

# Sony PS3 Cluster (IBM Cell BE)

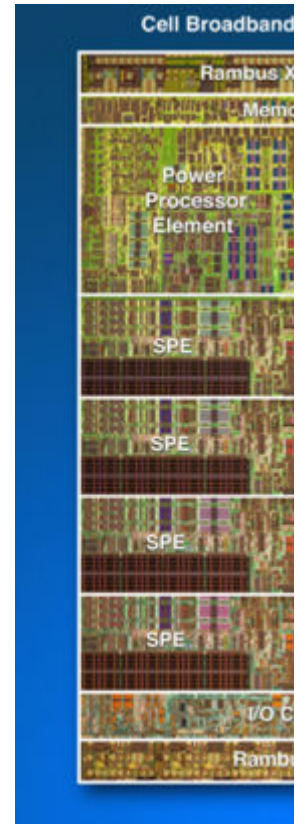
⇔ 8+1 nodes, ppc64 FC5 w/ CellSDK

[Official Annoucement of PS3 Cluster](#)

[NPR story on PS3 clusters](#) from "All things considered", Feb 21, 2009 ([local copy](#))

## Hardware

- 8+1 nodes
- IBM CellBE
  - one 2-way SMT PowerPC core (PPE), 3.2GHz, 32kB L1-I cache, 32kB L1-D cache, 512kB locking L2 cache, 256MB XDR DRAM
  - six SPEs (seventh is used by Sony virtualization software, eighth is hardware disabled in Sony PS3 version of CellBE), each at 3.2GHz w/ VMX vector unit, 128 SIMD GPRs, 256kB SRAM
  - 218 GFLOPS of total floating point performance (for 8 SPEs, that is)
  - 25.6GB/sec memory bandwidth (per interface to SPEs and XDR RAM)
- Gigabit Ethernet
- 20/60/120GB harddrive
- [full specs](#)
- operational since January 3, 2007



## Software



All software is 64 bit unless marked otherwise.

- [Fedora Core 5 Linux ppc64](#)
- [MPICH2](#) (also MPI API)  
locally: man MPI
- [OpenMP v2.5](#) (via gcc)
- gcc (C compiler)
- CellSDK 1.1



Notice that Linux is running on top of a virtualization layer provided by Sony. This layer is not open source. It shields many devices but should not affect performance on the processors. It may affect performance of the network and disk devices (unknown).

## Access

- Request a user id from Frank Mueller (server "optout" for cluster, server "opt" for research workstation). Please indicate your unity ID and student ID. Accounts on opt use your unity ID but do not share any unity file space.
- Install [~/.ssh/config](#) on your local Linux machine. If you're not using Linux, skip this step.
- On Linux, ssh into any psXX machine (XX=00..07).

```
ssh <your-username>@psXX
```

Your initial password is your student ID.

On windows, use your favorite ssh client to log in to optout.csc.ncsu.edu, port 22XX (XX=00..07). There may be a way to create aliases as well, check the manual of you ssh client.

For research workstation (only upon special request), use:

```
ssh <your-username>@ps3.csc.ncsu.edu
```

- Please change your password.

```
yppasswd
```

The new password will be effective immediately.



## Using OpenMP



- The "#pragma omp" directive in C programs works.

```
gcc -fopenmp -o fn fn.c
```

Gcc 4.1 for FC5 is back-patched with the 4.2 OpenMP support. It may not be the fastest but it works.

## Interactively running MPI programs



- Create the file [~/.rhosts](#) and set the access rights:

```
chmod 600 .rhosts
```

- Create a file [~/.mpd.conf](#) and set the access rights:

```
chmod 600 .mpd.conf
```

- Append to your file [~/.bashrc](#):

```
export PATH=".:~/bin:/usr/local/bin:/usr/bin:$PATH"
export MANPATH="/usr/local/man:/usr/share/man"
```

Log out and back in to optXX to activate the new settings.

- Check if MPD is running:

```
mpdtrace
```

This should return a list of available nodes to run MPI jobs on. If not, see below: user-level MPD

- Compile MPI programs:

```
mpicc -O3 -o pi pi.c
```

If you're using BLAS/ATLAS:

```
mpicc -O3 -o pi pi.c -L/usr/lib64/atlas -lcblas -latlas
```

- Execute the program on 2 processors (using MPICH2):

```
mpirun -np 2 pi
```

Try again with a different number of processors.

- Use the new MPICH2 run/execute interface (more options, see man page):

```
mpiexec -n 2 pi
```

Try again with a different number of processors.

## User-level MPD

If mpdtrace is not working, then you have not started your MPD ring yet. Follow the steps below.

- Figure out which nodes are alive: **ping optXX** for XX=00..15. Still, they may not work properly. Try: **rsh optXX pwd --** maybe working?
- Create a file like [~/mpd.hosts](#) but only include the nodes that are working!
- Edit your `~/mpd.conf` file to remove/comment out the line

```
MPD_USE_ROOT_MPD=1
```

- If you have 2 working nodes and you are logged in to optYY, issue:

```
mpdboot --ifhn=optmpiYY -n 2 -r rsh --ncpus=2
```

You can now run MPI jobs.

- Inquire about running jobs:

```
mpdlistjobs
```

- Stop MPD daemons:

```
mpdallexit
```

- Clean up MPD daemons and files:

```
mpdcleanup
```

- Vary your mpdboot parameters. In the above case, you'll run 2 jobs per node (since we have 2-way SMTs, i.e., 2 logical processors per node). Check out the man pages.

---

## Known Problems

Please let us know if you have solutions to these problems.

- Ganglia has not been installed yet
  - ATLAS has not been installed yet
  - PAPI has not been installed yet
-

## More Information on the PS3

- [Articles from CellBE Issue, IBM JR&D, 2007](#)
  - [README.fedora9](#) covering installation and programming for F9+SDK 3.1
  - [README.fedora7](#) covering installation and programming for F7+SDK 3.0
  - [all SDK 3.0 docs](#)
  - [UTK PS3 Cluster TR](#)
  - [Cell programming](#) by Terra Soft / Yellowdog
  - older docs:
    - [README](#) covering installation and programming for FC5+SDK 1.0
    - [Cell programming tutorial \(V1.1\)](#)
    - [Cell programming tutorial \(V2.0\)](#)
    - [Cell SKD Guide](#)
    - [C++ Language Extensions for Cell V2.2.1](#)
    - [SPE Library Calls \(V1\)](#)
    - [SPE Library Calls \(V2\)](#)
    - [SPE Library Calls \(V2.2\)](#)
    - [SPU Language Extensions \(V2.1\) by Sony \(local copy\)](#)
    - [Libraries SDK](#)
    - [IDL for RPCs to SPEs](#)
  - [Cell talks by IBM and others](#)
  - [Cell course by IBM \(local copy of slides\)](#)
  - [Cell workshop Cupertino 2006 \(local copy of slides\)](#)
  - [Linux on Cell \(Barcelona\)](#), incl. Cell SDK and kernel releases (but not PS3 specific)
  - [Sony PS3 Linux docs](#)
  - [Maximizing the power of the Cell Broadband Engine processor -- 25 tips to optimal application performance](#)
- 

## References:

- [A User's Guide to MPI](#) by Peter Pacheco
- Debugging: Gdb does not work with MPI, and we don't have totalview (an MPI-aware debugger). Hence, you are down to good old printf debugging, unfortunately.
- [Linux Cluster Info](#) (add cluster tools to the wish list)

### Additional references:

- [Intro to UNIX at NCSU](#)
- [Emacs reference card](#)
- [Gdb reference card](#)
- [Gnu manuals \(make, gdb and many more\)](#)