System Overview

WELCOME TO FROST

Your request for a logon on frost (Blue Gene/L) has been approved. Use your UCAR Central Authentication Server (UCAS, a.k.a. gatekeeper) password for access. Frost is currently open to a small group of selected friendly users (Principal Investigators, staff, and a few others).

END-USER ENVIRONMENT

The BlueGeneWiki is a central repository for end-user information about using frost, and Blue Gene/L systems in general: https://wiki.ucar.edu/display /BlueGene/Frost

You will find information about how to login, prepare and run your codes on Blue Gene/L, tips for performance and tuning, whitepapers and conference presentations, and other items of interest to the NCAR/CU Blue Gene/L community. This is a dynamic repository, and your requests for additions and changes are always welcome.

File systems

There are two file systems for users on Frost. The space in /home/username is an NFS filesystem and has a quota of 4GB. The space in /ptmp/username is a parallel GPFS filesystem and has a quota of 800GB.

The Blue Gene/L Runtime Environment

Codes for frost need to be compiled with one or more of the IBM Blue Gene/L XL compilers: `blrts_xlc`, `blrts_xlC`, `blrts_xlf`, or `blrts_xlf90`. These are cross-compilers, so codes compiled with them can only be run on the Blue Gene/L rack itself and not the front-end nodes.

To get started, use the following compiler and linker flags and options:

Compiler: `-g -O -qarch=440 -qmaxmem=64000 -I/bgl/BlueLight/ppcfloor/bglsys/include`

Linker: `-L/bgl/BlueLight/ppcfloor/bglsys/lib -lmpich.rts -lmsglayer.rts -lrts.rts -ldevices.rts`

You may also use the MPI wrapper scripts (`mpxlc`, `mpxlC`, `mpxlf`, `mpxlf90`) in the `/contrib/bgl/bin` directory for compiling your codes.

This and other information about compiling and running codes on Frost can be found on the BlueGeneWiki at: https://wiki.ucar.edu/display/BlueGene /Frost+Info

Additionally, the IBM mass and massv intrinsic libraries have been ported to Blue Gene/L and are available on Frost. NetCDF 3.6.1 is also available. Details are on the BlueGeneWiki at: https://wiki.ucar.edu/display/BlueGene/BGLLibraries

Profiling and debugging tools will be added as they become available, and usage instructions will be added to the BlueGeneWiki.

You will also find NetCDF, nco, ncarg, and ncl installed on the frost front-end nodes in the /contrib/fe_tools/gnu32 directory. To use these tools, add /contrib /fe_tools/gnu32/bin to your path and NCARG_ROOT=/contrib/fe_tools/gnu32/ncarg to your environment.

End-user information (including Redbooks) about Blue Gene/L from IBM is available on the BlueGeneWiki at: https://wiki.ucar.edu/display/BlueGene /Info+From+IBM

The Blue Gene/L Batch Job Scheduler

The batch job scheduler on frost is Cobalt from ANL. Frost currently supports partition sizes between 32 and 4096 nodes. A partition of the appropriate size will be automatically selected based on the number of nodes you request for your job. However, when using partitions of 512 nodes or less, the 3D-torus network is reduced to a 3D mesh (i.e. the wrap-around edges are inaccessible.) Cobalt on frost is currently configured with these partition sizes: 32, 64, 128, 256, 512, 1024, 2048, and 4096. Information about running code using Cobalt on frost can be found on the Frost Info page of the BlueGeneWiki: https://wiki.ucar.edu/display/BlueGene/Frost+Info

Blue Gene/L Support Contacts

- User support: frost-help@ucar.edu
- General Discussion: bglusers@ucar.edu

Please subscribe to the bglusers and bglstatus mailing lists via the following links:

BGLUsers: http://mailman.ucar.edu/mailman/listinfo/bglusers

Blue Gene/L Supercomputer (frost) users should use the bglusers mail list to discuss application issues.

BGLStatus: http://mailman.ucar.edu/mailman/listinfo/bglstatus

The bglstatus mail list is used to send email notifications regarding Blue Gene/L (frost) and the Front End Nodes uptime, downtime, emergency downtime, systems problems, etc.

HARDWARE OVERVIEW:

Frost is a four-rack, Blue Gene/L system with 4096 compute nodes. There is one I/O node for every 32 compute nodes (pset size of 32) for a total of 32 I/O nodes in each rack. Each compute node and I/O node is a dual-core chip, containing two 700MHz PowerPC-440 CPUs, 512MB of memory, and two floating-point units (FPUs) per core. Thus frost has a total of 8192 processors capable of sustaining a peak performance of 22.936 trillion floating-point operations per second (TFLOPs). By default, the compute nodes run in coprocessor mode (one processor handles computation and the other handles communication), but virtual node mode is also available, where both processors share the computation and communication load.

An IBM p630 service node controls the Blue Gene/L rack, and no state information in the Blue Gene/L rack itself. Rather, all of this information is housed in a DB2 database on the service node. A suite of daemons on the service node is responsible for monitoring and managing all aspects of the Blue Gene/L system (e.g., hardware events, job submission, execution, and completion, etc.). There is no direct end-user interaction with the service node.

User interaction with Blue Gene/L is performed via four front-end cluster nodes. Each of these is an IBM OpenPower720 server, with four POWER5 1.65 GHz CPUs with 8GB of memory and running SuSE Linux Enterprise Server Version 9 (SLES9). You will use the compilers, debuggers, and other end-user tools from these systems, and also submit all your jobs for the Blue Gene/L rack using Cobalt's `cqsub' tool.

The front-end nodes also provide the connection to the back-end disk subsystem. The back-end disk subsystem is comprised of two IBM DS4500 Fiber Channel storage servers with dual controllers and a combined total of ~6 TeraBytes of disk space spread across twenty-two RAID5 disk arrays. The disk subsystem is capable of ~600MB/sec of throughput.