

BLUE WATERS

SUSTAINED PETASCALE COMPUTING

3/17/19

Blue Waters User Monthly Teleconference



GREAT LAKES CONSORTIUM
FOR PETASCALE COMPUTATION

CRAY®

Agenda

- Moving Data: Nearline Best Practices
- Recent Events and Changes to Blue Waters
- Usage, Utilization and other Items
- Upcoming Opportunities
- Request for publications!

Nearline Use Best Practices

- Nearline will be offline 12/19/2019.
- For projects expiring 3/31/2019
 - Avoid putting new data into Nearline.
- If you need to upload files to Nearline
 - Consider short-term use such as for checkpoints.
 - Bundle or package many, smaller files into fewer, larger files using utilities like tar, or similar.
 - Prefer files in the GB range (1-1000).
- How much and how many in Nearline project space?
<https://bluewaters.ncsa.illinois.edu/usage-project-details?project=<psn>>

Globus Online Use Best Practices

- When moving lots of data off of Nearline
 - Use managed Globus endpoints hosted on robust data transfer nodes (DTN), when possible, such as XSEDE, DOE labs, larger institutes and organizations.
 - Globus Online (GO) retries on endpoint errors
 - Check your remote endpoint quota
 - Check for remote endpoint expiration
 - Too many remote endpoint errors
 - Consider transfer from Nearline to Blue Waters scratch.
 - Then from Blue Waters scratch to remote site.
 - Monitor your transfer activity <https://app.globus.org/activity>

Globus Online Use Best Practices

- When using Globus Connect Personal client
 - Usually on a smaller resource: desktop, laptop, VM.
 - Smaller network “pipe” and slower IO speeds.
 - LAN might not be in good shape.

Globus Python SDK or Python CLI

- Use Python SDK or Python CLI to script or automate transfer processing.
- <https://globus-sdk-python.readthedocs.io/en/stable/>
- <https://docs.globus.org/cli/>

Recent Events and Outages

- 2/23/2019 - High-speed network issue after hardware replacement.
- 3/2/2019 – Fan failed in single cabinet. System remained up.

Upcoming Changes or Outages

- 3/18/2019 – HPSS software update. Improved internal performance

Recent Changes

- No changes to report.

Outstanding Issues

- CUDA 9.1 and GCC 6.3
 - Patch to gcc/6.3.0, as a work-around.
 - Known issue with optimization flag: `-ftree-loop-vectorize`
 - Observed 20% performance impact when disabled.
 - Plan to disable by default.
 - User can re-enable via wrapper flag
`--craype-append-opt=-ftree-loop-vectorize`
 - gcc/5.3.0 works fine but 6.X is recommended for CUDA 9.1.

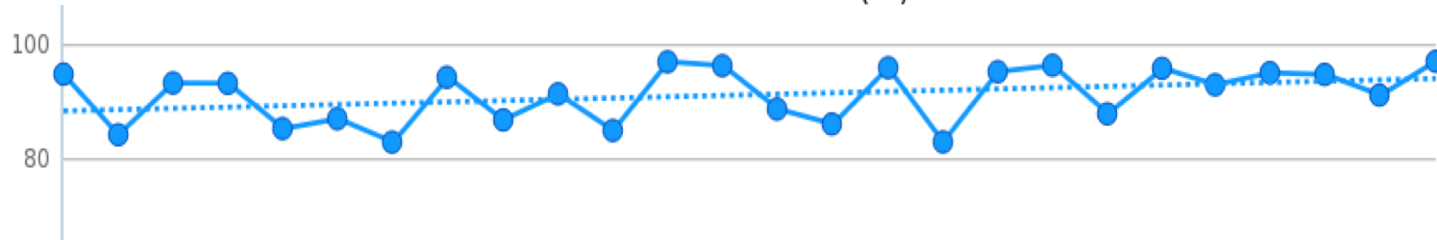
Upcoming Changes

- Preparing to make PE 18-12 default.
 - When available for testing we will send out a blog entry and update <https://bluewaters.ncsa.illinois.edu/pe-updates>.

