Octopus: A Multi-core implementation

Kalpesh Sheth
HPEC 2007, MIT, Lincoln Lab

Parameters to evaluate?

- **Many vendors have multi-core, multi-chip boards**
- **Characteristics of good evaluation**
  - How much memory BW among multiple cores? (i.e. core to core)
  - How much I/O BW between multiple sockets/chips on same board? (i.e. chip to chip on same board)
  - How much fabric BW across boards? (i.e. board to board)
- **CPU performance with I/O combined**
  - Data has to come from somewhere and go into memory

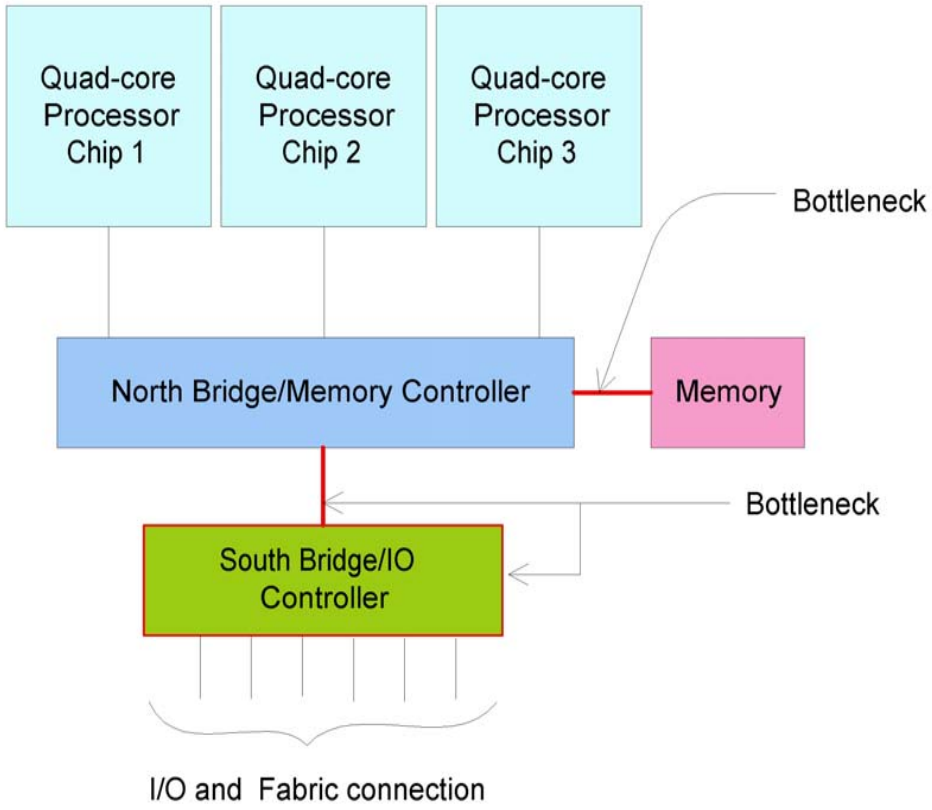
Multi-core Challenge

- **Current Benchmarks**
  - Most don't involve any I/O
  - Many are cache centric
  - Many are single core centric (no multi-threading)
- **Questions to ask?**
  - Interrupt handling among multiple cores
  - Inter-process communication
  - How many channels for DDR2 interface? 4 or better?
  - Size, Weight and Power (SWaP) when fully loaded?
  - How to debug multi-threaded programs? Cost of tools?
  - What is the cost of ccNUMA or NUMA? Memory R/W latency?

Traditional Intel Platform

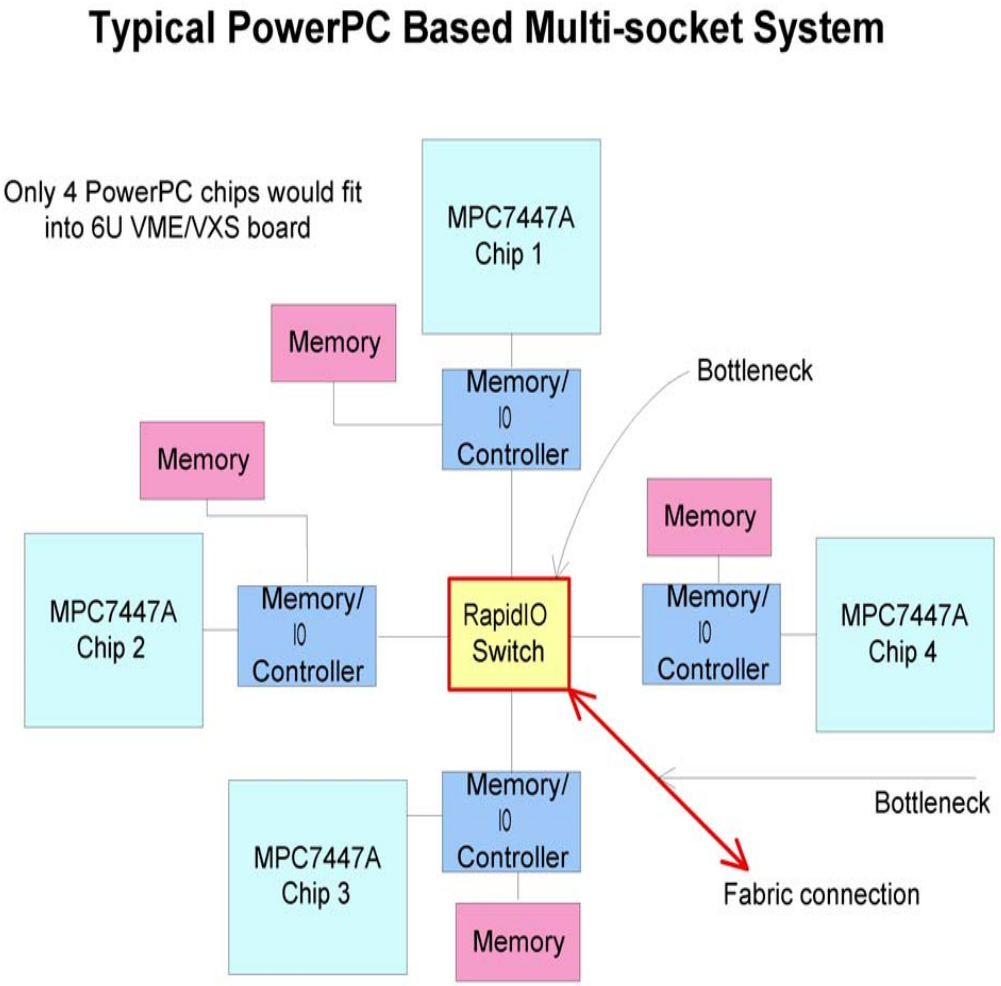
- **Single entry point for memory access**
- **All external I/O via Southbridge**
  - GigE, UART competes with Fabric
  - Always requires CPU cycles

