

A Low-Complexity Parallel System for Gracious, Scalable Performance. Case Study for Near PetaFLOPS Computing

Sotirios G. Ziavras Haim Grebel

Department of Electrical and Computer Engineering
New Jersey Institute of Technology, Newark, NJ 07102

Anthony Chronopoulos

Department of Computer Science
Wayne State University, Detroit, MI 48202

Abstract

This paper presents a "point design" for an MIMD distributed shared-memory parallel computer capable of achieving gracious 100 TeraFLOPS performance with technology that will definitely become feasible/viable in less than a decade. Its scalability guarantees a lifetime extending well into the next century. Our design takes advantage of free-space optical technologies, with simple guided-wave concepts, to produce a 1-D building block (BB) that implements efficiently a large, fully-connected system of processors. Designing fully-connected, large systems of electronic processors could be an immediate impact of optics on massively-parallel processing. A 2-D structure is proposed for the complete system, where the aforementioned 1-D BB is extended into two dimensions. This architecture behaves like a 2-D generalized hypercube, which is characterized by outstanding performance and extremely high wiring complexity that prohibits its electronic implementation. With readily available technology, a mesh of clear plastic bars in our design facilitate bit-parallel transmissions that utilize wavelength-division multiplexing and follow dedicated optical paths. Each processor is mounted on a card. Each card contains eight processors interconnected locally via an electronic crossbar. Taking advantage of higher-speed optical technologies, all eight processors share the same interface to the optical medium. Encouraging, preliminary results prove that our conservative design could have a tremendous, positive impact on massively-parallel computing in the near future. Another impressive property of our system is that its bisection bandwidth matches, within an order of magnitude, the performance of its computation engine. Our optical interconnection scheme is superior

*to other optical schemes because it is scalable, feasible, viable, fast, power efficient, point-to-point, and does not have an adverse effect on the system's size. We expect 2-D and 3-D implementations of our design to achieve gracious PetaFLOPS performance before the end of the next decade.*¹

1 Introduction

The demand for ever greater performance by many computation problems has been the driving force for the development of computers with thousands of processors. Two important aspects are expected to dominate massively-parallel processing. High-level parallel languages that support a shared address space (for distributed shared-memory computers), and point-to-point interconnection networks for workstation-like nodes.

Near PetaFLOPS performance and more is required by many applications, such as weather modeling, simulation of physical phenomena, aerodynamics, simulation of neural networks, simulation of chips, structural analysis, real-time image processing and robotics, artificial intelligence, seismology, animation, real-time processing of large databases, etc. Dongarra has pointed out that the world's top ten technical computing sites have peak capacity of only about 850 GFLOPS, with each site containing hundreds of computers. The goal of 1-TeraFLOPS parallel computer will be reached only later this year for the Sandia Labs, using the Intel P6 processor.

The PetaFLOPS computing objective seems to be

¹The work presented in this research is supported in part by NSF under a New Millennium Computing Point Design Grant, also sponsored by NASA and DARPA.

a distant dream primarily because of the, as currently viewed, unsurmountable difficulty in developing low-complexity, high-bisection bandwidth, and low-latency interconnection networks to connect thousands of processors (and remote memories). To quote Dally, “wires are a limiting factor because of power and delay as well as density” [22]. Several interconnection networks have been proposed for the design of massively-parallel computers, including, among others, regular meshes and tori [2], enhanced meshes, fat trees, (direct binary) hypercubes [4], and hypercube variations [3, 6, 8, 9]. The hypercube dominated the field in the 80’s and early 90’s because it provides low diameter and can emulate efficiently many topologies frequently employed in the development of algorithms [4, 10]. Nevertheless, these properties come at the cost of often prohibitively high VLSI (wiring) complexity due to a dramatic increase in the number of communication channels with any increase in the number of PEs (processing elements). Its high VLSI complexity is undoubtedly its dominant drawback [13], and does not permit the construction of powerful, massively-parallel systems. The versatility of the hypercube in emulating efficiently other important topologies constitutes an incentive for the introduction of hypercube-like, electronic interconnection networks of lower complexity that preserve to a large extent its topological properties [8, 9].

