Octopus: A Multi-core implementation



Kalpesh Sheth HPEC 2007, MIT, Lincoln Lab

Export of this products is subject to U.S. export controls. Licenses may be required. This material provides up-to-date general information on product performance and use. It is not contractual in nature, nor does it provide warranty of any kind. Information is subject to change at any time.



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?

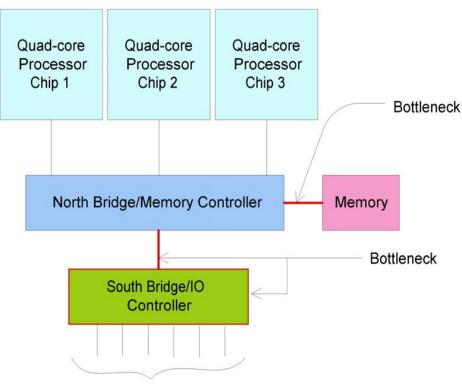
- Interrupt handling among multiple cores
- Inter-process communication
- How many channels for DDR2 interface? 4 or better?
- Size, Weight and Power (SWaP) when fully loaded?
- How to debug multi-threaded programs? Cost of tools?
- What is the cost of ccNUMA or NUMA? Memory R/W latency?



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

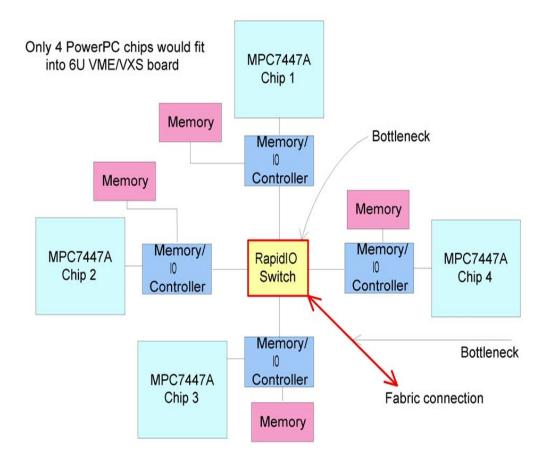


I/O and Fabric connection



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

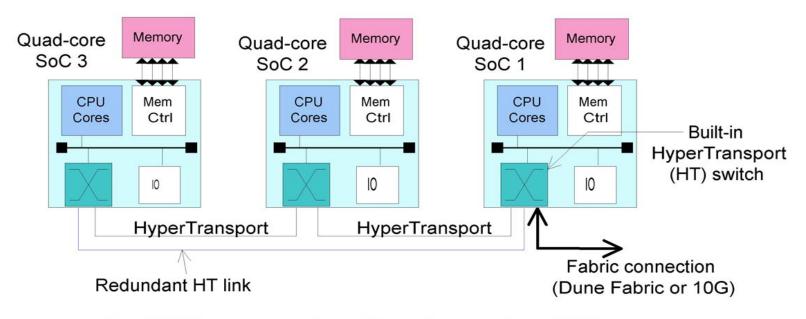


Typical PowerPC Based Multi-socket System



DRS Approach

DRS's MIPS based Multi-socket System

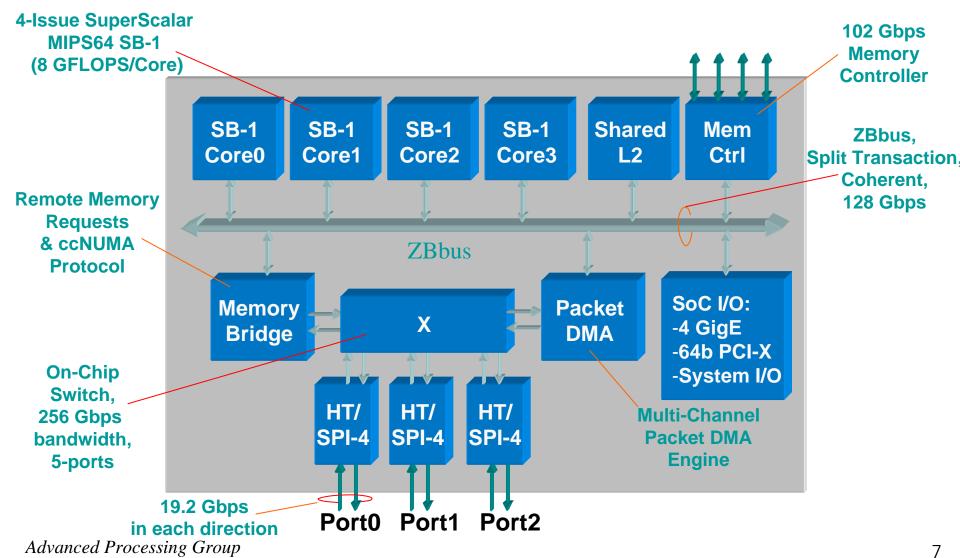


Zero CPU & memory cycles switching from any-to-any SoC

- 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 | | | | | | |
|--|---------|------------------------------------|----------|--|--|--|--|--|
| Ν | 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 <u>open source</u> development & run time environments
 - Including <u>Linux</u> OS (SMP), <u>Eclipse IDE</u>, <u>GNU</u> tools
 - Promotes <u>ease of portability</u> and <u>optimal</u> resource utilization
