: either 800 million bits per second—the minimum speed for a high-resolution, 30-frame-per-second movie—or twice that speed, 1.6 billion bits per second.) The HIPPI standard was adopted in a remarkably short time—a more typical time for acceptance of a new network standard is five to ten years. Today, HIPPI is the interface of choice for gigabit-per-second transmissions on most supercomputers and some high-performance workstations because it was the first standard at those speeds and because the technology is mature and stable. For instance, the Laboratory now has one HIPPI network that is in regular use and another that serves as a testbed for HIPPI hardware and software development. Newer gigabit network stan-

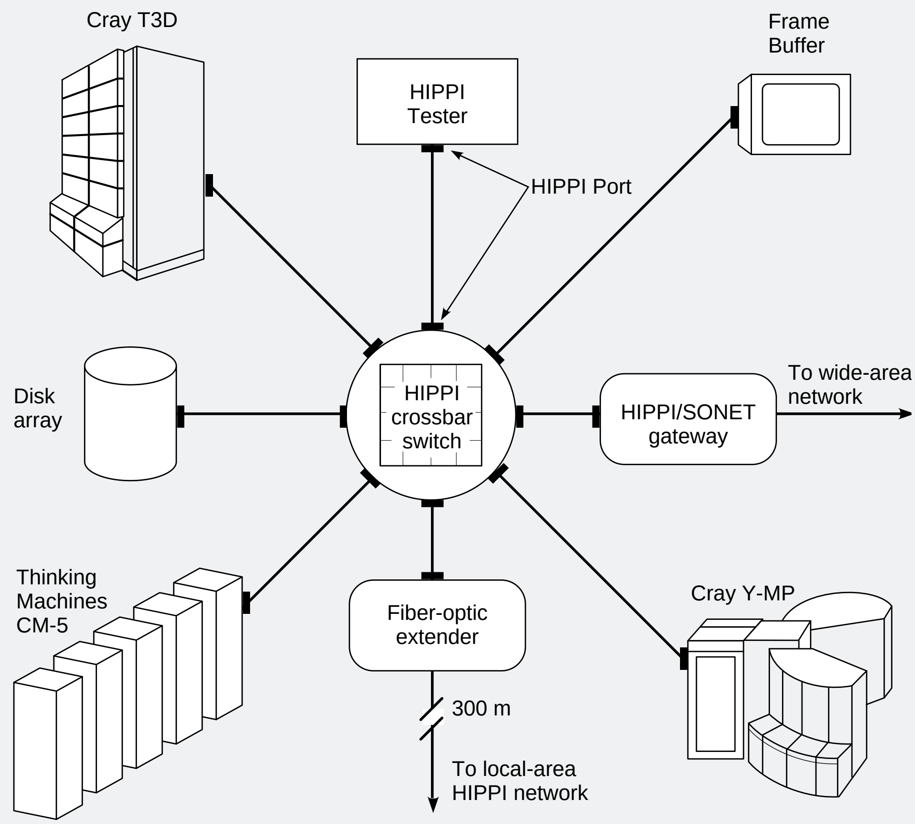

Figure 1. A Representative HIPPI Network
The figure shows a HIPPI crossbar switch connecting supercomputers, a high-performance disk array for file storage, a frame buffer, and a HIPPI tester (to monitor performance of the network). A fiber-optic extension connects this small network to another HIPPI switch, and a HIPPI/SONET gateway connects the local HIPPI network to a wide-area SONET network. HIPPI cables are sets of 50 twisted-pair copper lines (for 800 million bits per second) or 100 lines (for 1.6 billion bits per second). The data transmission is parallel; that is, many bits are sent simultaneously, one bit per twisted-pair line. Hence, HIPPI cables are similar to the buses, or ribbons of wires, connecting parallel ports on personal computers. HIPPI ports contain the circuitry that organizes and interprets the parallel transmissions, sending 32 bits of data on the 32 data lines and other necessary signals on the other lines. An 800-million-bit-per-second HIPPI port uses a 25-megahertz clock to synchronize transmissions, so a 32-bit word is sent every 40 nanoseconds. If a device cannot supply a continuous stream of data, HIPPI allows idle time between groups of words.

dards based on new technologies are emerging, for example, Asynchronous Transfer Mode (ATM), but such technologies are not yet readily available.

HIPPI Components

The physical elements of the HIPPI system are ports for the devices that are connected to the network, cables for the network links, and a switch to connect the devices as desired. Figure 1 shows a small HIPPI network. (The HIPPI Tester, the Frame Buffer, and the disk array are described below.) HIPPI was designed to be a "machine-room" network; the standards specify copper cables extending 25 meters or less. Vendors are now able to supply fiber-optic cables (governed by an implementors' agreement, not a HIPPI standard) that can extend HIPPI links to a maximum distance of 10 kilometers so that HIPPI networks can cover wider areas.

