

Requirements for Scalable Application Specific Processing in Commercial HPEC

Steven Miller

Silicon Graphics, Inc.

Phone: 650-933-1899

Email Address: scm@sgi.com

Abstract: More and more High Performance Embedded Computing (HPEC) leverages technology from commercial high performance computing systems. To date, HPEC has only tapped the lower end of commercial high performance computing technology. As more of the advanced commercial technology moves into the embedded space, this presents a unique opportunity to change the fundamentals of how HPEC solutions are addressed.

Within HPEC, two types of application specific processing elements, reconfigurable and custom are being used. Reconfigurable elements are comprised of Field Programmable Gate Array (FPGA) technology and custom elements are comprised of various devices as Digital Signal Processors (DSP), DARPA Polymorphic Computing Architectures (PCA) [1] and others. The integration of these devices presents significant challenges both to the system architecture and to the programming models. This presentation will describe a set of system requirements and methods to not only include these application specific processing devices but to allow effecting scaling of application specific processing devices.

Introduction and System Architectural Review

SGI's ccNUMA (cache coherent non-uniform memory architecture) [2] global shared memory system architecture is the basis for our general-purpose Origin and Altix HPC systems. The presentation will explore the architectural features used within both Origin and Altix systems which allows scaling in excess of one thousand commercial high performance processors. The architecture of SGI's systems which included 192 custom application specific processors (Tensor Processor Units) and 128 general-purpose processors will also be described.

System Architectural Features for Scalability

High-performance FPGAs represent an important architectural tool to provide total system performance while keeping both power and space requirements to a minimum. The use of FPGA elements have been limited to only the most computationally dense algorithms by the amount of data which can be pass through the device. But even with these limitations, significant performance and power improvements have been demonstrated [3]. This presentation will describe methods to increase the bandwidth to these devices and the communication primitives required to scale to hundreds of devices with out sacrificing performance.

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 Requirements for Scalable Application Specific Processing in Commercial HPEC				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) Silicon Graphics, Inc.				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 UU	18. NUMBER OF PAGES 17	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Systems that include hundreds of application specific processing devices pose a significant software challenge. This presentation will describe how to effectively allocate, manage, and decommission under changing workloads various application specific processing elements. Both software methods and APIs will be presented on ways to interface application specific processing targeted at HPEC applications.

References

- [1] <http://www.darpa.mil/ipto/programs/pca/index.htm>
- [2] <http://www.sgi.com/servers/altix/>
- [3] **W.D. Smith and A.R. Schnore. Towards an RCC-based Accelerator for Computational Fluid Dynamics Applications. In T.P. Plaks, eds., Proceedings of the International Conference of Engineering of Reconfigurable Systems and Algorithms, pp. 222-231. CSREA Press, Las Vegas, Nevada, 2003.**

