BLUE WATERS

SUSTAINED PETASCALE COMPUTING

7/17/17

Blue Waters User Monthly Teleconference


GREAT LAKES CONSORTIUM
FOR PETASCALE COMPUTATION
CRAY

Agenda

- Changes and Recent Events
- Utilization and Usage
- Opportunities
- PUBLICATIONS!

Recent Events and Outages

- 6/29 - Emergency Maintenance
  - Addressed security updates to software
    - Address “stack clash” vulnerability.
    - Non-user systems updated for older exploits.
- 6/20 - two cabinets emergency powered off (EPO)
  - Impacted only jobs using nodes on those cabinets.
  - Suspected cooling cabinet (XDP) issue.

Recent Changes

- Enabled memory resource control on login nodes using cgroups per user totals.
- Resolved issue with signal propagation from scheduler to applications in batch jobs.
  - Take advantage of `preemptee` and `minwclimit` **discounts** with less data loss.
  - Reduce losses if/when job runs into wallclock limit.
- Following slides from https://bluewaters.ncsa.illinois.edu/signal-handling

Signal handling
3+ step process

1. Enable a delay between SIGTERM and SIGKILL emitted by scheduler to job. **`#PBS -K seconds`**
2. Enable signal handling in the job shell script using **`trap`** (not in tcsh).
3. Enable signal handling in application using **`signal(2)`**.

```
1  #PBS -lnodes=2:ppn=2:xe,walltime=00:09:00
2  #PBS -e $PBS_JOBID.err
3  #PBS -o $PBS_JOBID.out
4  #PBS -K 135
5  #PBS -V
6  cd $PBS_O_WORKDIR
7  setenv APRUN_XFER_LIMITS 1
8  # trap SIGTERM in the batch script or it will exit and tear
9  down the job
10 trap '{
11   echo batch_script caught SIGTERM
12 }' TERM
   aprun -n 4 -N 2 ./mpi_signalcatcher
```

```
32      MPI_Init(&argc, &argv);
33      MPI_Comm_rank(MPI_COMM_WORLD, &rank);
34      MPI_Comm_size(MPI_COMM_WORLD, &size);
35
36      MPI_Get_processor_name(name, &len);
37      core= sched_getcpu();
38      printf ("rank %d of %d on %s core %d\n", rank, size,
39  // setup signal handler(s)
40       signal(SIGTERM, mysig);
41       while(1) // simulate main iteration loop
42       {
```

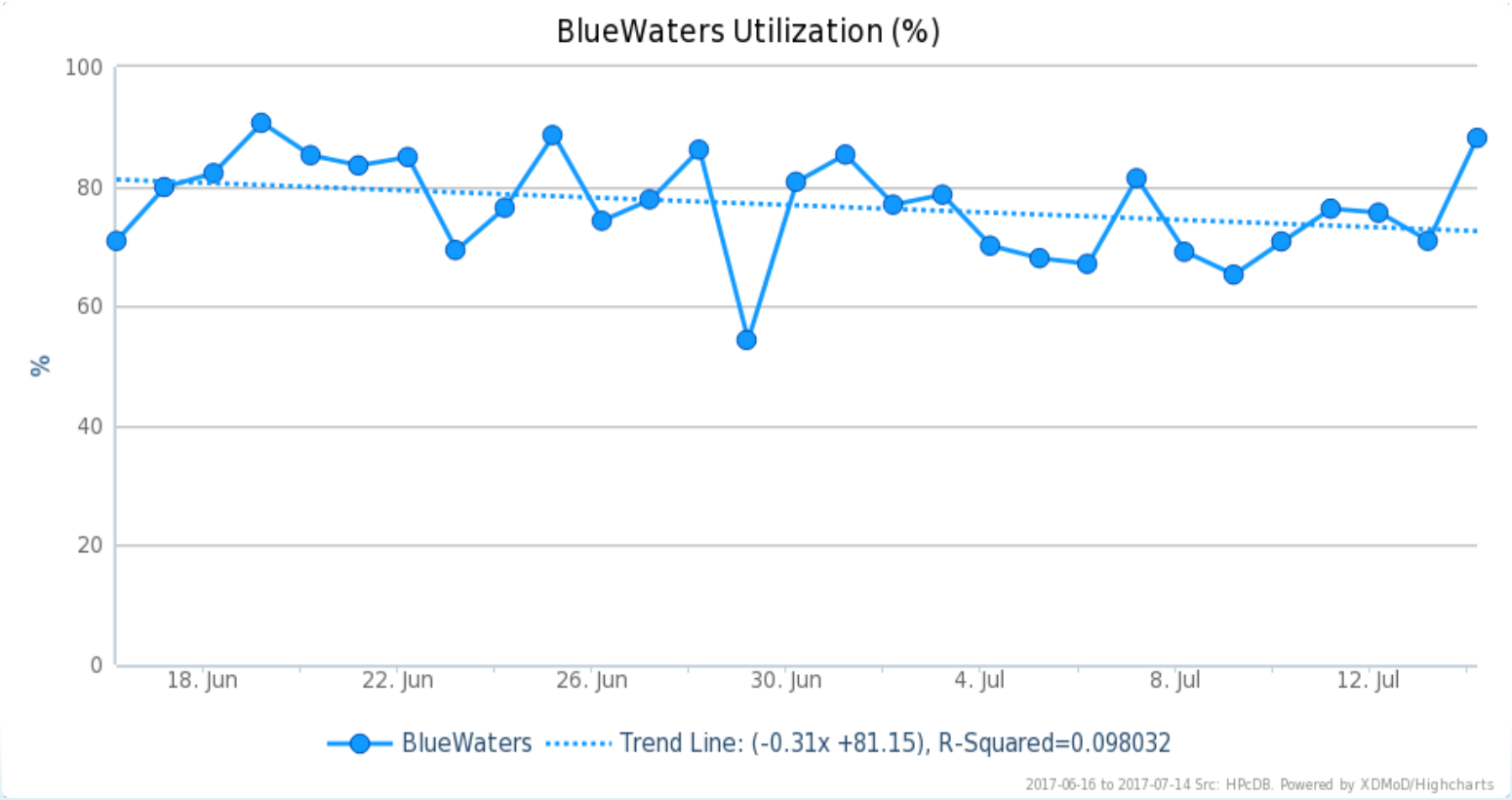
The + step

```
10  // signal handler , execution jumps here on receipt of SIGTERM anytime after signal()
11  has setup this function as the handler
12  void mysig(int sig) {
13   switch (sig)
14   {
15     case SIGTERM:
16      printf("myapp caught SIGTERM\n");
17      // some cleanup routine or checkpoint activity
18      break;
19     default:
20      printf("myapp caught signum %d\n", sig);
21   }
22  }
```

