Bringing Photonics to Broadband Switching

Stephen J. Hinterlong
Harry M. Hall, Jr.

Growing demand for high-bit-rate, large switching fabrics is forcing a reexamination of conventional electronic interconnection systems. A conclusion of this examination is that any large switching system will require thousands of connections at high rates to communicate the volume of data anticipated in future broadband switches. For example, a broadband switching fabric with 256 inputs and outputs each running at 622 Mbit/s would require a throughput of 160 Gbit/s. An approach to relieve the electronic interconnection bottleneck is to use optical connections. Photonics allows chip-to-chip and board-to-board communications at high rates to support building large switching fabrics. This paper discusses an approach that is being pursued to move the current state-of-the-art photonics work into a switching fabric capable of routing, in real-time, asynchronous transfer mode (ATM) cells received from 622 Mbit/s lines through a switching fabric that has 256 inputs and 256 outputs.

Photonic Connections

As system complexity increases, the demands on system construction increase as well. One of the reasons for reducing large circuit boards or standard interfaces to integrated circuits is to hide the complexity of the system construction within one or several easily connected packages. The complexity of the interface is reduced to the correct connection of the integrated circuit. Just as the drive toward consolidating interface circuitry on a chip has brought about integrated circuits with millions of transistors, further system complexity reductions will drive more and more functionality into the integrated circuit. The obvious benefit of more complex integrated circuits is the further masking of complexity from the designer. A difficulty with this evolution is that the devices could become so complex that to interface to these circuits would require more connections than are possible with electrical means. Although there have been impressive demonstrations of solder bump bonding in which 10,000 connections were achieved,[1] this level of electrical connectivity is not commonplace.

Another consideration is that these integrated circuits must be capable of processing the information presented to them at extremely high rates. For example, an examination of the trends from the earliest microprocessors to present processors shows an increase of nearly 100 times, to rates that are pushing 100 MHz. At these clock rates the interconnections must be treated as transmission lines capable of preserving the signal integrity. Transmission lines have fundamental physical requirements, such as spacing, thickness and width, dependent on the material underlying the transmission lines. Normally, the transmission lines must be dimensioned on a substrate to preserve a characteristic 50 ohm impedance. For effective transmission line performance, line density ranges from 100 lines per centimeter to several thousand per centimeter for silicon substrates.[2] With this range of interconnection density possibilities, the complexity of high-speed chips cannot be effectively served on either printed circuit boards or, possibly, even ceramic substrates. The effect is a connection bottleneck at the interface between the high-speed integrated circuit

and the printed circuit board.

An alternative approach to solve the electronics connection bottleneck is to provide optical connections of the integrated circuit to other integrated circuits. In this case, the interconnection density is limited by the imaging capability of the lens. This means that the interconnection density could be up to a million per square centimeter (i.e., one connection every 10 µm). In practice this number is greatly reduced, but densities of 250,000 connections per square centimeter were shown in a system at AT&T's 1992 Telecom demonstration. In the Telecom system, 4,000 simultaneous connections were made to the surface of an optical integrated circuit.[3] The capability to connect chips to chips at this density solves many interfacing problems.

Parallel Connections for High Throughput . As the numbers of components and connections of systems increase, the difficulty in assembling and maintaining these systems increases. Systems become impractical or excessively expensive to build. One problem encountered is the bandwidth required to move data between subsystems. Sufficient bandwidth cannot be supported by the state-of-the-art interconnection technology. One example is the lack of sufficient bandwidth on backplanes to support large broadband switching systems. An approach that has been used to attack the problem is the use of parallel data links to effectively multiply the bandwidth within any one channel. This is common in computer systems in which data and address busses are 16, 32 or 64 bits wide. In broadband communications equipment, the signal source is normally supplied as a single channel at extremely high bit rates — for example, 155 Mbit/s, 622 Mbit/s or 2.4 Gbit/s. Broadband switching systems may have tens to hundreds of these 155 Mbit/s or higher signals that all must be routed through a switching fabric. There are two approaches to this problem: de-multiplex all the high-speed signals to lower-bandwidth signals and send the data as parallel busses, as mentioned above, or use lower-density interconnections that are well engineered to support the high bandwidth of these connections. In each case there are physical limitations as to how so many high-speed connections can be made.

If preserving connections at their native rate, an engineer is faced with maintaining the transmission line integrity from printed circuit board to printed circuit

board. This is normally achieved by designing a backplane, circuit boards and connectors that are transmission lines. Unfortunately, there is a limit to the number of connections that can be made in this way. We have encountered a situation in which printed circuit boards can be designed that require more bandwidth than state-of-the-art backplane, circuit board and connector technology can support. To construct the system as designed, it must be partitioned into separate circuit boards, allowing all the bandwidth without exceeding the capabilities of the interconnection technology.

The other approach, de-multiplexing the signals to lower bandwidths, eases the problem, as the engineering of the connections is less stringent. This approach does not work, however, when the original bandwidth is large enough that even a substantial de-multiplexing results in a bus operating at 50 to 75 Mbit/s or above. For example, creating a 16-bit wide bus from a 2.4 Gbit/s signal results in each bus line operating at 150 Mbit/s. At this bit rate the bus design becomes a collection of 150 Mbit/s transmission lines. A critical factor for a bus is the disparity in the arrival times of the signals as they propagate from sender to receiver. This is sometimes referred to as the bus skew. In both solutions the system is limited in the aggregate bandwidth throughput. Many new experimental switching systems are being designed, and piece parts constructed, which will support aggregate throughputs of up to a few hundred Gbit/s. For these systems, a combination of detailed transmission line engineering and parallelism must be used to achieve the throughput.

