

# An introduction to the new OS on the Intel Jet system

Jet Management Team  
August 2004

Updated 24 Aug. 2004

# Introduction

- Several key changes
  - Reliability
  - Performance
  - Scalability
  - Security
- Changes are extensive
  - Kernel
  - Compilers
  - Tools
- User should not really see any difference

# Upgrade Schedule

- Current system
  - A (60 nodes), C (120 nodes), g0803 (front-end)
- Thu. Aug 19<sup>th</sup> and Wed. Aug 25<sup>th</sup>
  - Jet users meeting
- Thu. Aug 26<sup>th</sup>
  - System B (60 nodes)
- Wed. Sep 1<sup>st</sup>
  - System D (250 nodes), SGE server 2
- Wed. Sep 8
  - System E (250 nodes), cron servers, front ends, SGE server 1

# What you need to know

- Although there are many changes, there are 3 critical changes to the system
  - You must recompile your codes
  - You must change your batch scripts to submit to the *ncomp* parallel environment
    - - *pe ncomp <CPUS>*
  - The Intel Fortran compiler is now called *ifort*
- All of the system utilities are on the system path
- Their locations have changed (/opt)

# What prompted the change

- Update system and libraries for compatibility with new software
- Bug fixes, including issues with glibc and compiling with PGI and linking statically
- The new architecture allows security updates to be applied more easily
- Improved compiler feature and performance
- Improved interconnect performance
- Improved scaling for larger codes
- Better infrastructure for adding new features

# New or Updated Software

- Intel C++/Fortran 8.0
  - Version 7.1 is still available
- PGI C++/Fortran 5.1-6
- SMS-2.8.0
- GM 1.6.5
- MPICH 1.2.5..10
- Optimized BLAS library
- HDF 5
- Updates to Perl, Ruby, Python, Tcsh, Netcdf

# Performance Improvements

- Compiler performance has been improved
  - MM5 is running 40% faster with the Intel compilers
  - RUC20 is running 25% faster
- Kernel has been improved to reduce OS jitter
- Network drivers have been upgraded
  - Increased bi-directional bandwidth
  - Slightly lower latency

# Tips and Tricks

- Intel Compiler
- Portland Group Compiler
- Changes to MPIRUN
- Scratch Filesystem
- Other tips and tricks

# Intel Compiler 8.0

- Fortran has been renamed *ifort*
  - using *ifc* will still work
- You should compile with '-O3 -axN'.
  - The use of -xW is depreciated
- '-convert big\_endian' is supported
- Compile with '-g -traceback' to get tombstone-like information
- Compile with '-check all' to get run-time bounds checking

# Portland Group 5.1-6

- The new compiler provides better performance, improved features, and bug fixes.
- Use '-fastsse' to optimize for Intel system
- Code profiling supported directly by the compiler.
- Access the compilers same as now
  - `source /usr/local/bin/pgsettings.{c}sh`

# MPIRUN

- MPIRUN has changed slightly from previous versions
  - Environment variables are passed directly on the mpirun line
    - Ex: *mpirun -np 2 MY\_VAR=3 ./myexe*
    - This is done automatically for SMS variables and specifying endian-ness for the Intel compiler (F\_UFMTENDIAN)
  - Tombstone needs to be specified directly
    - *mpirun -np 2 gdb\_tombstone\_wrapper ./myexe*

# Scratch Filesystem

- /misc/scratch is a scratch space built on the Parallel Virtual Filesystem (PVFS)
- 1.6 TB, ~ 250 MB/s
- No quotes, no scrubbing
- Best for big data files
  - NETCDF performance is poor due to how it writes headers
- DO NOT put source code on this system
- The system is not backed up
- Submit to '*pvfs*' parallel environment to access

# Other tips and tricks

- qtop
  - Provides top like information for running job
- sgestat -why
  - Tells you why your job isn't running yet
- Trick to remove need to specify cpu count in batch scripts:

```
set ncpus=`cat $MACHINEFILE | wc -l`  
mpirun -np $ncpus wrf.exe
```

# Upcoming Features

- SGE 6.0
  - Upgrade to the batch system will provide
    - Better stability
    - Advanced reservations
    - Better control of jobs to improve system utilization and allow users to run more jobs during off-hours
- Attach debugging processes to running MPI jobs
- Hardware counters for code execution profiling

# Summary

- Please ask any questions you may have here or on the help list at:
  - [\*\*\*hpcshelp.fsl@noaa.gov\*\*\*](mailto:hpcshelp.fsl@noaa.gov)
- See the FAQ for a summary of this information and other facts about the system (ever changing):
  - [\*http://hpcs.fsl.noaa.gov/faq.html\*](http://hpcs.fsl.noaa.gov/faq.html)

# Reminder

- Alpha Jet is being decommissioned on September 10th. Please have all jobs migrated from that system.
- /ahome is not going away, but will be clean for other users. Please move your data.