Requirements for Scalable Application Specific Processing in Commercial HPEC

Steve Miller
Chief Engineer



The 3 Single-Paradigm Architectures

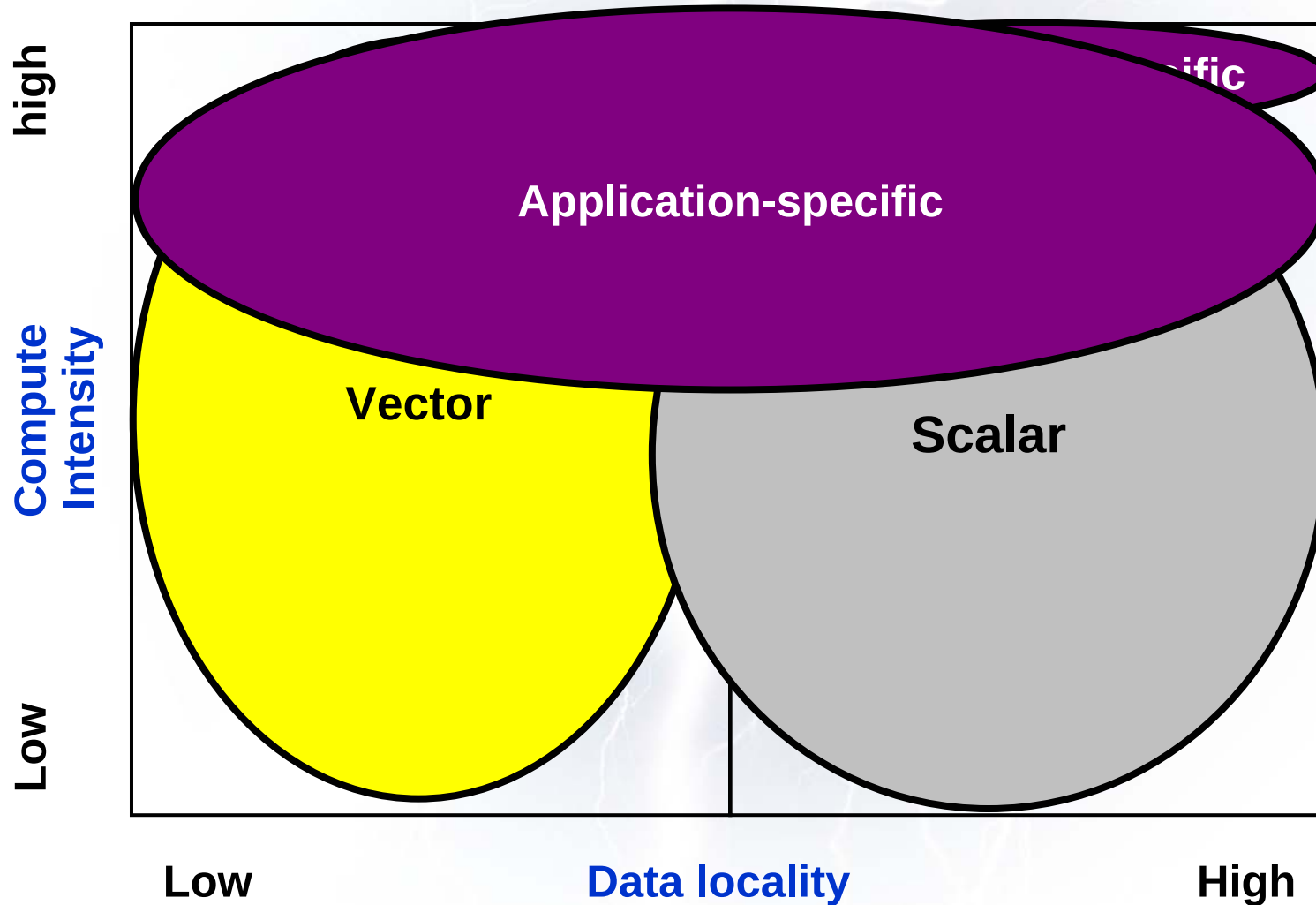


<u>Scalar</u>	<u>Vector</u>	<u>App-Specific</u>
Intel Itanium	Cray X1	Graphics - GPU
SGI MIPS	NEC SX	Signals - DSP
IBM Power		Prog'ble - FPGA
Sun SPARC		Other ASICs
HP PA		

