

Fiber-Optic Interconnection Networks for Signal Processing Applications

Magnus Jonsson

Halmstad University, Box 823, S-301 18 Halmstad, Sweden
Magnus.Jonsson@cca.hh.se

Abstract. In future parallel radar signal processing systems, with high bandwidth demands, new interconnection technologies are needed. The same reasoning can be made for other signal processing applications, e.g., those involving multimedia. Fiber-optic networks are a promising alternative and a lot of work have been done. In this paper, a number of fiber-optic interconnection architectures are reviewed, especially from a radar signal processing point-of-view. Two kinds of parallel algorithm mapping are discussed: (i) a chain of pipeline-stages mapped, more or less directly, one stage per computation node and (ii) SPMD (Same Program Multiple Data). Several network configurations, which are simplified due to the nature of the applications, are also proposed.

1 Introduction

The interconnection network is a very important part in a parallel computing system, especially in future radar signal processing systems where high amounts of data should be moved. In this paper, fiber-optic interconnection networks suitable for signal processing applications are reviewed and proposed. Two kinds of parallel algorithm mapping are discussed for the networks as follows. Both cases are taken from the field of radar signal processing, but the discussion holds for, e.g., other signal processing applications with similar mappings too.

The first case considered is a straight forward pipeline chain, shown in Fig. 1, where each box represents a computation node [1]. Each node runs one or several pipeline stages and the arcs only represent the dataflow between modules. The physical topology can, e.g., be a ring or a crossbar switch. For simplicity, the chain is viewed in a form without, e.g, multicasting between some of the nodes. A data cube, that initially comes from the antenna, contains data in three dimensions (channel, pulse, and distance). After the processing of one stage, the new data cube is forwarded to the next node in the chain.

The second case of mapping is to let all processors do the same operation but on different sets of data, i.e., SPMD (Same Program Multiple Data). All the PEs in Fig. 2 then work together on one of the algorithmic pipeline stages at the time. After the processing of one stage, the data cube is redistributed if needed. Incoming data from the antenna is not shown in Fig. 2, but is assumed to be fed into the PEs not effecting the communication pattern shown in the figure. For example, special I/O channels can

be used instead of the communication system carrying the traffic indicated by the arcs in the figure.

Although not considered here, inter-module communication during the processing of one stage might occur. The number of times it is needed to redistribute the data cube between the nodes might however be lower than the number of times the data cube is transferred to next module in a pipeline chain. Instead, it might be harder to overlap communication and computation, which implies extra bandwidth requirements [2]. When corner turning the data cube, one half of the data cube has to be transferred across the network bisection, essentially with all-to-all communication [2].

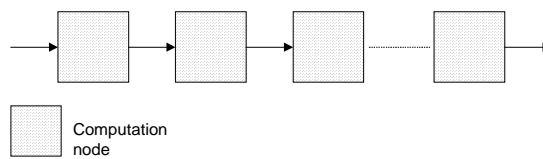


Fig. 1. A pure pipeline chain where one or several stages are mapped on each node

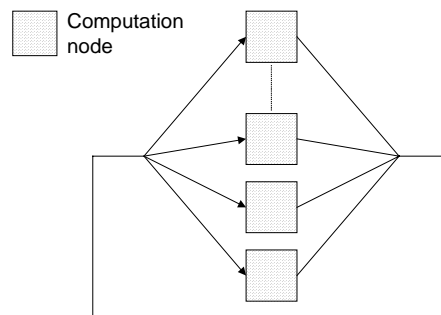


Fig. 2. If the SPMD model is used, all PEs work together on one stage at the time

The rest of the paper is organized as follows. In Section 2 through 6, the following technologies and interconnection architecture classes are reviewed, respectively: (i) fiber-ribbon ring, (ii) WDM star, (iii) WDM ring, (iv) integrated fiber and waveguide solutions, and (v) optical fibers and electronic crossbars. The paper is then concluded in Section 7.

2 Fiber-Ribbon Ring Networks

If fiber-ribbon cables are used to connect the nodes in a point-to-point linked ring network, bit-parallel transfer can be utilized. In such a network one of the fibers in each ribbon is dedicated to carry the clock signal. Therefore, no clock-recovery

circuits are needed in the receivers. Other fibers can be dedicated for, e.g., frame synchronization. An example of a fiber-ribbon ring network is the USC POLO Network, which is proposed to be used in COWs (Cluster of Workstations) and similar systems [3]. Integrated circuits have been developed for the network and tests have been performed [4] [5].