- Delivers <u>open standard</u> middleware & signal processing libraries
 - Including <u>VSIPL</u>, <u>MPI</u>, <u>FFTw</u> and <u>CORBA</u>
- Effectively <u>decouples</u> the application <u>software</u> from the <u>hardware</u>
 - Applications interface to the OS at a high level (layer 4)

- Utilizes standards based hardware
 - <u>VITA 41 (VXS)</u> backplane backwards compatible with VME
 - <u>VITA 42 (XMC)</u> mezzanine standards permits rapid insertion of new technology
- Implemented using <u>commodity</u> chips taken from <u>large adjacent markets</u>
 - Broadcom <u>dual/quad core processors</u> from Telecom/Datacom Network Processing
 - <u>HyperTransport</u> from commodity computing
 - Dune <u>fabric</u> from the <u>Tera-bit router</u> 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



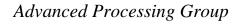


VITA 41 (OCTOPUS) System Overview

- High Performance Embedded Multi-computing system backwards compatible with VME
- High Speed scalable advanced switch fabric with <u>telecoms</u> grade reliability
- Separate and redundant out-ofband (VITA 41.6) <u>GigE control</u> <u>plane</u> to every processing element
- <u>Linux OS</u> (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

- <u>Measured</u> net <u>1.1 GBytes/s full duplex</u> between any payload boards after fabric overhead and encoding/decoding
- Provides <u>dynamic connectivity</u> (no more static route tables) and <u>single layer</u> of switching between payloads

Control Plane GigE Connectivity

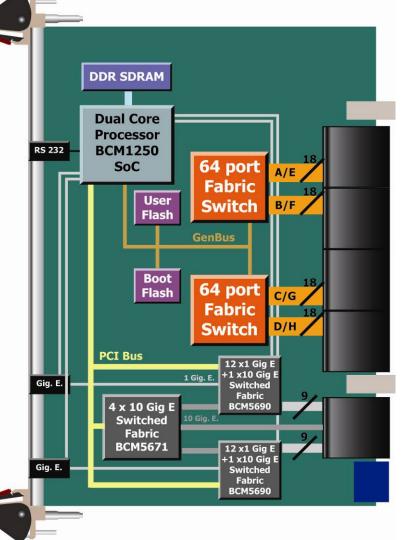
- <u>1 GigE</u> to every payload board and front panel
- <u>10 GigE</u> between switch cards