Partitioning Systems Effectively. State-of-the-art technology has certain physical limitations that allow a maximum throughput per printed circuit board, backplane or connector. A natural partitioning, in which functional subsystems would be combined on printed circuit boards, may need to be further broken down to "fit" the capabilities of the technology. One example of this limitation is the number of pins available on a card edge con-
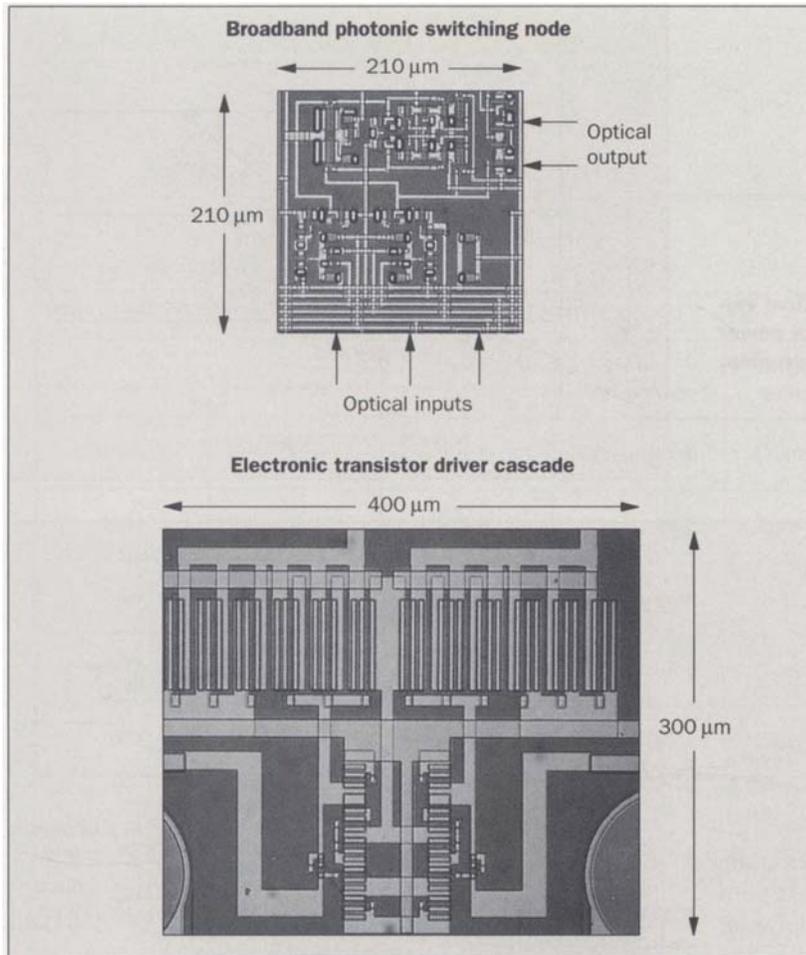
Broadband photonic switching node

210 μm

210 μm

Optical
output

Optical inputs

Electronic transistor driver cascade

400 μm

300 μm

nector. It is possible to design a subsystem that requires more pins than are available on state-of-the-art card edge connectors. The problem of having too few pins is only worsened by the need for many, if not all, of those pins to be correct impedance connections to maintain a transmission line.

A system in which these partitions have been forced by limitations in the connection technology leads to sub-optimal system performance. Usually the system will have more connections than the original design, and will be larger with more components. A particularly sensitive failure point is at the connections.

Technology Improvements Using Free-Space Systems

There are three obvious advantages of photonics in building large, complex systems: the increase in the number of potential connections, the decrease in the communications energy dissipated on chip, and the reduced skew in parallel systems. Each of these advantages is more and more important as the system achieves higher per-channel data rates, larger numbers of channels, or both. Just as the complexity of systems has been collapsed into integrated circuits, systems of complex integrated circuits will be interconnected with optics to mitigate the chip-to-chip interconnection bottleneck.

Optical Connections Through Imaging.
There are two reasons that the input/output (I/O) capacity of optics is superior to electronics. In optics, the entire area of an integrated circuit can be used to perform I/O functions. Optical I/O is theoretically limited by the resolution of the lens that is imaging the integrated circuit surface. Practical designs of optoelectronic circuits cannot, however, use the full resolving power of the lens system, because some of the circuit area is needed to perform processing. Optical connections can be made directly to the "processing" electronics. With optical I/O, no extraordinary provisions must be made to drive the large currents that are necessary for high-speed transmission lines. For example, an entire broadband, two-input one-output switching node requires the same surface area as the transistors used to drive a single output pin of a high-speed electronic circuit. Figure 1 shows one broadband photonic switching node and the cascade of transistors required to drive a single high-speed transmission line. The two are at the same scale. The switching node has two dual-rail optical inputs, processing electronics, and one dual-rail optical output. Everything that is required to select one or the other high-speed input, and to route that input to the output, is contained within the area of 210 μm x 210 μm. In contrast, the transistor cascade begins with the connection of smaller transistors to those of ever-increasing size, so that the current supplied at the output of the largest transistor is capable of charging a transmission line in the time allotted.