As seen in Fig. 3, aggregated throughputs higher than 1 can be obtained in ring networks with support for spatial bandwidth reuse (sometimes called pipeline rings) [6]. This feature can be highly useful in signal-processing applications with a pipelined dataflow, i.e., most of the communication is to the nearest down-stream neighbor. Two fiber-ribbon pipeline ring networks have been reported recently [7]. The first network has support for circuit-switching on 8+1 fibers (data and clock) and packet-switching on an additional fiber [8]. The second network is more flexible, but with a little more complexity, and has support for packet-switching on 8+1 fibers while using a tenth fiber for control packets (see Fig. 4) [9]. The control packets carries MAC (Medium Access Control) information for the collision-less MAC protocol with support for slot-reserving to get guaranteed bandwidth and worst-case latency.

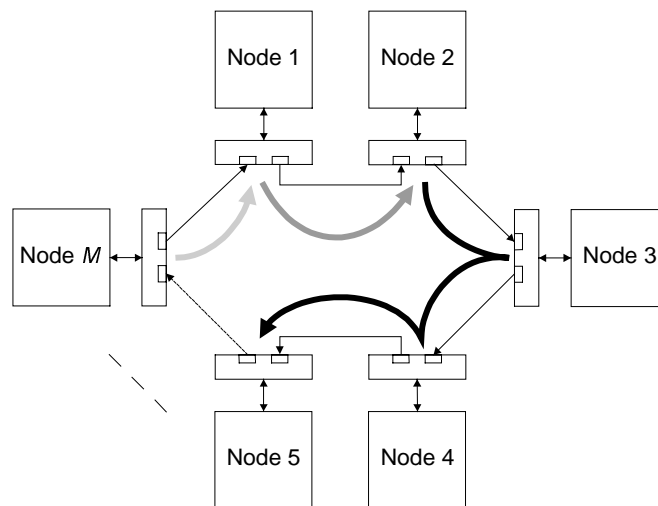


Fig. 3. Example of spatial bandwidth reuse. Node M sends to Node 1 at the same time as Node 1 sends to Node 2, and Node 2 sends a multicast packet to Node 3, 4, and 5

3 WDM Star Networks

A passive fiber-optic star distributes all incoming light, on the input ports, to all output ports. A network with the logical function of a bus is obtained when connecting the transmitting and receiving side of each node to one input and output fiber of the star, respectively. By using WDM (Wavelength Division Multiplexing),

multiple wavelength (color of light) channels can carry data simultaneously in the network [10]. To get a flexible network, we need tunable receivers and/or transmitters, i.e., it should be possible to send/listen on an arbitrary channel [11].

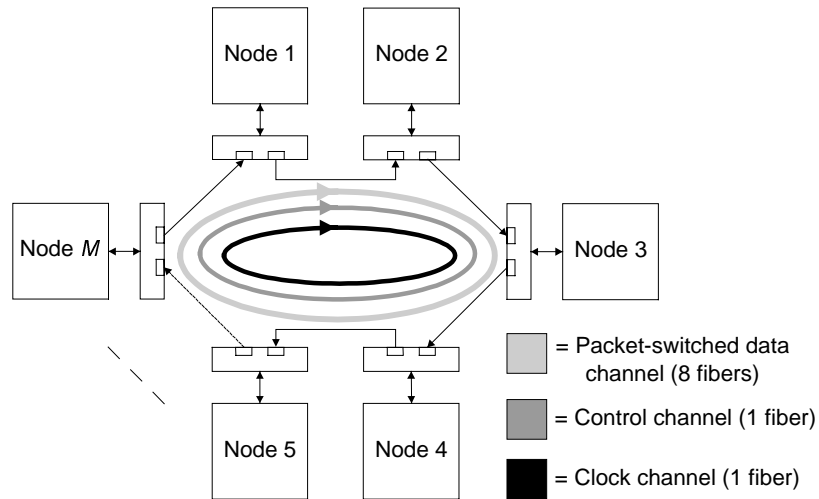


Fig. 4. Control-channel based network built up with fiber-ribbon point-to-point links

