

Application-Specific Optical Interconnects for Embedded Multiprocessors

Neal K. Bambha and Shuvra S. Bhattacharyya

Abstract—As transistor sizes shrink and we approach the “end of Moore’s law”, interconnects—both on-chip and off-chip—will represent the biggest bottleneck for embedded systems designers. Several groups are researching optical interconnects to cope with this trend. Optical interconnects enable new system architectures. These new architectures in turn require new methods for high-level application mapping and hardware/software co-design. In this presentation, we discuss high-level scheduling and interconnect topology synthesis techniques for embedded multiprocessors. We focus on designs that are streamlined for one or more digital signal processing (DSP) applications. That is, we seek to synthesize an *application-specific interconnect topology* for a multiprocessor DSP design. We show that flexible interconnect topologies that allow single-hop communication between processors offer advantages for reduced power and latency.

We have previously shown that multiprocessor scheduling algorithms can deadlock in the general case of a topology graph that is not strongly connected, or if communication is limited to be single hop. We have also demonstrated an efficient algorithm that can be used in conjunction with existing scheduling algorithms for avoiding this deadlock [1]. In this presentation we discuss the advantages of performing application scheduling and interconnect synthesis jointly, and present a probabilistic scheduling/interconnect algorithm utilizing graph isomorphism to pare the design space. We demonstrate the performance advantages that an application-specific interconnect topology can produce for several DSP benchmarks.

Index Terms—interconnect synthesis, multiprocessor, scheduling.

I. INTRODUCTION

Interconnect considerations are important for today’s embedded systems designs. As transistor density increases, more functional units can be placed on a single chip, and the number of possible interconnections (links) between them increases. The longest wires on the chip are usually due to these links. These wires contribute to delay and limit the maximum achievable clock rate. Also,

routing these interconnections is a significant challenge for the electronic design automation tools.

Embedded systems typically run a limited and fixed set of applications. We can use this application-specific information to optimize the interconnection network. For our purposes, an optimal network is defined in the context of a set of applications and constraints. The constraints may include the latency, throughput, and power consumption for the given applications, along with cost and area constraints of the overall system. A key distinguishing feature to our algorithm is that we perform the application scheduling and interconnect synthesis jointly.

A. Optical Interconnects

In recent years, optics have played an increasing role in multiprocessor systems. Commercial high-performance computers now use fiber ribbons to connect multiple processing nodes. Other examples include storage area networks using fiberchannel, and optical clock distribution to reduce clock skew across a chip. Programs such as the DARPA VLSI Photonics [2] program are pushing to integrate photonics technology on a single chip. Intel is currently backing an effort to bring “fiber-to-the-processor” [3]. The idea is to break the processor to cache bottleneck by using an optical waveguide integrated on the processor chip.

B. Connection Topologies

Electrically connected multiprocessor systems generally have a regular interconnection pattern, due to the physical constraints imposed by two-dimensional circuit board layout. Some examples include ring, mesh, bus, and hypercube interconnect topologies. Using these topologies, communication between remote processors requires multiple hops, which increases both latency and power, and increases contention throughout the network.

In contrast, optically connected multiprocessors, particularly those utilizing free space optics and three dimensions, are free to utilize arbitrarily irregular interconnection networks. Once the signal is in the optical domain, there is very little attenuation, so the energy

N. Bambha is with the US Army Research Laboratory.

S. Bhattacharyya is with the Department of Electrical and Computer Engineering and Institute for Advanced Computer Studies, University of Maryland, College Park.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 01 FEB 2005		2. REPORT TYPE N/A		3. DATES COVERED -	
4. TITLE AND SUBTITLE Application-Specific Optical Interconnects for Embedded Multiprocessors				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory; University of Maryland, College Park				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES See also ADM00001742, HPEC-7 Volume 1, Proceedings of the Eighth Annual High Performance Embedded Computing (HPEC) Workshops, 28-30 September 2004 Volume 1., The original document contains color images.					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Application	N	$\Delta(E)(\%)$	$\Delta(M)(\%)$
FFT1	7	16	8
Karp10	6	24	4
Irr	8	16	(2)
Qmf4	7	32	3
NN16-3-4	8	58	2
Sum1	6	1	4
Laplace	7	4	(3)
FFT2	7	12	2