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

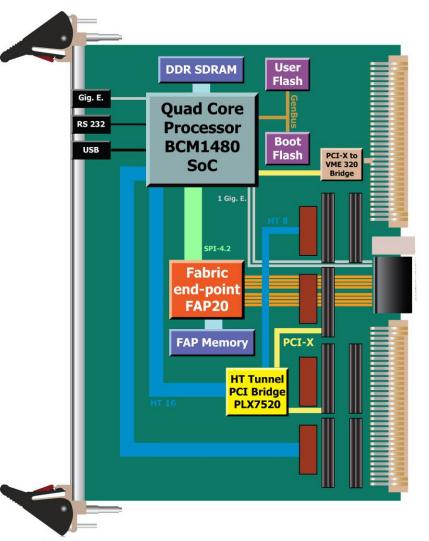
Temperature Sensors

1250 processor and each fabric switch



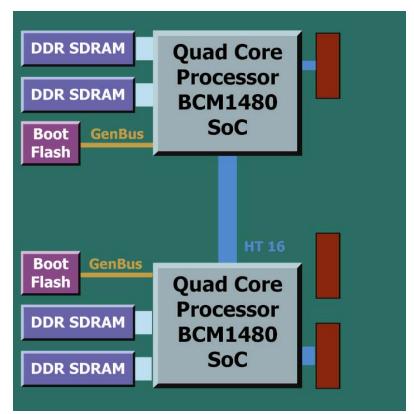
OCTOPUS Motherboard

- High performance <u>Quad Core</u> processor
 - MIPS64 SoC (System on a Chip) processor
 - 8 GFLOPS/processor (<u>32</u> <u>GFLOPS/Chip</u>) with <u>4.8 GByte/s</u> composite I/O
 - Running SMP Linux at 1GHz on each core (ccNUMA capable)
- <u>2 GB</u> 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



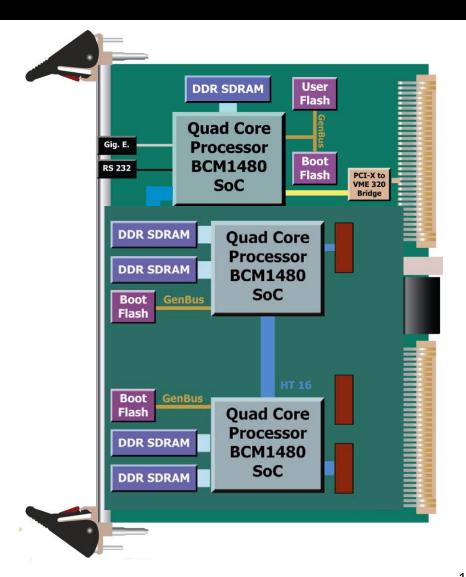
OCTOPUS Dual XMC

- 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:
 - <u>2 GB</u> of DDR2 SDRAM (3.2 GB/s)
 - 4 MB of flash
- Total of <u>8 x 1 GHz cores</u>, <u>4 GB</u> DDR2 SDRAM and <u>8 MB</u> boot flash!
- Power ~ 65 W
- Temperature Sensors
 - For each 1480 processor