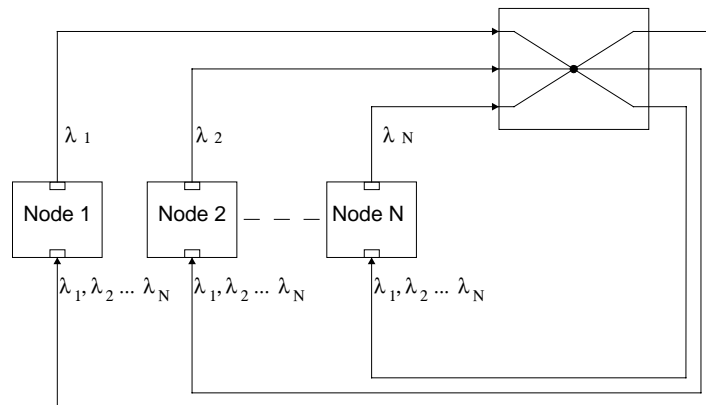


Fig. 5. WDM star network

In Fig. 5, an example of a WDM star network configuration is shown. Each node transmits on a wavelength unique to the node, while the receiver can listen to an arbitrary wavelength. The configuration is used in the TD-TWDM network (Time-Deterministic Time and Wavelength Division Multiple Access) [1], which has support for guaranteeing real-time services, both in single-star networks [12] and in star-of-stars networks [13]. One can say that this kind of networks implements a distributed

crossbar. The flexibility is hence high, and multicast and single-destination traffic can co-exist. The number of wavelengths is practically limited to 16-32 [14], but, as stated above, hierarchical networks with wavelength reuse can be built.

Tunable components (e.g., filters) with tuning latencies in the order of a nanosecond have been reported, but they often have a limited tuning range [15]. At the expense of longer tuning latencies, however, components with a broader tuning range exist [16]. Such components can be used to get a cheaper network for systems where much of the communication patterns remain constant for a longer period, e.g., during the processing of a data cube in a radar system with a pipelined mapping. To support for some more rapidly changing traffic patterns, the nodes can be extended with transmitters and receivers fixed-tuned to a special broadcast channel. This configuration can be compared to having support for both circuit- and packet-switching.

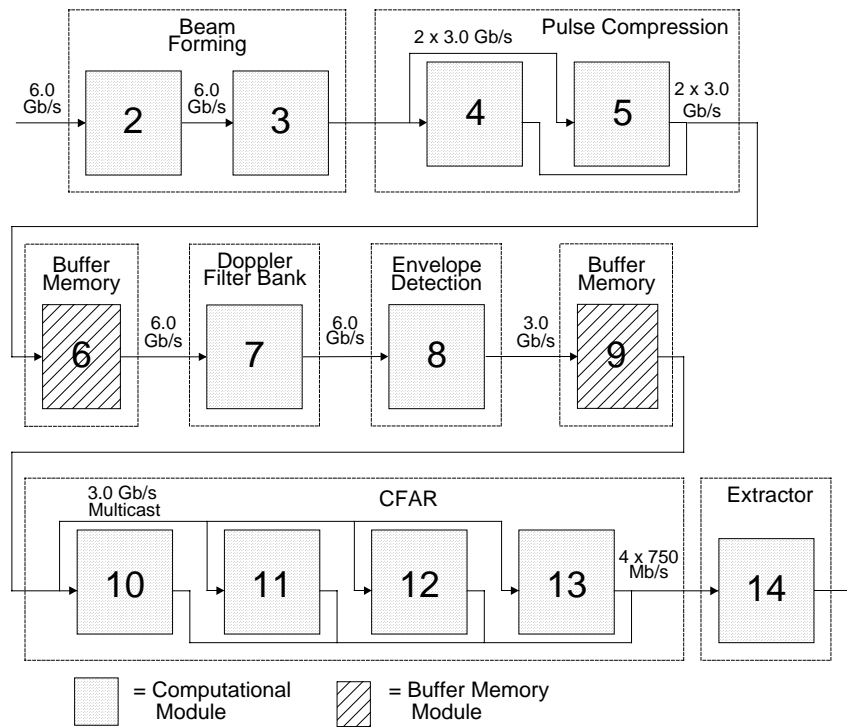


Fig. 6. Data flow between the modules in the sample radar signal processing chain