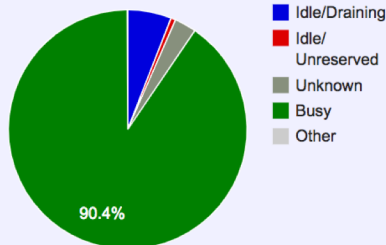
Usage, Utilization and other Items

- Overall Utilization since last user call

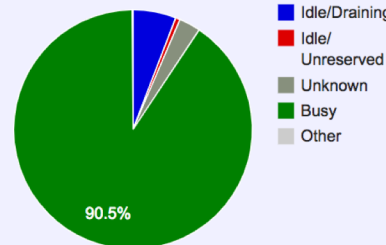
BlueWaters Utilization (%)



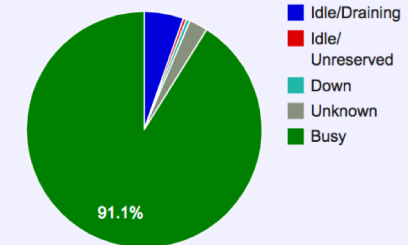
XE Utilization for this Period



Full System Utilization for this Period



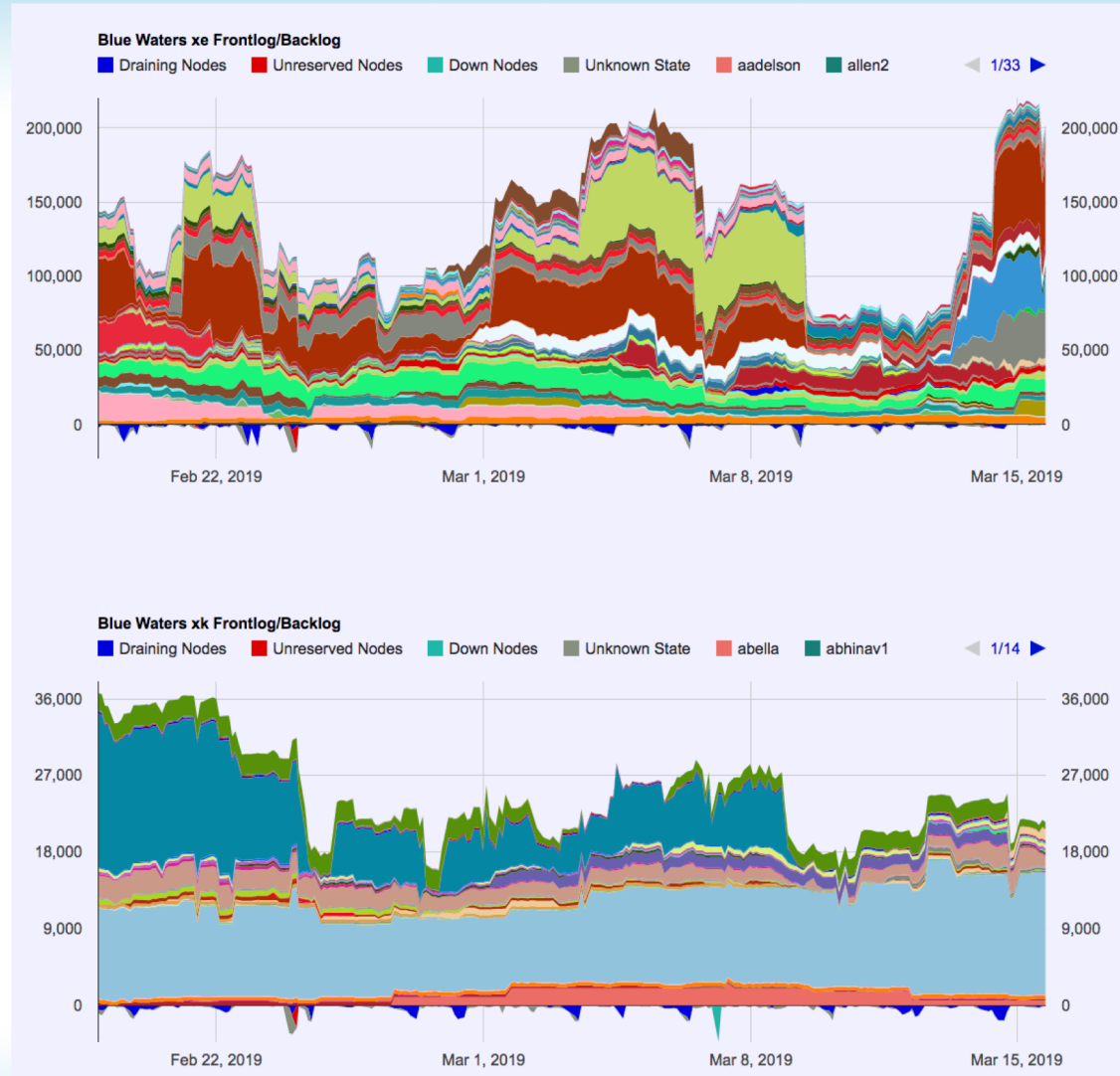
XK Utilization for this Period