- Need to decide what action to take when getting signal.
- Could set a counter to tell current iteration to checkpoint and resume.
- Could go directly to checkpoint routine if code supports exiting mid-step.

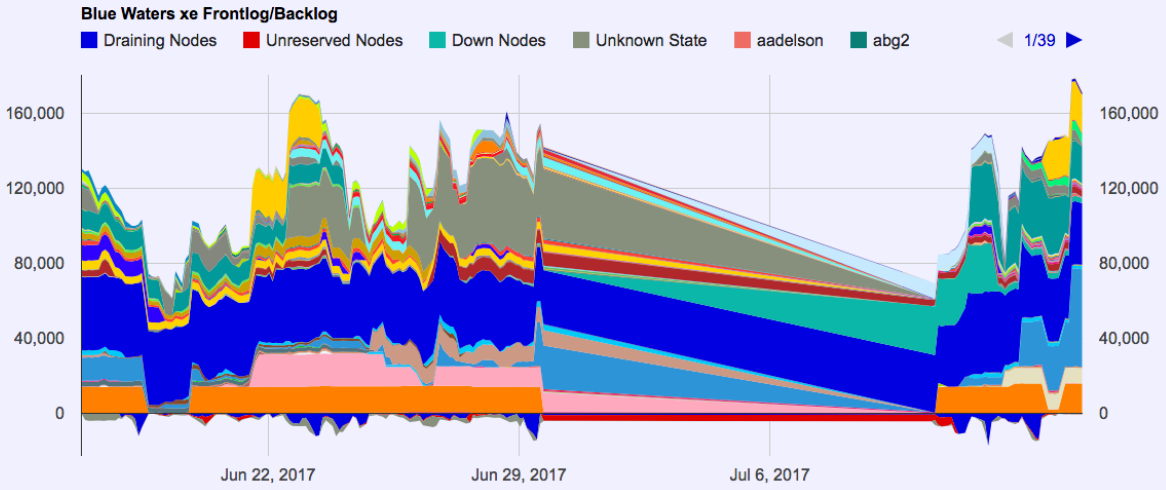
Usage, Utilization and other Items

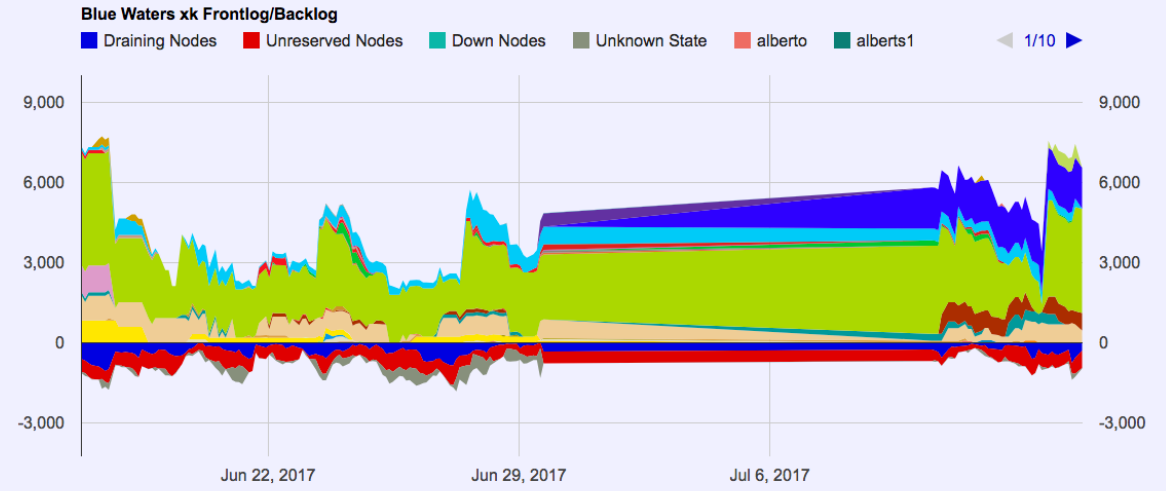