An example is given below to show that only one input and one output channel (in addition to the broadcast channel) are needed during the processing of a data cube, i.e., the normal minimum time running one pipeline chain (working mode of the radar). In Fig. 6, a signal processing chain similar to those described in [17] and [1] is shown together with its bandwidth demands. Each computational module in the figure contains multiple processors itself. The chain is a good example containing both multicast, one-to-many, and many-to-one communication patterns. The aggregated

throughput demand is 45 Gb/s. We leave out all details of the chain, referring to the two other papers, and do only treat the throughput demands here. In the figure, there are 13 nodes. In addition, the antenna is seen as one node (feeds the first node in the chain with data), Node 1, and there is one master node responsible for supervising the whole system and interacting with the user, Node 15. For simplicity we assume an efficient channel bandwidth of 6.0 Gb/s. A feasible allocation scheme of eight channels (in addition to the broadcast channel) is shown in Table 1.

Table 1. A feasible allocation scheme of the wavelength channels in a WDM star network for the sample radar system

Node	Input channel	Output channel
1	-	λ_1
2	λ_1	λ_2
3	λ_2	λ_3
4	λ_3	λ_4
5	λ_3	λ_4
6	λ_4	λ_5
7	λ_5	λ_6
8	λ_6	λ_7
9	λ_7	λ_7
10	λ_7	λ_8
11	λ_7	λ_8
12	λ_7	λ_8
13	λ_7	λ_8
14	λ_8	λ_8
15	λ_8	-

Totally removing the ability of tuning in a WDM star network leads to a multi-hop network [18]. Each node in a multi-hop network transmits and receives on one or a few dedicated wavelengths. If a node does not have the capability of sending on one of the receiver wavelengths of the destination node, the traffic must pass one or several intermediate nodes. The wavelengths can be chosen to get, e.g., a perfect-shuffle network [19], or a network with a pattern similar to the dataflow in a pipelined system. One can also choose to have a network with several topologies embedded, e.g., a ring and a hypercube.

4 WDM Ring Networks

Each node in a WDM ring network has an ADM (Add-Drop Multiplexer) to insert, listen, and remove wavelength channels from the ring. Spatial wavelength reuse can be achieved by removing the transmitted light at the destination node. At high degrees of nearest-downstream-neighbor communication (e.g., in pipeline mapping), throughputs significantly higher than 1 can be achieved with only a few wavelength

channels. In the WDM ring network described in [20], each node is assigned a node-unique wavelength to transmit on. The other nodes can then tune in an arbitrary channel to listen on. This configuration is logically the same as that for the WDM star network with fixed transmitters and tunable receivers.

As discussed for the WDM star network, components with long tuning latencies (ADMs in the case of a ring) can be used for traffic patterns that do not change rapidly. An additional broadcast wavelength dedicated for packet-switching keeps the network flexible. A single 6 Gbit/s channel (in addition to the broadcast channel) is enough for the signal processing chain shown in Fig. 6, if the ADMs in Node 1, 2, 3, 6, 7, 8, and 9 terminates the channel for wavelength reuse.

5 Integrated Fiber and Waveguide Solutions

Fibers or other kinds of waveguides (hereafter commonly denoted as channels) can be integrated to form a more or less compact system of channels. Fibers can be laminated to form a foil of channels, for use as intra-PCB (Printed Circuit Board) or back-plane interconnection systems [21] [22]. Fiber-ribbon connectors are applied to fiber endpoints of the foil. An example is shown in Fig. 7, where four computational nodes are connected in a ring topology. In addition, there is a clock node that distributes clock signals to the four computational nodes via equal-length fibers to keep the clock signals in phase. If one foil is placed on each PCB in a rack, they can be passively connected to each other via fiber-ribbon cables. Using polymer waveguides instead of fibers brings advantages such as the possibility of integrating splitters and combiners into the foil, and the potential for more cost-effective mass-production [21].