Typical Intel Based Multi-socket System



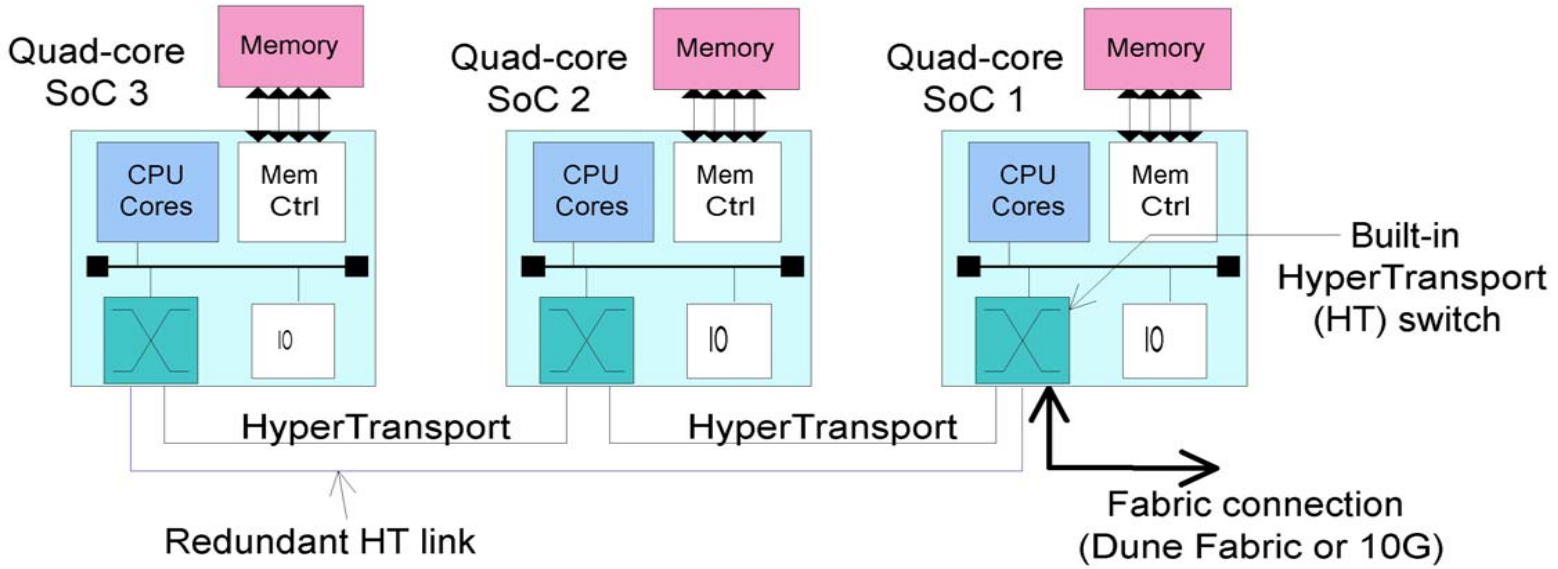
Traditional PowerPC Platform

- **Local memory access**
- **Mostly no I/O**
- **RapidIO (parallel or serial) switching**
- **Limited BW between chips**
- **Fabric bottleneck as all data comes via fabric only**



