

BLUE WATERS

SUSTAINED PETASCALE COMPUTING

5/17/15

Blue Waters User Monthly Teleconference



GREAT LAKES CONSORTIUM
FOR PETASCALE COMPUTATION

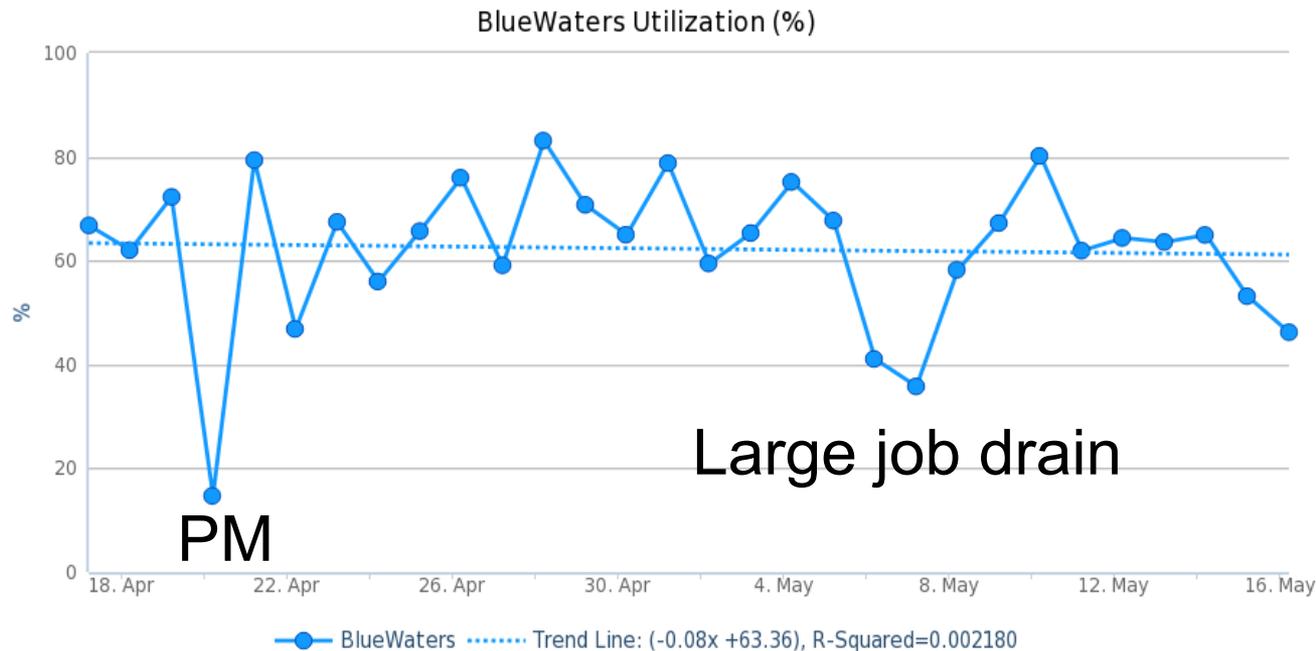
CRAY®

Agenda

- Utilization
- Recent events
- Recent changes
- Upcoming changes
- [Blue Waters Data Sharing](#)
- [2015 Blue Waters Symposium](#)
- Other Opportunities
- PUBLICATIONS!

System Utilization

- Utilization since last BW User Call (April 20)

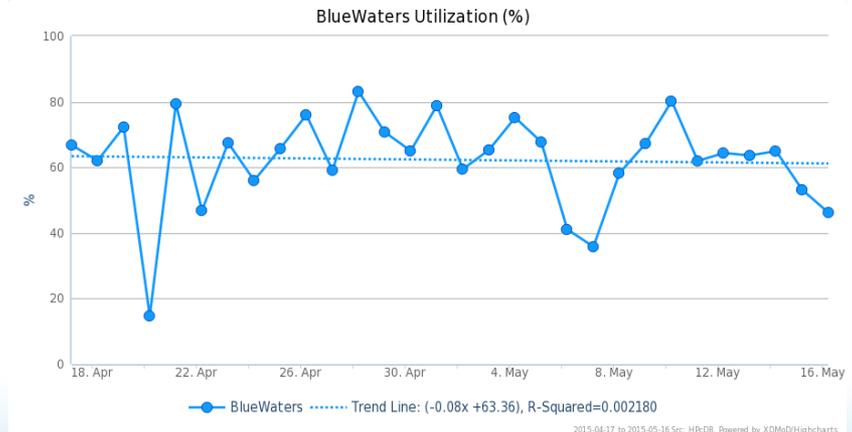
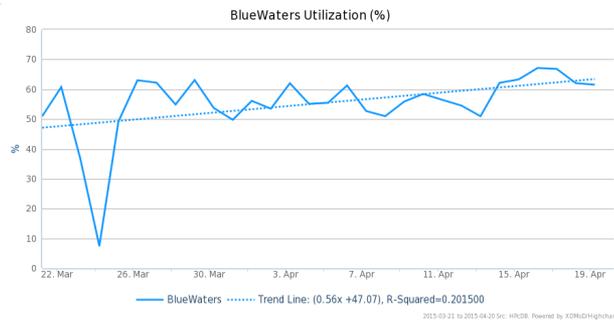


2015-04-17 to 2015-05-16 Src: HPCDB. Powered by XDMoD/Highcharts

- Utilization improving. We are still investigating.

