Build your own cluster: Session 2

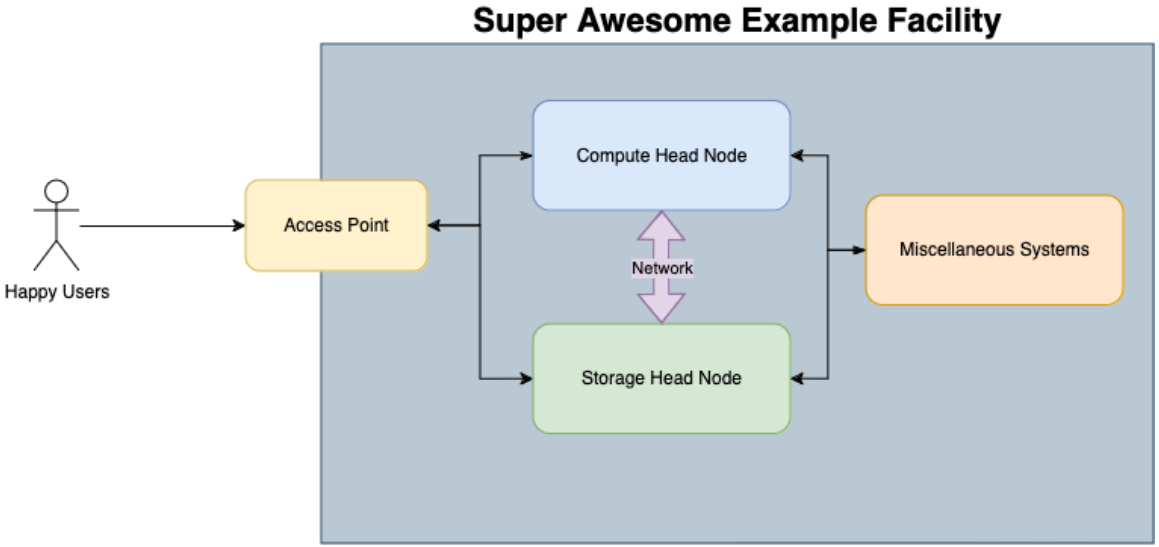
Farrukh Khan
Computational Science and AI Directorate (CSAID)
Fermi National Accelerator Laboratory
May 20, 2024

Session 1 recap

- We looked at what it means to be a system manager
- We looked at the heart of a system; operating system
- We did some exercises on Alma Linux 9
- In short:
  - You are system managers for the exercise!
  - You are now considered experts at Linux!

What is a facility?

- A facility consists of multiple systems working together to provide services and resources to users
- A simple facility:



HTCondor Software Suite (HTCSS)

- We will focus on compute systems during our sessions
- HTCondor software suite, or just HTCondor, is one of the leading batch systems in the world. It is developed by the center of high throughput computing (CHTC) at University of Wisconsin, Madison
- HTCondor at a high level consists of three main components
  - Access point (AP)
  - Central manager (CM)
  - Execution point (EP)

HTCondor Software Suite (HTCSS) – Access Point and Execution Point

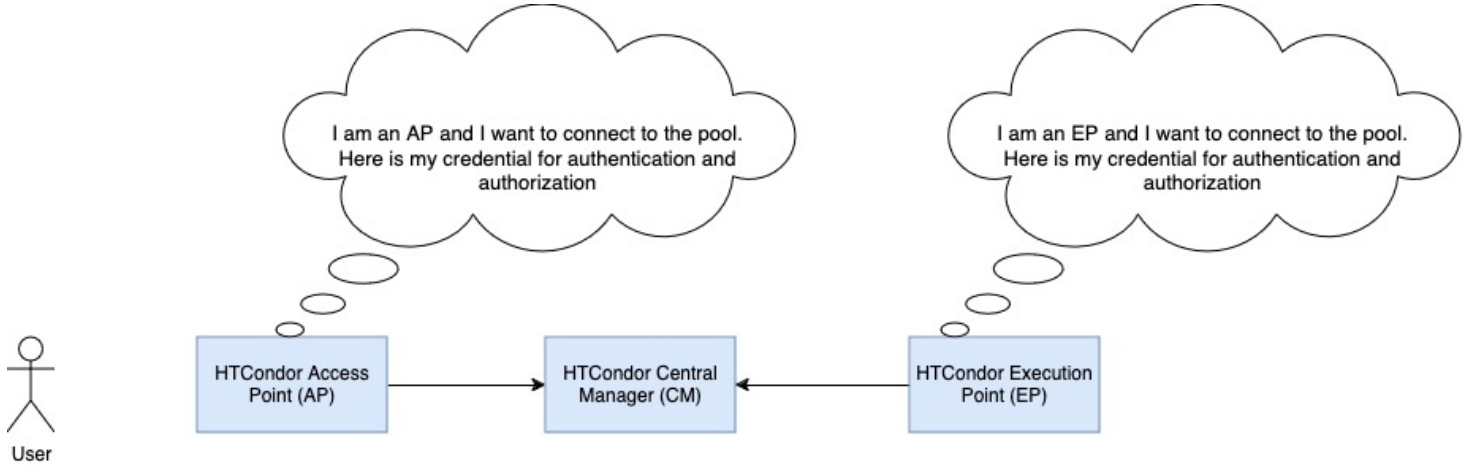
- HTCondor access point is essentially a job queue
- Access points are a user facing component and are responsible for scheduling jobs based on the configured priorities and policies
- They generally run the condor_schedd daemon and are often referred to as ‘schedds’
- HTCondor execution point is responsible for running the job
- Execution points run the condor_startd daemon and are generally referred to as ‘workers’