DRS Approach

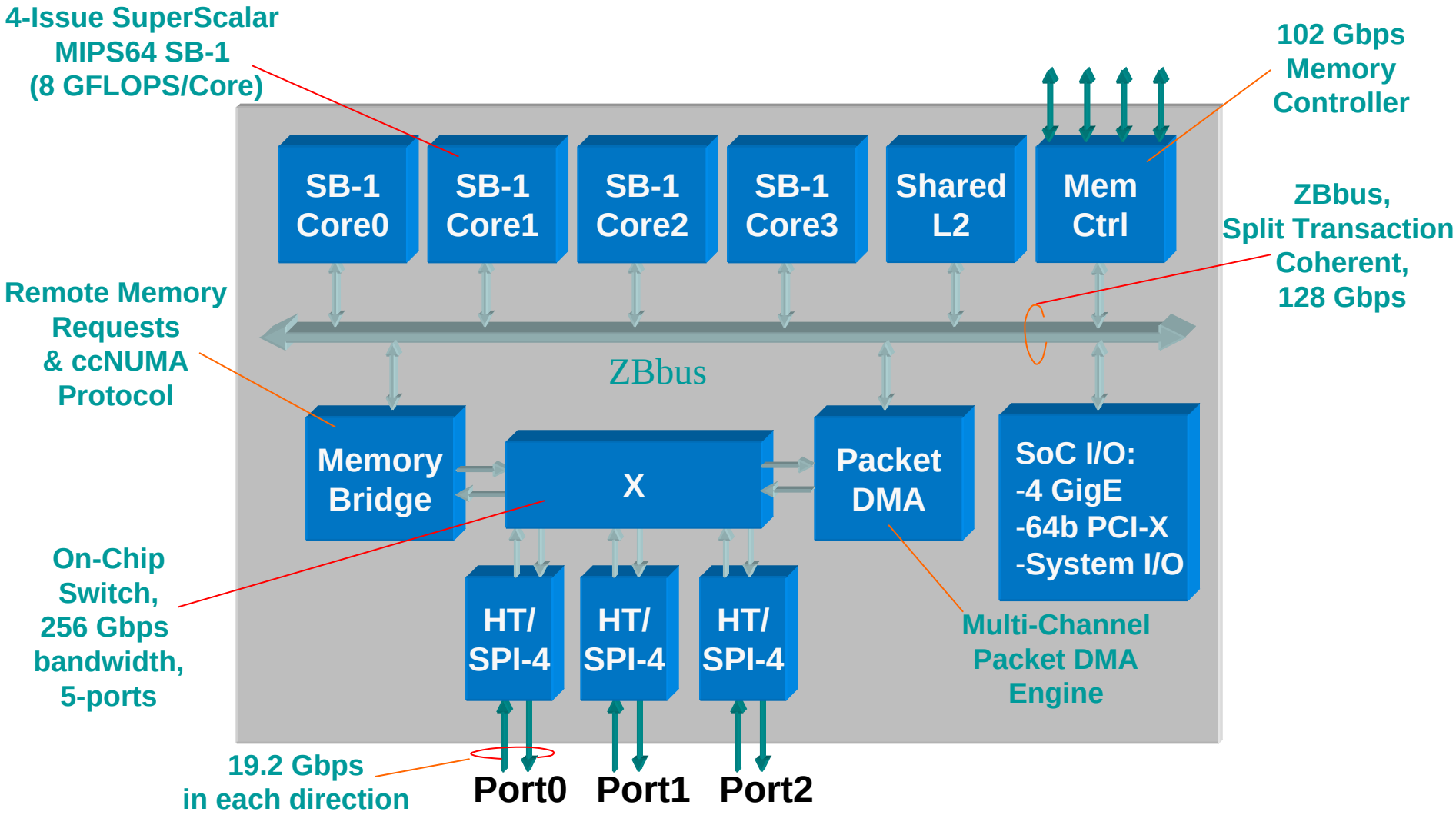
DRS's MIPS based Multi-socket System



Zero CPU & memory cycles switching from any-to-any SoC

- **Each System on Chip (SoC) has HyperTransport switch**
- **Local memory access with NUMA and ccNUMA capability**
- **Data can come locally or via fabric**

BCM1480 Architecture Overview



Performance measurement

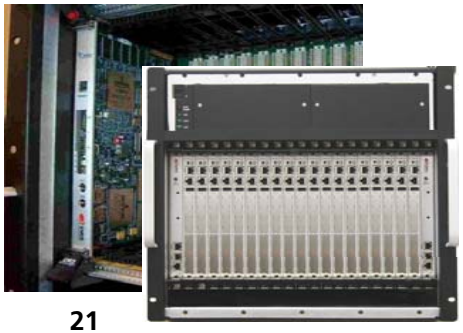
- **How do you measure performance of multi-core embedded system?**
  - Perform network I/O while doing number crunching
    - Examples uBench, netPerf with FFTw
  - Measure memory BW with streams benchmark
  - Measure intra-core and inter SoC BW performance
    - Examples openMPI (measures latency and BW)
  - How open standards are supported?
    - Examples CORBA, VSIPL, FFTw, RDMA, MPI etc.
  - Measure switch fabric BW and latency
  - XMC and PMC plug ability and their interface speed
  - Boot time (in seconds)

VSIPL Benchmarks

VSIPL Comparisons FFT Routines; Large Vectors Times in microseconds vsip_ccfftop_f:		1480: 1GHz MIPS64 G4: 1GHz 7447	
N	**32K**	**64K**	**128K**
1480	1120.00	2310.00	7640.00
G4	1660.00	11654.00	43197.00
Performance Improvement	1.48	5.05	5.65

Octopus as the Open Source/Standard and COTS commodity Solution

- **Leverages open source development & run time environments**
  - Including Linux OS (SMP), Eclipse IDE, GNU tools
  - Promotes ease of portability and optimal resource utilization
- **Delivers open standard middleware & signal processing libraries**
  - Including VSIPL, MPI, FFTw and CORBA
- **Effectively decouples the application software from the hardware**
  - Applications interface to the OS at a high level (layer 4)
- **Utilizes standards based hardware**
  - VITA 41 (VXS) backplane – backwards compatible with VME
  - VITA 42 (XMC) mezzanine standards - permits rapid insertion of new technology
- **Implemented using commodity chips taken from large adjacent markets**
  - Broadcom dual/quad core processors from Telecom/Datacom Network Processing
  - HyperTransport from commodity computing
  - Dune fabric from the Tera-bit router market



Portability & Inter-operability → Reduced Life Cycle Cost

Evaluation

- **Independent evaluation done (Summer 2007)**

- **Summary of Results (based on 8-slot chassis)**
  - 166 GFLOPS sustained throughput
  - 4.6 FLOPS/byte (2G DDR2 per BCM1480 SoC)
  - 136 MFLOPS/W computation efficiency
  - 4.9 GFLOPS/Liter computation density (not counting I/O)

- **Contact DRS for more details**

Thank You

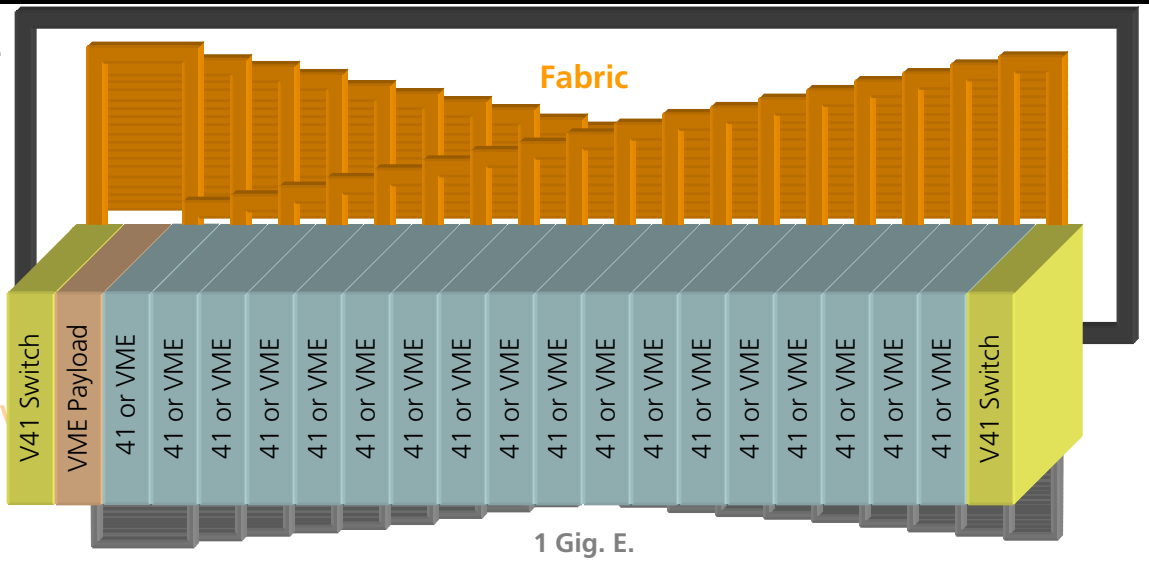
For more information:

Advanced Processing Group
21, Continental Blvd, Merrimack, NH 03054
Phone: 603-424-3750 x326
Octopus_support@drs-ss.com
http://www.drs-ss.com/capabilities/ss/processors.php

Backup Slides

VITA 41 (OCTOPUS) System Overview

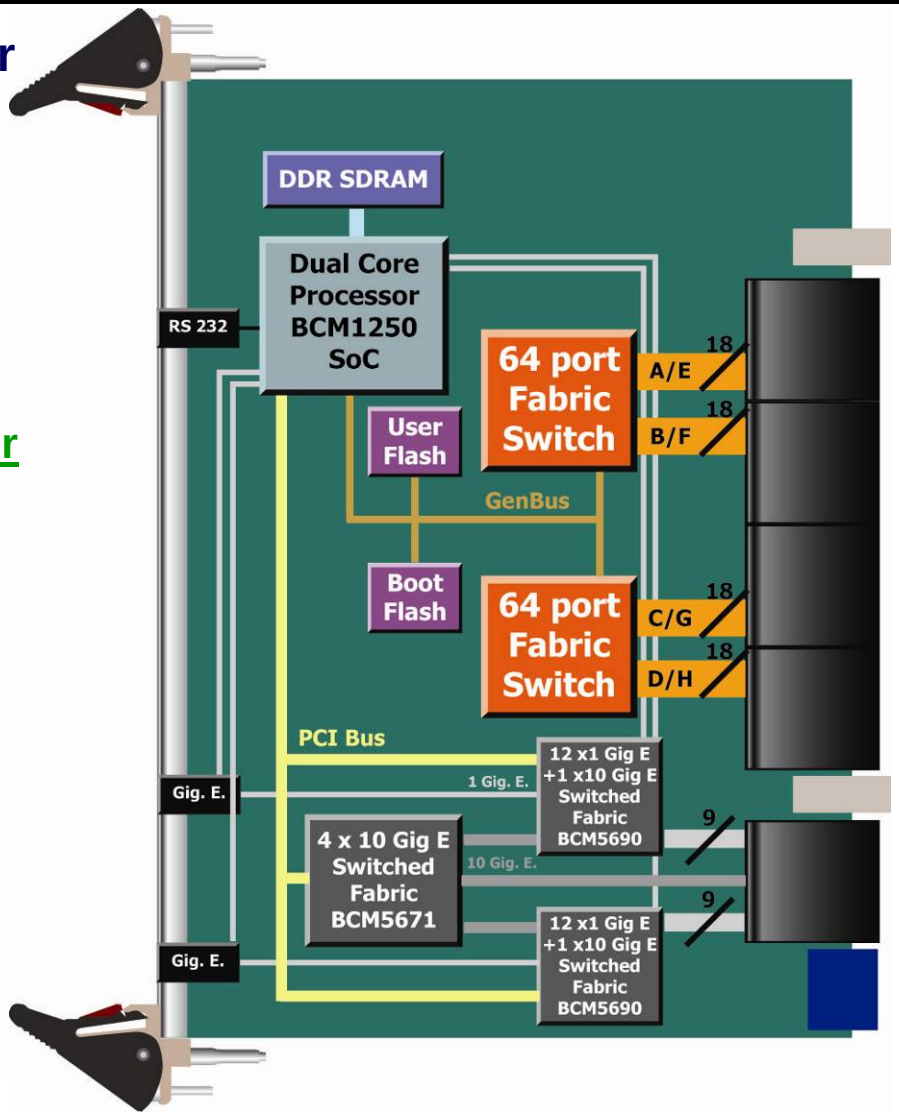
- **High Performance Embedded Multi-computing system backwards compatible with VME**
- **High Speed scalable advanced switch fabric with telecoms grade reliability**
- **Separate and redundant out-of-band (VITA 41.6) GigE control plane to every processing element**
- **Linux OS (SMP or ccNUMA)**



- **Octopus boards**
  - Shown in standard VITA 41 chassis'
  - Sourced from commodity chassis vendor
  - Alternative backplane configurations exist and are supported as defined in VITA 41