Current, feasible approaches to massively-parallel processing use bounded-degree networks, such as meshes or k -ary n -cubes, with low degree of connection (e.g., FLASH [11], Cray Research MPP, Intel Paragon, and Tera). However, low-degree networks result in large diameters and average internode distances, and small bisection bandwidth. Relevant approaches that employ reconfiguration (e.g., reconfigurable meshes) will not become feasible in the foreseeable future because of the requirements for long clock cycles and precharged switches to facilitate the transmission of messages over long distances.

The high VLSI complexity problem is very serious for generalized hypercubes that, contrary to nearest-neighbor k -ary n -cubes, implement fully-connected systems with k nodes in each dimension [1]. Figure 1 shows the $GH(2, 7)$ with 2 dimensions (i.e., $n = 2$) and $k = 7$. The n -D $GH(n, k)$ with k^n nodes has diameter equal to only n . For $n = 2$ and k an even number, the diameter of the generalized hypercube is 2 and its bisection width is the immense $k^3/4$. The increased VLSI cost of generalized hypercubes results in outstanding performance that permits optimal emulation of hypercubes and k -ary n -cubes, and effi-

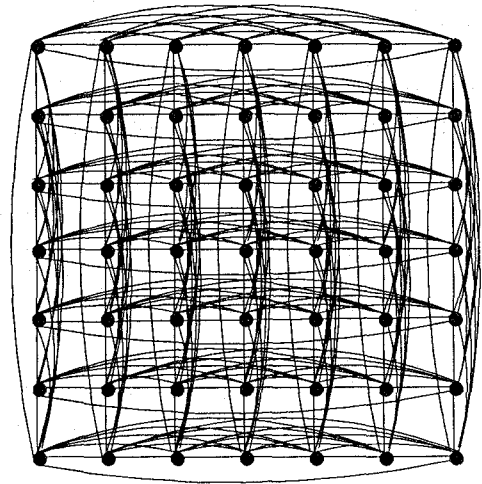


Figure 1: The generalized hypercube $GH(2, 7)$.

cient implementation of complex communication patterns. To reduce the number of its communication channels, the spanning bus hypercube uses a shared bus to implement a fully-connected system in each dimension [6]. Nevertheless, shared buses result in significant performance overhead imposed by the protocol that determines ownership of the bus. Hypergraph architectures implement all possible permutations of their nodes in each dimension by employing crossbar switches [5]. Reconfigurable generalized hypercubes interconnect all PEs in each dimension dynamically via a scalable mesh of very simple, low-cost programmable switches [14]. However, all these proposed reductions in complexity may not be sufficient for PetaFLOPS computing. To quote Patterson, “currently the most expensive scheme is a crossbar switch, which provides an explicit path between every communicating device. This becomes prohibitively expensive when connecting thousands of processors” [23].

To summarize, low-dimensional massively-parallel computers with full connectivity in each dimension, such as generalized hypercubes, are very desirable because of their outstanding topological properties (e.g., extremely small diameter and average internode distance, and immense bisection bandwidth), but their electronic implementation is a Herculean task because of packaging (and primarily wiring) constraints. Therefore, introducing pioneering technologies for the implementation of such systems could give life, for the first time, to scalable/feasible PetaFLOPS computing platforms. This is our main objective. We have chosen free-space optical technologies to satisfy this objec-

tive in the best possible way. There is another major drawback of pure electronic implementations of systems with thousands of processors; processor speeds increase much faster than memory and interconnection network speeds, and therefore there is an utmost need for the development of very advanced memory-latency hiding mechanisms, namely prefetching, cache coherence, multithreading, and relaxed memory consistency. However, *mitigation of the memory-latency problem is possible if free-space optical technologies are used for the implementation of large, almost fully-connected, point-to-point interconnection networks.*

Optical technologies have been enlisted before in parallel processing [5, 15, 16]. However, past efforts have not been successful because of large power consumption (often due to redundant broadcasts), mechanical components that do not match electronic speeds, inefficient reconfiguration schemes, unreliability, strict alignment requirements, large complexity that prohibits scalability, etc. Interconnections with wavelength selectivity for channel allocation have been under extensive study recently [15, 16].