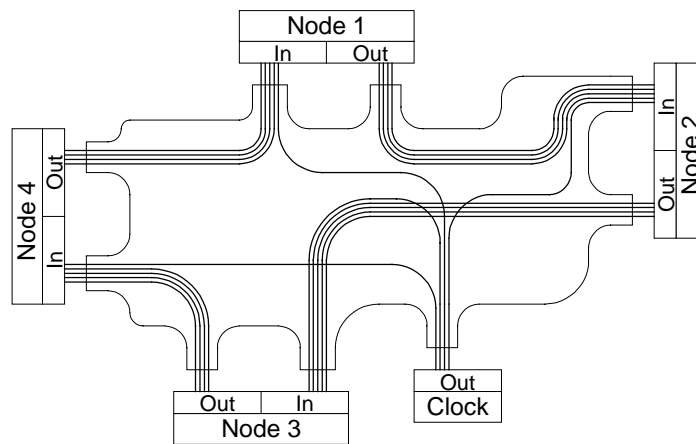


Fig. 7. A foil of fibers connects four computational nodes and a clock node

Integrated systems of channels can be setup and used in a number of configurations, for which some are discussed below. One way is to embed a ring with

bit-parallel transmission and the possibility of spatial bandwidth reuse as described in Section 2. Of course, this leads to the same good performance for pipelined data flows as the fiber-ribbon ring network, only changing the medium to a more compact form. Besides pure communication purposes, channels for, e.g., clock distribution (as seen in the example) and flow control can be integrated into the same system.

Another way is to follow the proposed use of an array of passive optical stars to connect processor boards in a multiprocessor system, via fiber-ribbon links, for which experiments with 6 x 700 Mb/s fiber-ribbon links were done (see Fig. 8) [23]. As indicated above, such a configuration can be integrated by the use of polymer waveguides. The power budget can, however, be a limiting factor to the number of nodes and/or the distance. Advantages are simple hardware due to bit-parallel transmission, and the broadcast nature. Many-to-many communication patterns, as used when corner-turning in SPMD mode, map easily on the broadcast architecture as long as the star-array not becomes a bottle-neck.

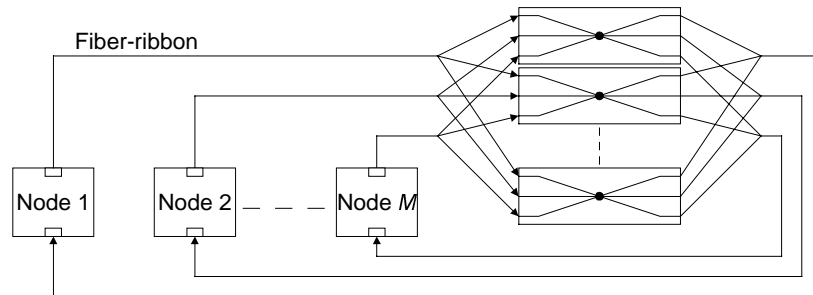


Fig. 8. Array of passive optical stars connects a number of nodes via fiber-ribbons

Electronic crossbars can be distributed on the PCBs and/or placed on a special switch-PCB in a back-plane system, and be connected by integrated parallel channels. A distributed crossbar, instead, can be realized by bit-serial transmissions over a fully connected topology, i.e., there is one channel between each pair of nodes (see Fig. 9). Broadcast is done by driving all the laser-diodes of a node with the same bit-stream. A simple solution is to always drive all the laser-diodes in the transmitting node and couple the photo-diodes together as one incoming channel. This brings an architecture with the support of a single broadcast at the time. By only sending on one fiber when performing single-destination communication will, however, give the opportunity of having multiple transmissions in the system at the same time. The number of nodes, N , in this configuration is limited because the number of fibers grows by N^2 . Also, clock-recovery circuits must be involved. After all, the distressed power budget brings an advantage over a system with passive splitters or stars (this holds for, e.g., the point-to-point connected ring also). The flexibility of a crossbar makes the network good for radar systems with the SPMD mode.

Other similar systems include the integration of fibers into a PCB for the purpose of clock distribution [24]. Distribution to up to 128 nodes was demonstrated. The fibers are laminated on one side of the PCB while integrated circuits are placed on the reversed side. The end section of each fiber is bent 90 degrees to lead the light through a via hole to the reversed side of the PCB.

