HTCondor Uses in Workspaces

Introduction

HTCondor in Workspaces

HTCondor is an integral part of the Workspaces system design and future. By leveraging HTCondor's features we are able to run workflow jobs in locations other than the local NRAO DSOC cluster, enabling processing of large datasets which our current resources cannot support.

Running Workspaces is currently possible in four environments: production (mcilroy), test (hamilton), dev (shipman), and local (for developers only).

HTCondor Clusters in NRAO

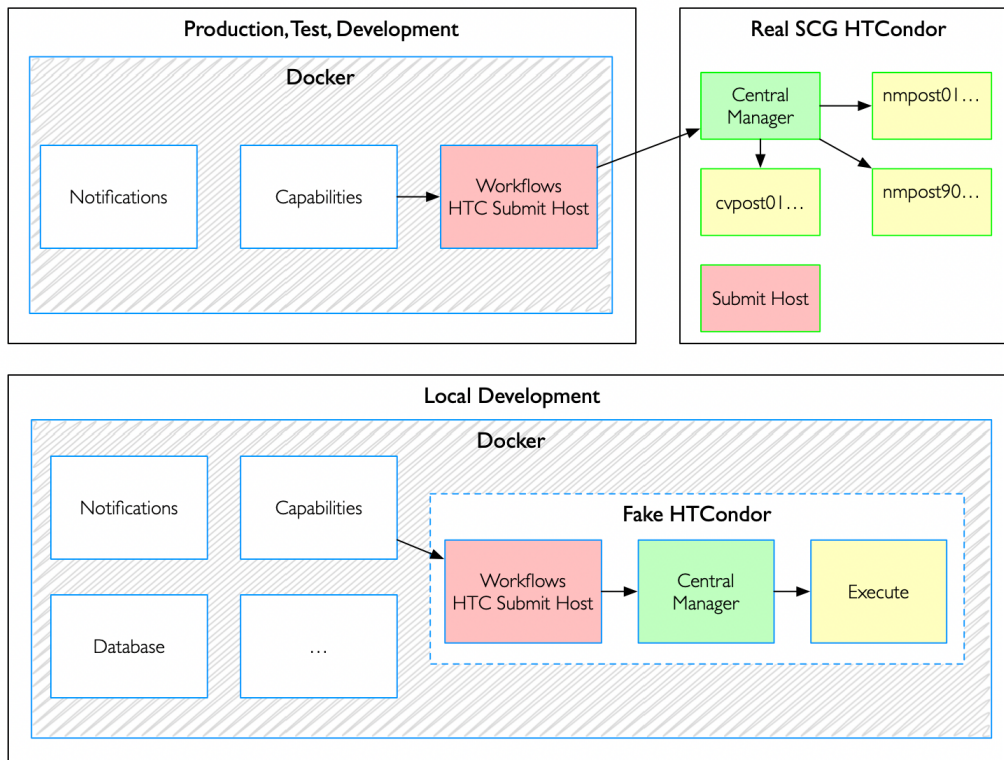There are currently two HTCondor clusters maintained by NRAO: *nmpost* and *testpost*.

NMpost is the production HTCondor cluster. All live (non-local) jobs should be running here regardless of the environment that originally created the job. The NMpost cluster nodes are then further divided into partitions identifying different processing sites. Currently there are only two other partitions with HTCondor nodes, the NMT cluster which is reserved for VLASS jobs and the CV cluster which currently only has one active HTCondor node (more are expected to be available soon).

TESTpost is the test HTCondor cluster. This cluster is smaller and meant for testing cluster changes before they are applied to the production system. Currently it is not possible to send jobs for remote execution (NMT or CV) when connected to the TESTpost cluster.

Workspaces Environments and their HTCondor Utilization

Environment Architecture

The basic Workspaces System services architecture is shown in the diagram below.

Production, Test, Development

Docker

Notifications Capabilities Workflows HTC Submit Host

Real SCG HTCondor

Central Manager nmpost01…

cvpost01… nmpost90…

Submit Host

Local Development

Docker

Notifications Capabilities

Database …

Fake HTCondor

Workflows HTC Submit Host Central Manager Execute

Workspaces is *dockerized*, i.e. all services comprising the system are running in Docker containers. From the perspective of Workspaces, HTCondor is an *external system* it must interact with. This delineation is marked in the diagram as either a separate box in the case of the live environments, and the dashed box in the case of the local environment. In both cases, the _Workspaces Workflow Service container is the submit host_ to the HTCondor cluster of choice.