Other proven architectural features, in addition to the chosen interconnection network technologies, are required to make any proposed system viable. Distributed shared-memory systems already dominate the massively-parallel processing field [11, 12], because the simultaneous incorporation of the message-passing and shared-memory communication paradigms introduces versatility [7]. In 2007, 16-way multithreaded microprocessors may be common. A system with thousands of processors may then be handling hundreds of thousands of threads simultaneously, thus making the problems of cache coherence, debugging, scheduling, and performance monitoring extremely difficult to handle. Hardware/software code-sign will be needed to develop relevant solutions for PetaFLOPS systems. Also, good programming environments are crucial to the success of such large systems, because of their complexity. An advanced programming environment should comprise an operating system, compilers for programming languages, application packages, and software tools for ease of program development, debugging, and performance monitoring. Performance visualization also is needed to identify bottlenecks in program execution. Finally, a parallel I/O capability and development of techniques to deal with resultant, non-sequential file systems are absolutely essential.

Our paper is organized as follows. Section 2 contains the objectives and significance of our “point design.” Section 3 contains a detailed description of

our design for a system capable of 100 TeraFLOPS. Impressive performance characteristics of this system and a feasibility analysis are also included. Section 4 contains performance results for some important computation- and/or communication-intensive problems. Finally, conclusions are presented in Section 5.

2 Objectives and Significance of this Project

The main objective of our project is to develop and evaluate a “point design” for a scalable, parallel computer that could become feasible in less than a decade and will be capable of 100 TeraFLOPS performance. However, our pioneering approach can also lead quickly to the development of gracious PetaFLOPS computing platforms. This project comprises the following major tasks:

- Design of a massively-parallel computer involving advanced, but at the same time readily available, electronic and optical technologies. Special attention to architectural and technological scalability guarantees a lifetime extending well into the next century.
- Feasibility and cost analysis of the proposed design.
- Outline of an associated programming environment, including the development of efficient mapping techniques for program creation in a manner highly independent of the underlying architecture.
- Performance evaluation of the proposed system based on theory, benchmarks, and simulations that can test its communications structure.
- Simulation of significant applications which are characterized by heavy computation and communication requirements.
- Further study of the system’s usability by illustrating that it satisfies the performance requirements of a wide range of diverse applications.
- Outline of an advanced I/O system.

Our “point design” could prove indispensable to the parallel-processing community because of the following reasons:

- Free-space optical technologies could have a significant influence on massively-parallel processing because of reduced packaging complexity that facilitates the construction of powerful systems with increased connectivity. Ours is a meticulous effort towards formulating a relevant, attainable objective and presenting a viable solution to fulfill this objective.
- The success of our effort could persuade other scientists and engineers of the feasibility of optics-based PetaFLOPS designs, and could contribute drastically to the impetus required to drive the computer industry in this direction. The low complexity, scalable performance, and ease of manufacturing attributes that characterize our system could undoubtedly become great incentives.
- High performance often will be the result of integrating effectively the message-passing and shared-memory communication paradigms in all phases of program execution. Our effort will facilitate a better understanding of the intricacies and inter-relationships between these paradigms because the fully-connected network treats them rather equally, independently of location (in contrast, distributed shared-memory systems with electronic networks inherently favor message passing).
- Finally, our success could provide much needed optimism to the parallel-processing community regarding the PetaFLOPS computing challenge.

