Fermilab

HEPCloud Project

Onboarding Experiments to HPC and Cloud Resources

Version 0.1

12/21/2018

PREPARED BY:

Anthony Tiradani

# Table of Contents

1. Purpose

The Fermilab HEPCloud Facility provides new capabilities to obtain resources on behalf of experiments and research groups. These resources are provided by a variety of sources using different types of currency to obtain access. Commercial clouds use dollars as their currency while HPC centers use allocations. In both cases, the Facility, in partnership with the experiments and research groups must establish procedures and guidelines to ensure proper, efficient, and impactful use of these resources.
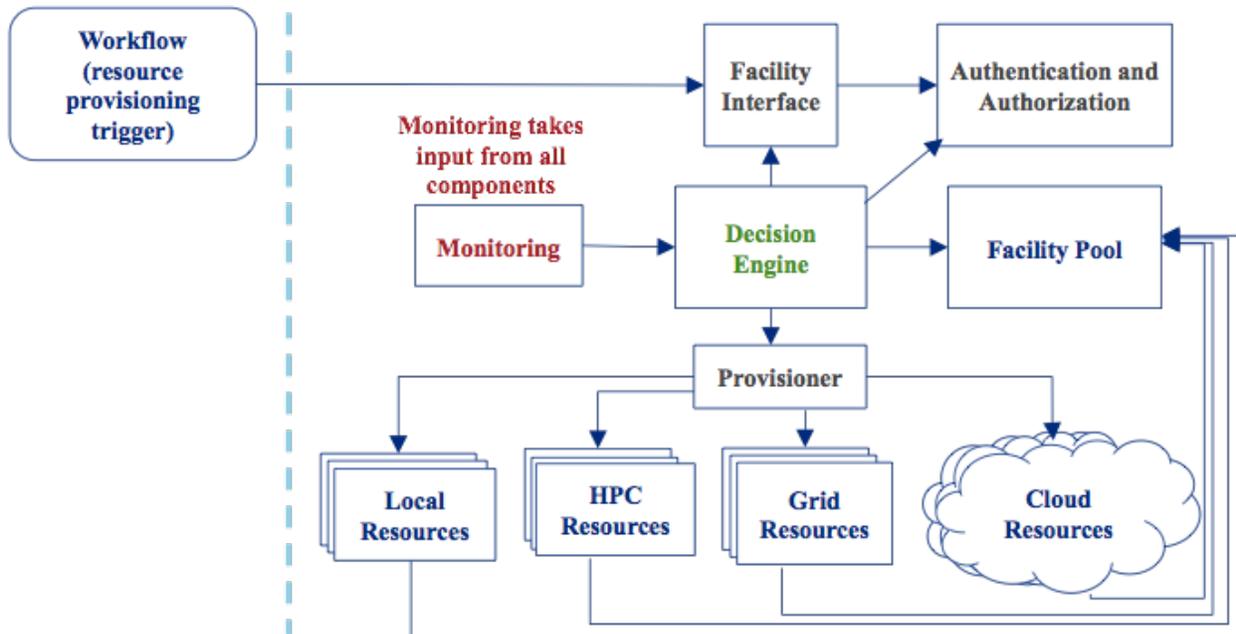
2. Intended Audience

The target audiences for this document are experiments, experiment support teams, facility support teams, and Computing Division management.

3. Terminology

All HEPCloud terminology and acronym definitions can be found in the HEPCloud Acronym Glossary List (http://cd-docdb.fnal.gov/cgi-bin/ShowDocument?docid=6592).

4. Scope

The HEPCloud Facility makes the distinction between "Analysis" jobs and "Production" jobs. Analysis jobs will be directed to local or OSG resources, presenting no change from what these users are accustomed to. Production jobs are jobs that are centrally controlled by an experiment, have a limited number of authorized submitters, have well understood behaviors, and are monitored by the experiment. Usually Production jobs are submitted with a "Production Role". However, it is understood

that certain corner cases exist. Those cases will still be considered "Production". Production jobs are the only flavor of jobs eligible to be directed to resources procured through budgets and/or allocations. Production jobs will also continue to be directed to OSG and local resources as appropriate.

This document outlines the expectations placed on the Facility, FIFE Support, and FIFE Experiments to be on-boarded for the new capabilities added by the HEPCloud Facility. Additionally, there is a short discussion about site support with regards to HPC and cloud sites.

The intended audience for this document is experiments that are officially supported by FIFE Support and use Jobsub.

5. Description of New Capabilities

5.1   HPC

NERSC is the first DOE Leadership Class Facility (LCF) that the HEPCloud Facility is targeting for HPC resources. The NERSC Facility is based on the Intel® Haswell™ and the Intel® Xeon Phi™ Knights Landing (KNL) processors. The primary use of the NERSC facility is parallel processing workflows. KNL provides new vectorization capabilities for code that is designed to use the new instruction sets.