Igniting Innovation
and Leadership



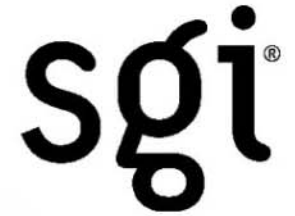
Paradigms to Applications



Igniting Innovation
and Leadership



Architectural Challenges

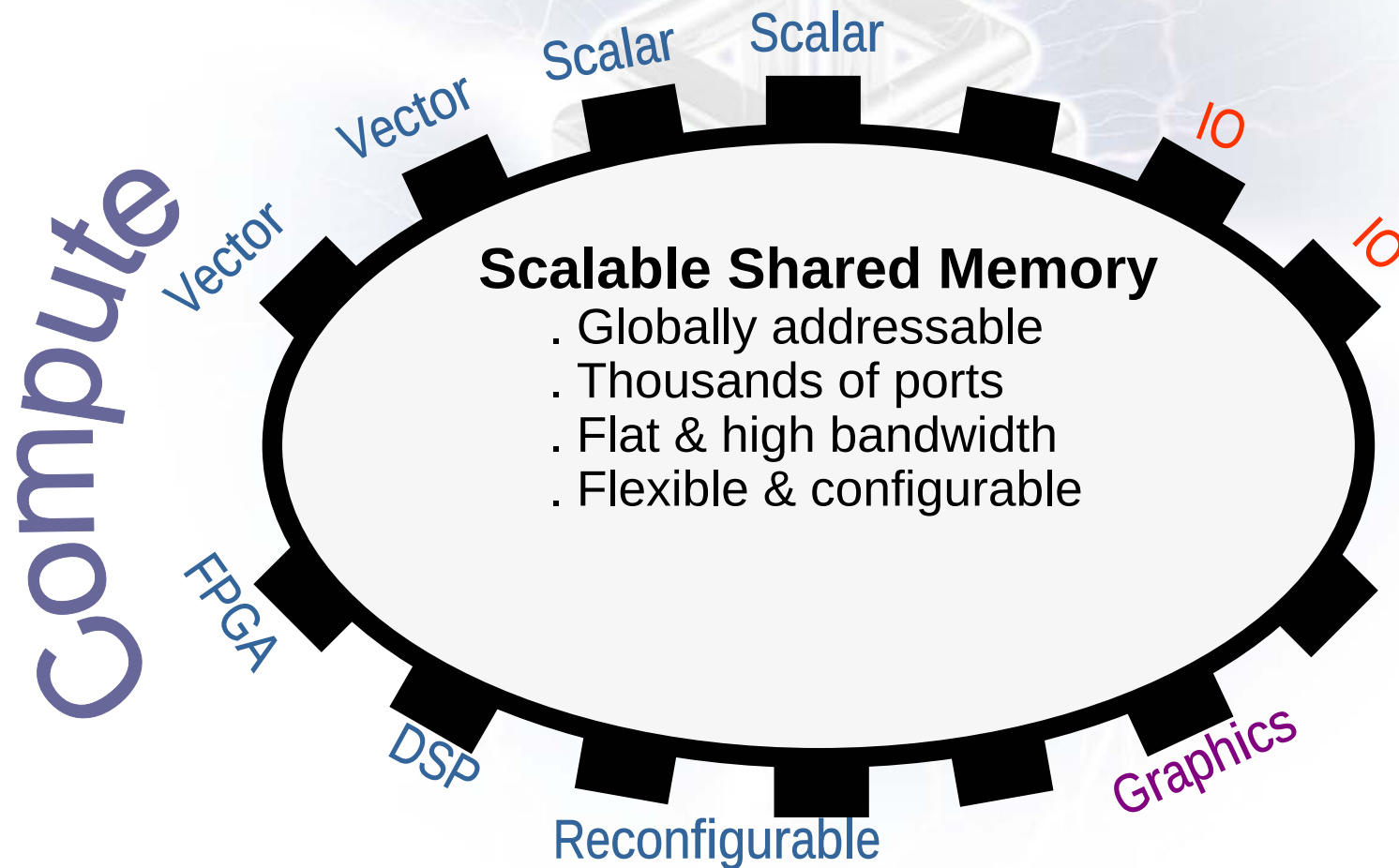


- Hardware
 - Bandwidth to/from System
 - Scalability
- Software
 - Compilers/Languages
 - Debuggers
 - APIs

