NAS and NCCS Moving to a Common Standard Billing Unit May 1, 2011

By Jarrett Cohen with NAS and NCCS Staff

The NASA Advanced Supercomputing (NAS) Facility and the NASA Center for Climate Simulation (NCCS) are moving to a common Standard Billing Unit (SBU) for allocating and tracking computing time usage. The new SBUs will appear in usage reports starting on May 1, 2011, the beginning of the allocation year for many computational projects. For projects that began November 1, 2010, usage through April 30, 2011 will be recalculated to reflect the change.

“We need one currency for an integrated supercomputing environment so that people can move from system to system and still get equitable credit for what they are allocated,” said Tsengdar Lee, NASA High-End Computing (HEC) Program Manager (currently on detail in the NASA Office of the Chief Information Officer). “We have several generations of scalable units based on different processor types and need to normalize usage across them.”

The new SBU represents the amount of work that a Minimum Allocatable Unit (MAU) can perform in one hour. The MAU is the smallest unit of each system that the scheduling software will allocate to a job. On the Pleiades (at NAS) and Discover (at NCCS) supercomputers, an MAU is one node; nodes on Pleiades and Discover vary from 4 to 12 processor-cores. On Columbia (at NAS), an MAU is 4 cores. “If an MAU runs faster, it gets a larger SBU; if it runs slower, it gets a smaller SBU,” explained Bill Thigpen, High-End Computing Capability Deputy Project Manager at NAS.

The Pleiades supercomputer at the NASA Advanced Supercomputing (NAS) Facility includes racks based on Intel’s latest Westmere processor. During 2010, a series of expansions brought Pleiades to 10,816 nodes (101,120 cores). Photo by Dominic Hart, NASA Ames Research Center.

MAU-based SBUs require a transition for both HEC facilities, akin to revaluing a currency. NAS has been using processor-based SBUs for more than 10 years, while NCCS has been tracking straight processor-hours for a similar period. Using MAU-based SBUs is both more realistic, because users cannot be charged for a smaller portion of the system when computing, and more consistent, because the facilities already charge computing time based on nodes on Pleiades and Discover.

Extensive Benchmarking

To determine the new SBUs, staff chose a new benchmark suite consisting of NASA-relevant applications. “We used many factors in choosing the suite, including code stability, portability, and verifiability,” said Piyush Mehrotra, NAS Applications Performance and Productivity Group Lead, who has spearheaded the effort to calculate the new SBUs. “However, the main focus was to identify applications whose characteristics best represented the wide spectrum of codes utilized by scientists and engineers on the NASA HEC systems.”

The benchmark suite includes six software codes:

- **Enzo**: An adaptive mesh refinement, grid-based hybrid code designed to simulate cosmological structure formation (University of California, San Diego, et al.).
- **FUN3D**: An unstructured computational fluid dynamics (CFD) code for analysis, adjoint-based error estimation, mesh adaptation, and aerospace design optimization extending into the hypersonic regime (NASA Langley Research Center).
- **Goddard Earth Observing System Model, Version 5 (GEOS-5)**: The atmospheric general circulation model from the GEOS-5 system of models that also integrates an atmospheric analysis to support climate and weather prediction, data analysis, observing system modeling and design, and basic research (NASA Goddard Space Flight Center, et al.).
- **OVERFLOW**: A CFD program for solving complex flow problems that is widely used for designing launch and re-entry vehicles, rotorcraft, ships, and commercial aircraft (NASA Langley Research Center).
- **USM3D**: An unstructured mesh code for calculating flows over complex geometries that is often used to analyze aerodynamic flow of aerospace vehicle designs (NASA Langley Research Center).
- **Weather Research and Forecasting (WRF) Model**: A next-generation mesoscale numerical weather prediction system designed to serve both operational forecasting and atmospheric research needs (National Center for Atmospheric Research, et al.).

NAS and NCCS staff ran the benchmarks on the full spectrum of processor types available at each facility. “It is a good time for us to be doing this since it is only recently that we have diverse architectures running at the same time,” said Ellen Salmon, NCCS User Liaison for Data Services.

For each code, the staff determined an appropriate core count and then ran the code on each processor type using that core count. Next, they applied a mathematical formula that compares the running time of the code on each system to the running time on a baseline system. The formula then takes a weighted average across all codes to determine the SBU rate for each processor type. The weights used in this calculation were proportional to the code’s usage on NASA systems. For the new SBUs, a Westmere (Intel Xeon 5670 series) node on Pleiades serves as the baseline system. Thus, the Westmere’s SBU rate is equal to 1. Multiplying the SBU rate by wall clock hours and the number of MAUs used determines the SBU usage for any given run. The new SBU rates for NAS and NCCS follow.