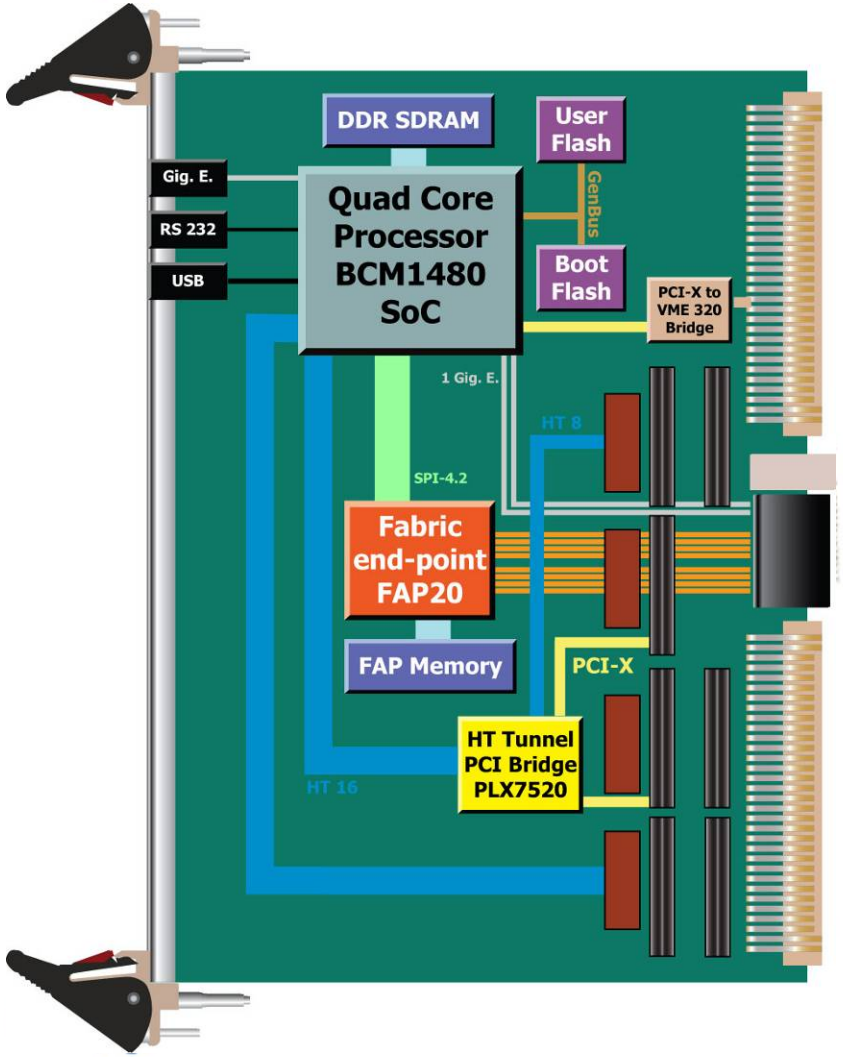
OCTOPUS Switch Card

- **"Smart Switch" acts as Host controller**
  - Controls boot sequence
  - Provides services to system
- **Data Plane Connectivity**
  - Measured net 1.1 GBytes/s full duplex between any payload boards after fabric overhead and encoding/decoding
  - Provides dynamic connectivity (no more static route tables) and single layer of switching between payloads
- **Control Plane GigE Connectivity**
  - 1 GigE to every payload board and front panel
  - 10 GigE between switch cards
- **Dual Core Processor on board**
  - Each core running Linux at 600 MHz
  - 256 MB DDR SDRAM at 125 MHz DDR memory bandwidth
  - User (64 MB) and boot (4 MB) flash
- **Temperature Sensors**
  - 1250 processor and each fabric switch



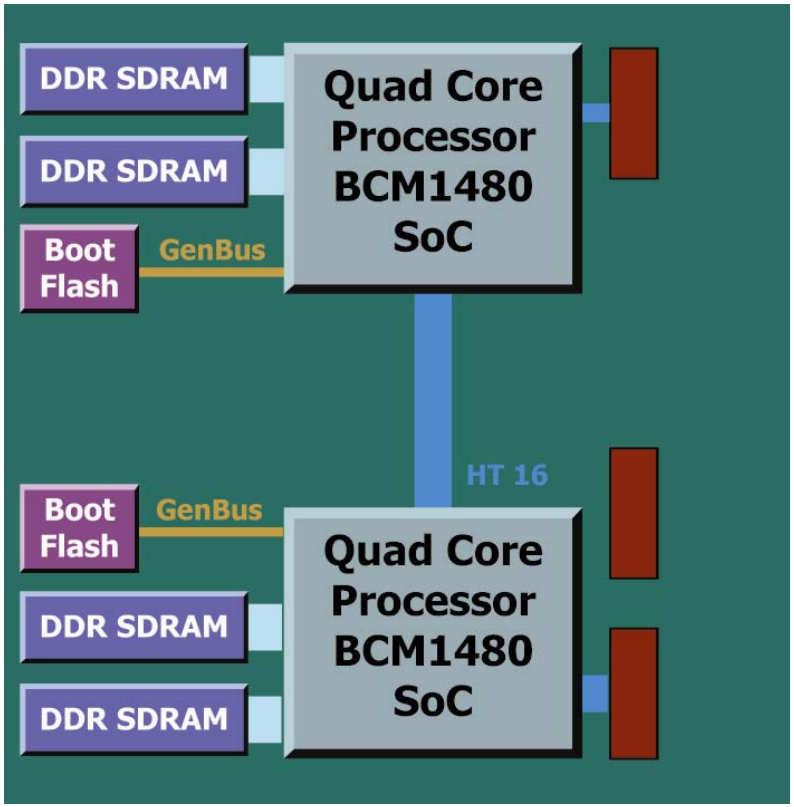
OCTOPUS Motherboard

- **High performance Quad Core processor**
  - MIPS64 SoC (System on a Chip) processor
  - 8 GFLOPS/processor (**32 GFLOPS/Chip**) with **4.8 GByte/s** composite I/O
  - Running SMP Linux at 1GHz on each core (ccNUMA capable)
- **2 GB of DDR2 SDRAM**
  - @200MHz
- **Flash**
  - User (128MB) and boot (4MB)
- **Front panel GigE and USB**
- **Power ~ 45W (Max)**
- **Temperature Sensors**
  - 1480 processor and fabric end-point