Igniting Innovation
and Leadership



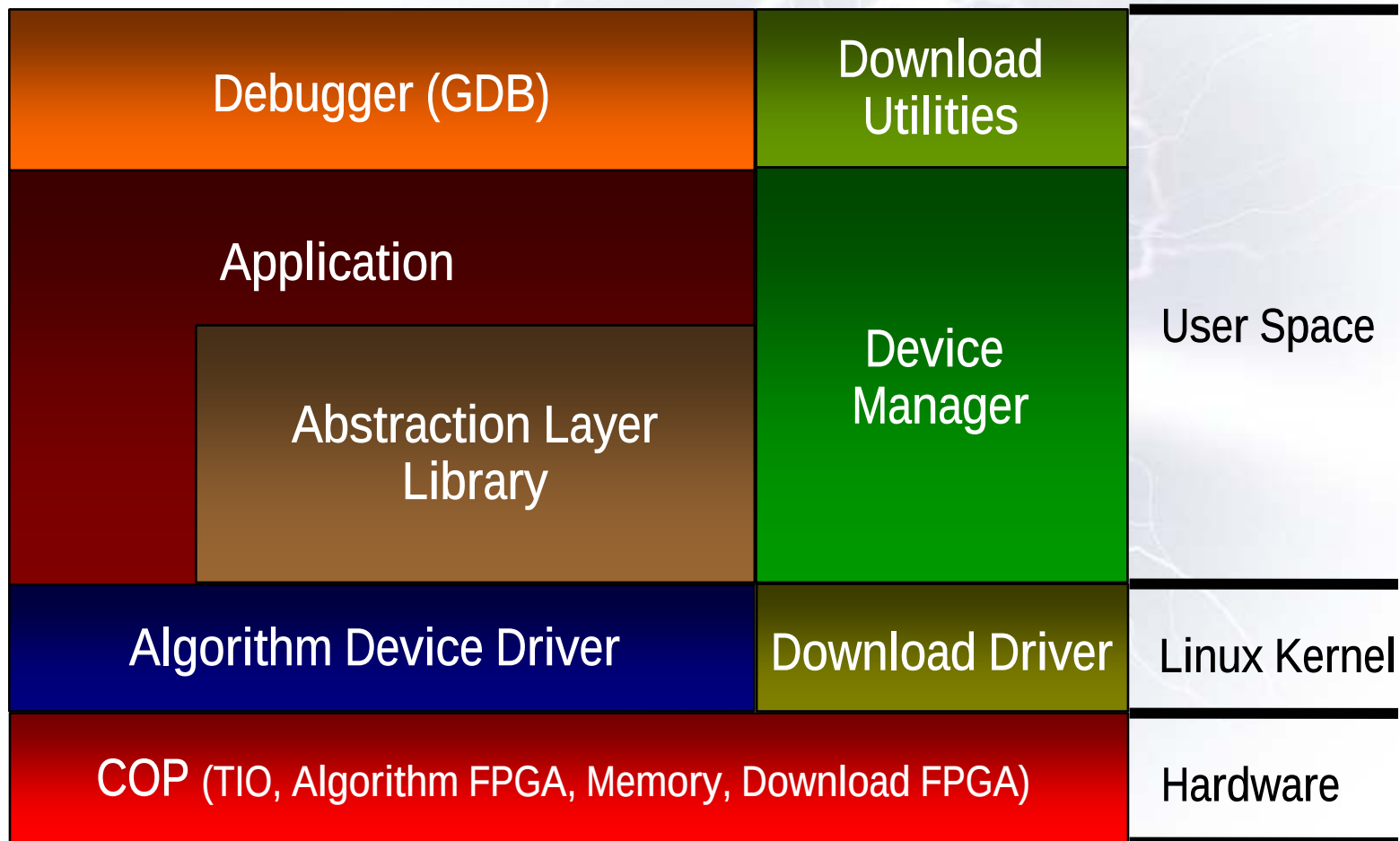
Terascale to Petascale Data Set : Bring Function to Data



- Provide for HDL modules
 - Integrated environment with debugger
 - Highest performance
- Leverage 3rd Party Std Language Tools
 - Celoxia, Impulse Acceleration, Mitrion, Mentor Graphics
- Developed an FPGA aware version of GDB
 - Capable of debugging the FPGA and System Software
 - Capable of multiple CPUs and multiple FPGAs
- Developed RASC Abstraction Layer (RASCAL)