The capability to have many inputs and outputs greatly enhances the alternatives that an engineer may use to build connection-intensive systems. An apt comparison is to a microprocessor system that has time-multiplexed data and address busses of a fixed size — for example, 32 bits. The limited number of inputs and outputs available on the package requires time-multiplexing data and address busses. If the I/O limitation were eliminated, it would be possible to have non-multiplexed data

Figure 2. Optical versus electrical power dissipation on chip.
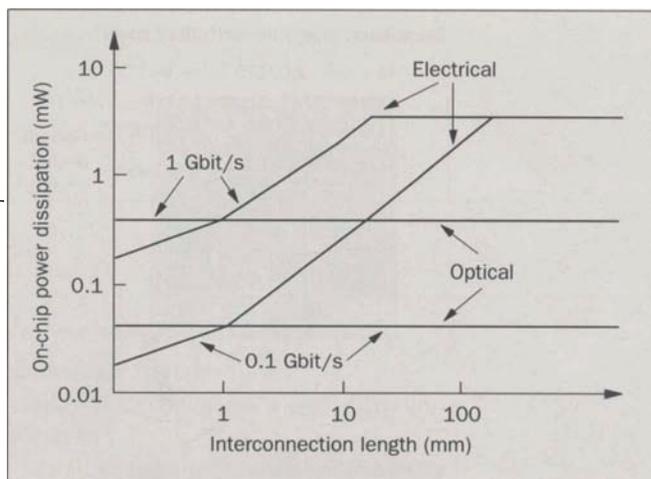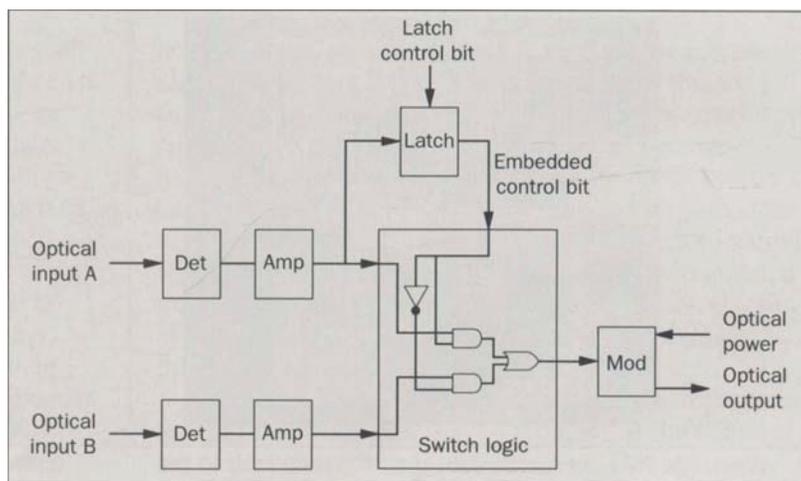

Figure 3. Schematic view of single FET-SEED broadband (2,1,1) switching node.

and address busses. Thus, a microprocessor architecture could be built in which the address and data busses are separate and distinct for memory, low-speed I/O, high-speed I/O and cache memory. With separate busses, the processor could be executing in cache with the cache-dedicated bus, while simultaneously updating memory from disk using the high-speed I/O-dedicated bus and the memory-dedicated bus. While multiple-bus hybrid processors can be built using multiple integrated circuits, a photonically interconnected single chip could be built to perform all these functions.

High-Speed Optical Signals Cut Power Dissipation. Another important factor in the construction of high-speed integrated circuits is that they generate more heat than equivalent circuits that run at lower speeds. A significant part of the difference is in the amount of power used to drive the transmission line interfaces. Shown in Figure 2 is a comparison of the power used to drive a transmission line versus the power used to change the state of an optical modulator used for photonic connections. The curves represent the power that is dissipated on the chip in order to send the data electrically, or to send the data optically, for 100 Mbit/s and 1 Gbit/s. In the electrical case there are two different calculations, very short distances and longer distances. For the shortest distances the power is increasing linearly with distance. This is because of a capacitance-per-length factor. Once the electrical communication distance is larger than a few centimeters, the calculation assumes a lumped transmission line. In this instance, the power is dissipated in the transmission line termination resistor. For this situation, the power is constant. In the case of the optical modulator, the capacitance is fixed. Thus, the power dis-

sipated is independent of the distance. For distances of more than about a millimeter, and data rates greater than 100 Mbit/s, less power is expended on-chip for photonics than for electronics. Thus, for distances of 1 mm and more, it is more energy-advantageous to send the data optically for bandwidths of 100 Mbit/s or more. This is particularly attractive when a circuit requires a large number of I/O. Consider the case in which the circuit design requires one thousand 100-Mbit/s I/O to communicate over a distance of 1 cm. The power savings of optical communications over electrical communications is over 150 mW. If the distance had been 10 cm, the savings would be over 1.5 W.

