Using Millennium and Seaborg

E. Jason Riedy

February 2, 2004
More Hardware Information

- Seaborg – Power3
- Millennium – IA64
- Itanium2 Info
- Millennium – x86

Accessing the Systems

Storage Resources

Running Serial Jobs

Available Tools, Compilers, Debuggers

Diagnosing and Reporting Problems
Seaborg – Power3

- 380 POWER 3 16-way SMP compute nodes; 6,080 processors
  - Partitioned: Interactive, debug, and regular nodes
  - 4 have 64 GB; 64 have 32 GB; 312 have 16 GB memory
  - Two funky network adapters
- 4.5 Tflops peak
- 5 Tbytes memory (1.7 Gbytes/processor)
- 20 Tbytes disk (6.65 Gbytes/processor)
Power3 Information

- Clock speed: 375MHz
- Two fused multiply-add units
  - For *software divide*, not speed
- Two fmas yield two results, but perform “four” operations
  - “Peak performance” is 1.5 Gflop/s, real is 750 Mflop/s
- L1 insn cache: 32 KiB
- L1 data cache: 64 KiB
- L1 line size: 128 bytes
- L2 cache: 8 MiB (≥6.4 GiBps)
- Memory bandwidth: 1.6 GiBps (1.3 GiBps daxpy)
- Tuning info:
  www.redbooks.ibm.com/abstracts/sg245155.html
Issue 4, complete 8 per cycle

Swiped from an IBM summary page
Figure 2. POWER3 processor die photo

Swiped from an IBM summary page
Millennium – Itanium2

- Mix of nodes.
  - All dual-processor
  - Some 900 MHz Itanium2s, some 1.3 GHz
  - I’ll maintain a list

- Storage discussed later

- Don’t know “expansion” plans
Itanium2 Info

- Clock speed: 1.3 GHz, 800 MHz
- Two fused multiply-add units, two misc units
- “Peak”: 5.2 Gflop/s, real: 2.6 Gflop/s
- Also has SSE insns, mini-vectors
  - four single-precision entries
  - single-issue
- Caches: 32 KiB (d&i), 256 KiB, 3 MiB
- Line sizes: 64 b, 128 b, 128 b
- Cache bandwidth: 32 GiBps
- Memory bandwidth: 6.4 GiBps (don’t know practical)
- (See hw1)
Millennium – x86

- Rack-mount Pentium 3s
  - Some dual, some quad
  - Mix of 500, 550, 700MHz nodes
  - 512 KiB, 1 MiB caches
- Pretty heavily used
- Many “background” jobs
- Use it for testing / fun.
More Hardware Information

Accessing the Systems

Millennium

Seaborg

Storage Resources

Running Serial Jobs

Available Tools, Compilers, Debuggers

Diagnosing and Reporting Problems
Millennium

- Login nodes:
  - IA64: \{lime, lemon\}.millennium.berkeley.edu
  - x86: \{napa, sonoma\}.millennium.berkeley.edu

- Access with ssh (shell), sftp (file transfer)
  - Free clients listed at http://www.freessh.org

- Connect from ???
Seaborg

- Login name: seaborg.nersc.gov
  - Multiple nodes answer.
  - `uname -n` gives node name.
  - Specific node: go to seaborg and *then* node.

- Again, ssh and sftp work ([http://www.freessh.org](http://www.freessh.org))
More Hardware Information

Accessing the Systems

Storage Resources
   Millennium
   Seaborg

Running Serial Jobs

Available Tools, Compilers, Debuggers

Diagnosing and Reporting Problems
Millennium Storage

- “Small”, slow home directory (NFS)
- Much larger workspace: /work
  - mkdir /work/username
  - Nine day deletion policy
- Node-local storage: /scratch
  - Fast, but hard to manage.
  - Ten day deletion policy.
Seaborg Storage

- Ten GB home directory (GPFS), 15k inodes
- Scratch space: $SCRATCH
  - 256 GB quota, 50k inodes
  - Fuzzy deletion policy
  - Will not delete during a run.
- Do not use /tmp or /var/tmp.
More Hardware Information

Accessing the Systems

Storage Resources

**Running Serial Jobs**
- Millennium – IA64
- Seaborg
- Parallel Jobs

Available Tools, Compilers, Debuggers

Diagnosing and Reporting Problems
Running Jobs on Millennium

- Don't run on frontends lemon or lime.
- `/usr/mill/bin/gexec -n 5 uname -n`
  - `-n 5`: Run on 5 nodes. Default: All nodes.
  - Output will have relative node number prepended.
  - Specify nodes through `GEXEC_SVRS` env. var.

```
env GEXEC_SVRS="c10 c11" /usr/mill/bin/gexec -n 5 uname -n
env GEXEC_SVRS="'cat nodes-fast'" gexec cat /proc/cpuinfo
```