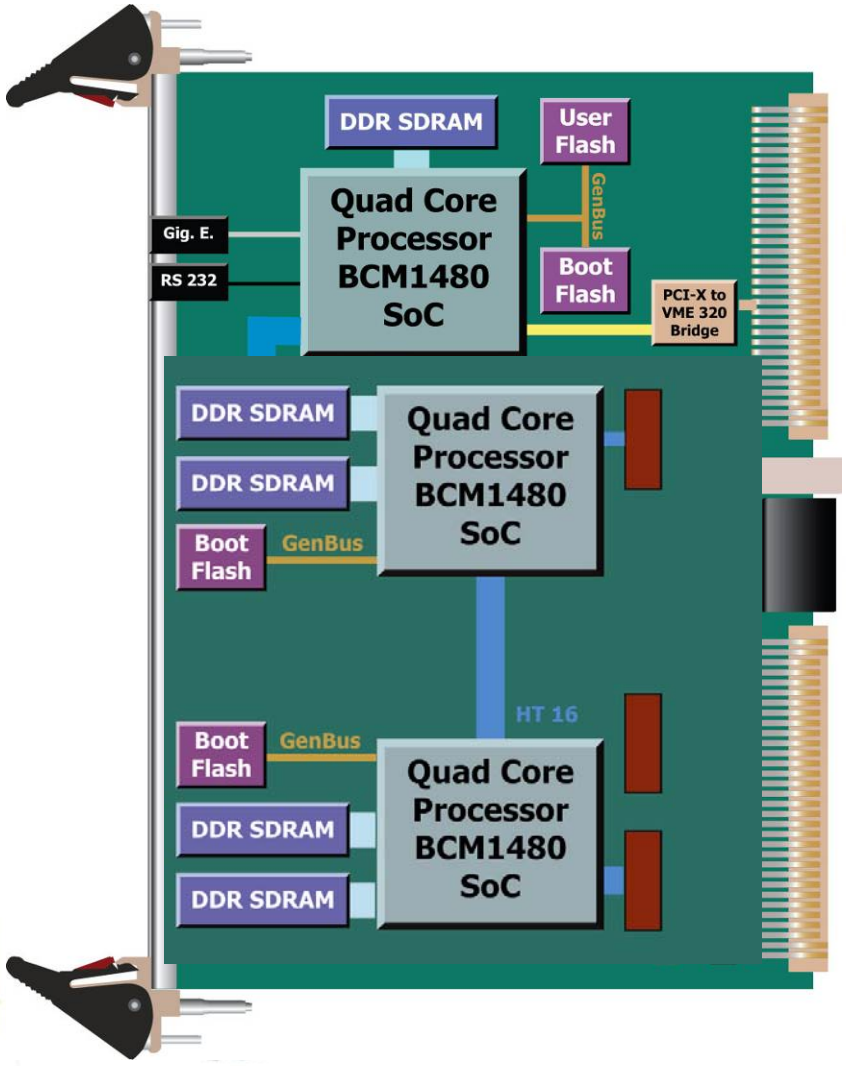
OCTOPUS Dual XMC

- **VITA 42.4 compliant**
- **2 x High performance Quad Core processors**
  - MIPS64 SoC processors
  - Running SMP Linux at 1GHz on each core (ccNUMA capable)
- **With each Quad Core processor:**
  - **2 GB** of DDR2 SDRAM (3.2 GB/s)
  - 4 MB of flash
- **Total of 8 x 1 GHz cores, 4 GB DDR2 SDRAM and 8 MB boot flash!**
- **Power ~ 65 W**
- **Temperature Sensors**
  - For each 1480 processor



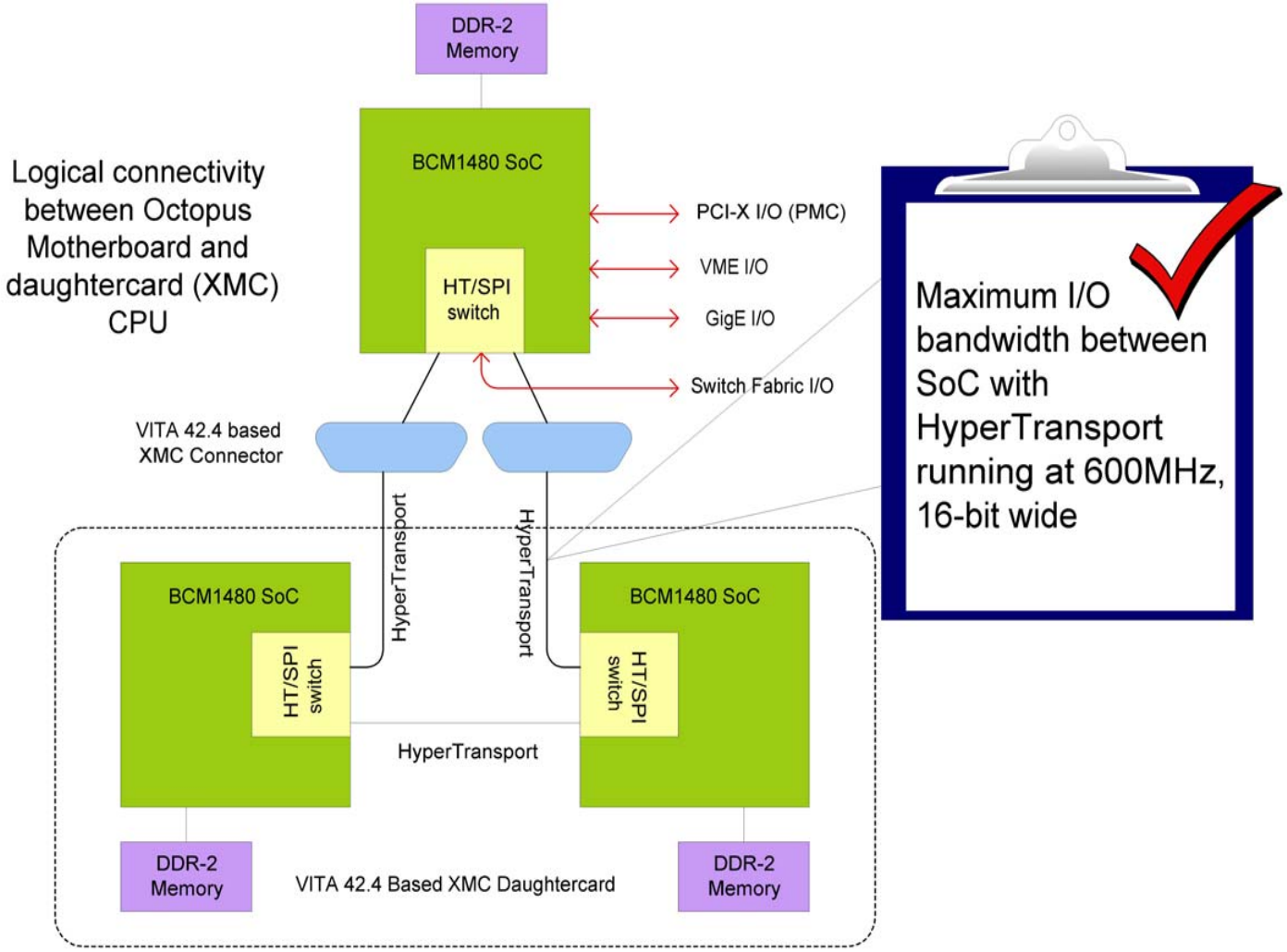
OCTOPUS Payload Slot

- **Fully loaded payload slot (Motherboard plus Dual XMC) provides:**
- **96 GFLOPS = 3 x High performance MIPS64 Quad BCM1480**
- **6 GB of DDR2 SDRAM**
- **Flash**
  - User 128 MB
  - Boot 12 MB
- **Node aliasing and ccNUMA**
- **On board PCI-X and HyperTransport**
- **Off board VITA 41 serial fabric**
- **Power ~ 110W (Max)**