Ensuring That Signals Arrive at the Correct Time. The third advantage of optics over electronics is in minimizing arrival time differences for signals sent simultaneously on a bus. As electrical signals propagate, some encounter the load of receivers (connections) on circuit boards plugged into the bus. Other bus signals may not encounter the loading of receivers. Because of the differences in the load presented to the bus, some of the signals will propagate faster than others. The signals arrive non-simultaneously. If the bus signals are to all be used

Table 1: History of free-space photonics system for switching

Year	System	Device technology	Optics technology	Mechanics technology	System throughput
1988	System $_1$	S-SEED	Commercial	Commercial	10 kbit/s
1989	System $_2$	S-SEED	Commercial	Semi-custom	320 kbit/s
1990	Supercomm 90 System $_3$	S-SEED	Commercial	Custom	6.4 Mbit/s
1991	Telecom 91 System $_4$	S-SEED	Semi-custom	Custom	100 Mbit/s
1993	System $_5$	FET-SEED	Custom	Custom	2.5 Gbit/s

simultaneously, the result of this loading is that the wrong data will be used. This situation can be minimized through careful engineering. As the signal rates increase, the margin of error diminishes and the engineering of the electrical bus becomes critical. The optical approach to the problem is to propagate all the signals through the same medium, that is, through the glass lenses, beam splitters and "free-space" between the optical components. In this way, all the beams will incur the same delay; no subset of the beams will be delayed more than others. Here, again, optical systems benefit the designer at high data rates.

With the advantages of large input/output densities, lower on-chip power dissipation and reduced signal skew, AT&T has been experimenting with building photonics into broadband switching systems. Such systems offer the advantages of potentially smaller size and cost. The next section describes the system experiments that were performed. The latest experimental system constructed at AT&T to test this technology for applications such as broadband switching will be discussed in detail.

Visiting the Photonic Systems Museum

The development of free-space photonics technology for switching at AT&T has been accomplished in stages. Experimental systems were built to evaluate construction techniques, device technology, optomechanics design and alignment techniques and tolerances. There have been five experiments since 1988 to develop an understanding of free-space photonics for switching. Systems developed are listed in Table 1.

The earliest experiment was very large. The simple task of optically interconnecting two S-SEED (symmetric self electro-optic effect device) chips to form a switching node consumed an entire optical bench. This experiment was an existence proof that integrated circuits could be interconnected using beams of light.

The next experiment in the series was more aggressive. Four S-SEED chips were interconnected. Not only were more integrated circuits involved, but more S-SEEDs per chip were used in the circuit. This experiment was also intended to teach us how to build some optomechanical components in anticipation of the day when smaller systems would require fully customized optomechanics. This first attempt was somewhat successful because the system was assembled, aligned and operated. The optomechanics performed as expected, but the system was still too big.

The next attempt at building a system (System$_3$) employed fully customized optomechanics. While part of the scheme for mounting worked well (a V-groove component mounting), adjustment screws on the moveable components were very difficult to adjust. This system interconnected three S-SEED integrated circuits with 64 inputs and 64 outputs. System$_3$ was demonstrated at Supercomm '90.

System$_4$ was the first system to use the fully customized no-screw adjustment approach. The idea was to pre-build the sub-assemblies so that the optical axis of the component closely matched the physical axis. In that way, when the component was placed into the pre-aligned slots, the optical axis would also be pre-aligned.

Although this did not produce exact alignment, the optical alignment was close enough so that final alignment could begin almost immediately. For final alignment, optical wedges were used to move the beams of light instead of the optical components. A simple twist of an optical wedge resting in the pre-aligned slot was enough to place the spots onto the detectors and modulators. This system interconnected five S-SEED circuits, with more than 1,000 S-SEEDs used on each integrated circuit. This still stands as the largest number of S-SEED devices addressed in a switching system. This system was demonstrated at Telecom '91.

The latest system experiment is called System$_5$. The System$_5$ experiment was constructed to test the capabilities of the FET-SEED (field-effect transistor self electro-optic effect device) technology while allowing further development of the optics, optomechanics, lasers and overall packaging. In the earlier system experiments, the per-channel bit rates were limited by the switching energy of the S-SEED devices. The FET-SEED devices, however, incorporate a more sensitive receiver circuit to minimize switching energy. Thus the limited available laser energy used to switch the devices could switch more quickly the data inputs to the FET-SEED circuits.

Devices Technology: S-SEEDs vs. FET-SEEDs. The most notable new technology component is the FET-SEED[4] arrays, which are co-integrated optical receivers, transmitters and processing electronics that operate at a rate of 155 Mbit/s in the system. FET-SEEDs are also called smart pixels. This is because they are organized into a regular array of optical devices that is capable of processing data, locally, to any element in the array. Functionally, the individual devices are two-input, one-output switching nodes that will pass data at the 155 Mbit/s rate.