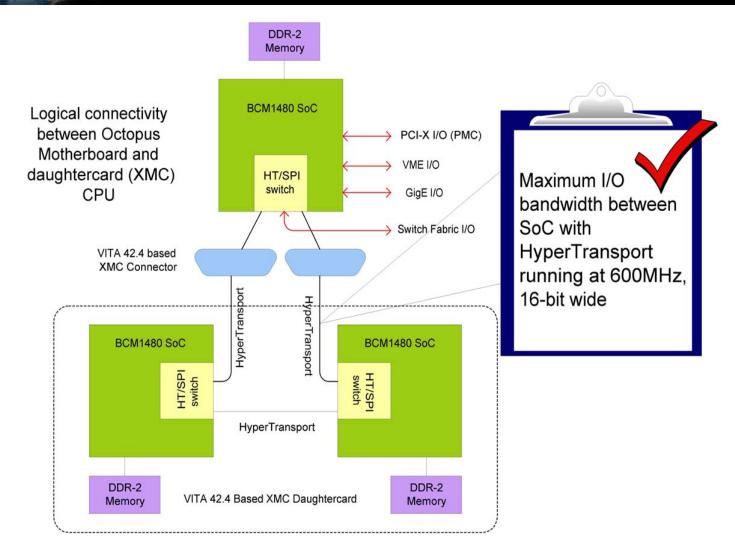


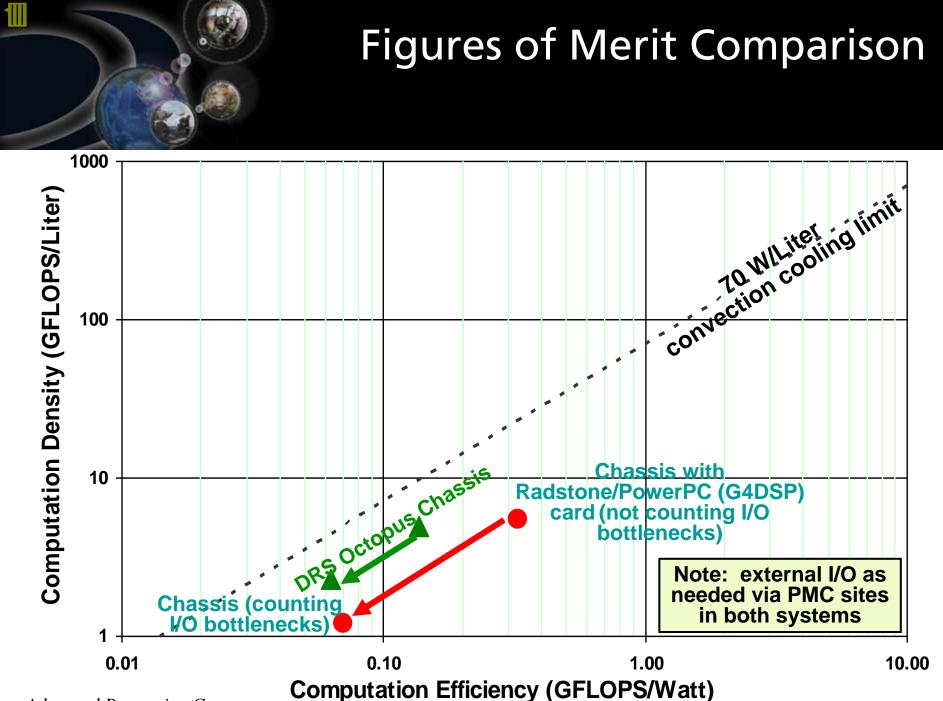
OCTOPUS Payload Slot

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



Inter-SoC Connectivity





Octopus SAR Benchmark Comparison Data Presented at Processor Technology Symposium on 10/3/06

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

Timing Results Normalized by Volume

Cell Processor software effort in process

Advanced Processing Group

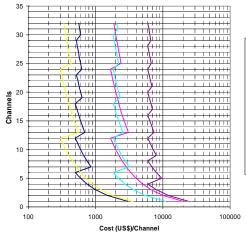
Best SWaP
Performance



Signal Recognizer

Octopus Shows SWaP Advantage over Blade Servers

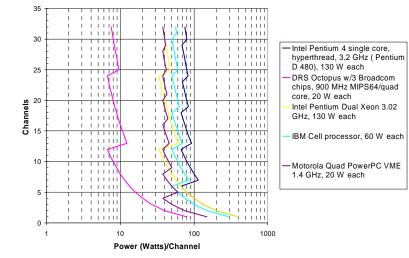
Recognizer Channel Performance for various processors



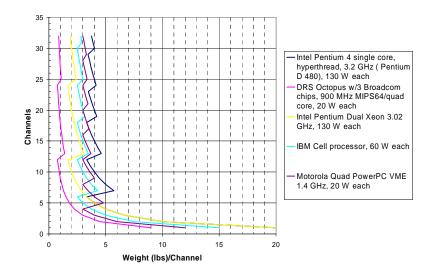
 Intel Pentium 4 single core, hyperthread, 3.2 GHz (Pentium D 480), 130 W each DRS Octopus w/3 Broadcom chips, 900 MHz MIPS64/quad core, 20 W each Intel Pentium Dual Xeon 3.02 GHz, 130 W each IBM Cell processor, 80 W each Motorola Quad PowerPC VME