● BlueWaters Trend Line: $(0.22x + 88.16)$, R-Squared=0.129281

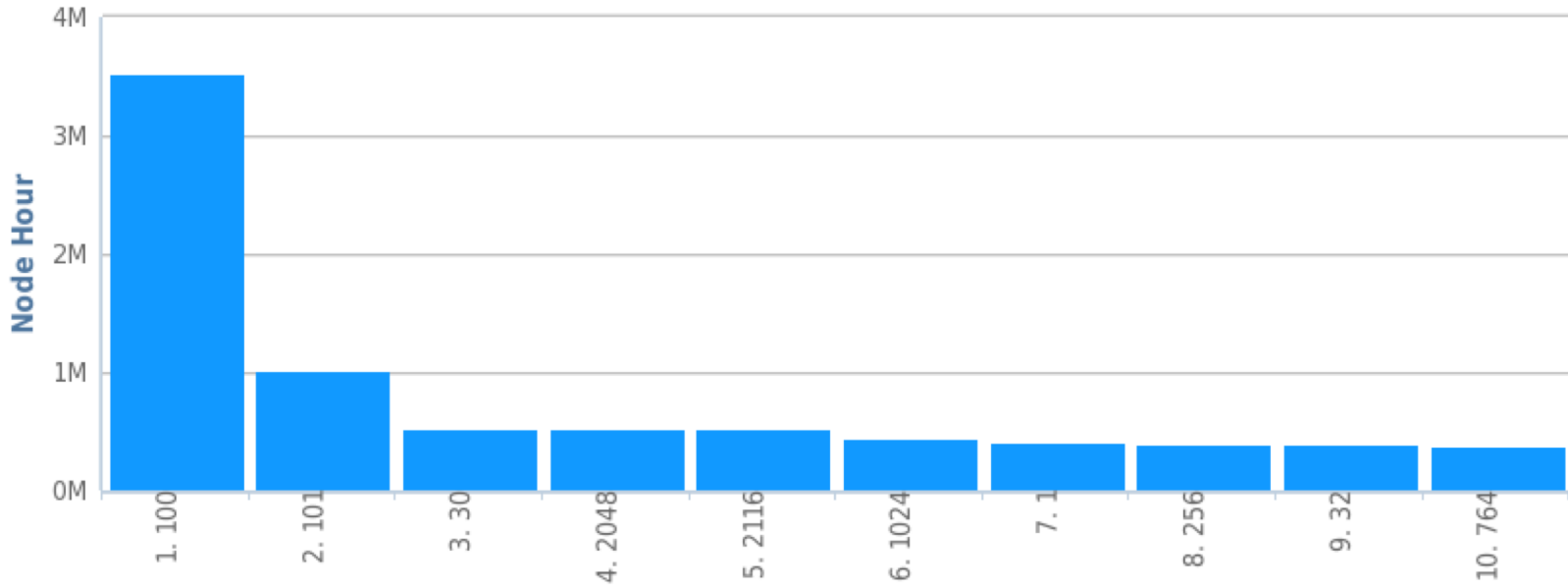
Backlog

- Since last user call.
- Vertical axis in units of **nodes**. Colors are different users.
- Red **below** the x-axis indicates unreserved nodes. Blue **below** the x-axis indicates draining.
- **Increase in backlog.**
- **Many projects expire 3/31.**