3 Case Study: 100 TeraFLOPS

The performance objective in this phase of our project is 100 TeraFLOPS in 2007 (New Millennium Computing Point Design awards objective of NSF, NASA, and DARPA). Based on SIA projections [19], commodity microprocessors will be capable of 10 GFLOPS in 2007, and therefore we need at least 10,000 processors in the system. For a system of 100 TeraFLOPS performance to be viable, its physical volume must be reasonably small, and its communications and I/O capabilities should match (within an order of magnitude) the speed of its computation engine. Free-space optical technologies eliminate the need for wires in the implementation of communication channels and could realize fully-connected BBs with large numbers of processors and small physical volume.

Our design takes advantage of free-space optical technologies to produce a 1-D fully-connected, scalable building block (BB). Since the complete system is a 2-D configuration of 8-PE cards, we need a BB with 36 cards; then the total number of PEs is $(36)^2 \times 8$ or 10,368 (for 103.68 TeraFLOPS peak performance). *This BB is actually a fully-connected system of 288 PEs*, because the eight PEs on each card are fully-connected via an electronic crossbar network and the bandwidth of the optical interface is such that all eight PEs on the card can be involved simultaneously in inter-card data transfers without any performance degradation. All communication channels are bit-parallel.

The communications capabilities of our complete system resemble those of the extremely powerful 2-D generalized hypercube, which is by far much better than any interconnection network that has ever been built for massively-parallel processing. Also, the incorporation of advanced cache schemes (e.g., coherence) then becomes a viable task because of the system's extremely small diameter (i.e., 2), immense bisection width, and high-speed network. Similar tasks will be of extraordinary difficulty for electronics-based PetaFLOPS-oriented designs. Our complete, distributed shared-memory design is coherent in terms of inter-PE data-transfer speeds and connectivity patterns throughout, thus supporting scalability and ease of mapping [23] application tasks to the system.

Other designs are characterized by limited bandwidth and substantial latencies that result in unpredictable performance. In contrast, the very efficient, rather uniform interconnection of resources in our system makes performance prediction much more accurate. Also, we expect algorithms for our system to be developed rather easily, by assuming an MIMD, fully-connected target architecture.

3.1 Building Block Implementation

Guided, planar, optical interconnects offer a robust system with built-in optical filtering at the expense of system complexity [17]. On the other hand, free-space interconnections are relatively simple to implement. However, they are more prone to system vibrations and source/detector misalignment. Free-space systems possess an energy-bandwidth product which is larger than their electronic counterpart [18]. Yet, both systems may take advantage of wavelength-division multiplexing (WDM) to enhance their performance. Our objective here is to produce a low-cost, powerful, free-space, reliable, communications system of low packaging complexity that incorporates some

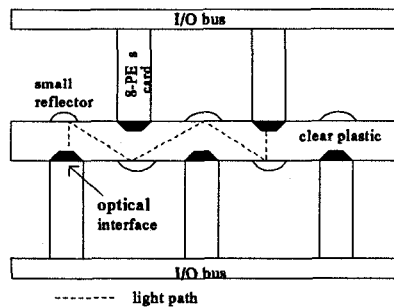


Figure 2: Block diagram for part of the building block.

guided-wave concepts.

Part of the BB is schematically shown in Figure 2. It is a 1-D stack of 8-PE cards, attached to an inexpensive, clear plastic bar that provides alignment. Each card carries the entire processing and memory power of eight processors fully interconnected via an electronic crossbar. This approach was chosen because of the high efficiency of small electronic crossbars. Each card is interfaced with optical transmitter/receiver modules and attached prismatic elements, and the destination address for a data transfer is decoded to determine the prismatic element and associated modules to be used for the appropriate path.