1.4 GHz, 20 W each





Recognizer Channel Performance for various processors



Development Environment

| C/C++-ports.c-Eclipse SDK Eile Edit Refactor Navigate Search Project Run Window Help | | | | | | | | | |
|--|-----------------------------|---------------------|--------------|-------------------|--------------|--------------|----------------------------|--------|---|
| ° ▼ 🛛 🚔 🖬 🔤 🛣 ▼ 🕃 ▼ 🚱 | •] 🅸 • 🔘 • 🂁] 🥭 🔗 |] ĝ - ĝ - | *\$ | ⇔ - | | | | | 😭 🔽 C/C++ |
| C/C++ Projects 🕱 🔭 🗖 🗖 | 🖻 ports.c 🛿 🚺 ports.h | lc hr.c | 🖻 hr.h | spi_defs.h | 尾 pac | ket_mgr.c | | - 🗆 🔡 | E Outline 😫 |
| | /** | | | | | | | • | |
| bitwiseops.h 1.1 (ASCII -kkv | ✓ Properties for HelloWorld | | | | | | | | |
| Compioctl.h 1.2 (ASCII -kkv) | type filter text 🖌 | C/C++ Build | l | | | | ¢ | • - | |
| ▷ 🕞 drsdrv.h 1.1 (ASCII -kkv) | Info | | | | | | | | + ⁸ ports_AssignSpiBuffers |
| ▷ 🔄 drsioctl.h 1.4 (ASCII -kkv) | Builders | Active confi | guration | | | | | | ^{⊕ s} ports_Configure |
| ▷ Benv_defs.h 1.1 (ASCII -kkv) | C/C++ Build | Project Typ | e: Octopus | MIPS64 Execut | | | | \sim | ^{⊕ s} ports_ConfigureHt |
| Ap_defs.h 1.20 (ASCII -kkv) | C/C++ Documentatio | Configuratio | n: Debug | | | | ✓ Ma | nage | ⁺ ⁸ ports_ConfigureSpi4 |
| ▷ ⓑ fe_defs.h 1.8 (ASCII -kkv) | C/C++ File Types | | | | | | | | ⁺ ^s ports_ConfigureSpi4Calend |
| ▷ Insp_pkt.h 1.8 (ASCII -kkv) | C/C++ Indexer | Configuratio | n Settings | | | | | | ⊕ ⁸ ports_Discover |
| ▷ ⓑ id_defs.h 1.1 (ASCII -kkv) | Project References | Tool Settin | gs Build S | ettings Build Ste | eps Error F | arsers Binar | ry Parser Environment M | acros | |
| ▷ Bikm.h 1.2 (ASCII -kkv) | . iget received | | C C++ Com | - | | mips64_fp_l | | | ⁺ ^s ports_DumpHtBufferAssign |
| ▷ [h] packet_mgr.h 1.14 (ASCII -} | | | Preprocesso | ipiler | | | ll -c -fmessage-length=0 - | | ⁺ ⁸ ports_DumpPktBufferAssig |
| RamAllocNumEvents.h 1.1 | | | Directories | и . | All options. | | arch=mips64 -mtune=sb1 | | |
| ▷ ⓑ spi_defs.h 1.7 (ASCII -kkv) | | 1000 | Optimization | | | | 3 | | ↔ ^s ports_MapRegisters |
| ▷ ⓑ switch_mii_defs.h 1.1 (ASC | | | Debugging | | | | | | # BAR1_ADDRESS |
| ▷ | | | Varnings | | | | | | # HSP_RX_BUFFER_SIZE |
| Makefile 1.8 (ASCII -kkv) | | | Miscellaneo | | | | | | <pre># HSP_TX_BUFFER_SIZE</pre> |
| 7 😂 helloworld | | _ | C C++ Link | | | | | | # HT_RX_ACK_PHITCNT |
| Includes | | | General | ei | | | | | # HT_RX_CFILL_PHITCNT |
| 🗳 HelloWorld | | | ibraries | | | | | | # HT_RX_CRD_PHITCNT |
| hsp_cmd | | | Miscellaneo | | | | | | # HT_RX_NPC_CMD_PHIT |
| 🔓>hsp_net [cvs.cri-dsp.com] | | | Shared Libra | | | | | | # HT_RX_NPC_DATA_PHI1 |
| 、合>hsp_pkt [cvs.cri-dsp.com] レ | | | C Assemble | | | | | | # HT_RX_PC_CMD_PHITC |
| | | | General | | | | | | * BU BC BITT BUITC |
| | | | Seneral | | | | | | Make Targets 😫 |
| BCM1X80_HT.h 1.3 (AS) Buffer.h 1.12 (ASCII - kky) | | | | | | | | | viake raigets as |
| Common.h 1.6 (ASCII -kk | | | | | | | | | |
| · · · · · · · · · · · · · · · · · · · | Pr | | | | | | | | ²⁶ hdrs |
| | 2. | | | | | | Restore Defaults A | pply | ≝hsp_net |
| End test.h 1.4 (ASCII -kkv) | | | | | | | | PP'7 | ⊨ bsp_pkt |
| buffer.c 1.9 (ASCII -kkv) | | | | | | | OK Car | rel | |
| Durier.C 1.9 (ASCII - kkv) Market Structure (ASCII - kkv) | | | | | | | | icei | |
| Main.c 1.12 (ASCII - kkv) | | | | | | | | | |
| ▷ Packet_mgr.c 1.55 (ASC) | | | | | | | | | |
| P is >packet_ingite 1.55 (ASC >ports.c 1.24 (ASCII -kk) - | | | | | | | | | |
| | | | | | | | | | |
| /HelloWorld | | | | | | | | | |