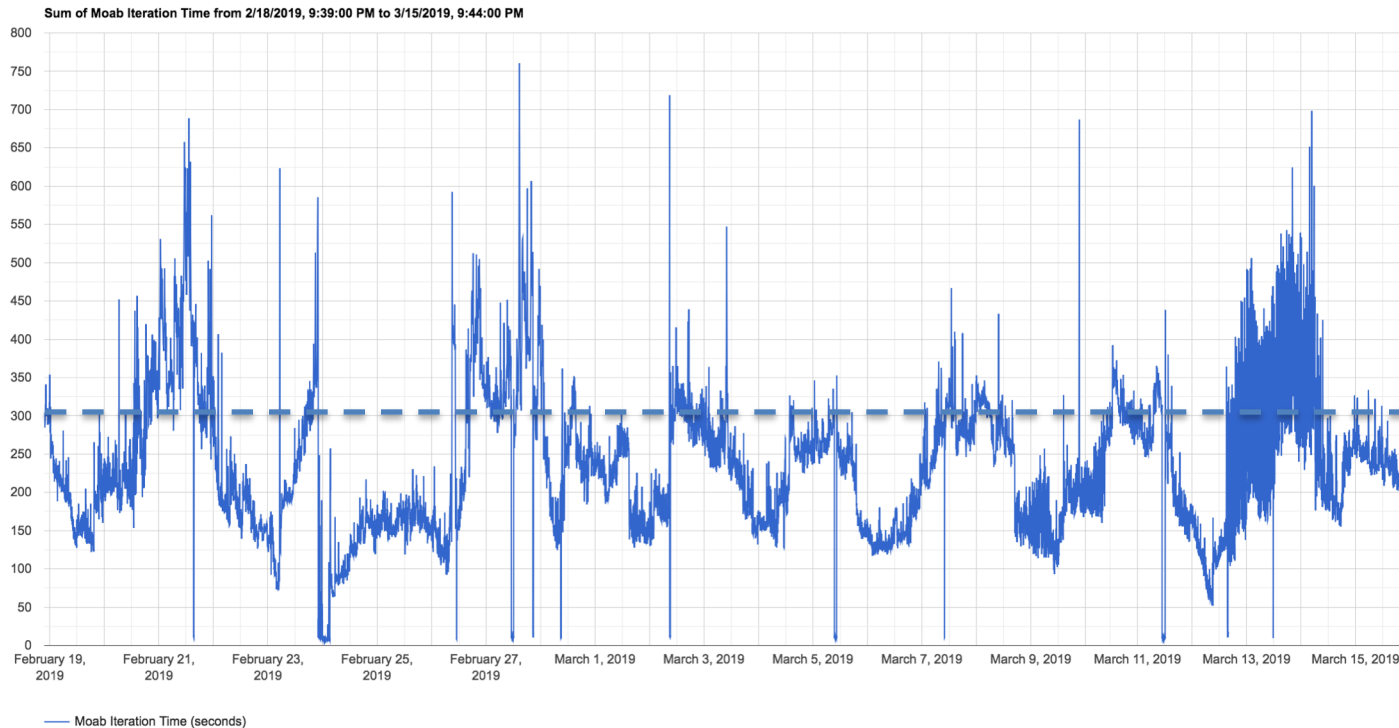
Workload Details

Node Hours: Total: by Node Count



- Data combines XE and XK jobs.
- Since last user call.

Job scheduler iteration time



- Keeping iteration time mostly under 5 minutes.

XSEDE Training Events

- April 2-3 - Big Data
- <https://portal.xsede.org/course-calendar>

Petascale Computing Institute

- August 19-23, 2019
- Looking for host sites.
- Keynote speaker of note.

Blue Waters Symposium

- When: June 3-6, 2019
- Where: Sunriver, OR.
- Serves as the PRAC annual meeting.
- Registration emails sent February 15th.


Blue Waters Weekly Webinar Series

The next scheduled webinar is:

[Reproducible data analysis with Snakemake](#) - Johannes Köster March 20, 2019

- For more information about the webinar series, including registration, abstracts, speakers, as well as links to Youtube recordings, please visit the [Blue Waters webinar series webpage](#).
- Make sure to [RSVP on Facebook](#) and on the [Blue Waters Portal](#)
- We welcome suggestions for topics that will benefit the petascale community. Send your suggestions to bw-eot@ncsa.illinois.edu.

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