

OLCF Best Practices



Bill Renaud
OLCF User Assistance Group



Overview

- This presentation covers some helpful information for users of OLCF
 - Staying informed
 - Some aspects of system usage that may differ from your past experience
 - Some common errors
 - Common questions/Other tips on using the systems
- This is by no means an all-inclusive presentation
- Feel free to ask questions

Staying Informed

Staying Informed

- OLCF provides multiple layers of user notifications about system status and downtimes
 - OLCF Weekly Update
 - OLCF Status Page
 - Status indicators on olcf.ornl.gov
 - Opt-in email lists
 - Android/iPhone Apps
 - Twitter
- A summary of these items can be found at http://www.olcf.ornl.gov/kb_articles/communications-to-users/

Staying Informed-Weekly Update

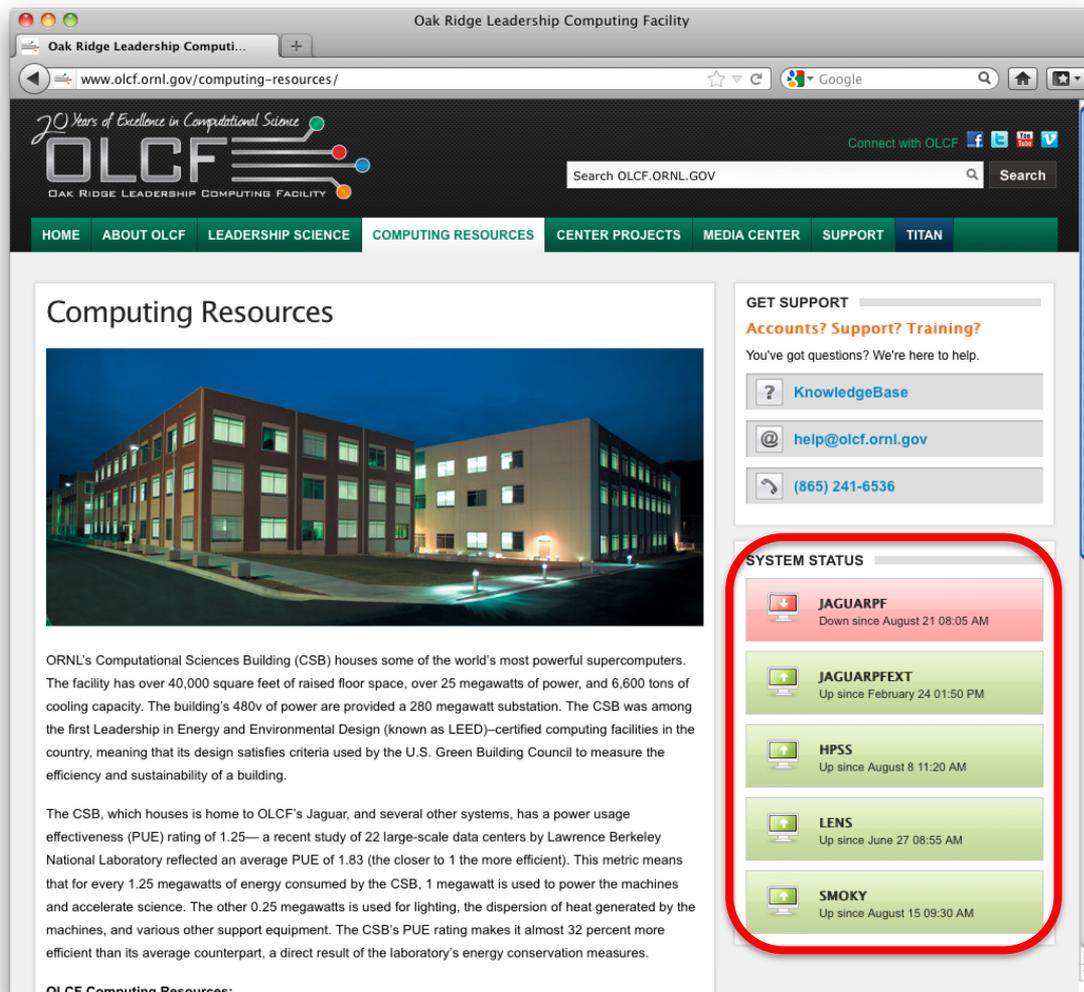
- Sent weekly on Thursday or Friday
- Contains several items
 - Announcements about upcoming training
 - Announcements about upcoming system changes
 - Planned outages for the next week
- **All OLCF users should receive this email**
 - Let us know (help@olcf.ornl.gov) if you're not receiving it!