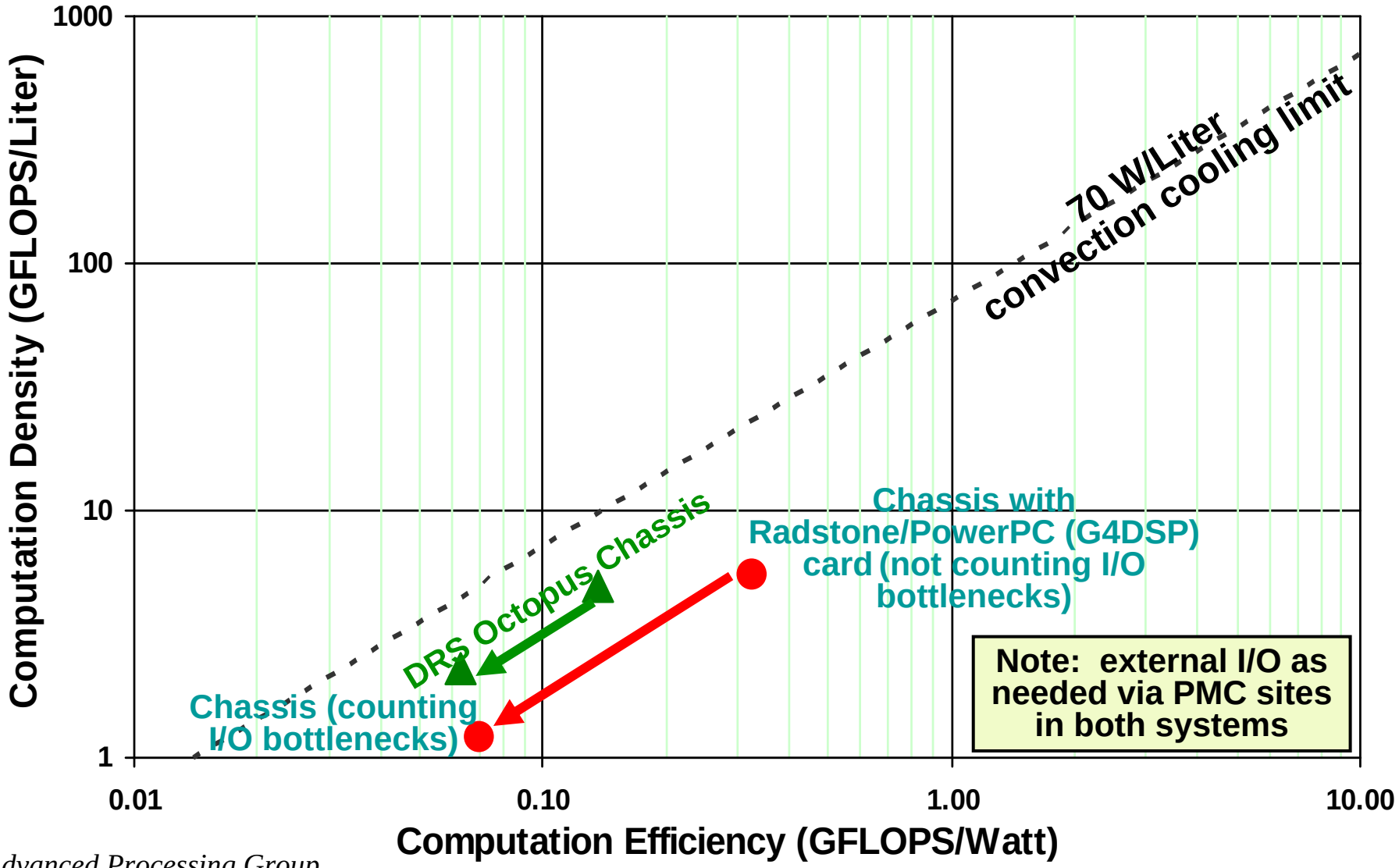


Inter-SoC Connectivity

Logical connectivity between Octopus Motherboard and daughtercard (XMC) CPU



Figures of Merit Comparison



Octopus SAR Benchmark Comparison

Data Presented at Processor Technology Symposium on 10/3/06

Multi-core Board	Volume per Board Cu. Inches	Gcc/Linux Multi-threading	Gcc/Linux Multi-threading Per cu. In.	Gcc/Linux Multi-processes	Gcc/Linux Multi-processes Per cu. In.	Vendor Compiler Multi-threading	Vendor Multi-threading Per cu. In.	Vendor Compiler Multi-processes	Vendor Multi-processes Per cu. In.
Intel Sossaman	756	26.2	2,817,157	19.45	3,794,834	28	2,636,054	19.8	3,727,754
Intel Dempsey	756	23.2	3,181,445	12.2	6,049,961	20.5	3,600,465	9.93	7,432,983
Intel Woodcrest	756	12.4	5,952,381	8.5	8,683,473	12.1	6,099,961	6.45	11,443,337
Intel Montecito	3671	26.8	567,172	10.75	1,413,974	17.9	849,174	6.7	2,268,689
DRS-IT MIPS64	186.6	28	11,196,601	13.44	22,249,655				
Fabric7 Opteron	2766	16.3	1,237,640	5.75	3,508,441				