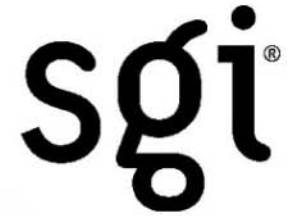
Software Overview



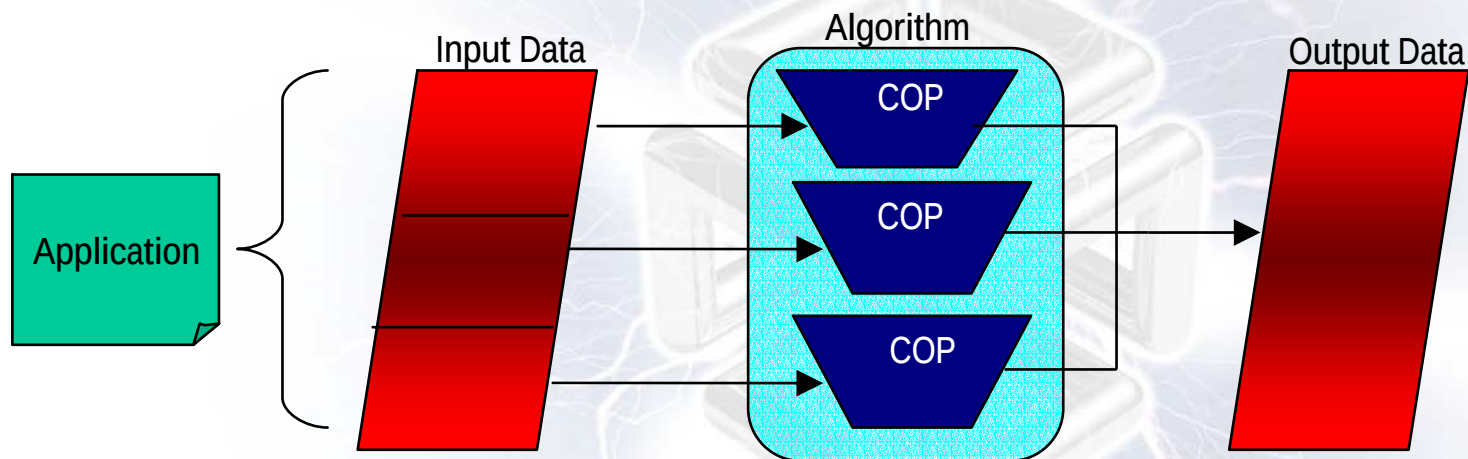
Igniting Innovation
and Leadership



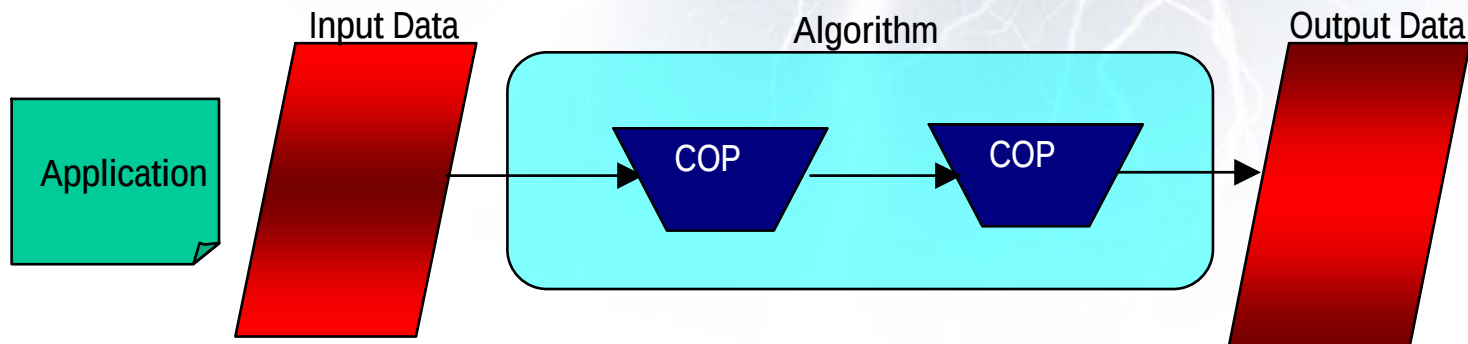
Abstraction Layer: Algorithm API



The Abstraction Layer's algorithm API mirrors the COP API with a few additions that enable wide scaling,



and deep scaling.