Due to the use of Docker containers, there are some side effects that must be worked around when interfacing any environment's WS system with the live HTCondor cluster, most notably the differing HOME environment variable settings. For the WS containers, HOME is /home/vlapipe, however, for live HTCondor, HOME is /users/vlapipe. There was significant effort put in to making the containers use /users/vlapipe instead, however this was found to cause other problems with the system and was eventually abandoned. The issue is instead solved by incorporating the HTCondor submit host into the Workflow Service container. This allows for proper translation of the HOME location and lets the jobs setup and submit for cluster processing without complications from the HOME setting.

However, it was discovered, while testing authorization token use with the testpost cluster, that bypassing the Workflow Service container's submit node breaks this fix, resulting in submitted jobs being unable to transfer necessary files such as the transfer plugin's ssh key. Per SCG:

> When you submit to testpost you are submitting to testpost-master because of the line `CONDOR_SCHEDD = testpost-master.aoc.nrao.edu [in the submit config]` and testpost-master cannot access /home/vlapipe.

The possible ramifications of this issue (if any) for when token use is put into production are still being investigated.
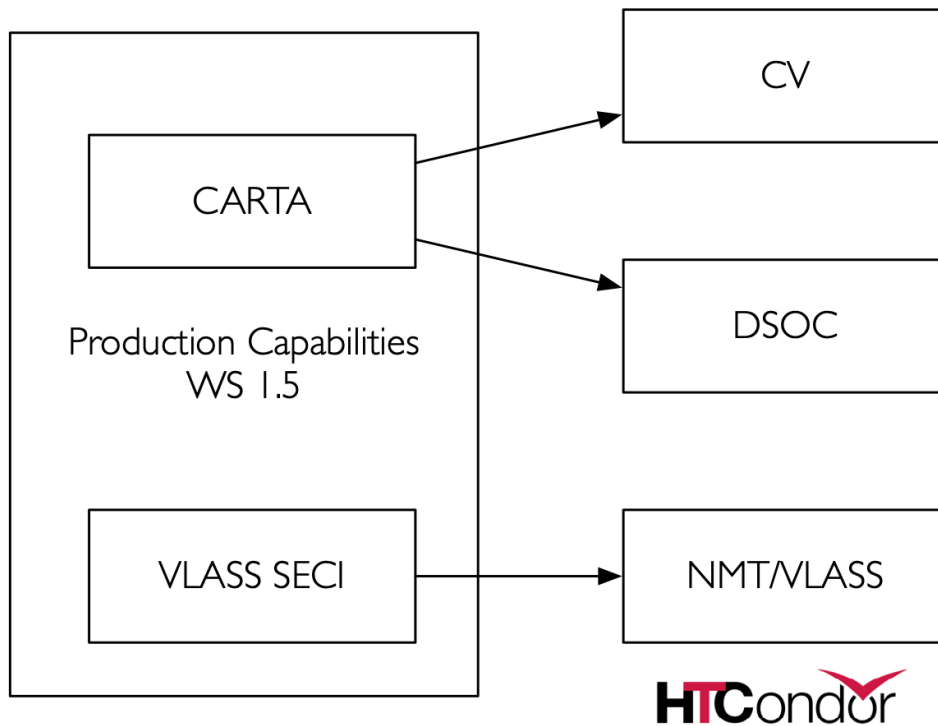
Current Status of Environments

Production Environment

As of November 12 2021, Workspaces 1.5 is deployed to production and utilizing all three NMpost execution sites.

In production, we are using:

- SCG Production HTCondor deployment at DSOC
- VLASS HTCondor deployment at NMT
- Charlottesville 1-node HTCondor carve-out

Development & Test Environment

In test/dev, we are using:

- The same as above

Local Environment

In our local development environment we are using:

- Dockerized 1-execute node HTCondor composition maintained by SSA/Workspaces

Future Expectations

Workspaces 2.0 / February 2022

Workspaces 2.0 is expected to enter TRR on February 8th and should release to production a few weeks later. Version 2.0 will bring additional usage of the production HTCondor cluster for:

- Standard calibration
- Standard imaging

These are currently serviced by the archive workflow system which is Torque-based. So this will increase the load.