SGI’s Approach to Multi-paradigm Computing

Ron Renwick
Product Manager
Reconfigurable Computing
SGI Scalable ccNUMA Architecture
Basic Node Structure and Interconnect

![Diagram of the basic node structure and interconnect in a ccNUMA architecture. The diagram shows two CPU/CACHE pairs connected by a NUMAlink Interconnect, with physical memory interfaces.]
SGI Scalable ccNUMA Architecture
Basic Node Structure and Interconnect

- Global Shared Memory
- CPU
- Interface Chip
- NUMAlink Interconnect
- Cache
SGI Scalable ccNUMA Architecture
Scaling to Large Node Counts
SGI Scalable ccNUMA Architecture
Multi-Paradigm Computing Architecture

SGI Scalable ccNUMA Architecture

RASC (FPGA)

Scalable GPUs

General Purpose I/O Interfaces

NUMAlink Interconnect Fabric
SGI® RASC™ Technology

• RASC is *Reconfigurable* Application Specific Computing
• First instantiation of RASC is FPGA’s *Field Programmable Gate Arrays*
Why FPGAs...?

In a word...**performance**

- Some apps...orders of magnitude perf. increases

- Low power

- Broader market use ensures that technology will progress from 90nm to 65nm and beyond...
HPC Use of FPGAs...Challenges

• Performance
  – Bandwidth to/from system
  – Scalability

• Ease of Use
  – Languages
  – Compilers
  – Debuggers
  – APIs
Addressing the Performance Challenge
Integration into NUMAlink

- Direct connect to Numalink (6.4 GB/sec)
- Fast System Level Reprogramming of FPGA
- Atomic Memory Operations
- Hardware Barriers—load balancing
- Configurations to 128 NUMA/FPGA Nodes
MOATB- Early Access System Configuration

Itanium2

PC2700 DDR SDRAM

SHUB

NUMAlink

Itanium2

PIC

PCI-X

BASE I/O

Altix 350

Algorithm FPGA

TIO

2MB QDR SRAM

Addr & Ctrl

PCI 66MHz

Loader FPGA

36

Select Map Programming Interface

SRAM 0

SRAM 1

MOATB

SRAM 2
Current Customers

- Classified HLS customer (2)
- U of FL
- NCSA (2 units)
- Classified DoD customer
- GWU
- PNNL (2 units)
- Exegy
- Starbridge Systems
MOATB Sample Application Speedup

Bit Manipulation (Crypto)
79x 1.5GHz Itanium-2 (single MOATB)
119x 1.5GHz Itanium-2 (dual MOATB)

Graphics Edge Detection
42x 1.5GHz Itanium-2 (single MOATB)…
demo’d at NAB

Customer Application
10**5x speedup on scalar microprocessor
Current Product – SGI® RASC™ Technology

NUMAlink Connectors

TIO

SSP

Algorithm FPGA

2 – 8 MB QDR SRAM

2 – 8 MB QDR SRAM

2 – 8 MB QDR SRAM

PCI 66MHz

Loader FPGA

Algorithm FPGA  Virtex2 6000 -6
Development Environment
RASC Software Stack

- Debugger (GDB)
- SpeedShop™
- Download Utilities

- Application
- Abstraction Layer Library

- Algorithm Device Driver
- Download Driver

- COP (TIO, Algorithm FPGA, Memory, Download FPGA)

- User Space
- Linux® Kernel
- Hardware
FPGA Aware Debugger (GDB-FPGA)

- Based on Open Source Gnu Debugger (GDB)
- Uses extensions to current command set
- Can debug host application and FPGA
- Provides notification when FPGA starts or stops
- Supplies information on FPGA characteristics
- Can “single-step” or “run N steps” of the algorithm
- Can HLL line step per C-line source
- Dumps data regarding the set of “registers” that are visible when the FPGA is active
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**

Working with industry/customers ([openfpga.org](http://openfpga.org)) on API stds…
Verilog / VHDL Module Support

• Templates for Verilog and VHDL
  – Fast start to algorithm coding
• Provide a system simulation stub
  – Allows both simulation debug or system debug
• Provide source code for core service
  – Allows user to modify to meet special needs
• Extractor tools supports GDB meta-data
  – Application and FPGA debugging
Multi-Paradigm Computing
Other Non-traditional Processing Initiatives

• GPU-based processing
  – High potential performance (200-300GF peak today) and performance/price on single precision floating point applications...clear roadmap to future semiconductor process technologies
  – SGI working with SI on scaling to multiple GPUs and on development environment/programming paradigms...initial focus on signal processing apps

• Specialized processors...Clearspeed processors, custom processors (MD-GRAPE, classified chip)
  – High potential performance/watt on certain apps
  – Not clear that market exists which would fund semiconductor process migration roadmap...?
Multi-Paradigm Computing
Terascale to Petascale Data Set...Bring Function to Data

Scalable Global Shared Memory Structure

- Globally addressable, universally accessible, high-bandwidth flat address space
- Addressing the transition from Terascale to Petascale computing
- Tightly coupling computation, I/O, and visualization functions to memory

This slide contains forward-looking statements. The results and forecasts as stated may vary. Other risks and uncertainties relating to this slide may be