Shown in Figure 3 is a schematic view of a single FET-SEED switching node. Data is delivered to the node via beams of light at the receivers A and B. Depending on the state of the control memory, the received data is routed from either A or B to the output modulator. The shorthand naming convention for these kinds of nodes is (2,1,1) — that is, two inputs, one output from the node, and one path through the node. Despite its apparent simplicity, an array of nodes of this type can be used to build complex switching fabrics. The designer trades a more complicated switching node design for a more complex interconnection. Fortunately, the interconnection com-
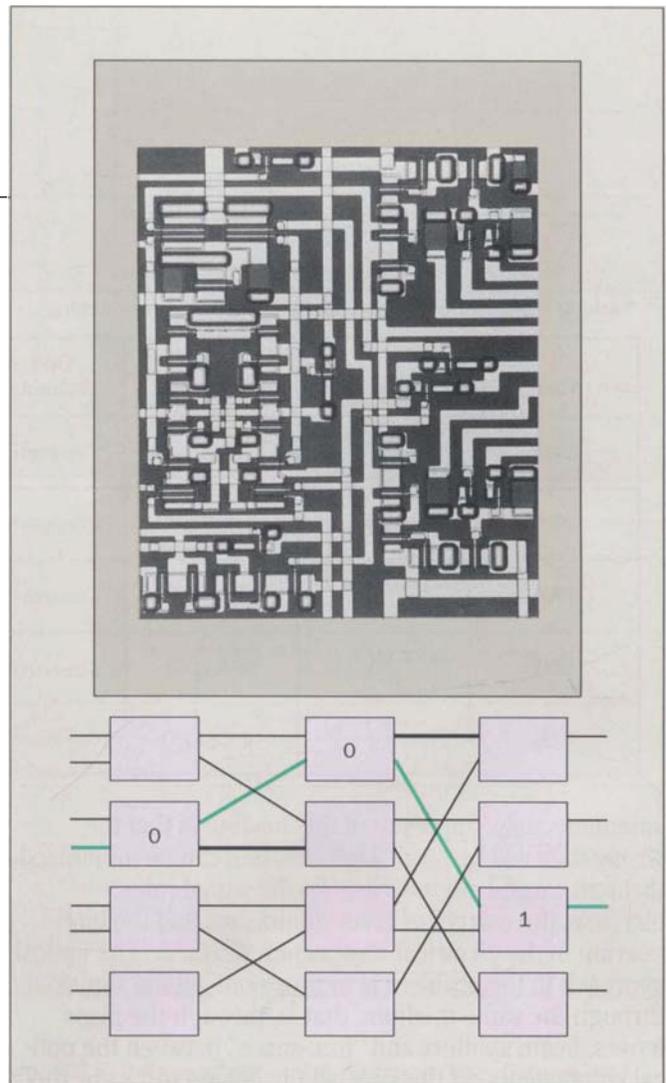

Figure 4. A picture of the surface of a FET-SEED chip in which one pixel or node is defined by power supply and control busses. Below is the interconnection pattern for a Banyan network with 12 nodes and three stages.

plexity problem is conveniently solved by the advantages of free-space photonics, as discussed earlier.

The node acts as a programmable (2,1,1) node because a control memory has been integrated into the node itself. The control memory selects the data from input A or input B for delivery to the output. As shown in Figure 3, if the state of the latch is set to a logical "1," data from input A are routed to the output. If the state of the latch is a logical "0," data from input B are routed to that output. Because the control memory is integrated into the switching node, it may be changed at the native rate of the data that is flowing through the node. Only one bit-time is required to change any node, and thus the entire multistage network can be reconfigured in twice the number of stages times the bit period. First, the net-
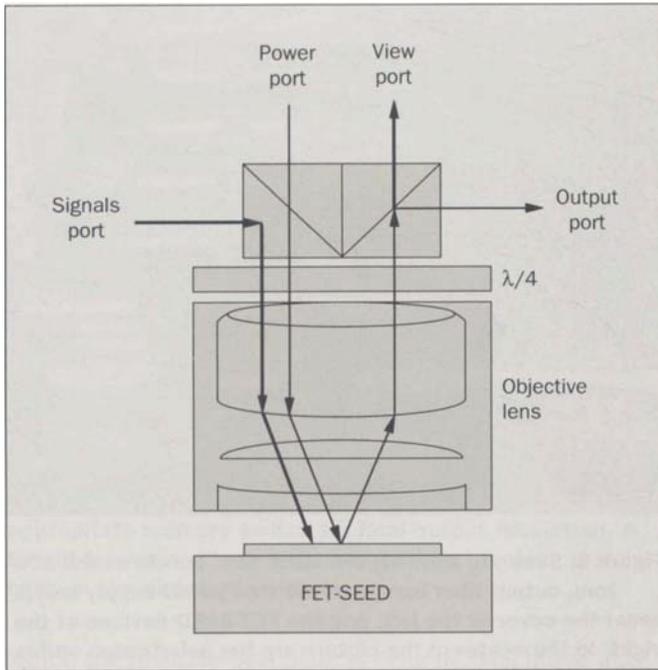
Figure 5. Optical data flow through a FET-SEED showing two input ports and two output ports.

work must be configured so that control data can reach each control latch. The configuration of the network can begin at the last stage and progress to the first stage. For example, if the native bit rate in the fabric is 155 Mbit/s (bit period of 6.5 ns) and the fabric has five stages, the reconfiguration time is,

Reconfig Time = 2 (number of stages x bit period)

Thus, for System$_5$,

Reconfig Time = 2 (5 x 6.5 ns) = 65 ns.