The crossbar switch used in HIPPI networks is shown in more detail in Figure 2. The switch has multiple inputs and outputs and a separate connection path between every input and every output so that many simultaneous connections can be made. The original crossbar switches had sets of vertical and horizontal conducting strips as inputs and outputs—hence the term "crossbar."

As the standards were being developed, Laboratory engineers were also developing equipment to help test and demonstrate HIPPI. The HIPPI Tester, for example, was developed to help ensure compatibility as well as to measure the performance of HIPPI ports developed by various computer manufacturers. The Tester, a briefcase-sized unit containing a lap-top computer and a special version of a HIPPI port, receives test signals from computers and checks to see if they were transmitted correctly; it also generates and transmits signals so that the received versions can be checked. Since the Tester is portable, Laboratory personnel were able to travel around the country to test the HIPPI ports being developed for various supercomputers. By 1990 the technology for the HIPPI Tester had been transferred to private industry. The Tester is easy to use, and now that it is manufactured in quantity, manufacturers and system designers can test their equipment themselves.

Another device developed at the Laboratory was the HIPPI Frame Buffer, a device for the high-speed display of computer graphics—which was the original motive for the development of HIPPI. A frame buffer consists of a large memory split into two buffers that each store a complete frame of display data. One buffer receives new data from a computer while the other buffer sends data to the display. When the "screen" buffer is empty, the roles of the two buffers are reversed. As long as the frame buffer contains data, the movie is not interrupted. The HIPPI

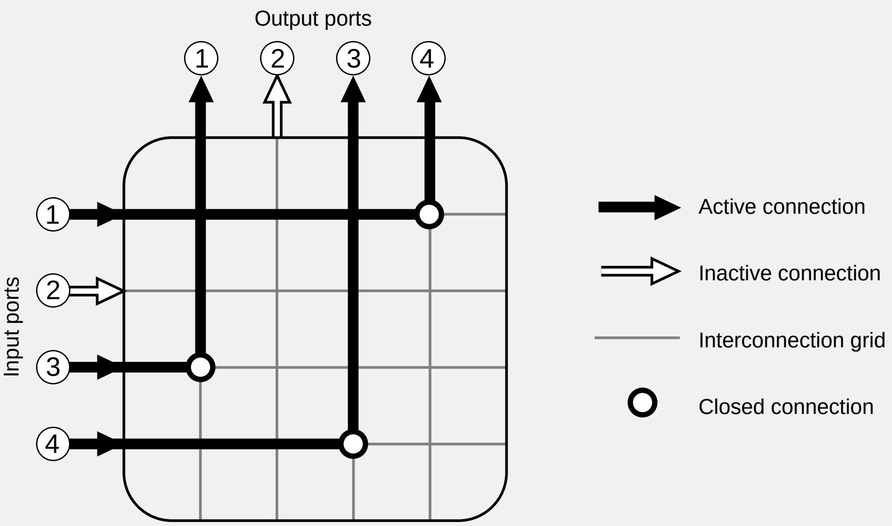

Figure 2. HIPPI Crossbar Switch

The figure illustrates a 4 x 4 crossbar switch, that is, one with four inputs and four outputs. There is a separate connection path between each input and each output, so the switch can connect any member of the network to any other as long as it makes no more than four simultaneous connections. (Similarly, a telephone switchboard, which is also a crossbar switch, can connect any telephone on the system to any other as long as the number of calls is not too great.) Three active one-way connections are shown: 1→4, 3→1, and 4→3. A typical HIPPI switch can connect up to 32 members in a HIPPI network (32 inputs and 32 outputs); several switches can be connected in a chain to make a larger network. Since one HIPPI connection can transmit data at 800 million bits per second, 32 simultaneous connections can transfer data at a maximum rate of 25.6 billion bits per second.

Frame Buffer was designed to be connected both to a supercomputer and to a high-resolution color monitor. The technology for this system was transferred to private industry in 1992.

As HIPPI networks were being tested, researchers realized that the bottlenecks in the networks were the supercomputers, not the HIPPI links. For example, a 30-second HIPPI framebuffer movie requires about 24 billion bits of data. On the one hand, a typical Cray supercomputer could send data fast enough, but it had a memory capacity of only a few billion bits—less than what even a short movie requires. On the other hand, the new massively parallel computers had memory capacity several times the Cray capacity, but were inefficient at protocol processing—packaging data for transmission and checking the data for errors. For example, when the massively parallel CM-2 was doing the protocol processing, it could supply data to a HIPPI port at only about 80 million bits per second or less.

In the last several years two new HIPPI components, the high-performance disk system (HPDS) and the crossbar interface (CBI), have been developed to solve these problems. An HPDS is an array of high-speed disks connected directly to a HIPPI port and controlled by a workstation. The array can hold more data and send data faster than the memory of a supercomputer—a typical disk array has a capacity of tens of billions of bits and can transfer data at about 500 million bits per second. Data are transferred from a supercomputer or file storage to the disk array and then sent through the HIPPI network for analysis or display.