Recent Opportunities

- Discount Period Completed
 - Friday, April 17th at 12:00AM - Midnight May 15th
 - Charge factors reduced by 50% for all queues.
 - Resulted in good job mix and good utilization.



Recent Issues

- Outages
 - 5/8 – scratch availability due to OST outage. Scheduler paused for 1 hour.

Recent Changes

- CUDA 6.5
 - NVIDIA UNIX x86_64 Kernel Module 340.87
 - Provides unified memory (UM), performance improvements, and improved debugging.
 - cudatoolkit/6.5.XYZ
 - Existing CUDA 5.5 binaries shown to work
 - If issues are encountered, swap to previous model
 - cudatoolkit/5.5.51-1.0502.9594.3.1.

Recent Changes

- April 2015 Programming Environment
 - Default to support CUDA 6.5
 - cce/8.3.10
 - cray-mpich 7.2.0
 - cray-libsci_acc/3.1.1
 - PGI and GNU to remain at current defaults
 - PGI 15.X to come in June.
 - GCC 4.9.0 (or later) are not supported with CUDA Toolkit 5.5 or CUDA Toolkit 6.5.

Upcoming Changes

- Modules
 - Continue to clean out unused and old modules.
 - Removing deprecated modules: old `xt-xyz` replaced with `cray-xyz`.

Improving Job Turn-around time

Improvement of job wall clock accuracy

Wall clock accuracy is defined as actual elapsed wall clock time divided by requested wall clock time. Providing accurate wall clock times makes it easier for the scheduler.

Job backfills available nodes that are draining for other jobs

Backfilling improves job turn-around time and utilization and happens most often for jobs requesting less than several hundred nodes and wall clock times less than 6 hours. Backfill opportunities vary based on system load and are not guaranteed.

Communication specifiers

Use of flags `commintolerant` or `commglobal` will cause jobs to take longer to be placed. Avoid where ever possible.

Unbundling Jobs

Prior to topology aware scheduling (TAS), bundling was recommended. With TAS for workloads that don't require large, convex shapes we don't recommend bundling of jobs.

Improving Job Turn-around time

Flexible wall clock time

Flexible wall clock times can improve job turn-around time and utilization. By providing a minimum wall clock time for a job to start as well as a maximum time, the scheduler can start a job early and try to keep the job running by extending the working wall clock time in increments of 30 minutes for as long as the nodes remain available. Wall clock accuracy will be based on the extended wall clock time which is the mintime plus the total of the 30 minute extensions given to the job. The syntax to enable flexible wall clock time in a job is

```
#PBS -l minwclimit=[mintime] -l walltime=[maxtime]
```

Job Preemption

Job preemption is possible in any queue and is no longer default for the low queue as was announced previously. Job preemption will not occur before 4 hours of wall clock time is reached. Preemption is based on relative job priority and can be used in conjunction with flexible wall clock. Job starts when specified walltime is available. The PBS syntax for preemption specification is

```
#PBS -l flags=preemptee
```

Signal handling (COMING SOON)

To complement the use of the above options it is possible to have a signal sent from the scheduler to the application that can be used to alert the application that it will run out of wallclock time or be preempted soon. At the moment this feature has issues and is not available. We are working with Adaptive on this issue.

Blue Waters Data Sharing Service

- Prototype service for data sharing.
- <https://bluewaters.ncsa.illinois.edu/data-sharing>
- Sharing of data sets from on-line or Nearline / projects/sciteam/xyz/share using Globus Online or web services (http).
- Metadata includes contact information, description, website, size and count, DOI*.

2015 Blue Waters Symposium

- Thanks to all who participated.
- 3 audio interviews with Blue Waters Symposium speakers Arden Bement, Satoshi Matsuoka, and Steve Scott are now available online
 - <https://bluwaters.ncsa.illinois.edu/invited-speakers>

NSF CAREER Program

- Faculty Early Career Development Program (CAREER)
 - Awards in support of junior faculty who exemplify the role of teacher-scholars through outstanding research, excellent education and the integration of education and research..
- Webinar 1:00 - 3:00 pm EDT May 26,2015
 - http://www.nsf.gov/events/event_summ.jsp?cntn_id=134940

CSESSP Challenges

- Computational Science & Engineering Software Sustainability and Productivity (CSESSP)
- An inter-agency workshop sponsored by the Networking and Information Technology Research and Development (NITRD) / Software Design and Productivity (SDP) Coordinating Group (CG).
- October 15th-16th 2015, Washington DC, USA
- <https://www.nitrd.gov/csessp>

XSEDE 2015

- St. Louis, MO – July 26-30
- Tutorials
- BOFs
- Technical Program
- Visual Showcase
- Student Participation



Request for Science Successes

- We need to be current on products that result from time on Blue Waters such as:
 - Publications, Preprints (e.g. [arXiv.org](https://arxiv.org) ), Presentations.
 - Very interested in data product sharing.
- Appreciate updates sooner than annual reports.
 - Send to gbauer@illinois.edu
- NSF PRAC teams send information to PoCs.
- See the [Share Results](#) section of the portal as well.
- Be sure to include [proper acknowledgment](#)
 - Blue Waters - National Science Foundation (ACI 1238993)
 - NSF PRAC – OCI award number

Future Topics?

- Please send us your suggestions on topics for future teleconferences / webinars