NAS Standard Billing Unit (SBU) Rates					
System		**Columbia** SGI Altix 4700	**Pleiades** SGI ICE 8400EX	**Pleiades** SGI ICE 8400EX	**Pleiades** SGI ICE 8400EX
Processor		Intel Itanium2	Intel Xeon Harpertown	Intel Xeon Nehalem	Intel Xeon Westmere
MAU		4 Cores	8 Cores	8 Cores	12 Cores **BASELINE**
SBU Rate		**0.18**	**0.45**	**0.80**	**1.00**

NCCS Standard Billing Unit (SBU) Rates					
System	**Discover** Linux Cluster Custom Supersystem	**Discover** Linux Cluster Custom Supersystem	**Discover** IBM iDataPlex	**Discover** IBM iDataPlex	**Discover** Dell PowerEdge C6100
Processor	Intel Xeon Dempsey	Intel Xeon Woodcrest	Intel Xeon Harpertown	Intel Xeon Nehalem	Intel Xeon Westmere
MAU	4 Cores	4 Cores	8 Cores	8 Cores	12 Cores
SBU Rate	**0.15**	**0.20**	**0.38**	**0.69**	**0.95**

These new Standard Billing Unit (SBU) rates will be used for calculating computing time usage on NAS Facility and NASA Center for Climate Simulation (NCCS) supercomputers. The number of cores represents the size of a Minimum Allocatable Unit (MAU) on the system. For Xeon-based systems, it is the number of cores in a node; for Itanium2-based Columbia, it is four cores.

Getting the Most Return on SBU Dollars

While usage calculation is based on new SBU rates, principal investigators (PIs) request computing time under the same procedures as before (see Related Links). All NASA Mission Directorates now use the e-Books online submission system. For each supercomputer they want to access, PIs estimate the number of processors per run, the wall clock hours per run, and the number of runs. Calculators within e-Books total the processor-hours from these estimates and convert the processor-hours to SBUs. PI requests then go through an approval process that results in the HEC Program granting computing time allocations measured in SBUs. Although there is no monetary cost for using NASA supercomputers, PIs will have what amounts to SBU “bank accounts” good for the allocation period.

For PIs and users of the supercomputers, NAS and NCCS offer several tools for tracking usage such as acct_ytd at NAS and allocation_check at NCCS.
In addition, NCCS has web-based User, PI, and Sector Head Portals for tracking individual and group usage (see Related Links). To ease the transition to SBUs, NCCS usage will be shown in both SBUs and processor-hours through October 31, 2012.

Beyond mechanics, moving to the new SBU rates may alter users' perceptions of NASA's processor offerings. "The value of the user's allocation dollar is different depending on where they spend it," said Tyler Simon of NCCS User Services. "Users should be charged to compute based on performance. If they compute on a Nehalem node, it will be more expensive, because it is faster. If they compute on a Dempsey node, it will be cheaper, because it is slower. The Nehalem job will take less time, however, so the total cost of the job may be similar."

After more than a decade of using SBUs at NAS, Thigpen has seen two types of users develop. "Some want to run as fast as possible; they won't change," he said. "Others will look for the processor family that gives them the best return on their dollars."

By changing from charging for cores to charging for MAUs, the true cost of some jobs will be more apparent. "Often the best way to run a job uses a different number of cores per node than you might expect," said Robert Hood, NAS Application Performance Lead. "For example, using only half the cores on Harpertown is often the most efficient use of resources. Similarly, using hyperthreading on Nehalem and Westmere can be an efficiency win. The trouble is that when a job uses only half the cores on a Harpertown node, it still pays for all of them. The hope is that the new charging scheme will encourage users to find the best way to use each node that they are paying for."

NCCS recently doubled the capacity of its Discover supercomputer by adding a Dell PowerEdge C6100 cluster with 1,200 Westmere nodes (left). Discover now has a combined 2,864 nodes (29,368 cores). Photo by Pat Izzo, NASA Goddard Space Flight Center.

Resource availability is an important factor in determining the best processor type to use for a run. "It boils down to rate of consumption versus turnaround time," Salmon said. The NAS System Status and NCCS Job Monitor online tools show up-to-the-minute statistics on current utilization organized by type of node. When submitting jobs, users can request specific node types by entering information into a Portable Batch System (PBS) script. On the NCCS Discover supercomputer, for example, proc=neha specifically requests Nehalem nodes. On the NAS Pleiades supercomputer, model=neh requests Nehalem nodes. (For more details on system monitoring and job submission, see Related Links.)

Although the benchmark codes are broadly representative of NASA work, Salmon urged users to experiment. “It is probably worth doing runs on different processors to see how well codes perform,” she said. “You might get better performance or more work done than the SBU would indicate.”

By charging for the amount of work performed, SBUs help the HEC facilities run more efficiently. Thigpen said that NAS wants to encourage people to run on all the systems, but “with older equipment, if there is no reason to use it, users won’t.” In 2010, NAS turned off the oldest parts of Columbia because it was cheaper to replace them than pay the electricity cost to run them. They expanded Pleiades with new Westmere-based racks, which supply 13 times more SBUs per kilowatt-hour than the retired Columbia racks.

Similarly, SBUs are beneficial “when we allocate resources and buy new equipment,” Lee said. “We are not buying peak performance. We are buying throughput, and SBUs are a way to measure how much the system can deliver.”

Questions
Users with questions should contact one of the following:

NAS User Services
Toll-free: 1-800-331-8737
Local: 650-604-4444
E-mail: support@nas.nasa.gov

NCCS User Services Group
Phone: 301-286-9120
E-mail: support@nccs.nasa.gov

Related Links:

- NASA Advanced Supercomputing (NAS) Facility: New Common SBU Rates
- NASA Center for Climate Simulation (NCCS): New Common SBU Information
- NAS: System Status
- NCCS: Job Monitor
- NAS: SGI Altix ICE Running Jobs (Pleiades)
- NAS: SGI Altix Running Jobs (Columbia)
- NCCS: Discover User Guide
- NASA High-End Computing Program: Request Computing Time

NCCS Portals (password required)
- User Portal
- PI Portal
- Sector Head Portal