In an optical cycle, 32 bits of data are sent in parallel from one card to another via a card-to-card, color-coded interconnect, using 32 distinct colors (i.e., WDM); these colors are the same for all the cards. Each inter-card channel is viewed as being 128 bits wide, and therefore 128-bit information is transmitted each time by utilizing four (i.e., $128/32$) optical cycles. All eight PEs on a card share the same lasers, so that a PE P_{ij} on a given card, i , and with position j in the card, uses the same color set of 32 λ_m s, however with a different RF carrier frequency (i.e., time multiplexing). Assuming 128-bit data transfers and that the communication frequency for each PE is 375 MHz (SIA projection), each laser source of color hue, λ_m , has to operate at 12 GHz (i.e., $8 \times \frac{128}{32} \times 375$ MHz), in order to facilitate simultaneous inter-card data transfers involving all eight PEs. Such lasers already exist. The chosen prismatic element determines a specific optical path via the set of reflectors used between the transmitting and the receiving cards. Since any two cards communicate via dedicated prismatic elements, multi-access node communication is available. Common colors from different cards are detected by different detector arrays at different locations on the card's interface. Separation among the messages sent to a

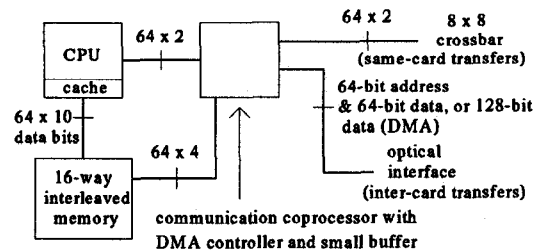


Figure 3: PE schematic diagram.

given card from other cards is made by separating the fields of view, and therefore activating different detectors on the receiving card.

A special module attached to the card's interface schedules data transmissions for the implementation of complex, frequently-used communication patterns, such as multicasting, broadcasting, gather, scatter, etc. In addition, a DMA controller is part of each PE in our detailed design, as shown in Figure 3; large numbers of DMA controllers are essential to sustained PetaFLOPS performance. We assume memory modules with two ports that permit simultaneous load/store by the CPU and the communication coprocessor.

Our optical design employs a subcarrier-based multiplexing/demultiplexing scheme. The receiver demultiplexes the information and sends it to the destined PE on the given card. The receiver may utilize a coherent detection system to increase its sensitivity by employing a distributed optical clock. The clock frequency may be transmitted on a color different from the information or may be incorporated in the address field.

A typical clear plastic has a transmission factor of 0.25 dB/cm. A typical coherent detection system is able to detect -30 dB of optical signals, which is translated to a maximum optical distance of $30/0.25 = 120$ cm. This is larger than 90 cm, the approximately largest optical distance for 36 cards, assuming that the distance between adjacent cards on the same side of the bar is about 5 cm (1.97 inches) and the thickness of the plastic bar is equal to 2.5 cm.

3.2 Complete System

Extension of the 1-D fully-connected BB of 288 PEs into a 2-D configuration is now in order. In addition to interfacing a horizontal clear plastic bar (used for 1-D interconnects), each 8-PE card now also belongs to a similar 1-D structure in the second dimension by in-

Table I: Bandwidth of communication links, including all associated latencies. All CPU to/from memory values are for single data transfers.

Type of Data Transfer	Bandwidth
CPU/local-memory	30 GBytes/s
CPU/RMSC	1.540 GBytes/s
CPU/RMAC	1.075 GBytes/s
Memory/memory via DMA	6 GBytes/s
Bisection bandwidth	31.104 TeraBytes/s

RMSC: remote memory on the same card.

RMAC: remote memory on another card.

terfacing a vertical clear plastic bar. The clear plastic columns are patterned with small metallic reflectors and prismatic interfaces, as for the horizontal bars. All in all, the system may be viewed as a 2-D mesh of rows and columns of fully-connected PEs, and therefore it is equivalent to a versatile 2-D generalized hypercube. This complete system contains $(36)^2 = 1,296$ 8-PE cards and 10,368 processors. Our system is scalable in terms of both its architecture and the optics technology, and therefore further performance improvement is possible, if desired.