Staying Informed-System Status

- Automated scripts parse logs from our monitoring software and make an educated guess as to system state
- This status is then sent to multiple destinations: websites, Twitter, smartphone apps, and email lists
- While this is fairly accurate, it is a fully automated process so there is a possibility of both false positives and false negatives.
 - We do take some measures to mitigate this

System Status-Websites

- Computing Resources tab of olcf.ornl.gov



The screenshot shows the OLCF Computing Resources website. The main navigation bar includes links for HOME, ABOUT OLCF, LEADERSHIP SCIENCE, **COMPUTING RESOURCES**, CENTER PROJECTS, MEDIA CENTER, SUPPORT, and TITAN. The page title is "Computing Resources". A large image of the ORNL Computational Sciences Building (CSB) is featured. Below the image, text describes the facility's power and cooling capacity, and its LEED certification. A "GET SUPPORT" section provides links to KnowledgeBase, help@olcf.ornl.gov, and a phone number. A "SYSTEM STATUS" section, highlighted with a red border, lists the status of several supercomputers:

System Name	Status	Up Since
JAGUARPF	Down	August 21 08:05 AM
JAGUARPFEXT	Up	February 24 01:50 PM
HPSS	Up	August 8 11:20 AM
LENS	Up	June 27 08:55 AM
SMOKY	Up	August 15 09:30 AM

System Status-Websites

- Knowledgebase on olcf.ornl.gov

The screenshot shows the Oak Ridge Leadership Computing Facility (OLCF) KnowledgeBase website. The browser address bar displays www.olcf.ornl.gov/support/knowledgebase/. The page features a navigation menu with links for HOME, ABOUT OLCF, LEADERSHIP SCIENCE, COMPUTING RESOURCES, CENTER PROJECTS, MEDIA CENTER, SUPPORT, and TITAN. A search bar is located at the top right, and social media icons for Facebook, Twitter, YouTube, and LinkedIn are visible. A red box highlights a row of system status indicators for jaguarpf, jaguarpfext, hpss, lens, and smoky, each with a status icon and a timestamp. The main content area is titled "KnowledgeBase" and includes a search bar, a "SEARCH" button, and radio buttons for filtering by resource (All Resources, DTN, HPSS, Jaguar, Lens, Sith, Smoky). Below the search bar, there are two columns of article counts: "Browse Articles by Resource" and "Browse Articles by Category".

System	Status	Last Update
jaguarpf	Offline (Red)	Since 8/21/12 08:05 am
jaguarpfext	Online (Green)	Since 2/24/12 01:50 pm
hpss	Online (Green)	Since 8/8/12 11:20 am
lens	Online (Green)	Since 6/27/12 08:55 am
smoky	Online (Green)	Since 8/15/12 09:30 am

Browse Articles by Resource

- DTN (40)
- HPSS (33)
- Jaguar (139)
- Lens (98)
- Sith (69)
- Smoky (78)

Browse Articles by Category

- Accounts (11)
- Allocations (3)
- Compiling (7)
- Connect & Login (10)
- Data Management (29)
- Getting Help (3)
- Modules (2)
- Optimization (2)
- Policies (4)
- Running Jobs (25)
- Software (163)

System Status-Websites

- Status Page on users.nccs.gov/statuspages/summary



The screenshot shows a web browser window displaying the NCCS System Status page. The page title is "NCCS System Status". Below the title, there is a navigation bar with tabs for Jaguar, Lens, Smoky, Frost, HPSS, and Spider. The Jaguar tab is selected and highlighted with a red box. Below the navigation bar, the page is titled "Site Summary" and includes a "Page Last Updated" timestamp: "21 Aug 2012 at 13:25 EDT". The main content area is titled "Current Status" and contains a table with the following data:

System	State	Last State Change	Comments
Jaguar	⚠️	21 Aug 2012 08:05 EDT	Jaguarpf is currently down.
	✅	24 Feb 2012 13:50 EST	Jaguarpf-ext is currently up.
Lens	✅	27 Jun 2012 08:55 EDT	Lens is currently up.
Smoky	✅	15 Aug 2012 09:30 EDT	Smoky is currently up.
Frost	✅	20 Jun 2012 07:00 EDT	Frost is currently up.
HPSS	✅	08 Aug 2012 11:20 EDT	HPSS is currently up.