The HEPCloud Facility interfaces with NERSC to provide an HTCondor overlay that reports back as part of the FermiGrid pool. This allows standard production workflows to access these resources. Unless code is optimized for the KNL architecture, KNL processors tend to require more wallclock time than Xeon processors to accomplish the same tasks.

5.2   Cloud

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. Users pay only for capacity actually used with no long-term commitment or up-front cost. Amazon EC2 enables you to increase or decrease capacity within minutes, not hours or days. Users can commission one, hundreds or even thousands of server instances simultaneously via web service APIs, enabling automatic scaling.

The Google cloud platform enables customers to resize clusters, create machine images, virtualize the network, use Preemptible VMs for batch workloads, and create Custom Machine Types to optimize for specific needs. Network storage, up to 64 TB in size, can be attached to VMs as persistent disks in HDD or SSD formats. Global load-balancing technology helps customers distribute incoming requests across pools of instances across multiple regions.

Commercial cloud providers operate at scale and constantly refresh their hardware with the latest technologies including GPUs, FPGAs, and quantum computing resources. The HEPCloud Facility provides a standard interface with which to request these types of resources. The first iteration of the HEPCloud Facility focuses on standard computing resources. Future enhancements will include more specialized capabilities.

Each Cloud provider has their own storage solutions available for use. In Amazon it is called S3; in Google, it is called Cloud Storage. These services are offered at a price per unit of storage and also charge per write/read. Therefore, use of Cloud storage services requires careful consideration of amount of storage used, access patterns, and length of expected storage.

6. Facility Expectations

The HEPCloud Facility implements aggregate monitoring of costs and allocation usage to manage the overall facility budget(s) and allocation(s). This includes metrics like, "how much money has the Facility spent in the last hour?", or "how much allocation has the Facility used in the last hour?". This is in addition to tracking the overall budget or allocation. HEPCloud Facility Tier-3 support team will scale back resource requests if budgets or allocations are being used too quickly, or are close to exhaustion. If the usage rates are too severe, HEPCloud Facility Tier-3 support team may start killing jobs that are running on HPC or cloud providers to preserve the budget or allocation at its discretion.

The HEPCloud Facility will handle resource provisioning for jobs via resource provisioning requests based on the requirements and descriptions specified in the job manifests. The Decision Engine (DE) is the component of the HEPCloud Facility that contains the "intelligence" to figure out the optimal combination of resources for the jobs in the system based on these provisioning requests.

The HEPCloud Facility makes use of the Landscape service to provide monitoring of resources, resource usage, and other Facility metrics. Some of these metrics are publicly available and some are restricted to authorized personnel. For example, all financials are restricted, while statistics concerning job operations are public.

The HEPCloud Facility personnel will maintain documentation that explains what characteristics make workflows or jobs suitable for HPC or Cloud providers. As additional resource providers are added, this documentation will be updated.

HPC resources tend to require containers to provide the runtime environment for jobs. The HEPCloud Facility will maintain a general-purpose container for experiments to run their jobs within. The HEPCloud Facility Tier-3 support team will work with experiments that have requirements which differ from the standard container in order to deploy a custom container.

7. Onboarded Experiment Expectations

Onboarded Experiments will be expected to describe their workflows at submission time via a job manifest file to inform the Decision Engine on how to request resources. These parameters are detailed in the Job To Resource Mapping spreadsheet located on the HEPCloud Fermipoint site.

Onboarded Experiments must work with the FIFE Support team to run test workflows on HPC resources and commercial clouds. These workflows should be representative of actual work that will be directed to these resources. These workflows will be used to ensure that the job execution environment contains all necessary information, and that all required software is available to the jobs. Benchmarks should be taken to ensure that expected efficiencies are met, and that data input and output are within

expected boundaries. Testing must be completed successfully before a given workflow will be allowed to run at scale on HPC and cloud resources.

In addition to the test workflows FIFE Experiments must provide small, short running workflows with known outputs to act as unit tests for the facility. These workflows will be run periodically on HPC resources and cloud providers to ensure continued functionality at the remote facilities.

Onboarded Experiments must monitor their workflows executing on remote resources for behaviors that deviate from the expectations outlined in the job manifest. Careful monitoring of Onboarded Experiment workflows is especially critical for workflows running on commercial cloud resources where misbehaving jobs can burn through the cloud budget very quickly.

The following table contains a sample of the metrics to monitor from a workflow.