System$_5$ has fewer stages than would be needed to interconnect all of the inputs and outputs in a large fabric. A fully interconnected fabric would require no more than three times the number of stages in System$_5$, and its reconfiguration rate would be ~200 ns. This is a very fast reconfiguration rate, because all the nodes comprising an array are simultaneously being reconfigured.

Shown in Figure 4 is a photograph of the FET-SEED (2,1,1) switching node chip. One entire switching node is evident on the chip (with power supply and control lines defining the node). Surrounding this node are multiple identical switching nodes. In the center of the picture, on the right side of the single node, are the two receivers for the node data input. On the left side of the node is the latch circuitry and the modulator output. Both the receiver inputs and the modulator output are dual-rail signals, and thus they have two optical windows each. In operation, when the signals from the receivers have propagated through the electronics to the modula-

tor output, the modulator windows are illuminated to extract the output signal in the optical domain. The modulator windows are uniformly illuminated with optical power beams, and the state of the bit at the modulator will be encoded as beams of light of unequal intensity and sent to receivers on a succeeding array.

Optoelectronic broadband switching nodes like those in Figure 3 are integrated on a single gallium arsenide (GaAs) chip in arrays of 16 elements. The nodes are arranged as a 4 x 4 array.[5] With 16 nodes, each chip can receive 32 input beam pairs and route 16 output beam pairs. The beam pairs are routed in a Banyan interconnection scheme, like that in Figure 4. Just as in Figure 4, in which 8 beam pairs were routed to any of 4 places, with 32 input beam pairs and five stages of FET-SEED devices, any input may be routed to any of the 16 outputs. As a demonstration of routing in this fabric, consider the highlighted path that passes entirely through the Banyan network in Figure 4. In order to route as depicted in this path, the nodes involved in the path must be configured as shown. In the left-most stage, the node must have a "0" stored in the control latch to select the lower, or "B," input to the node. At the output of this node, the optical data will be split and sent to two other nodes in the center stage. For this path, the center-stage node must also be programmed to route the lower input to the output. The other highlighted path from the left-side node has the same data as the completed path, and cannot be used for another data flow. Routing from the center stage to the right-most stage occurs as in left-to-center-stage routing. In the right-most node, the control latch must be set to the "1" state to select the "A" input to route to the output.

Just as the example above has illustrated the strategy to route from an input to an output for a single-path call, multiple calls can be established simultaneously in a fabric. When the fabric needs to be set up and an entire array reconfigured, an external controller asserts a latch-control signal. The latches all capture the control data. The center stage nodes are all configured simultaneously in the same way. An array of data is loaded into the center stage, and the center-stage control latch signal is asserted. Finally, the left-most stage is configured with an array of data, and the latch - control signal is asserted.

The fabric in this example is not representative of a desired fabric configuration. It is evident that the fabric in Figure 4 is a blocking fabric. A call could not be
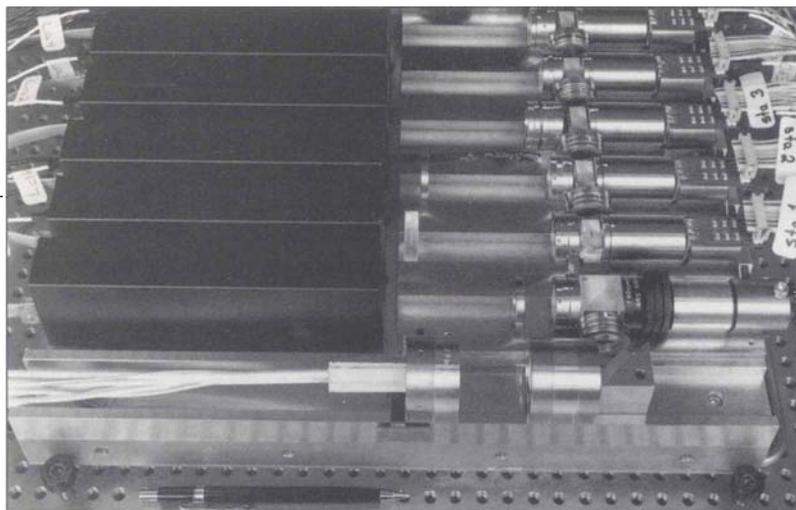
routed through any of the nodes that were currently being used by the highlighted, completed call. With the capacity of optical networks, however, it is possible to build an entirely non-blocking Banyan interconnected fabric in the same way as described in Figure 4.

Learning to See with Improving Optics. Because the chips are configured in arrays, the data that emanates from the surface of the chip is itself an array of spots. This array of spots is identically reproduced, shifted and imaged with the original array of spots onto the subsequent FET-SEED array. Thus with an array of 16 beam pairs from the modulators on the previous FET-SEED array, 32 beam pairs are produced to send to the 32 inputs of the subsequent FET-SEED array. It is this fan-out by two which enables the routing of all possible inputs to all possible outputs in the five-stage network.