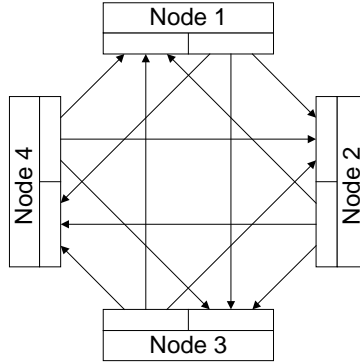


Fig. 9. Fully connected topology where each node has an array of $N - 1$ laser diodes and an array of $N - 1$ photo diodes, where N is the number of nodes

6 Optical Fibers and Electronic Crossbars

Communication systems like, e.g., Myrinet [25], where arbitrary switched topologies can be built, can support a number of different traffic patterns possible in radar systems. Fiber-ribbons can be used to increase bandwidth while still sending in bit-parallel mode. Bit rates in the order of 1 Gbit/s over each fiber in the ribbon is possible over tens of meters using standard fiber-ribbons. As noted in Section 5, foils of fibers or waveguides (e.g., arranged as ribbons) can be used to interconnect nodes and crossbars on the PCB and/or back-plane level.

An alternative to ribbons is bit-parallel transfer over a single fiber by using WDM. In such configurations each bit in the word, plus the clock signal, is given a dedicated wavelength. Wavelengths can also be dedicated to other purposes like frame synchronization and flow control. Significantly higher bandwidth-distance products can be achieved when using bit-parallel WDM over dispersion-shifted fiber, instead of fiber-ribbons [26]. If, however, only communication over shorter distances exist (e.g., a few meters), the bandwidth-distance product is not necessarily a limiting factor. Transmission experiments with an array of eight pie-shaped VCSELs arranged in a circular area with a diameter of $60 \mu\text{m}$, to match the core of a multimode fiber, has been reported [27]. Other works on the integration of components for short-distance (non telecom) WDM links have been reported, e.g., a 4×2.5 Gbit/s transceiver with integrated splitter, combiner, filters, etc. [28].

The switch itself can also be modified to increase performance or packing density. A single-chip switch core where fiber-ribbons are coupled directly to optoelectronic devices on the chip is possible [29]. Attaching 32 incoming and 32 outgoing fiber-ribbons with 800 Mb/s per fiber translates to an aggregated bandwidth of 204 Gbit/s through the switch when eight fibers on each link are used for data.

A 16×16 crossbar switch-chip, with integrated optoelectronic I/O, have been implemented for switching of packets transferred using bit-parallel WDM [30]. Each node has two fibers (input and output) coupled to the switch.

Another switch chip with integrated optoelectronic I/O is intended to be attached to each node in a static topology like a multidimensional torus [31]. A special feature of the chips is that potential deadlocks is handled by a global mechanism. The mechanism brings mutual exclusion to let one packet at a time use dedicated hardware to recover from a potential deadlock. Finally, it can be noted that both Myricom (Myrinet) and Mercury (RACEway) are looking at optical technologies for their future products [32].

7 Conclusions

Some of the mentioned networks are summarized in Table 2, where remarks on suitability/performance for the two cases of mapping are made. The increasingly good price/performance ratio for fiber-ribbon links indicates a great success potential for several of the networks discussed. On the other hand, the WDM technique offer flexible multi-channel networks that can be passively implemented. Integrated fiber and waveguide solutions makes the building of compact systems possible, especially for networks like those using fiber-ribbons.

Table 2. Summary of some of the discussed networks

Network	Pipeline	SPMD	Notes
Fiber-ribbon pipeline ring	Good	Moderate	Good for SPMD too if enough bandwidth
WDM star with rapid tuning	Good	Good	Flexible passive network
WDM star with slow tuning and broadcast channel	Good	Poor	WDM star alternative that might be cheaper
Multi-hop WDM star	Moderate	Moderate	Can be optimized for pipelined mapping
WDM ring with rapid tuning	Good	Good	More channels might be needed for SPMD
WDM ring with slow tuning and broadcast channel	Good	Poor	WDM ring alternative that might be cheaper
Fiber ribbons and array of stars	Moderate	Moderate	Power and bandwidth limited
Fully connected topology with broadcast driving	Moderate	Moderate	Bandwidth limited. Grows with N^2
Fully connected topology with flexible driving	Good	Good	Grows with N^2
Optical fibers and electronic crossbars	Good	Good	Optoelectronics needed in switch too

8 Acknowledgement

This work is part of the REMAP project, financed by NUTEK, the Swedish National Board for Industrial and Technical Development, and the PARAD project, financed by the KK Foundation in cooperation with Ericsson Microwave Systems AB.