TABLE I

REDUCTION IN COMMUNICATION ENERGY ($\Delta(E)$) AND
MAKESPAN INCREASE ($\Delta(M)$) OF SINGLE HOP SCHEDULE OVER
THREE-HOP SCHEDULE.

required to transmit a unit of data is essentially independent of distance. The required energy instead is a function of the number of electrical-to-optical conversions that must be performed [4], which in turn is determined by the number of hops. With single-hop schedules the overhead associated with routing data through intermediate processors is eliminated. Furthermore, due to the flexibility of the communication medium, it is generally possible to avoid multi-hop communication operations by simply activating direct communication channels between the source and destination processors. Together, these properties make it desirable to limit the number of hops per communication operation when exploring configurations (interconnection patterns and task graph mappings) for an optically connected, embedded multi-processor.

In order to quantify this effect, we scheduled several DSP benchmark applications using our modified scheduling technique, which takes the number of hops as an input parameter. We scheduled the benchmarks with hop constraints of one hop and three hops, and compared the communication energy required. For our purposes, we assumed all communication tasks transferred the same number of bits, so the energy cost of all IPC actors was equal. Table I shows the reduction in the required communication energy for single-hop schedules over three-hop schedules for the benchmark applications. For these benchmarks, we found that any undesirable effect on the makespan of the additional constraint for single-hop schedules was very small, as can be seen in Table I. In two of the benchmarks (Irr and Laplace), the makespan was in fact better (lower) when we limited the scheduler to single hops.

C. Interconnect Synthesis Algorithm

We present a genetic algorithm for synthesizing efficient interconnection networks for embedded multipro-

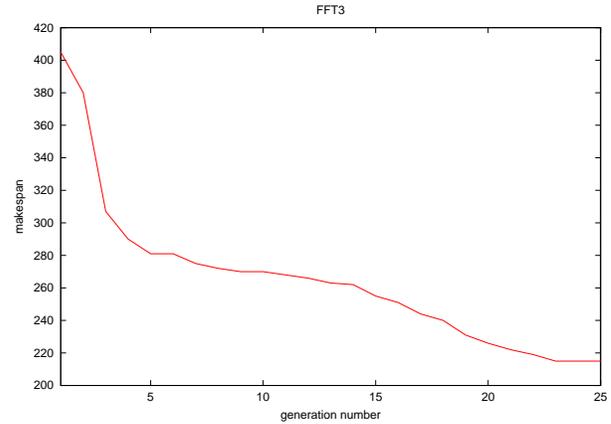


Fig. 1. GA output versus generation.

cessors. The algorithm works in conjunction with a list scheduling algorithm to jointly optimize both the schedule and the interconnect topology. The algorithm is able to account for different distributions of local vs. global (long) interconnect routing tracks via a processor *fanout constraint*. It uses graph isomorphism to significantly pare the search space in order to search more efficiently.

We evaluated our interconnect synthesis algorithm on several DSP benchmark application graphs. Figure 1 shows the convergence of the GA vs. generation number for an FFT application with a maximum allowed fanout of 4. In this plot the y-axis refers to the schedule makespan of the best interconnection topology found for a given generation. We see that GA was able to reduce the makespan by almost a factor of two over the best topology in the initial population.

We calculated how the makespan improves as the maximum fanout constraint is increased. This amounts to an area/performance tradeoff in the system. We also compared the performance of systems with topologies available with electrical interconnects vs. optical interconnects. These topics will be described in the presentation.