As the 32 beam pairs from the previous array arrive at the next array, they are imaged to the appropriate receivers via polarization optics. The polarization optics are used to provide two input ports to send images of arrays onto the FET-SEED and two output ports to receive images of arrays of beams. Shown in Figure 5 are the optics needed to provide these two input imaging ports and the two output imaging ports. In operation, data from the previous FET-SEED or the switching fabric input is imaged onto the surface of the FET-SEED devices. These 32 beam pairs propagate as the input data to the next chip, where the electronic selection process routes 16 data signals to the output modulators. The 16 output beam pairs exit, via the output port, to be used by the next FET-SEED array as input data. The data is moved from the chip surface by illuminating the output modulators with an array of 16 beam pairs of equal intensity sent via the power port. The light reflected from the modulators carries the information from the chip, encoded as low and high-intensity beams. These beams emerge from the output port ready to be imaged onto the next array. The imaging step to the next array includes an optical split by two, performed with a computer-generated hologram. This split reproduces each of the 16 beams

twice. In this way, two copies of the data are sent to two different locations on the next array. Of course, the split will produce the 32 beams to match the 32 inputs on the next chip.

The view port is used in constructing the system. By using infrared sensing cameras, the surface of the FET-SEED chip can be observed. The camera's view sees two sets of spots registered on the surface of the integrated circuit. One set of spots is from the signals port, and one is from the power port. The array of spots from the signals port carries the input data to the FET-SEED. These signal beams must be properly registered on the receiver windows. The second array of spots is from the power port. It is these spots that provide the power to illuminate the modulators. In a fashion similar to the alignment of the input signals, the power signals must be aligned to the modulator windows.

Optomechanics: "Don't Turn that Screw!" One of the most critical considerations for optical systems is the need to maintain alignment of the light beams on the device array. For the high-density connection scheme employed here, the registration of the spots to the device windows must not vary by more than a micron or two. Not only is the registration a matter of movements in the X or Y direction, but, because all the connections are made simultaneously, rotation must also be maintained. The approach that we have used in designing optomechanical components is to over-engineer to remove as many of the variables of misalignment as possible.

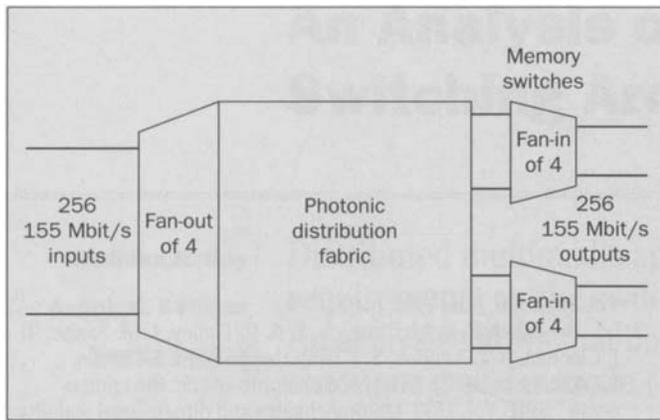Shown in Figure 6 is a picture of System₅ in

Figure 7. Growable packet switch architecture in which the photonic fabric routes packets from the inputs to the appropriate memory switch for final output resolution. A fanout is required to ensure satisfactory blocking probability in the distribution fabric.

which the massive baseplate is evident. Machined into the baseplate are grooves or channels containing all the optical components. It is the precision machining of these grooves that ensures alignment of the nominal mechanical axis of the components. Before the components are placed into the groove the optical and mechanical axis of each component is aligned. Pre-aligning the optical and mechanical axis of the components simplifies assembly of the system. Although the pre-alignment is not sufficient to ensure that the system can be optically aligned by placement of the components into the system, it does allow the image of the device under alignment to be seen quickly. Once the image is seen in the observation system, the final careful alignment can be completed by rotating the devices and sliding the elements along the grooves to ensure proper focus. Although this system of mechanical pre-alignment works well for systems experiments, a future task is to engineer a new method to simplify the alignment further.

Early attempts at building such systems used adjusting screws that never worked well. Our experience with these systems resulted in experiments in which the alignment was never completed. Once alignment was finished with the earlier systems, a new re-alignment was needed because something moved. The optomechanics used in the System$_5$ experiment has remained stable for more than six months.

Planning for the Future

The potential for photonic technology is to build large, high-bandwidth, high-throughput systems with natural partitioning of subsystems, avoiding the constraints of current electrical interconnection technology. An application that can benefit from photonic technology is the construction of a large ATM switching system. Large sys-

tems will have hundreds of user connections at 155 Mbit/s or more. As an example, a switch with 256 inputs and 256 outputs, all operating at 155 Mbit/s, has an aggregate throughput of close to 40 Gbit/s and more than 4,000 internal connections within the switching fabric. This is the size of switch fabric that we are planning for the future. Building a fabric with this throughput will test the photonic device technology. If the experiment is successful, switching fabrics with throughputs of hundreds of Gbit/s are not inconceivable.

An ATM Switch Architecture for High Throughput. The approach that has been developed to build a large ATM switching fabric exploits the capabilities of current and future memory switches and adds to them the distribution function of a photonic fabric. In Figure 7 is a schematic of the architecture that will be built. The architecture is called a growable packet switch architecture. It uses memory switches capable of performing the routing of packets from a small number of inputs to any of a small number of outputs. In addition, the memory switch is responsible for buffering the ATM cells destined for a particular output. The function of the distribution fabric is simply to route the packet, without a buffering function, to the group of outputs associated with the memory switch.