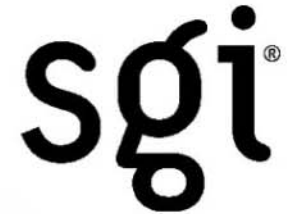
Igniting Innovation
and Leadership



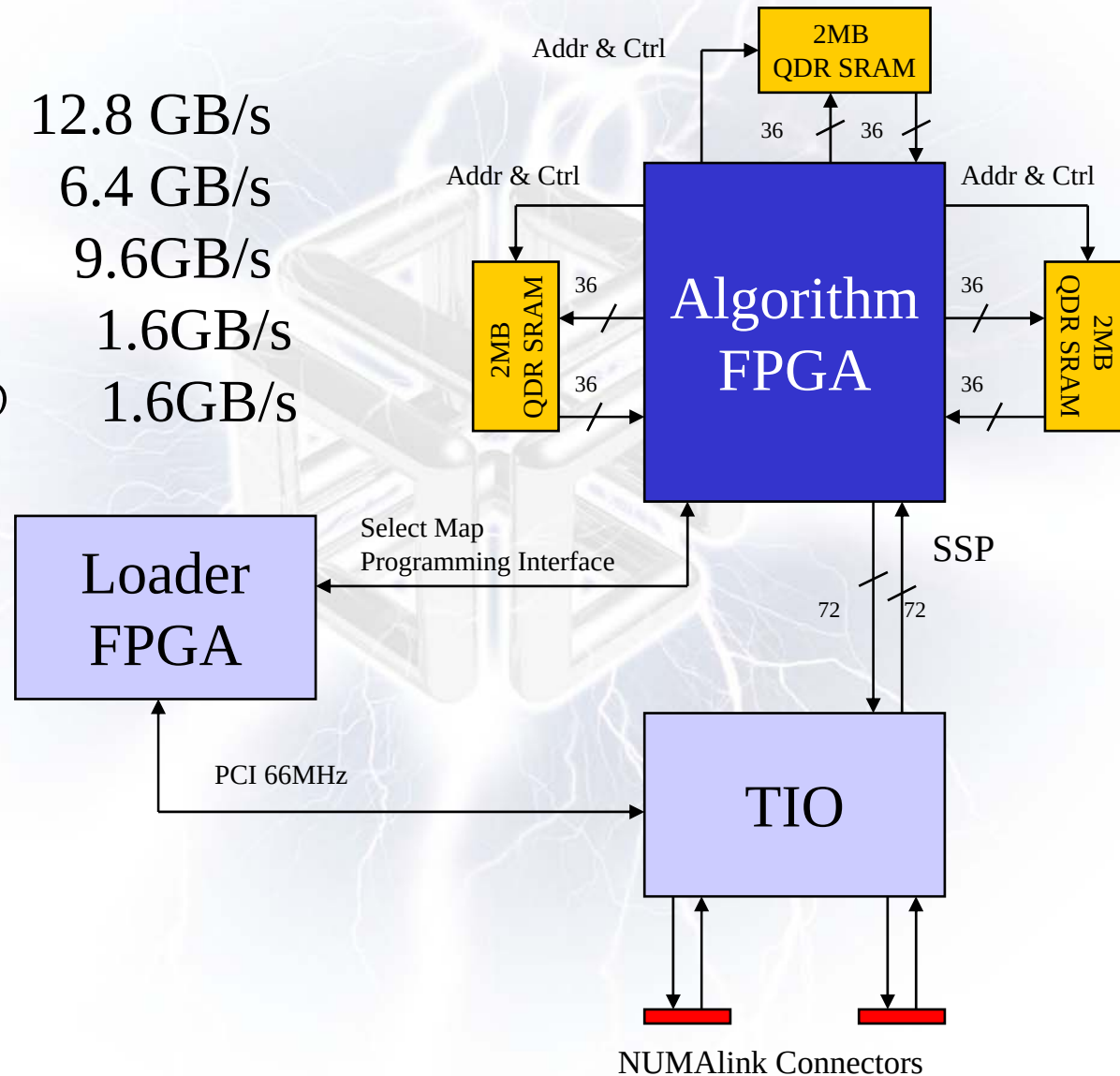
- Direct Connection to NUMALink4
6.4GB/s/connection
- Fast System Level Reprogramming of FPGA
- Atomic Memory Operations
Same set as System CPUs
- Hardware Barriers
- Configurations to 8191 NUMA/FPGA connections



MOATB Block Diagram



NUMAlink 12.8 GB/s
SSP 6.4 GB/s
QDR SRAM 9.6GB/s
3 reads @ 1.6GB/s
3 writes @ 1.6GB/s