REFERENCES

- [1] N. K. Bambha, S. S. Bhattacharyya, and G. Euliss, "Design considerations for optically connected systems on chip," in *Proceedings of the International Workshop on System on Chip for Real Time Processing*, Calgary, Canada, June 2003, pp. 299–303.
- [2] "http://www.darpa.mil/mto/vlsi."
- [3] (2004) The Primarion website. [Online]. Available: <http://www.primarion.com>
- [4] R. Kostuk, J. Goodman, and L. Hesselink, "Optical imaging applied to microelectronic chip-to-chip interconnections," *Applied Optics*, pp. 2851–2858, 1985.

Abstract

As transistor sizes shrink and we approach the "end of Moore's law", interconnects, both on-chip and off-chip, will represent the biggest bottleneck for embedded systems designers. Several groups are researching optical interconnects to cope with this trend. Optical interconnects enable new system architectures. These new architectures in turn require new methods for high-level application mapping and hardware/software co-design. In this presentation, we discuss high-level scheduling and interconnect topology synthesis techniques for embedded multiprocessors. We focus on designs that are streamlined for one or more digital signal processing (DSP) applications. That is, we seek to synthesize an *application-specific interconnect topology* for a multiprocessor DSP design. We show that flexible interconnect topologies that allow single-hop communication between processors offer advantages for reduced power and latency.

We have previously shown that multiprocessor scheduling algorithms can deadlock in the general case of a topology graph that is not strongly connected, or if communication is limited to be single hop. We have also demonstrated an efficient algorithm that can be used in conjunction with existing scheduling algorithms for avoiding this deadlock. In this presentation we discuss the advantages of performing application scheduling and interconnect synthesis jointly, and present a probabilistic scheduling/interconnect algorithm utilizing graph isomorphism to pare the design space. We demonstrate the performance advantages that an application-specific interconnect topology can produce for several DSP benchmarks.

Deadlock and Flexibility

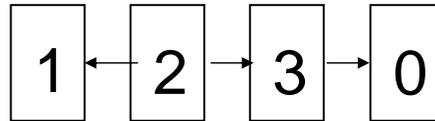
- Existing scheduling algorithms assume every pair of processors can communicate
- Scheduling not well studied for irregular interconnection networks
- Can *deadlock* for arbitrary topologies
 - Developed algorithms to adapt existing list scheduling algorithms to avoid deadlock
- Some scheduling moves have greater *flexibility*

Partial Schedule

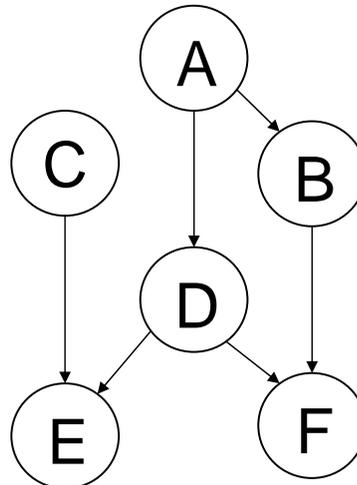
A on proc. 2
B on proc. 1

Constraint Sets

$F[A] = \{2\}$
 $F[B] = \{1\}$
 $F[F] = \{1\}$
 $F[D] = \{1,2\}$
 $F[E] = \{1,2,3\}$
 $F[C] = \{1,2,3\}$
 Flexibility = 11/24