System Status-Twitter

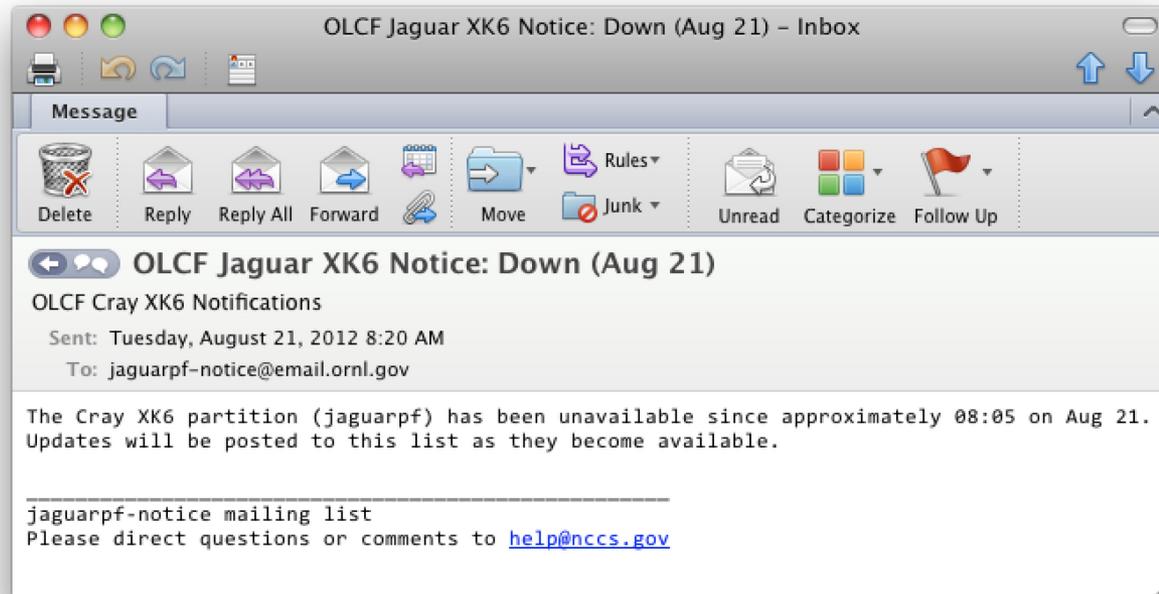
- @OLCFStatus on Twitter

The screenshot shows the Twitter profile for @OLCFStatus. The profile header includes the name "OLCF Status", the handle "@olcfstatus", and the location "Oak Ridge, TN" with a link to "http://olcf.ornl.gov". There are 180 tweets, 0 following, and 8 followers. The "Tweets" section displays a list of system status updates:

Tweet Text	Time
System status change (21 Aug @ 08:05 EDT): jaguarpf is down	6h
System status change (15 Aug @ 09:30 EDT): smoky is up	15 Aug
System status change (15 Aug @ 08:00 EDT): smoky is down	15 Aug
System status change (14 Aug @ 20:35 EDT): jaguarpf is up	14 Aug
System status change (14 Aug @ 15:35 EDT): dtn is up	14 Aug
System status change (14 Aug @ 15:10 EDT): dtn is down	14 Aug
System status change (14 Aug @ 08:05 EDT): jaguarpf is down	14 Aug
System status change (08 Aug @ 11:20 EDT): hpss is up	8 Aug

System Status-Email Lists

- We also send status up/down notices via email
- These are available on an opt-in basis
 - See http://www.olcf.ornl.gov/kb_articles/system-notification-lists/
 - Subscribe only to lists of systems of interest to you
- Other notices are sent to these lists, so you may want to sign up



System Status-Smartphone Apps

- System status apps are available for smartphones
 - Search for “OLCF StatusApp” in Google Play
 - Search for “OLCF System Status” in the iTunes Store
- Choose which systems you monitor
- Automated notifications of system changes
- Usage instruction on olcf.ornl.gov



The left screenshot shows the 'OLCF StatusApp' interface. It displays a list of systems with their status and last update time. The systems listed are jaguarpf, jaguarpfext, lens, smoky, frost, hpss, widow2, and dtn. The status of each system is indicated by a red arrow (down) or a green arrow (up).

System	Status	Last Update
jaguarpf	Down	08:05 on 21 Aug 2012
jaguarpfext	Up	13:50 on 24 Feb 2012
lens	Up	08:55 on 27 Jun 2012
smoky	Up	09:30 on 15 Aug 2012
frost	Up	07:00 on 20 Jun 2012
hpss	Down	15:15 on 21 Aug 2012
widow2	Up	22:25 on 19 Jun 2012
dtn	Up	15:35 on 14 Aug 2012

The right screenshot shows the 'Systems' app interface. It displays a list of systems with their status and last update time. The systems listed are jaguarpf, widow1, widow2, widow3, lens, jaguarext, hpss, and dtn. The status of each system is indicated by a green circle (up) or a red circle (down).