The CBI can boost the rate at which a massively parallel computer supplies data to a network. A CBI, which is a small, special-purpose computer, was originally developed to manage HIPPI networks. However, a CBI can also perform protocol processing. If data are sent out "raw" from the HIPPI port on the CM-2 and the protocol processing is done in a CBI, the effective transmission rate is about 400 million bits per second—a five-fold increase.

HIPPI and Wide-Area Networks

In 1991, as the official HIPPI standards were being completed, the National Science Foundation and the Advanced Research Projects Agency were beginning to build and evaluate wide-area gigabit networks. The Center for National Research Initiatives in Reston, Virginia, coordinated the establishment of five testbeds for gigabit-per-second communication and asked the Laboratory to join one of the projects, the Casa Gigabit Testbed. The other participants in the Casa testbed are the San Diego Supercomputer Center (SDSC) and the supercomputer centers at CalTech and NASA's Jet Propulsion Laboratory (JPL). The goal of the project was to construct a HIPPI-based network connecting the four supercomputer centers and to use the network to investigate the applicability of distributed computing—several computers working together on the same calculation—to large computational problems. The problems designated were global climate modeling, seismic modeling, and chemical-reaction dynamics.

When the project started, each of the supercomputer centers had a local HIPPI network. To send HIPPI data over distances of hundreds or thousands of miles, the Casa network used another technology, the Synchronous Optical Network, or SONET, which the telephone companies had developed and were installing across the country as the next generation of telecommunication transmission. SONET is a network of fiber-optic transmission lines designed to transmit data for long distances at various rates of up to several billion bits per second. Engineers at the Laboratory developed a HIPPI/SONET gateway that moves data over the SONET network at the HIPPI rate of 800 million bits per second. The links between CalTech, JPL, and the SDSC were in place by August 1993; the Los Alamos link was brought up in June 1994 and connects a HIPPI switch at the Laboratory Data Communications Center at Los Alamos with a similar switch at the San Diego Computer Center. The project has now turned to distributed-computing research.

Current Research and Prospects

Laboratory researchers are now developing and evaluating new methods for gigabit-per-second transmission, developing links from HIPPI to other networks, and improving the performance of HIPPI networks. We are testing technology for a transmission standard called ATM, or Asynchronous Transfer Mode, which is under development and has the potential of reaching speeds of several gigabits per second (though the speed of most ATM transmissions in present testbeds is 155 million bits per second or less). When ATM transmissions can be made at speeds comparable to those of HIPPI, ATM will have a great advantage since HIPPI can transmit only computer data, but ATM fully supports voice and video as well as data traffic. Moreover, ATM transmissions can be sent on SONET lines. Therefore ATM is being viewed as the network technology for the future national information highway. We are planning to develop a HIPPI/ATM gateway that will allow us to link the current HIPPI network to any future wide-area ATM network.

Another new technology uses optical transmission techniques with no electronic switching; optical switches direct light from a source to a destination. The Laboratory is helping to develop an optical method called wave-division multiplexing, which can send thousands of messages simultaneously on a single fiber-optic cable. Each message has a carrier frequency and is encoded as modulations of the light at that frequency; this method is similar to radio or television transmission. Current research on such networks is directed toward increasing the transmission speed to gigabit-per-second rates and increasing the speed of switching.

At present, though, HIPPI is the only well-developed technology for gigabit-per-second networks, and as an ANSI standard it will continue to be useful for at least the next five or ten years. When the use of fiber-optic networks becomes widespread, HIPPI will probably serve as a local-network backbone that connects to long-distance links for special applications. The Laboratory has had a prominent and crucial role in the development of HIPPI and continues to refine its capabilities. We are also helping to develop new methods of data transmission for the coming tenfold increase in network speeds. □

Further Reading

Don E. Tolmie and John Renwick. 1993. HIPPI: Simplicity yields success. *IEEE Network* 7: 28–32.

D. E. Tolmie. 1994. High performance parallel interface (HIPPI). In *High Performance Networks: Technology and Protocols,* edited by Ahmed N. Tantawy. Kluwer Academic Publishers.

Stephen C. Tenbrink is a technical staff member in the Network Engineering Group of the Laboratory's Computing, Information, and Communications Division. Since 1980 his focus has been on the development of high-performance computer networks. Tenbrink was the section leader for the group of engineers that developed the HIPPI-based systems at the Laboratory. Tenbrink received his B.S. from the University of Denver in 1970 and his M.S. in electrical engineering from the University of California, Los Angeles, in 1977.

Donald E. Tolmie has been involved in the networking of supercomputers since 1972. He was the chairman of the American National Standards Institute Task Group X3T9.3, which was responsible for HIPPI and other high-performance-transfer standards. Tolmie is a technical staff member in the Network Engineering Group of the Laboratory's Computing, Information, and Communications Division. He received his B.S. from New Mexico State University in 1959 and his M.S. in electrical engineering from the University of California, Berkeley, in 1961.