Topology graph



Application graph

Partial Schedule

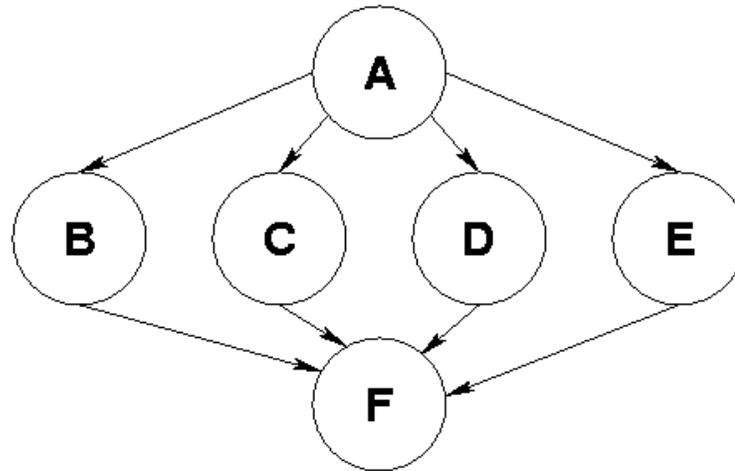
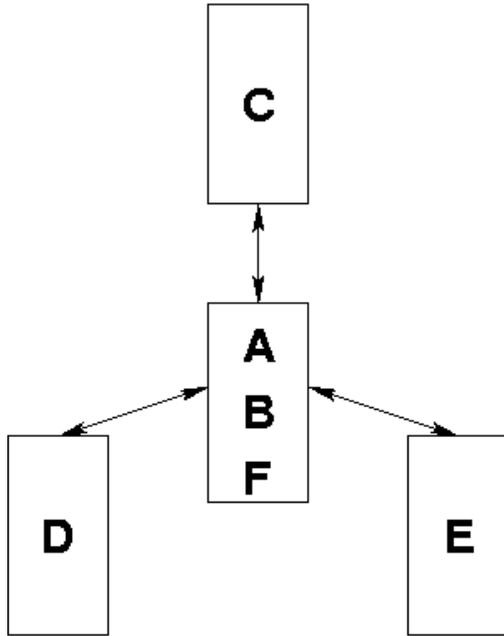
A on proc. 2
B on proc. 3

Constraint Sets

$F[A] = \{2\}$
 $F[B] = \{3\}$
 $F[F] = \{0,3\}$
 $F[D] = \{1,2,3\}$
 $F[E] = \{0,1,2,3\}$
 $F[C] = \{0,1,2,3\}$
 Flexibility = 15/24

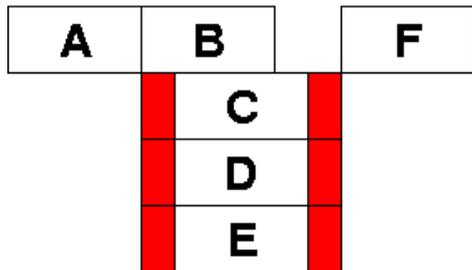
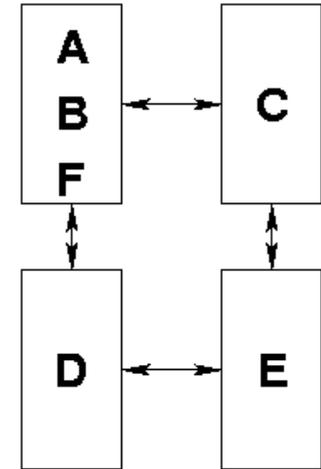
Effect of Topology