System	Status	Last Update
jaguarpf	Up	5:10 PM on Oct 10, 2011
widow1	Up	4:50 PM on Sep 28, 2011
widow2	Up	4:50 PM on Sep 28, 2011
widow3	Up	4:50 PM on Sep 28, 2011
lens	Up	4:50 PM on Sep 28, 2011
jaguarext	Up	4:50 PM on Sep 28, 2011
hpss	Up	1:00 PM on Sep 30, 2011
dtn	Up	4:50 PM on Sep 28, 2011

Both screenshots show the time as 3:48 PM and the carrier as AT&T 3G. The right screenshot also shows a notification that the app was updated on 10/11/11 at 3:45 PM.

Using the Systems at OLCF

- Software
- Compiling
- Common Error Messages
- Common Questions

Finding Software

- Some software is part of the default environment
 - Basic commands
 - Text editing utilities
- Larger packages are typically managed via the ‘modules’ utility
 - Software is actually installed in /sw
 - To list available software, use “module avail”
 - To use a package, use “module load”
 - More information is available on the OLCF website
- “Important” items, such as compilers, are also available via modules

Software Installation/Updating

- We are moving to a model of updates to software packages at certain intervals (or in the case of major revisions)
 - This means not all minor versions will be installed
 - We'll move towards adding build instructions on the website so that you can build minor revisions/slightly different versions
- Look for information on the OLCF website and via Weekly Update emails

Software Installation

- You are free to install software in your directories (including your project directory)
 - Subject to terms of license agreements, export control laws, etc.
- If you think a piece of software would be of general interest, you might ask us to install it for general use
 - Preferred method:
<http://www.olcf.ornl.gov/support/software/software-request/>, but email to help@olcf.ornl.gov works, too.
 - This will be reviewed by our software council

Compiling At OLCF

- The compilers on the XT/XE/XK line of systems may differ (significantly) from your previous experience
- Combination of `xt-asyncpe` and `PrgEnv-?` modules
 - `xt-asyncpe` provides compiler wrapper scripts
 - `PrgEnv-?` loads modules for back-end compilers, math libraries, MPI, etc.
- Regardless of actual compiler being used (PGI, Intel, GNU), invoke with `cc`, `CC`, or `ftn`
- MPI, math, and scientific libraries included automatically
 - No `-lmpi`, `-lscalapack`, etc.
 - This can be challenging when dealing with some build processes

Compiling at OLCF

- You are actually cross-compiling...processors (& instruction sets) differ between login and compute nodes
 - *It is very important to realize this...utilities like “configure” often depend on being run on the target architecture, so they can be challenging to use on the XK6*
- Compiling for login/batch nodes is occasionally necessary
- There are three ways to do this
 - module swap xtpe-interlagos xtpe-target-native
 - Add `–target=native` to `cc/CC/ftn`
 - Call the compilers directly (e.g. `pgcc`, `pgf90`, `ifort`, `gcc`)

Common Runtime Errors

- Illegal Instruction
 - A code was compiled for the compute nodes but executed on login nodes
- request exceeds max nodes alloc
 - The number of cores required to satisfy the aprun command exceeds the number requested
 - Also happens when your request is correct, but at launch time a node is discovered to be down

Common Runtime Errors

- relocation truncated to fit: R_X86_64_PC32
 - The static memory used by your code exceeds what's allowed by the memory model you're using
 - Only the “small” memory model is available (static size \geq 2GB)
 - Solution: use dynamic memory allocation to the greatest extent possible

Common Questions

- *Is my data backed up?*

- NFS directories: Yes, to an extent. Take a look at `/ccs/home/.snapshot/`

```
$ ls /ccs/home/.snapshot  
hourly.0 hourly.1 hourly.2 hourly.3 hourly.4 hourly.5 nccsfiler3(0151729160)_home.1  
nightly.0 nightly.1
```

- Lustre directories: No
- HPSS: No. While you might use it as a backup of your directories, HPSS itself is not backed up. If possible, it's a good idea to have another level of backup at some other site.