- Utilization since last BW User Call (June 16)



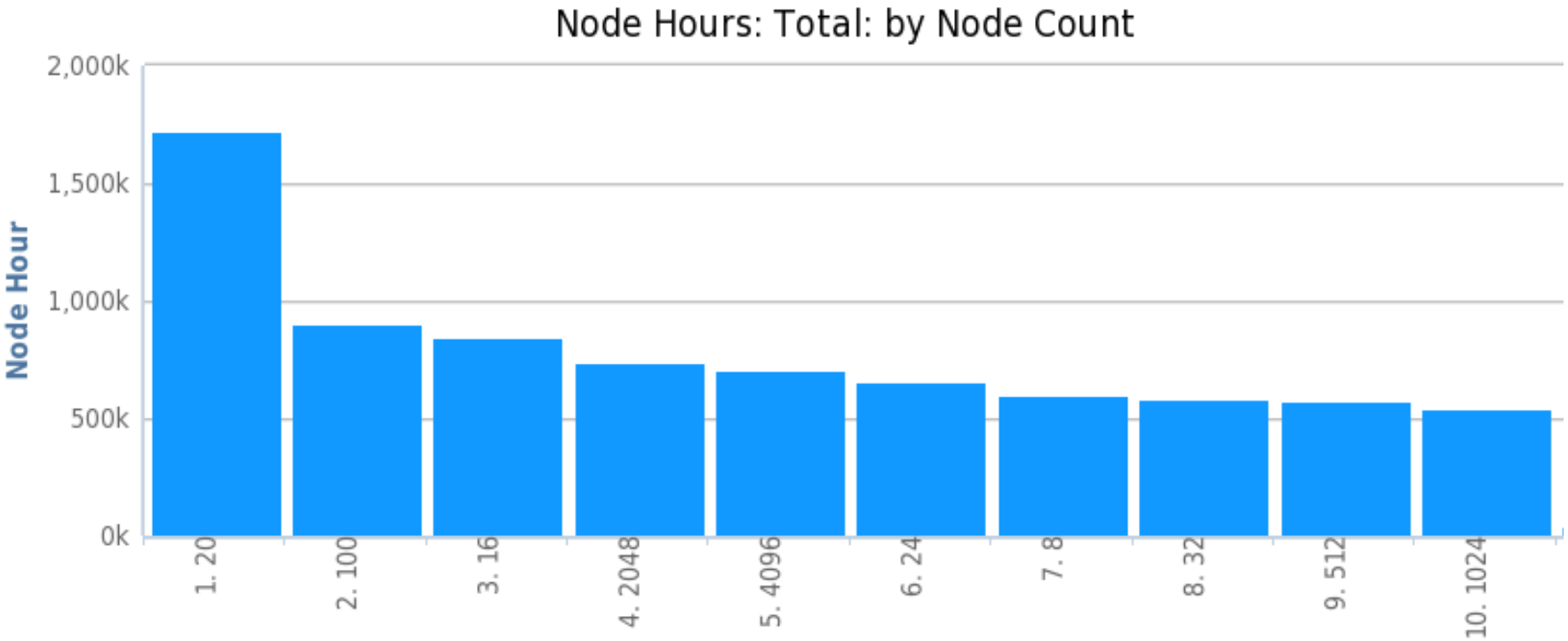
Backlog

- Since BW User Call 06/16
- 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.
- Data feed dropped after emergency maintenance.
- XE region continues with substantial backlog.
- XK region has lower backlog, less potential backfill.