A. Performance Characteristics

Table I summarizes the performance characteristics of the BB and the bisection bandwidth of the complete system. It is easy to see the outstanding performance of the system's interconnection network. There can not be any realistic electronic implementation of an interconnection network that could match these characteristics.

We assume 10,368 processors of 10 GFLOPS and 1 GHz each (SIA projections). Each 10-way multithreaded processor operates on 64-bit integers and 128-bit floating-point numbers. A message directed to another processor card contains 64-bit data, and a 64-bit field with the source and destination PE addresses and the address of a memory location in the distributed shared-memory system; some control information also may be included in the latter field.

An intelligent, very high-performance decoding system has been designed for directing the data from/to the communication coprocessor to/from the right set of LDs/detectors. This design is based heavily on associative-memory technology. Finally, the final product of our effort will definitely include hardware support, at numerous levels of the memory hierarchy, for performance monitoring and data reassignments, that can be taken advantage of by the compiler and

the operating system.

3.3 Feasibility Analysis

Although the required number of 8-PE cards in the BB is 36, our analysis here is for 40 cards. The four additional cards in each BB could be used for other services (e.g., I/O and fault tolerance), or to increase the size of the proposed system to 12,800 processors for 128 TeraFLOPS peak performance.

A. Optical Interface

The optical interface is composed of prismatic elements, and two dedicated arrays of laser diodes (LDs) and detectors arranged underneath each element. The element is acting like a collimating lens that directs the light out of 32 lasers to the detector array on the destination card through the appropriate reflectors. The same optical element focuses the incoming light from the sender onto the dedicated detector array. Each detector in the 32-element array is equipped with a color filter that allows only a particular color hue to pass through. In this way, we separate further the 32 bits from one another. With current technology, each LD or detector could occupy a square area of 0.5 mm \times 0.5 mm (including its electrodes), and the entire area occupied by the 32 LD/detector pair array will be 1 mm \times 16 mm. Since the interface extends throughout the width of the card (about 20 cm), we may divide the 39 LD/detector pair arrays into ten groups. Therefore, the width of the optical interface, as viewed from above, will be about 4 mm.

B. Light Sources and Detectors

The light sources will be LDs made of GaAs at wavelengths between 0.8 and 0.9 μ m. The GaAs technology is a mature technology which is able to produce high-speed lasers at a reasonable cost. The detectors will make use of silicon technology. Each detector is equipped with a thin layer which serves as an optical filter. The filter for an array of detectors is easily made in an incremental manner. Chromatic dispersion is negligible at these distances (only 10^{-11} sec for 1 meter of propagation).

C. Optical/Electrical Power Consumption

Our analysis here is very conservative, even for current optical technologies. Each LD puts out an average of 250 μ W of optical power. This power is smaller or larger for short or long data transmissions, respectively. Each card transmits information to each of the 39 other cards via 32 dedicated LDs, and radiates, on the average, $39 \times 32 \times 250 \mu\text{W} = 312 \text{ mW}$ of optical power. The electrical-to-optical power conversion ratio is normally 30%, thus the RMS (root-mean-square) value of the electrical power consumption for each card

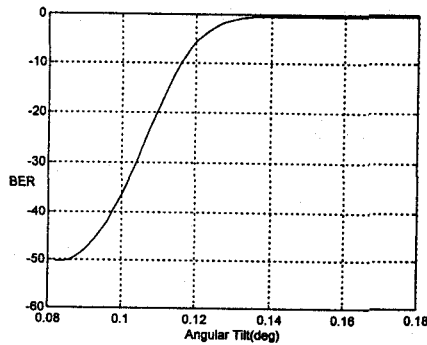


Figure 4: The bit-error rate (\log_{10} BER) as a function of the angular tilt (AT), for the most distant data transfers.

is about 1 W. Expected improvements in optoelectronic devices will further enhance these numbers.

