

A Possible OLCF Operational Model (2019+)

HPC cross-experiment discussion
CERN, May 10th 2019

Jack C. Wells, Valentine Anantharaj
National Center for Computational Sciences, Oak Ridge National Laboratory

Shantenu Jha, Alexei Klimentov
Brookhaven National Laboratory

Kaushik De
University Texas at Arlington

ORNL is managed by UT-Battelle, LLC for the US Department of Energy

Highly Competitive Process. Summit User Program Update – May 2019

1. Early Science Program (ESP) on Summit

- 25 proposals have been awarded time and work began January 2019
- Early Science Program terminates at the end of June.

2. INCITE Program on Summit

- 64 INCITE proposals requested Summit resources; 30 proposals were accepted
- 9 INCITE proposals reviewed by new “learning panel”; 2 of these were awarded projects
- 31 INCITE projects have been awarded time and work began January 2019

3. Nine 2019 ACM Gordon Bell nominee submissions from work on Summit

- Diverse topics spanning modeling & simulation, data analytics, and AI

4. 2019-2020 ALCC program on Summit will begin by 1 July 2019.

5. 2020 INCITE proposal call issued 15 April, closes 21 June, 2019

- <http://www.doeleadershipcomputing.org/proposal/call-for-proposals/>

A Possible Operational Scenarios on Summit at OLCF?

Option A: Create a new center-wide service/queue to which multiple user projects may submit jobs.

- Support by OLCF, in collaboration with ATLAS/PanDA team.
 - We need a conversation about details.
- What are the advantages of having access to a pool of tasks from multiple user projects external to Summit's queue from which one could proactively backfill Summit?
- Is there an execution strategy that would benefit from access to a pool of "backfill tasks"?

We are not ready to move straight to option A.

Option B: Support user projects from ATLAS and other science projects in using their WLMS of choice

- Kubernetes/OpenShift container orchestration ("Slate" service) is available, but still in "pilot" development.
- Each project would be responsible for deploying WLMS/WFMS middleware upon Slate.
- Enables access to wide-area, distributed task management and proactive backfill of Summit's queues (as demonstrated by BigPanDA project @ Titan.
- Normal queue policies apply
 - Queue policy special requests can be considered.

We should begin with option B on Summit.

Option C: Support user projects from LHC using PanDA or/and other science projects in using their WLMS of choice

- PanDA instance @ORNL
- Kubernetes/OpenShift container orchestration ("Slate" service) is available, but still in "pilot" development.
- Support by OLCF in collaboration with ATLAS/PanDA team an instance (including harvester and NGE)
 - Projects will have a choice use PanDA or be responsible for deploying WLMS/WFMS middleware upon Slate.
- Enables access to wide-area, distributed task management and proactive backfill of Summit's queues (as demonstrated by BigPanDA project @ Titan.
- Normal queue policies apply
 - Queue policy special requests can be considered.

LHC Community input is needed

Considerations for implementing Option B:

- The implementation and deployment will be facilitated by OLCF, in collaboration with ATLAS/PanDA team.
 - Identify individuals who will develop an implementation strategy.
 - Contribute to the knowledge base by documenting the experience.
 - Develop and document a *recipe* for deploying the essential services.
 - Harden the process by enlisting friendly users to test the *recipe* for a set of use cases.
 - How do we make it as easy as possible for diverse user projects?
 - The KN8/Openshift platform at OLCF is still maturing.
 - How do we develop an automated test suite?
 - Identify risks and mitigation strategies.
 - We are ready to work on implementation ‘today’