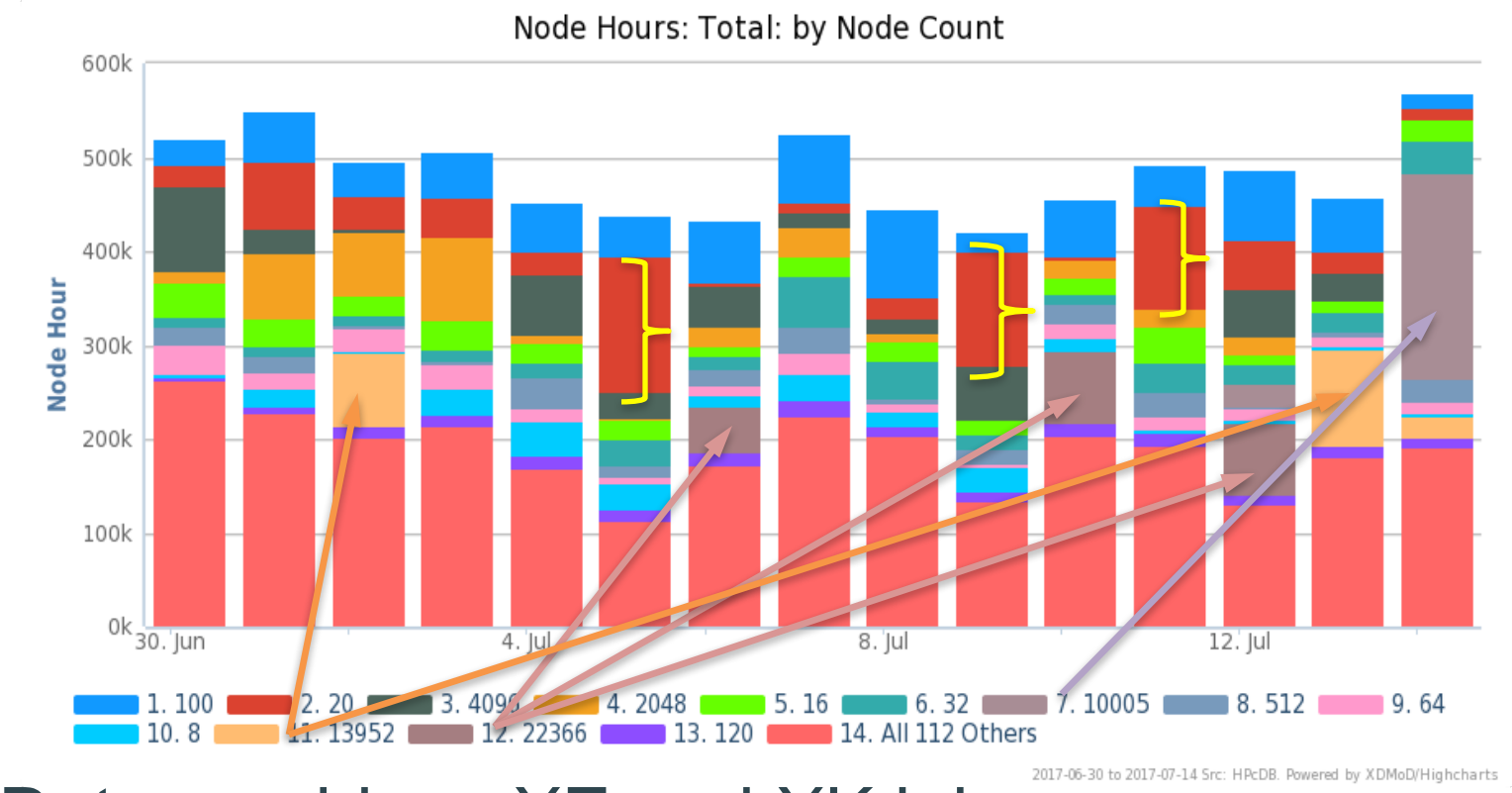


Workload Details



- Data combines XE and XK jobs.
- Since last BW User Call (June 16).

Workload Details of last two weeks



- Data combines XE and XK jobs.
- Last two weeks (6/30 – 7/14)

Blue Waters Weekly Webinar Series

- Started Wednesday, March 1. Webinars will begin at 10 a.m. Central Daylight Time.
- 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.

July 26, 2017	Scientific Visualization in Houdini	Kalina Borkiewicz, A. J. Christensen	--	--	--
August, 2017					
August 2, 2017	Awarding Badges as Incentives for Self-Assessment of Scientific Visualization Expertise	Jeff Sale	--	--	--

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 ), 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