



NERSC-6 Machine Component Characteristics

July 2008



Memory Summary

- **Requested Features**

- Large amount of aggregate user addressable memory with automated error detection and correction
- Minimal intrusion upon memory available to application data structures by system libraries, daemons, operating system and/or kernel

- **Benchmarks**

- **Bandwidth:** *Stream, NPB MPI, MILC*
- **Latency:** *MAESTRO*
- **Retire rate/performance:** *MAESTRO, AMR Elliptic*
- **Memory access patterns**
 - Unit stride: *GAMESS, PARATEC*
 - Random access and non-unit stride: *MAESTRO, IMPACT-T, GTC*
- **Capacity:** *PARATEC, MAESTRO, GAMESS*



CPU Summary

- **Requested Features**

- Must commit to minimum SSP
- Must provide correct, consistent and reproducible computation results
- Must provide consistent and reproducible execution times in dedicated and production mode
- Must comply with IEEE 754 64-bit floating point arithmetic

- **Benchmarks**

- **Sustained Performance on Workload:** *SSP Metric*
- **Sustained Performance on Subset of Workload:** *Full Applications*
- **CPU Performance Without Interconnect:** *Serial NPBs*
- **Low Computational Intensity:** *STREAM, AMR_Elliptic, MAESTRO, NAS-CG (packed/unpacked)*
- **High Computational Intensity :** *MILC, PARATEC*



Interconnect Summary

- **Requested Features**

- Require a high-performance, high-bandwidth, low-latency, fault-tolerant, scalable interconnect
- Must be able to run a single application instance over all compute nodes
- Consistent and reproducible execution times for communication-intensive applications in dedicated and production mode
- Support for advanced programming languages eg. UPC, CAF, shared memory abstractions such as Global Arrays through one-sided messaging, efficient RDMA support, and/or global addressing

- **Benchmarks**

- **Ability to run across entire machine:** *Full Configure Test (FTC)*
- **Point-to-Point Bandwidth:** *Multipong*
- **Bisection Bandwidth:** *Full Configure Test (FCT)*
- **Point-to-Point Latency:** *Multipong*
- **One sided communication:** *UPC, GAMESS*
- **MPI Collective communications:** *IMPACT-T, MAESTRO, GTC, MILC*
- **MPI Point-to-point communications:** *PARATEC, CAM, GTC, MILC*



I/O Summary

- **Requested Features**

- Require external parallel filesystem, administered independently from computational nodes
- Require consistent and reproducible execution times in dedicated and production mode.
- Global shared storage system capacity of at least 1 PB, providing at least 70 GB/s of measurable, sustained aggregate filesystem I/O bandwidth
 - Final size negotiated during SOW phase
- High efficiency for both large and small block I/O for shared file access with high concurrency.
- Tightly integrated filesystem software to MPI-IO and POSIX layers.
- High sustained parallel and single stream filesystem I/O bandwidth to and from global shared storage system.

- **Benchmarks**

- **Raw I/O Bandwidth of Disk Subsystem:** *IOR large-block I/O test*
- **Sustained I/O Bandwidth for Applications:** *IOR for small block I/O (size based on survey of application profiles)*
- **Interface:** *IOR with POSIX, MPI-IO tests*
- **Parallel Strategy:** *IOR tests for one-file per proc and to shared file*
- **Metadata Server Performance:** *Metabench file create, stat, and delete tests*
- **Full Application Benchmark:** *MAESTRO with I/O capability enabled*



OS/Software Stack

- **Requested Features**

- Require an application development environment consisting of at least: standards compliant Fortran, C, and C++ compilers, and MPI and MPI-IO libraries.
- High performance, well-integrated MPI (MPI-2) implementation
- Support within MPI (and underlying hardware) for accelerated collective operations
- Advanced resource management functionality, high utilization, backfill, reservations, etc.
- Minimal intrusion upon memory available to applications or CPU resources by system libraries, daemons, operating system and/or kernel.
- Architectural features to improve application scaling and decrease system jitter
- User access to performance counters and PAPI interface
- Fully featured application development environment, vendor optimized serial and parallel scientific libraries (e.g., LAPACK, BLAS); MPMD MPI; Python; GNU tools and utilities; parallel debugger, performance analysis tools

- **Benchmarks**

- **OS Noise:** *PSNAP, Chombo AMR, Consistency Metric*
- **Vendor Tuned Libraries:** *PARATEC makes heavy use of FFTW/BLAS/BLACS/SCALAPACK*
- **Build and Runtime System Robustness:** *CAM, GAMESS*
- **Fortran Compiler Robustness:** *IMPACT-T (object oriented f90), and CAM (complex link and fortran modules)*
- **C/C++ Compiler Robustness:** *CHOMBO AMR (C++)*
- **Environment support for PGAS and/or Global Memory:** *GAMESS, UPC*
- **Resource management, batch scheduler:** *Effective System Performance (ESP) Test*