Timing Results Normalized by Volume

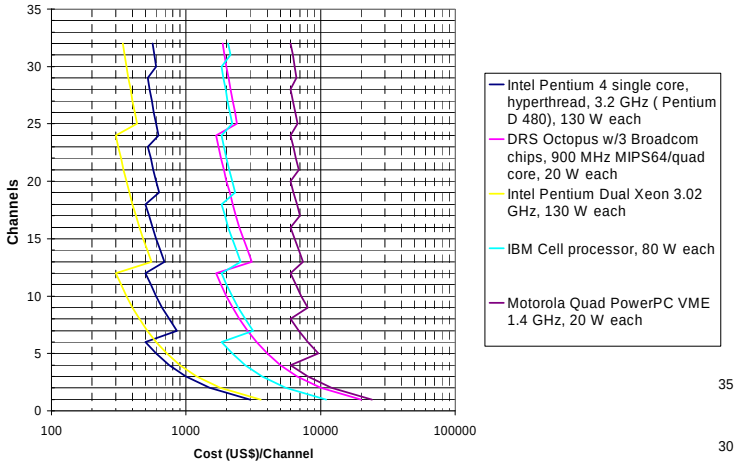
Best SWaP Performance

Cell Processor software effort in process

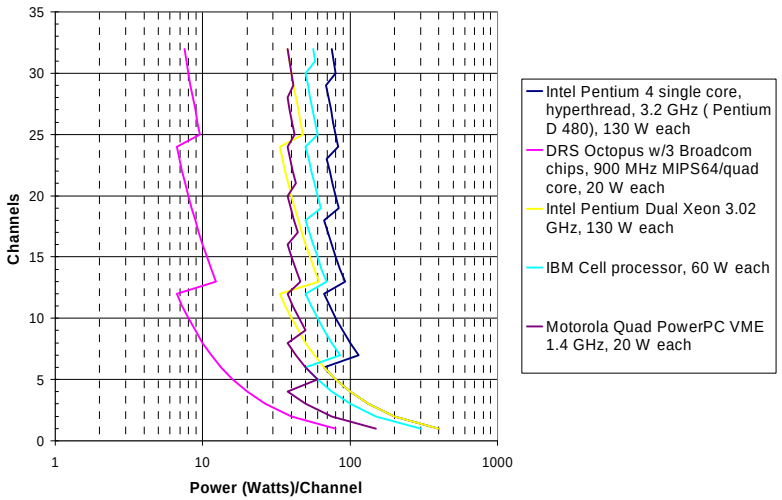
Signal Recognizer

- **Octopus Shows SWaP Advantage over Blade Servers**

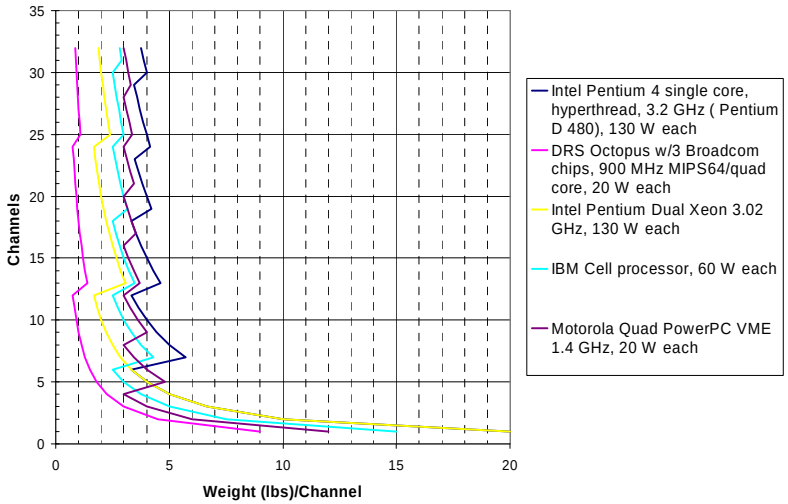
Recognizer Channel Performance for various processors



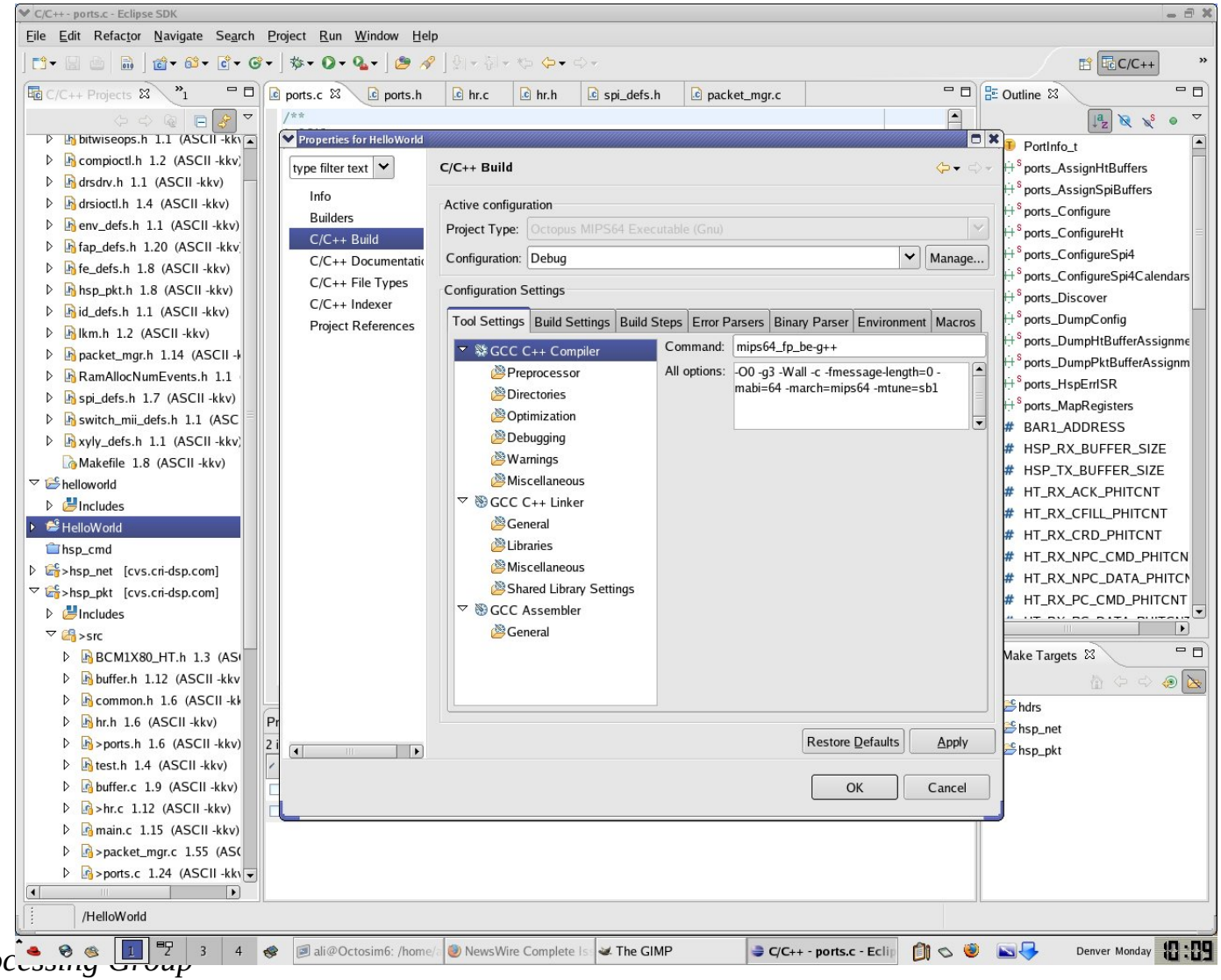
Recognizer Channel Performance for various processors



Recognizer Channel Performance for various processors



Development Environment



Advanced Debug Tools