References

1. Jonsson, M., A. Åhlander, M. Taveniku, and B. Svensson, "Time-deterministic WDM star network for massively parallel computing in radar systems," *Proc. Massively Parallel Processing using Optical Interconnections, MPPOI'96*, Lahaina, HI, USA, Oct. 27-29, 1996, pp. 85-93.
2. Teitelbaum, K., "Crossbar tree networks for embedded signal processing applications," *Proc. Massively Parallel Processing using Optical Interconnections (MPPOI'98)*, Las Vegas, NV, USA, June 15-17, 1998, pp. 200-207.
3. Sano, B. J., and A. F. J. Levi, "Networks for the professional campus environment," in *Multimedia Technology for Applications*. B. Sheu and M. Ismail, Eds., McGraw-Hill, Inc., pp. 413-427, 1998, ISBN 0-7803-1174-4.
4. Sano, B., B. Madhavan, and A. F. J. Levi, "8 Gbps CMOS interface for parallel fiber-optic interconnects," *Electronics Letters*, vol. 32, pp. 2262-2263, 1996.
5. *POLO Technical Summary*. University of Southern California, Oct. 1997.
6. Wong, P. C., and T.-S. P. Yum, "Design and analysis of a pipeline ring protocol," *IEEE Transactions on communications*, vol. 42, no. 2/3/4, pp. 1153-1161, Feb./Mar./Apr. 1989.
7. Jonsson, M., "Two fiber-ribbon ring networks for parallel and distributed computing systems," *Optical Engineering*, vol. 37, no. 12, pp. 3196-3204, Dec. 1998.
8. Jonsson, M., B. Svensson, M. Taveniku, and A. Åhlander, "Fiber-ribbon pipeline ring network for high-performance distributed computing systems," *Proc. Int. Symp. on Parallel Architectures, Algorithms and Networks (I-SPAN'97)*, Taipei, Taiwan, Dec. 18-20, 1997, pp. 138-143.
9. Jonsson, M., "Control-channel based fiber-ribbon pipeline ring network," *Proc. Massively Parallel Processing using Optical Interconnections (MPPOI'98)*, Las Vegas, NV, USA, June 15-17, 1998, pp. 158-165.
10. Brackett, C. A., "Dense wavelength division multiplexing networks: principles and applications," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 6, pp. 948-964, Aug. 1990.
11. Mukherjee, B., "WDM-based local lightwave networks part I: single-hop systems," *IEEE Network*, pp. 12-27, May 1992.
12. Jonsson, M., K. Börjesson, and M. Legardt, "Dynamic time-deterministic traffic in a fiber-optic WDM star network," *Proc. 9th Euromicro Workshop on Real Time Systems*, Toledo, Spain, June 11-13, 1997.
13. Jonsson, M., and B. Svensson, "On inter-cluster communication in a time-deterministic WDM star network," *Proc. 2nd Workshop on Optics and Computer Science (WOCS)*, Geneva, Switzerland, Apr. 1, 1997.
14. Brackett, C. A., "Foreword: Is there an emerging consensus on WDM networking?," *Journal of Lightwave Technology*, vol. 14, no. 6, pp. 936-941, June 1996.
15. Kobrinski, H., M. P. Vecchi, E. L. Goldstein, and R. M. Bulley, "Wavelength selection with nanosecond switching times using distributed-feedback laser amplifiers," *Electronics Letters*, vol. 24, no. 15, pp. 969-971, July 21, 1988.