Topology 1



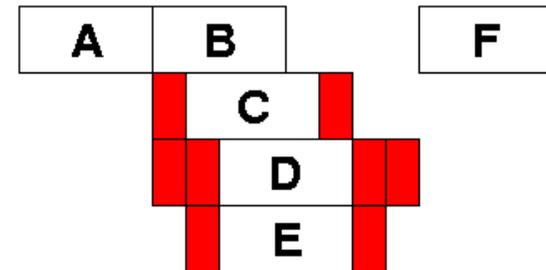
Application Graph

Topology 2



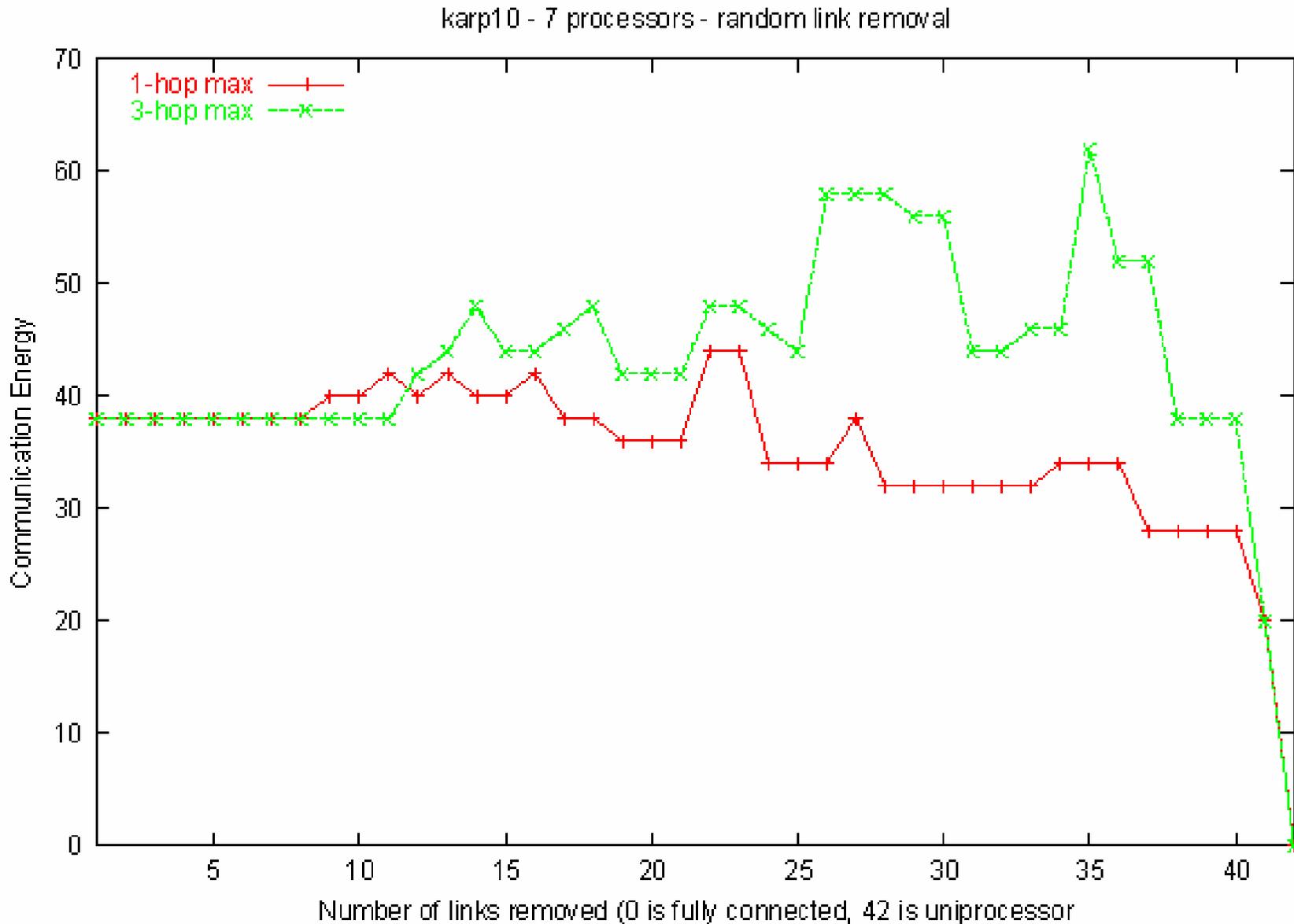
← Latency 1 →

Lower latency and communication energy for topology 1



← Latency 2 →

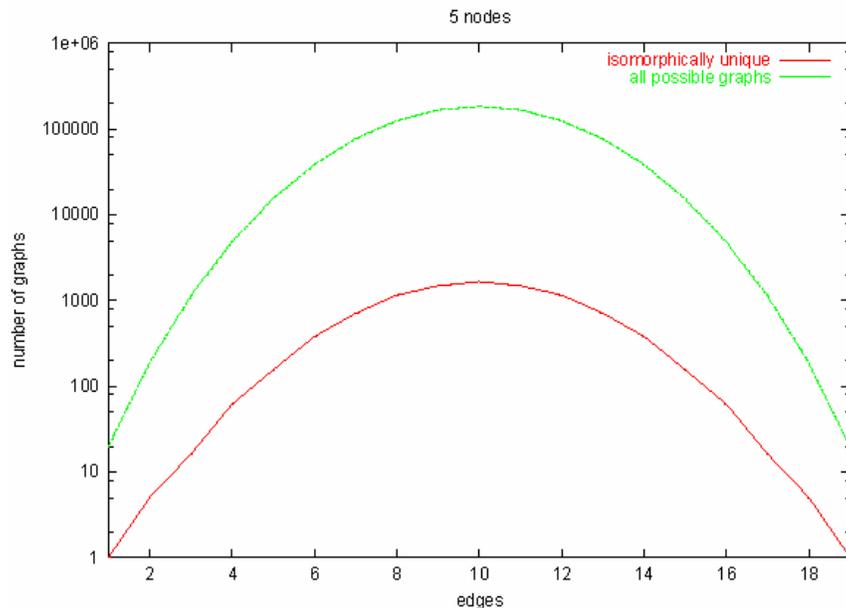
Low Hop Communication Saves Energy