D. Bit-Error Rate (BER)

We have simulated the optical system within the BB. We present here BER results for the longest transmission, namely 1 meter. With a 12GHz LD of 0.5 mW power for this largest distance, the angular tilt (AT) allowed before crosstalk occurs is 0.1° . Figure 4 shows the BER as a function of AT. For $AT=0.1^\circ$, the BER is much better than what is required (e.g., 10^{-15}).

4 Performance Evaluation

Our general-purpose MIMD system targets the majority of the computation-intensive applications. Further usability analysis was carried out by investigating the efficiency of emulating topologies widely used for the development of algorithms. The results are impressive, as expected and shown to some extent for generalized hypercubes in [21]. Wide usability is further substantiated through additional performance evaluation.

Theoretical analysis and simulations are used for a highly accurate performance evaluation of the proposed system. For a system to potentially have a niche in the massively-parallel processing field, it must provide direct support for some very frequently used communication operations that are very costly to implement by repeating some of the basic communication primitives. Such operations are: multicasting, broadcasting, reduction using associative operators, prefix computations, and barrier synchronization. Other less frequent operations are one-to-all personalized and its

Table II: Performance results.

Algorithm	Timing
Alg. I	7.08 μ s
Alg. II	51.46 μ s
Alg. III	25.76 μ s
Alg. IV	10.17 ms
Alg. V	135.78 ms

Alg. I: SAXPY loop. Alg. II: large-stride vector fetch and store. Alg. III: irregular gather/scatter. Alg. IV: Jacobi 1. Alg. V: Jacobi 2.

dual single-node gather communications, and total exchange. Our results show that these communications can be carried out consistently and efficiently throughout the entire system.

A. Algorithms

The efficient mapping of application algorithms onto the proposed system is vital to its success. This task benefits tremendously from the versatility of our system's communications structure. For highly accurate evaluation of the system, its performance will be estimated for the kernels of the following set of computation-intensive applications: weather forecasting, image benchmark, multivariable spline-blending approximation, grid interpolation, and traffic-flow simulation of freeway networks.

In addition, the implementation of the kernels for some important algorithms that were assigned during the PetaFLOPS Architecture Workshop (April 1996) was investigated. The expected execution times are shown in Table II. The total number of element assignments for each of the first three algorithms is 10^8 . Alg. I has a sequential loop $c(i) = \pi * a(i) + b(i)$. Alg. II has $b(121 * i) = a(131313 * i)$. Alg. III has two loops: $n(i) = \text{mod}(13131313 * i, n)$ and $b(n(i)) = a(i)$. The Jacobi kernels in Algs. IV and V have five nested loops. The outermost three loops count from 1 to 1000 and the innermost two loops count from 1 to 5 in Alg. IV. The corresponding loop counts for Alg. V are 100 (outermost three loops) and 150 (innermost two loops). The Jacobi kernels were restructured to allow for parallel operations within the diagonals of a 2-D matrix. Algs. IV and V then use 5,000 and 10,000 PEs, respectively. These results further prove the suitability of our case-study system for near PetaFLOPS computing.

5 Conclusions

We have proven in this paper the suitability of our "point design" for near PetaFLOPS, and more, computing. The complete system is characterized by immense bisection bandwidth and other outstanding topological properties. Not only can our proposed system graciously achieve its performance objective, but also its dramatically low interconnect complexity renders it viable. Such a dramatic reduction in the system interconnect complexity is not possible with any other existing or expected technology. Preliminary performance results were also employed to support our claim of outstanding performance.