Igniting Innovation
and Leadership



Requirements for Scalable Application Specific Processing in Commercial HPEC

Presented by:
Steve Miller
Chief Engineer

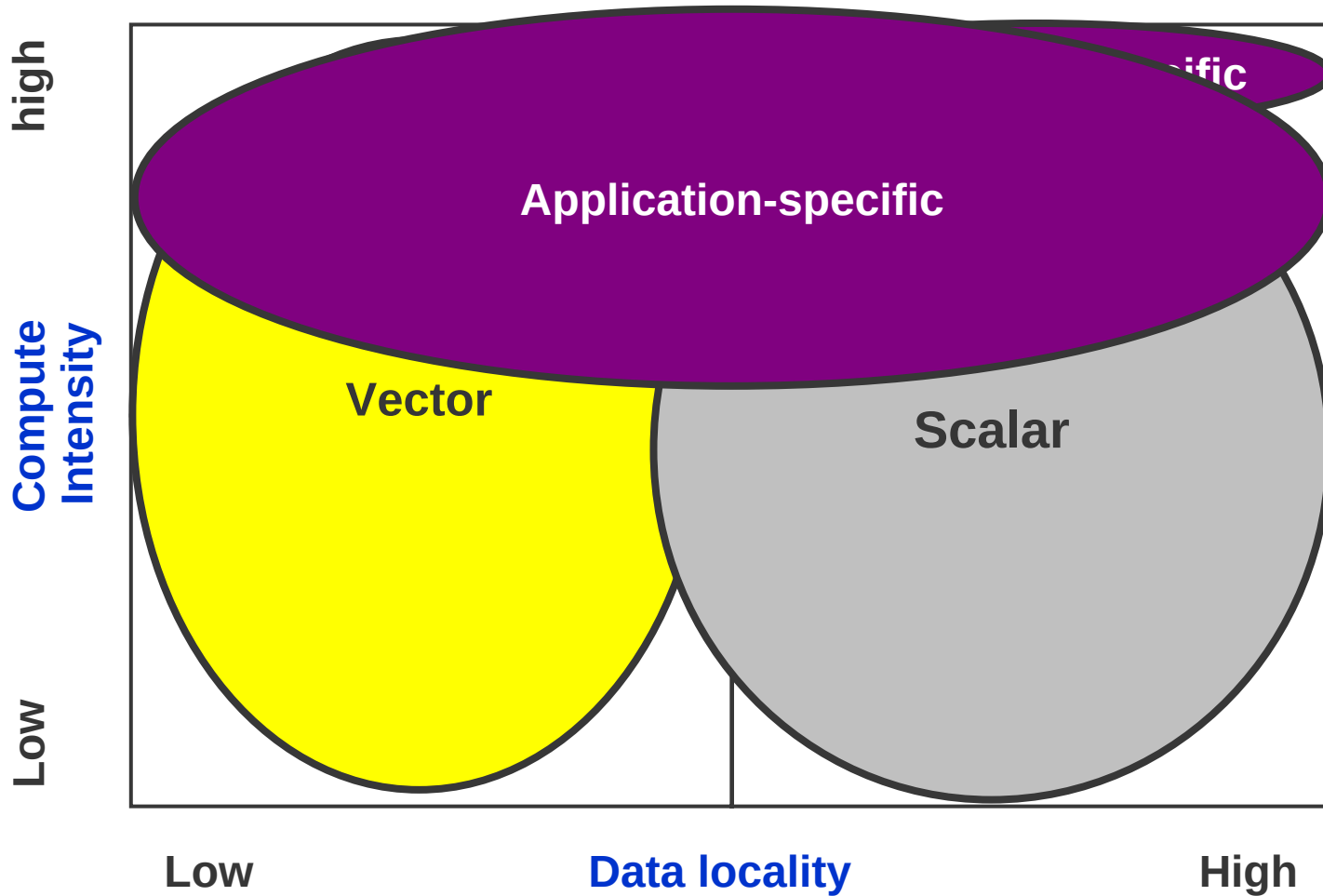
Silicon Graphics, Inc.