Link Synthesis Algorithm

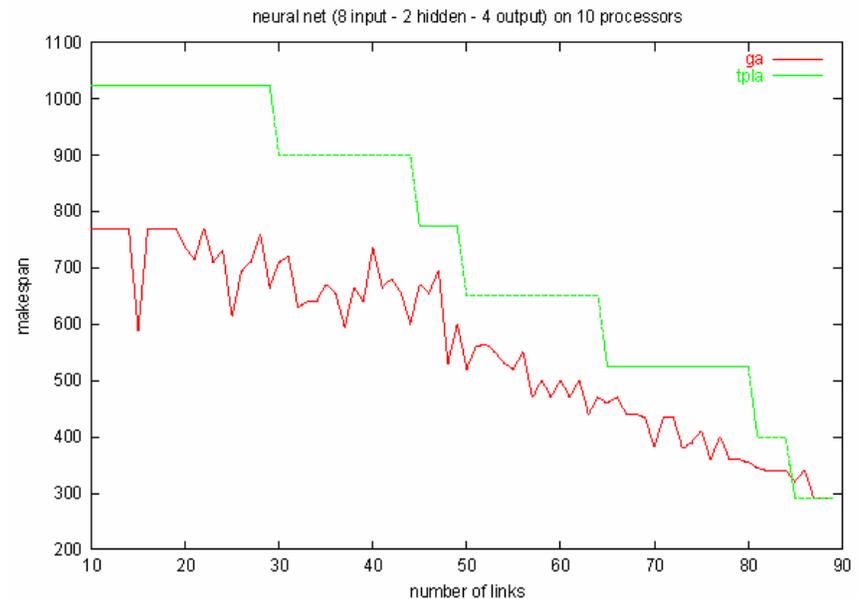
- Developed both deterministic and evolutionary (GA) algorithms
- GA objective utilizes DLS scheduling modified with flexibility metric
 - Crossover operators allow fan-out constraints to be preserved
 - Use graph isomorphism to pare the design space