Common Questions

- *What project am I on, and what's its allocation?*
 - Use showproj to list your projects
 - Use showusage to display utilization
 - Both commands have a “help” option...run them with -h for usage info

```
$ showproj
```

```
brenaud is a member of the following project(s) on jaguarpf:  
stf007
```

```
$ showusage
```

```
jaguar usage in CPU hours:
```

Project	Allocation	Project Totals Usage	brenaud Remaining	Usage
stf007	600001	562227.60	37773.40	12968.42
stf007de1	500000	0.00	500000.00	0.00

Common Questions

- *What happens when my project overruns its allocation?*
 - Most importantly, we do **not** disable the project...jobs simply run at lower priority
 - If slightly over allocation (100-125%), jobs have a 30-day priority reduction
 - If well over (>125%), jobs have a 365-day priority reduction
 - This allows a degree of “fairshare” while still allowing people to run when the system is quiescent
- *My project has lost X hours due to system issues...can I get that time reimbursed?*
 - Since we don't disable projects for going over allocation, we also don't deal with refunds *per se*
 - If many jobs are affected, the priority reduction can be delayed. This is basically a refund but is much easier to manage.

Common Questions

- *I changed permissions on /tmp/work/\$USER, but they changed back...why?*
 - Permissions in the lustre filesystem are controlled by settings in our accounts database
 - These settings only affect the top-level permission
 - Permissions are automatically (re-)set regularly
 - Most users can request they be changed
 - Send email to help@olcf.ornl.gov
 - Note that you need to email us to change them “back”
 - Of course, you can always just `chmod` everything under the top-level directory
 - We can't change permissions on directories associated with sensitive data

Important Support Systems at OLCF

- HPSS
 - Mass storage system
 - Accessed via hsi & htar
- dtn01/dtn02
 - Data Transfer Nodes
 - Preferred system for handling data transfer
- <http://www.olcf.ornl.gov>
 - Technical info, user guides, knowledgebase, known issues, forms, etc.
- <https://users.nccs.gov>
 - Project information, usage, etc.

Data Storage Practices

- HPSS is the proper location for long-term storage
- Project areas (NFS and lustre) offer a common area for shared data files, executables, but should not be considered long-term storage
 - Need to keep an eye on disk usage
 - Should still be backed up
- User scratch areas are intended for use during computations
 - Regularly purged
 - Store files to HPSS as soon as practicable
 - File cleanup is important

Dealing With the Scratch Purge-Conditional Transfers

- Many codes use files from previous iterations of the code
- Sometimes, needed files can be deleted by the scratch purge
- This can present challenges:
 - Pulling from HPSS every time is inefficient
 - Multiple scripts (one that assumes data is there, one that transfers data) are cumbersome
 - Using touch to preserve a file when you won't really need it for weeks isn't ideal
- Conditional transfers help with this (i.e. check for file's existence and transfer only if it's not there)

Conditional Transfer

```
#!/bin/bash
...
if [[ ! -a /tmp/work/brenaud/some_important_file ]];
then
  hsi -q get /home/brenaud/data/some_important_file
fi

aprun -n 4096 ./a.out
...
```

Interacting with HPSS

- HPSS is a somewhat complex system
- HPSS prefers a small number of large files and not a large number of small files-htar is your friend in this regard
 - htar is (much) faster than a tar followed by hsi put
 - Limited disk space is no problem...data is streamed directly to HPSS so there is no “intermediate” local storage
- Running many transfers at a time can be problematic
 - Multiple transfers may not give you parallelism
 - Limiting the number of per-user transfers helps the system operate more efficiently (& therefore can be more efficient for you)
- Usage examples are on the OLCF web site

Running Jobs at OLCF

- Batch job information is available on the OLCF Web Site
- Due to our designation as a “leadership-class” facility, queuing policy heavily favors large jobs
- Special requests for temporary high priority/quick turnaround are considered
 - Don’t wait on an answer to submit your job...many times jobs start more quickly than expected
 - Allow plenty of lead time when making a request...discussion may be necessary prior to a decision on approval

Running Jobs at OLCF

- From a user perspective, titan has three major parts
 - The system proper
 - External login nodes
 - MOAB server
- Often, only the system proper is affected by outages
 - External login nodes and the MOAB server node remain up
 - This means you can compile/submit jobs/etc while titan is down
 - Jobs will be queued and will run when the system returns

Debugging/Optimization at OLCF

- Several software tools are provided for debugging and optimizing your applications
 - DDT
 - Vampir
 - CrayPAT
- Information on these tools is available on the web; you can also contact the OLCF User Assistance Center if you have questions

Support Best Practices

- Send as many error messages as possible
 - Or, place them all in a file and direct us to it
- When sending code, create a .tar file & tell us where it is
 - More efficient than sending through email
- When possible, reduce error to a small reproducer code
 - We can assist with this
 - If the error has to go to the vendor, they'll want this

Finally...

- We're here to help you
- Questions/comments/etc. can be sent to the OLCF User Assistance Center
 - 9AM – 5PM Eastern, Monday-Friday exclusive of ORNL holidays
 - help@olcf.ornl.gov
 - (865) 241-6536