The distribution function will be performed in optics. The approach that is planned is to build a single optoelectronic integrated circuit capable of one fourth of the total distribution function bandwidth throughput. In particular, this integrated circuit will be an extension to the FET-SEED work that has been used in System$_5$. Each such FET-SEED chip will have 4,096 optical inputs and 256 optical outputs. Each of these inputs and outputs will operate at the nominal switch input bandwidth, plus some overhead bandwidth to accommodate switch setup. For example, at a 155 Mbit/s switch input rate, the fabric rate would be between 175 and 200 Mbit/s. This overhead includes bandwidth to set up connections in the fabric and to encode the data so that it can be received by an AC-coupled optical detector at the memory switch. Although the switch has 4,096 optical inputs, in normal operation only 256 inputs would be active. At the lower required rate of 175 Mbit/s and 256 inputs, the total input bandwidth could be greater than 40 Gbit/s. Since data is neither generated nor lost in the switch fabric, the output bandwidth is also greater than 40 Gbit/s. The total I/O bandwidth for one FET-SEED integrated circuit is greater than 80 Gbit/s, with throughput of greater than 40 Gbit/s. This is a very large

throughput for a complex, although specialized, integrated circuit. Only with free-space photonics can the individual circuit elements routinely be addressed.

Performance Expectations. A complete system will have four optoelectronic integrated circuits that will perform the entire distribution function for the system. Attached to the output of the distribution fabric will be 16 memory switches. Each of the four distribution fabrics will have 64 connections to each of the 16 memory switches. In this way, any packet from the switch input could be routed, in one of four possible distribution fabrics, to any of the memory switches. The 40 Gbit/s throughput of this switch seems like a large bandwidth, but there are needs for switches with even greater throughput. The approach that we intend to follow is to increase the per-channel data rate from 155 Mbit/s to 622 Mbit/s, or 2.4 Gbit/s. The throughput of the fabrics would then be 4 or 16 times greater. Fabrics of this size are comparable and beyond the latest experimental electronic fabrics.

Conclusion

Theoretical assertions about the capabilities of free-space photonics have been validated in the success of all the system experiments. The results of the experiments have shown that free-space photonics is a viable technology to build broadband switching fabrics. Large numbers of broadband connections from chip to chip are exactly what is needed in a broadband switching system, and can be supplied best by free-space photonics. The decision to build a large ATM switching fabric will highlight photonics in an application in which it is at parity with or beyond what can be done with electronics. To achieve the success that this technology has enjoyed, a dedicated team is required.

The photonics team spans AT&T Research, the Engineering Research Center, and Switching. The results reported in this paper are attributable to the hard work and dedication of the photonics team.

References
1. R. R. Tumala and E. J. Rymaszewski, eds., *Microelectronics Packaging Handbook*, Van Nostrand Reinhold, 1989, p. 391.
2. R. A. Nordin, A.F. J. Levi, R. N. Nottenburg, J. O'Gorman, T. Tanbun-Ek, and R. A. Logan, "A System Perspective on Digital Interconnection Technology," *Journal of Lightwave Technology*, Vol. 10, No. 6, June 1992, p. 811.
3. J. L. Brubaker, F. B. McCormick, F. A. P. Tooley, J. M. Sasian, T. J. Cloonan, A. L. Lentine, S. J. Hinterlong and M. J. Herron, "Optomechanics of a free space photonic switch: the components," SPIE Vol. 1533, Optomechanics and dimensional stability (1991).
4. A. L. Lentine, R. A. Novotny, T. J. Cloonan L. M. F. Chirovsky, L. A. D'Asaro, G. Livescu, S. Hui M. W. Focht, J. M. Freund, G. D. Guth, R. E. Leibenguth, L. E. Smith and T. K. Woodward, "4 x 4 arrays of FET-SEED embedded control 2 x 1 optoelectronic switching nodes with electrical fan-out," submitted to *IEEE Photonic Technology Letters*.
5. L. A. D'Asaro, L. M. F. Chirovsky, E. J. Laskowski, S.S. Pei, T. K. Woodward, A. L. Lentine, R. E. Leibenguth, M. W. Focht, J. M. Freund, G. G. Guth and L. E. Smith, "Batch Fabrication and Operation of GaAs-AlxGAl-xAs Field-Effect Transistor-Self Electro-optic Effect Device (FET-SEED) Smart Pixel Arrays," *IEEE Journal of Quantum Electronics* Vol. 29, No. 2 (1993).

Stephen J. Hinterlong *is a technical manager in the Broadband and Photonic Switching Department at AT&T Bell Laboratories in Naperville, Ill. He works to develop high-bandwidth photonic switching fabrics. Mr. Hinterlong joined AT&T in 1983. He holds a B.S. degree in physics and mathematics from Illinois Benedictine College and a Ph.D. degree in physics from the University of Notre Dame in South Bend, Indiana.*

Harry M. Hall, Jr., *is technology planning director of the Broadband Photonic Switching Department at AT&T Bell Laboratories in Naperville, Ill. He is responsible for developing broadband and photonic technologies and products to support the introduction of new video and multimedia services. Mr. Hall holds M.S. and Ph.D. degrees in electrical engineering from Stanford University in Palo Alto, California.*