Paring the design space



- Consider only isomorphically unique graphs
- Reduction by orders of magnitude

Link synthesis results



GA (red) outperforms deterministic over a range of topologies

Application-Specific Optical Interconnects for Embedded Multiprocessors

Neal K. Bambha

US Army Research Laboratory

Shuvra S. Bhattacharyya

University of Maryland, College Park

`nbambha@eng.umd.edu`, `ssb@eng.umd.edu`



Introduction

- Develop software tools and algorithms to efficiently map digital signal and image processing (“DSP”) applications onto Systems on Chip.
 - Joint scheduling/interconnect synthesis optimization
 - Scheduling for low-hop communication on arbitrary topologies
 - Synthesize an optimal *application-specific* interconnect topology

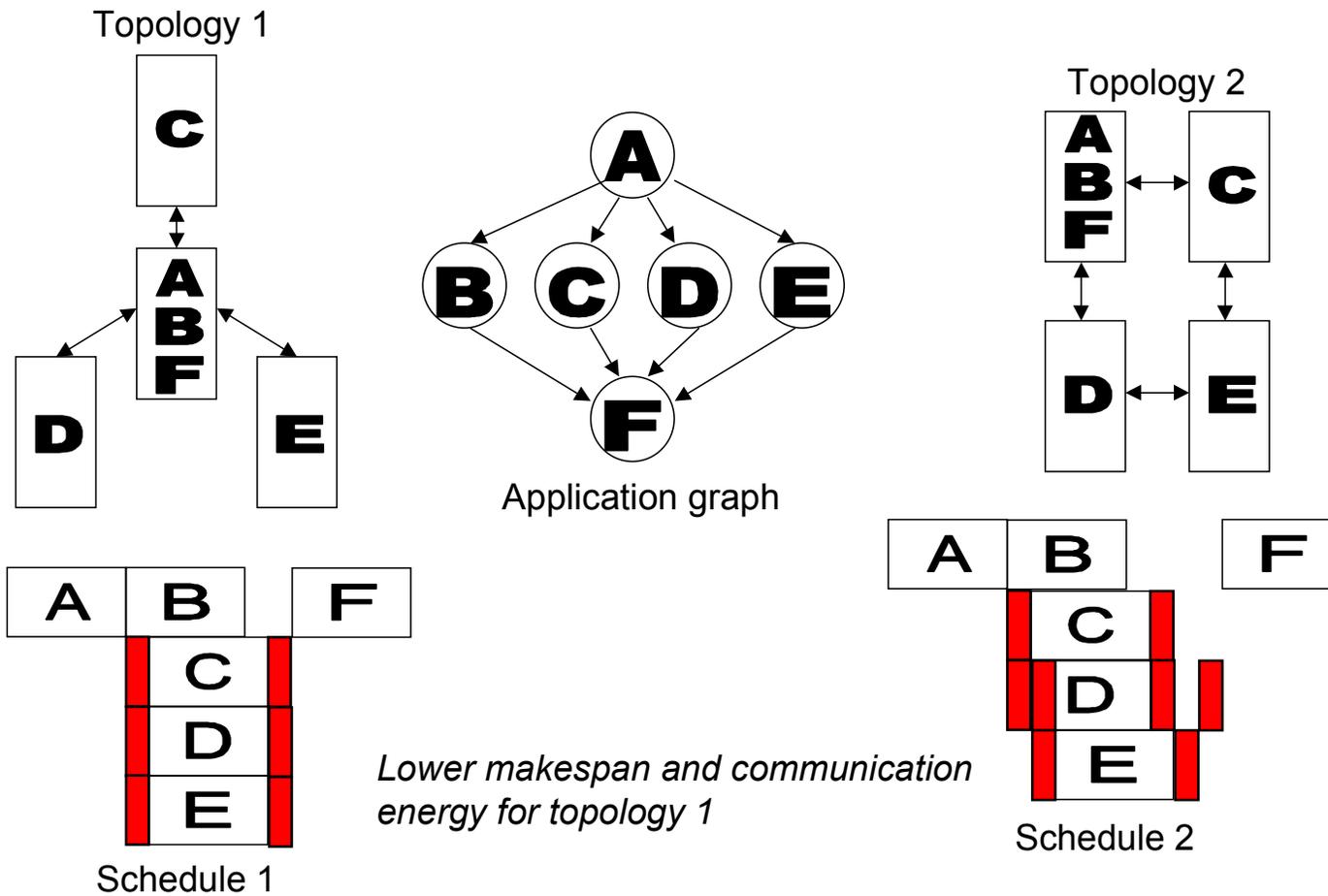


Scheduling

- Task graph $G(V, E)$, $\nu \in V$, $e \in E$
 - Dataflow specification
 - General point-to-point networks
- Topology graph $T(P, L)$, $p \in P$, $l \in L$
 - Link constraints