- Caveats:
  - Free-for-all, no queues.
  - “Load balancing” on stale info.
  - Each node broadcasts state every 30 secs.
  - Spawn 10 jobs in <30 secs, all land on same node.
  - Need different option to run on all x86 nodes
Running Jobs on Seaborg

- Debug on login nodes and special debugging queue.
- Full docs at http://hpcf.nersc.gov/computers/SP/running_jobs/
- LoadLeveler queue system:
  - Batch jobs use commented shell file.
  - `llqs`: Lists full queue
  - `llqs -u username`: list of your jobs
  - `llsubmit file`, `llcancel job number`
- Pretty pictures at http://hpcf.nersc.gov/cgi-bin/qstat/llq_seaborg
  - `module load www` to get `uname/passwd`
#@ job_name = Some job name
# llqs shows "Some job name"
#@ account_no = mp309
# mp309 is the class repo
#@ output = file-to-hold-stdout
#@ error = file-to-hold-stderr
# Default outputs to /dev/null
#@ job_type = serial
# Charges for whole node.
#@ notification = complete
# Email on completion
#@ class = regular
# Also debug, low, premium
#@ wall_clock_limit= 00:01:00
# Run for one minute
#@ queue

./a.out
Parallel Jobs

- mpirun on Millennium
  - Will include with later MPI information.

- poe on Seaborg
  - More LoadLeveler options.
  - Options not orthogonal...
More Hardware Information

Accessing the Systems

Storage Resources

Running Serial Jobs

Available Tools, Compilers, Debuggers
  Millennium – IA64
  Seaborg

Diagnosing and Reporting Problems
Millennium Tools – IA64

- Located in /usr/mill
  - Typical bin, include, lib, man hierarchy
  - Also see /usr/mill/pkg

- Compilers:
  - gcc-3.3, g77-3.3 (and -3.0)
  - Has the “famous” broken default gcc
  - Intel 7.1: /usr/mill/bin/{ecc, efc}
  - Will put latest in /home/cs/ejr/cs267/ia64
  - Also various perl, python, tcl versions

- Debugger: gdb (I’ll build more recent one), ddd

- BLAS, ATLAS, LAPACK in /usr/lib
  - needs /usr/lib/libg2c.so.0
Using Millennium and Seaborg

Available Tools, Compilers, Debuggers

Seaborg

Seaborg Tools

- See all modules: `module avail`
- Use IBM’s C compiler: `module use xlc`
- Check additional defs: `module info gcc/3.2.1`
- An appended `_64` implies a 64-bit library
  - Don’t worry about 64-bit apps for now.
- Compilers: (all separate modules)
  - `gcc 3.3, 3.2.1 (incl g77)`
  - `xlc 5, 6 (xlc_r: thread-safe libraries)`
  - `xlf 7, 8`
  - Also various perl, python, tcl versions
- Debuggers: `gdb, totalview, ddd, dbx`

http://hpcf.nersc.gov/computers/SP/programming.php
More Hardware Information

Accessing the Systems

Storage Resources

Running Serial Jobs

Available Tools, Compilers, Debuggers

Diagnosing and Reporting Problems
  Network Problems
  Seaborg
  Millennium
Problems

- Unreachable hosts
- Hung logins
- No home directory
Network Problems

- Try ping, traceroute, tracepath
- If those fail, contact intermediate network admin.
  - (Probably not by the broken network...)
- Otherwise, target host has problems.
Seaborg Administration

- Few problems.
- Check status:
  - http://hpcf.nersc.gov/
- Contacts:
  - Help page: http://hcpf.nersc.gov/help/
  - operator@nersc.gov, 1-800-66-NERSC (8-5 Pacific time)
Seaborg Administration

- Sometimes file system will hang nodes.
  - Don’t run `sync`
  - Contact operator@nersc.gov
- Password “doesn’t work” or needs reset
  - Try logging into sadmin.nersc.gov
  - Wait an hour after password changes
Millennium Administration

- Easy problems: support@millennium.berkeley.edu
- Real problems: 505 Soda Hall
- Semi-frequent issues:
  - Home directory gone: File server problems, check EECS
  - gexec hangs: Nodes down, “easy” problem
  - Programs don’t work: “Real”, just install them yourselves