Advanced Processing Com

Advanced Debug Tools

| V Debug - main.cc - Eclipse SDK | _ 8 × |
|--|---|
| <u>F</u> ile <u>E</u> dit Refac <u>t</u> or <u>N</u> avigate Se <u>a</u> rch <u>P</u> roject <u>R</u> un <u>W</u> indow <u>H</u> elp | |
|] 📬 🕶 📾 🖬 🛯 🏇 🛛 🐼 🖓 🖉 🧶 🥔 🖋] 회 문화로 🖏 🗇 - 오고 | ■ お ジ Debug ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ ・ |
| 🎋 Debug 🛛 🚺 🗰 🗰 🕼 🐐 🦃 🗈 🗮 🚿 🛊 ヤ 🗖 | 🗢 Variables 🗴 Modules Registers Signals Breakpoints 🛛 🏂 🐗 🕞 🗳 💥 🎇 🌄 🗖 🗖 |
| | (×)=i = 2948256 |
| | ▷ ➡ plnt = 0x08048650 |
| ∑ 🗗 Thread [0] (Suspended) | |
| ≡1 main() at/main.cc:4 | |
| Debugger Process (11/21/05 10:14 AM) | |
| 📓 /home/ali/work/HelloWorld2/Debug/HelloWorld2 (11/21/05 10:14 AM) | |
| | |
| | |
| | Ξ |
| | |
| | |
| i i ports.c i ports.h i i hr.c i i hr.h i spi_defs.h i packet_mgr.c i main.co | |
| <pre>#include <iostream></iostream></pre> | iostream |
| <pre>int main(){</pre> | an ain an |
| <pre> int i=0; int *pInt=&i </pre> | |
| | |
| <pre>std::cout << "Hello Octopus World\n"; i+=2;</pre> | |
| | |
| return 0; | |
| 3 | |
| | |
| | |
| | |
| | |
| | |
| Console Tasks 🟮 Memory 🛿 | |
| Memory Monitors 🛛 🖕 💥 Memory Renderings | ÷ x |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| 🌯 😌 🍥 👖 🕎 3 4 🛷 🗊 ali@Octosim6: /h 🥘 NewsWire Comple 🕊 The GIMP | 🔿 Debug - main.cc 🦷 BuildOptions.jpg-1 🎒 🐟 🥘 🔊 🥪 Denver Monday 📳 : 🕤 |