 - Processor fanout constraints
 - $l = (p_i, p_j)$ assigned weights—delay and power
 - $\mathcal{E}(G, T, n) = \sum_{e \in E} \left(\text{IPC}(e) \sum_{l \in \text{route}(e)} \epsilon_{\text{bit}}(l) \right)$
- Communication hop limit

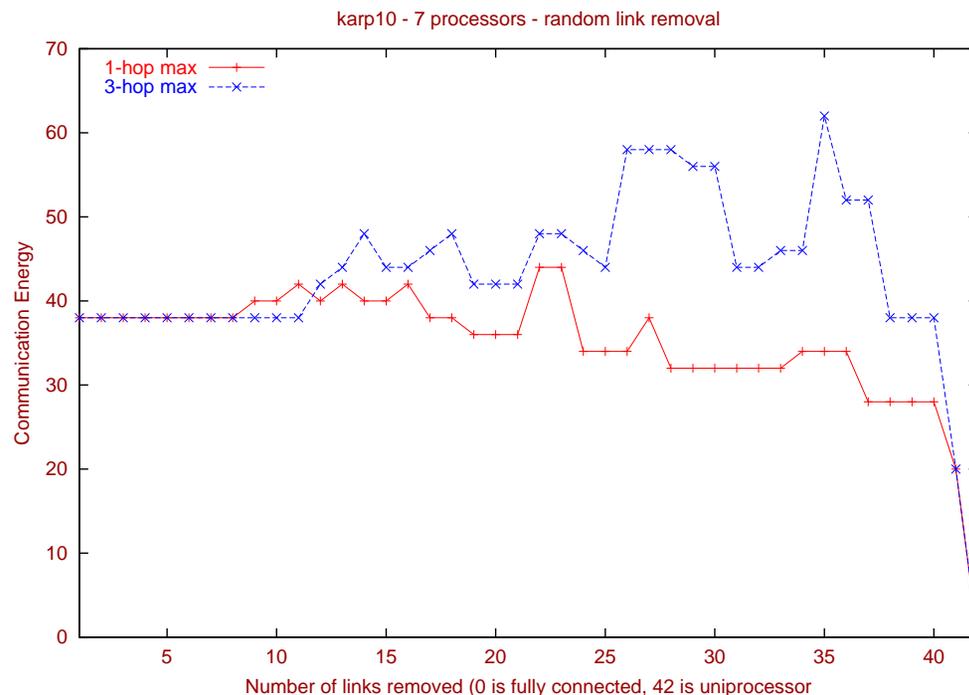


Effect of Topology





Low Hop Communication Saves Energy



Compare communication energy across a range of randomly generated topologies with single-hop and 3-hop limit.



Application-Specific Interconnect Topologies

- Design constraints for optical interconnects
 - Topology—total links, maximum fanout
 - Performance—throughput, power
- Joint schedule/topology optimization
 - GA generates population of solution candidates $T(P, L)$
 - Scheduler evaluates fitness of each T
 - * DLS adapted for arbitrary topologies
 - * Avoids deadlock, calculates *flexibility*
 - * Constructs hop-limited schedules
 - Given constraints on T , maximize performance
 - Given constraints on performance, optimize T