Things to Monitor	Description
Data transfer out to storage	Experiment will be cut off from cloud resources if experiment exceeds the data transfer threshold. The experiment will be supplied with a formula to determine what that threshold is.
Job Failure Rates	More than 10% failure rate on a resource – Experiment will be cut off from that resource until the experiment fixes issue(s)
Execution Time	If jobs are running much longer than anticipated, or much shorter than anticipated (as defined by the experiment inserted values in the job manifest) on a particular resource, the experiment may be cut off from the resource

8. FIFE Expectations

Primary monitoring and debugging responsibility rests with experiments. FIFE support will assist the experiments in debugging activities when experiments cannot determine the cause of issues.

FIFE support will be expected to be the interface between Onboarded Experiments and the HEPCloud Facility. It is expected that the Onboarded Experiments contact FIFE Support for all questions, support requests, and debugging/troubleshooting requests.

FIFE Support will assist the Onboarded Experiments in their testing and benchmarking activities described in section 6.

FIFE Support will provide debugging and troubleshooting assistance to the Onboarded Experiments. FIFE Support will work with HEPCloud Tier-3 on these debugging and troubleshooting activities. FIFE Support will work with the HEPCloud Tier-3 Support to develop tools, processes, and policies that enable efficient communication and support for Onboarded Experiments.

9. Support Models

Onboarded Experiments are supported by the FIFE Support team. The HEPCloud Facility Tier-3 support team will assist the FIFE Support team in debugging activities, including providing access to logs and/or machines when and where necessary. The HEPCloud Facility Tier-3 support team will work with the FIFE Support team to develop tools, processes, and policies that enable efficient communication and support for Onboarded Experiments.

CMS has its own support teams, including the HEPCloud Facility Tier-3 support team which will assist the CMS Production team(s) in debugging activities. The HEPCloud Facility Tier-3 support team will work with the CMS Production team(s) to develop tools, processes, and policies that enable efficient communication and support for CMS.

9.1   LCF/HPC Support

In many ways, NERSC is presented as a very large OSG site. Other than computing capabilities, the main difference between an OSG site and NERSC is that NERSC provides allocations to access their site, whereas OSG sites are purely opportunistic. This presents a unique challenge for supporting NERSC and any other HPC site that provides access via allocations.

Fermilab obtains the aggregate allocation(s) for the Onboarded Experiments as well as an aggregate allocation for CMS. The HEPCloud Facility manages both sets of allocation.

It is understood that experiments may obtain their own allocations at various HPC sites, and may wish HEPCloud to manage the allocation(s) on their behalf. This is not supported at this time. This document will be updated when the support model is understood and agreed upon.

Onboarded Experiments may submit requests to SCMPT to get an initial share of the Fermilab HPC allocation for Onboarded Experiments. Additional requests during the course of the year may be submitted to SPPM for approval. The HEPCloud Facility Tier-3 support team will implement the decisions made by SCPMT and SPPM. Total usage measured against the managed allocations will be a factor in determining whether to send jobs to an HPC facility.

Although Fermilab obtains separate allocation(s) on behalf of CMS, CMS may still be requested to present their needs for the year at SCPMT.

The HEPCloud Facility Tier-3 support team will work with the HPC site support team(s) to troubleshoot and resolve issues with the site, and HEPCloud services. Additionally, the HEPCloud Facility Tier-3 support team will respond to issues raised by the HPC site support team(s).

The HEPCloud Facility provides a standard supported container for use on HPC systems.

9.2   Cloud Support

Cloud Providers rent resources and infrastructure services. It is up to the customer to configure the cloud services in a secure manner, then stand up any additional services needed to run application(s) in the Cloud Provider's infrastructure. The HEPCloud Facility manages the Cloud Provider configurations

and stands up additional services such as Frontier or Squid services, for expansion of the HEPCloud Facility into the Cloud. These resources cost money, and usage is carefully monitored to ensure the cost does not exceed the budget.

Fermilab provides a budget for obtaining cloud resources for FIFE Experiments and the HEPCloud Facility manages the FIFE Experiment's budget. CMS provides a budget that is managed by the HEPCloud Facility for the work CMS wishes to run in the cloud.

It is understood that experiments may obtain their own grants and/or budgets for cloud resources and may wish HEPCloud to manage the budget(s) on their behalf. This is not supported at this time. This document will be updated when the support model is understood and agreed upon.

FIFE Experiments may submit requests to SCPMT to get an initial share of the Fermilab cloud budget for FIFE experiments. Additional requests during the course of the year may be submitted to SPPM for approval. The HEPCloud Facility Tier-3 support team will then implement the decisions made by SCPMT and SPPM. Usage against the managed budgets will be a factor in determining whether to send jobs to cloud resources.

Fermilab obtains separate budgets on behalf of CMS. CMS may be requested to present their needs for the year at SCPMT.

The HEPCloud Facility Tier-3 support team will work with the Cloud Provider(s) to troubleshoot and resolve issues with the site and HEPCloud services. Additionally, the HEPCloud Facility Tier-3 support team will respond to issues raised by the Cloud Provider(s).