16. Cheung, K.-W., "Acoustooptic tunable filters in narrowband WDM networks: system issues and network applications," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 6, pp. 1015-1025, Aug. 1990.
17. Taveniku, M., A. Åhlander, M. Jonsson, and B. Svensson, "A multiple SIMD mesh architecture for multi-channel radar processing," *Proc. International Conference on Signal Processing Applications & Technology (ICSPAT'96)*, Boston, MA, USA, Oct. 7-10, 1996, pp. 1421-1427.
18. Mukherjee, B., "WDM-based local lightwave networks part II: multihop systems," *IEEE Network*, pp. 20-32, July 1992.
19. Acampora, A. S., and M. J. Karol, "An overview of lightwave packet networks," *IEEE Network*, pp. 29-41, Jan. 1989.
20. Irshid, M. I., and M. Kavehrad, "A fully transparent fiber-optic ring architecture for WDM networks," *Journal of Lightwave Technology*, vol. 10, no. 1, pp. 101-108, Jan. 1992.
21. Eriksen, P., K. Gustafsson, M. Niburg, G. Palmskog, M. Robertsson, and K. Åkermark, "The Apollo demonstrator – new low-cost technologies for optical interconnects," *Ericsson Review*, vol. 72, no. 2, 1995.
22. Shahid, M. A., and W. R. Holland, "Flexible optical backplane interconnections," *Proc. Massively Parallel Processing using Optical Interconnections (MPPOI'96)*, Maui, HI, USA, Oct. 27-29, 1996, pp. 178-185.
23. Parker, J. W., P. J. Ayliffe, T. V. Clapp, M. C. Gear, P. M. Harrison, and R. G. Peall, "Multifibre bus for rack-to-rack interconnects based on opto-hybrid transmitter/receiver array pair," *Electronics Letters*, vol. 28, no. 8, pp. 801-803, April 9, 1992.
24. Li, Y., J. Popelek, J.-K. Rhee, L. J. Wang, T. Wang, and K. Shum, "Demonstration of fiber-based board-level optical clock distributions," *Proc. 5th International Conference on Massively Parallel Processing using Optical Interconnections (MPPOI'98)*, Las Vegas, NV, USA, June 15-17, 1998, pp. 224-228.
25. Boden, N. J., D. Cohen, R. E. Felderman, A. E. Kulawik, C. L. Seitz, J. N. Seizovic, and W.-K. Su, "Myrinet: a gigabit-per-second local area network," *IEEE Micro*, vol. 15, no. 1, pp. 29-36, Feb. 1995.
26. Bergman, L., J. Morookian, and C. Yeh, "An all-optical long-distance multi-Gbytes/s bit-parallel WDM single-fiber link," *Journal of Lightwave Technology*, vol. 16, no. 9, pp. 1577-1582, Sept. 1998.
27. Coldren, L. A., E. R. Hegblom, Y. A. Akulova, J. Ko, E. M. Strzelecka, and S. Y. Hu, "Vertical-cavity lasers for parallel optical interconnects," *Proc. 5th International Conference on Massively Parallel Processing using Optical Interconnections (MPPOI'98)*, Las Vegas, NV, USA, June 15-17, 1998, pp. 2-10.
28. Aronson, L. B., B. E. Lemoff, L. A. Buckman, and D. W. Dolfi, "Low-cost multimode WDM for local area networks up to 10 Gb/s," *IEEE Photonics Technology Letters*, vol. 10, no. 10, pp. 1489-1491, Oct. 1998.
29. Szymanski, T. H., A. Au, M. Lafrenière-Roula, V. Tyan, B. Supmonchai, J. Wong, B. Zerrouk, and S. T. Obenaus, "Terabit optical local area networks for multiprocessing systems," *Applied Optics*, vol. 37, no. 2, pp. 264-275, Jan. 10, 1998.
30. Krisnamoorthy et al., "The AMOEBA chip: an optoelectronic switch for multiprocessor networking using dense-WDM," *Proc. Massively Parallel Processing using Optical Interconnections (MPPOI'96)*, Maui, HI, USA, Oct. 27-29, 1996, pp. 94-100.
31. Pinkston, T. M., M. Raksapatcharawong, and Y. Choi, "WARP II: an optoelectronic fully adaptive network router chip," *Proc. Optics in Computing (OC'98)*, Brugge, Belgium, June 17-20, 1998, pp. 311-315.
32. Lund, C., "Optics inside future computers," *Proc. Massively Parallel Processing using Optical Interconnections (MPPOI'97)*, Montreal, Canada, June 22-24, 1997, pp. 156-159.