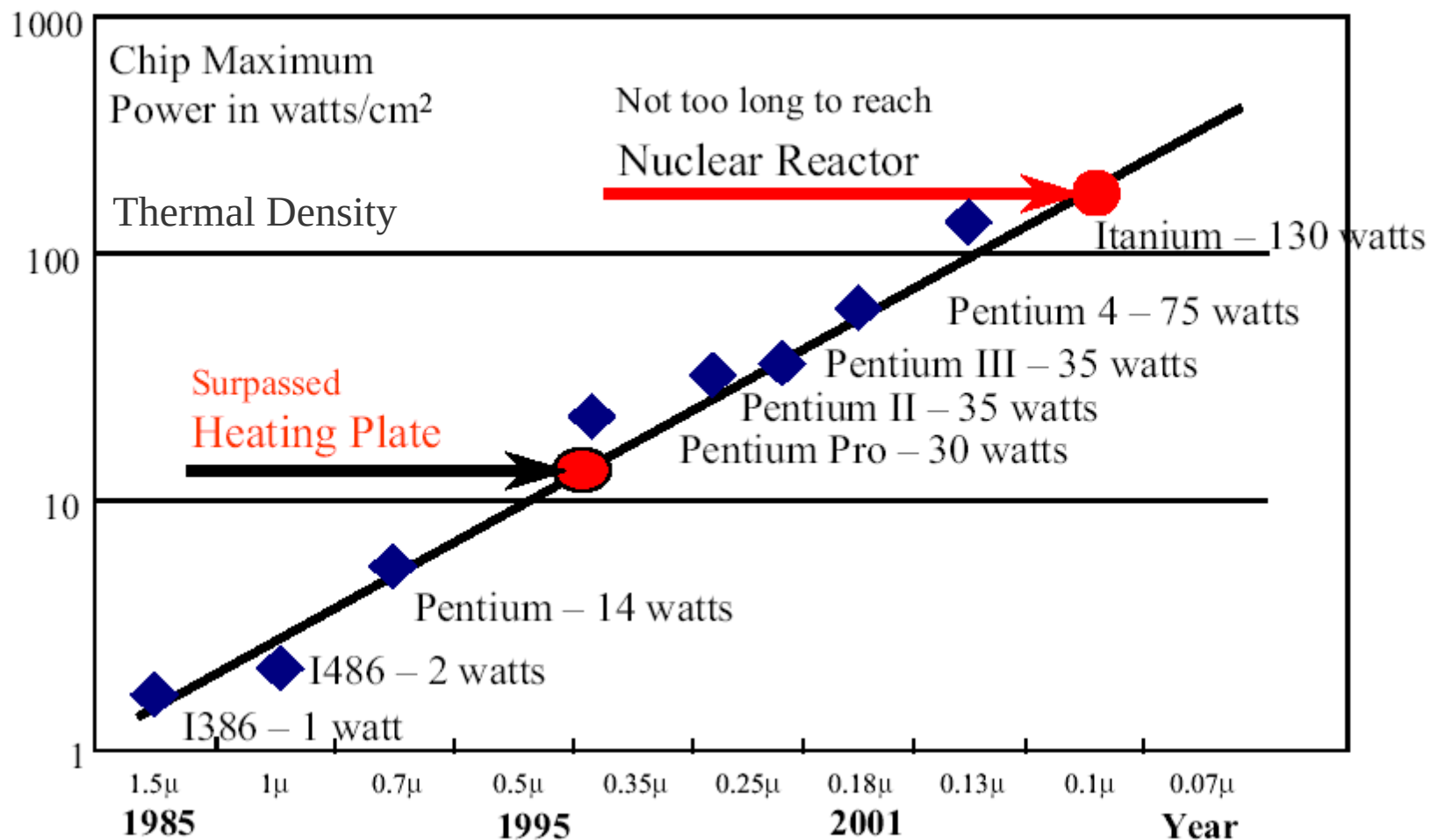
The 3 Single-Paradigm Architectures

<u>Scalar</u>	<u>Vector</u>	<u>App-Specific</u>
Intel Itanium	Cray X1	Graphics - GPU
SGI MIPS	NEC SX	Signals - DSP
IBM Power		Prog'ble - FPGA
Sun SPARC		Other ASICs
HP PA		

Paradigms to Applications



Microprocessors & Heat



Architectural Challenges

- **Ease of Use**
 - Languages
 - Compilers
 - Debuggers
 - APIs
- **Performance**
 - Bandwidth to/from System
 - Scalability