References

- [1] L.N. Bhuyan and D.P. Agrawal, "Generalized Hypercube and Hyperbus Structures for a Computer Network," *IEEE Trans. Comput.* 33 (4), 1984, pp. 323-333.
- [2] W.J. Dally and C.L. Seitz, "The Torus Routing Chip," *Distr. Comp.* 1, 1986, pp. 187-196.
- [3] M.C. Pease, III, "The Indirect Binary n-Cube Microprocessor Array," *IEEE Trans. Comp. C-26*(5), 1977, pp. 458-473.
- [4] C.L. Seitz, "Concurrent VLSI Architectures," *IEEE Trans. Comp. C-33*(12), 1984, pp. 1247-1265.
- [5] T. Szymanski, "A Fiber Optic Hypermesh for SIMD/MIMD Machines," *Supercomp. Conf.*, 1990, pp. 103-110.
- [6] L.D. Wittie, "Communication Structures for Large Networks of Multicomputers," *IEEE Trans. Comput. C-30*(4), 1981.
- [7] X. Li, S. G. Ziavras, et al., "Parallel DSP Algorithms on TurboNet: An Experimental Hybrid Message-Passing/Shared-Memory Architecture," *Conc.: Pract. Exp.* 8(5), 1996, pp. 387-411.
- [8] S.G. Ziavras, "RH: A Versatile Family of Reduced Hypercube Interconnection Networks," *IEEE Trans. Paral. Distr. Syst.* 5(11), 1994, pp. 1210-1220.
- [9] S.G. Ziavras, "Generalized Reduced Hypercube Interconnection Networks for Massively Parallel Computers," in: *Networks for Parallel Computations*, Amer. Math. Soc., D.F. Hsu, A. Rosenberg, and D. Sotteau (Eds.), 1995, pp. 307-325.
- [10] S.G. Ziavras and A. Mukherjee, "Data Broadcasting and Reduction, Prefix Computation, and Sorting on Reduced Hypercube Parallel Computers," *Parallel Computing* 22, 1996.
- [11] D. Lenoski, et al., "The Stanford FLASH Multiprocessor," *IEEE Comp.* 3, 1992, pp. 63-79.
- [12] A.L. Cox, et al., "Software Versus Hardware Shared-Memory Implementation: A Case Study," *Int'l Symp. Comp. Arch.*, 1994, pp. 106-117.
- [13] S.G. Ziavras, "On the Problem of Expanding Hypercube-Based Systems," *Journ. Paral. Distr. Comp.* 16(1), 1992, pp. 41-53.
- [14] S.G. Ziavras, "Scalable Multifolded Hypercubes for Versatile Parallel Computers," *Paral. Proc. Letts.* 5(2), 1995, pp. 241-250.
- [15] E. Frietman, et al., "Parallel Optical Interconnects: Implementation of Optoelectronics in Multiprocessors Architecture," *Appl. Opt.* 29, 1990.
- [16] K. Aly and P. Dowd, "Parallel Computer Reconfigurability Through Optical Interconnects," *Int'l. Conf. Paral. Proc.*, 1992, pp. I:105-137.
- [17] S.C. Tsay and H. Grebel, "Design of Transverse Holographic Optical Interconnect," *Appl. Optics* 33, 1994, pp. 6747-6752.
- [18] L. Camp, et al., "Guided-Wave and Free Space Optical Interconnect for Parallel Processing Systems: A Comparison," *Appl. Opt.* 33, 1994, pp. 6168-6180.
- [19] *The National Technology Roadmap for Semiconductors*, Semicond. Industry Assoc. (SIA), 1994.
- [20] H. Grebel and W. Zhong, "Holographic Interconnects: Transverse Bragg Waveguides," *Opt. Letts.* 18, 1993, pp. 1123-1125.
- [21] P. Fragopoulou, et al., "Optimal Communication Primitives on the Generalized Hypercube Network," *Journ. Paral. Distr. Comp.* 32, 1996.
- [22] W. Dally, "Network and Processor Architecture for Message-Driven Computers," in: *VLSI and Parallel Computation*, R. Suaya and G. Birtwistle (Eds.), Morg. Kauf. Publ., 1990, pp. 140-222.
- [23] D.A. Patterson, "Teraflops Design in Eight Easy Steps," *Scient. Amer.*, Jan. 1991, pp. 104-105.