HTCondor Software Suite (HTCSS) – Central Manager

- HTCondor central manager is the “brain” of the pool
- The central manager generally runs two daemons; condor_collector and condor_negotiator
- HTCondor collector is a memory hungry process responsible for collecting information. Every other daemon in the HTCondor ecosystem reports to the collector
- HTCondor negotiator is responsible for ‘matchmaking’. It decides how jobs are routed to different EPs in your pool based on job requirements and machine constraints
- HTCSS uses a classAd mechanism to express requirements and constraints

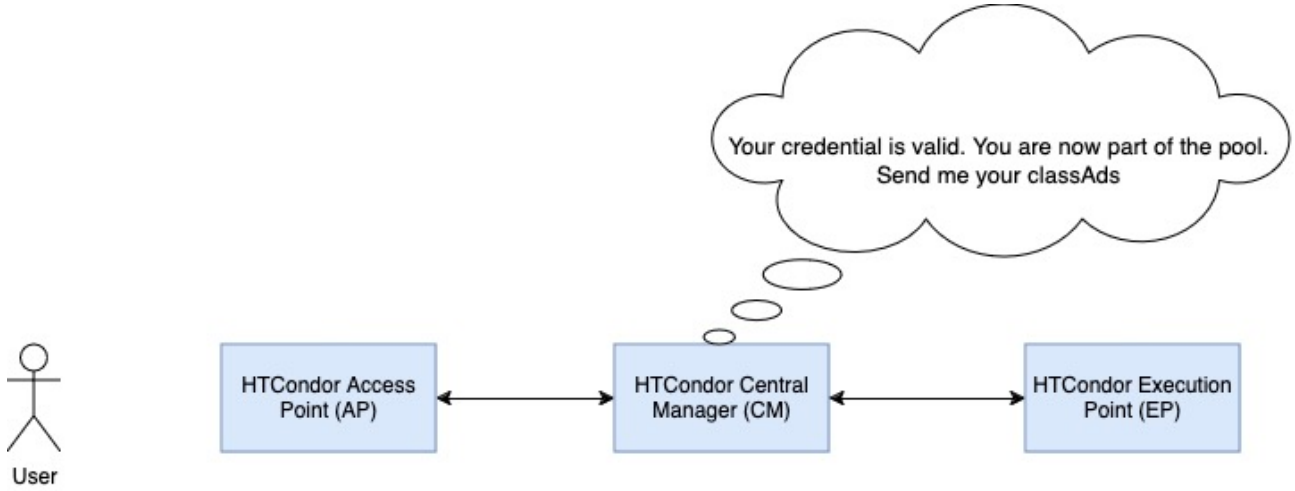
HTCondor Software Suite (HTCSS) – A simple pool

- A simple HTCondor pool setup looks like this:



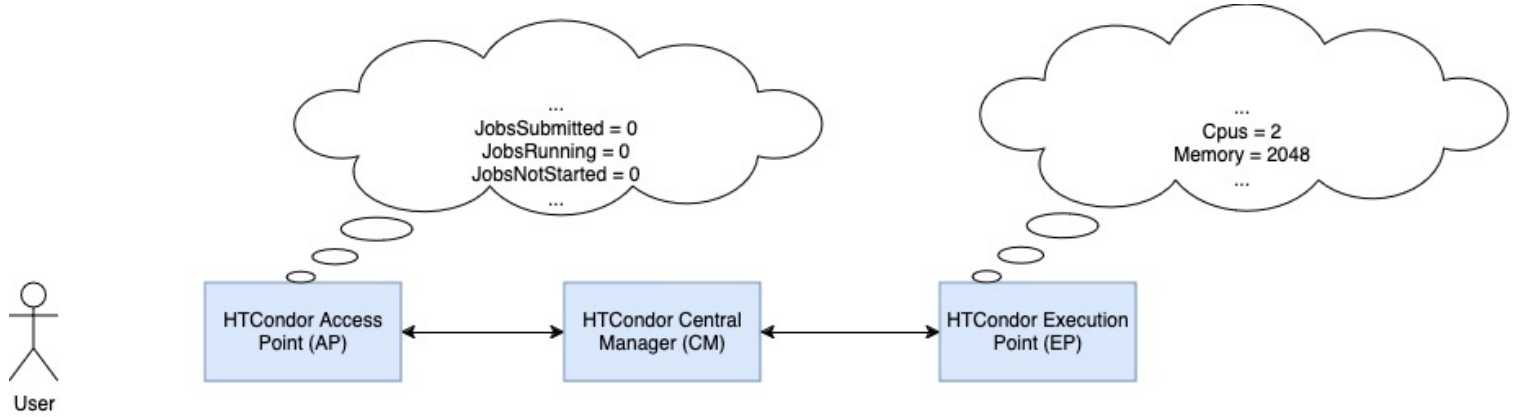
HTCondor Software Suite (HTCSS) – A simple pool

- A simple HTCondor pool setup looks like this:



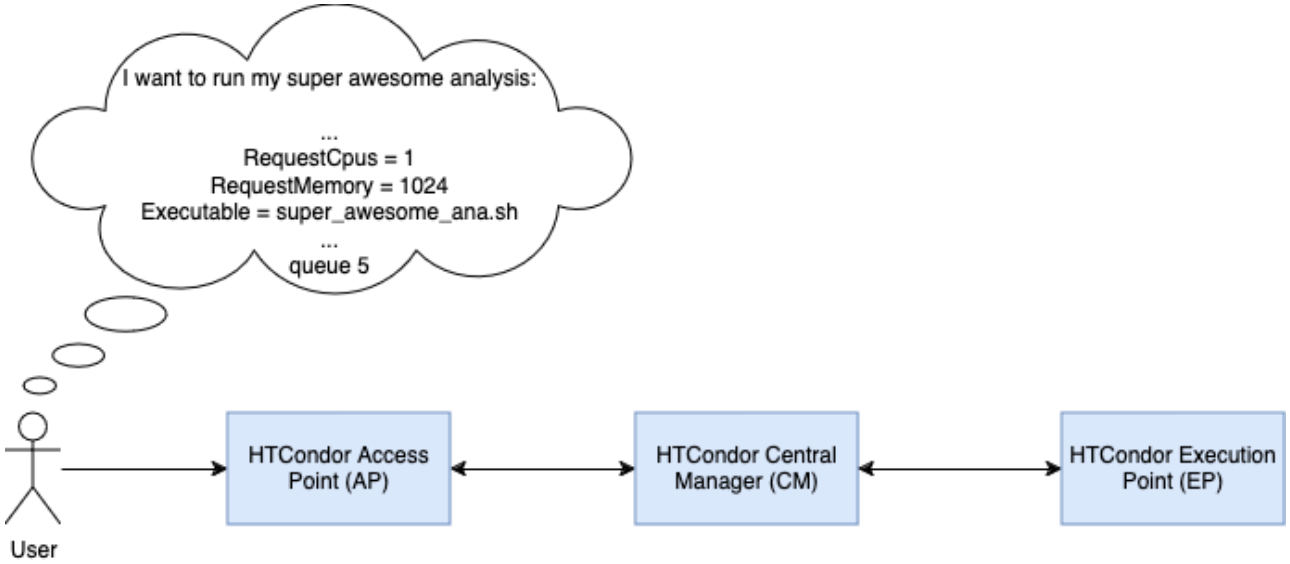
HTCondor Software Suite (HTCSS) – A simple pool

- A simple HTCondor pool setup looks like this:



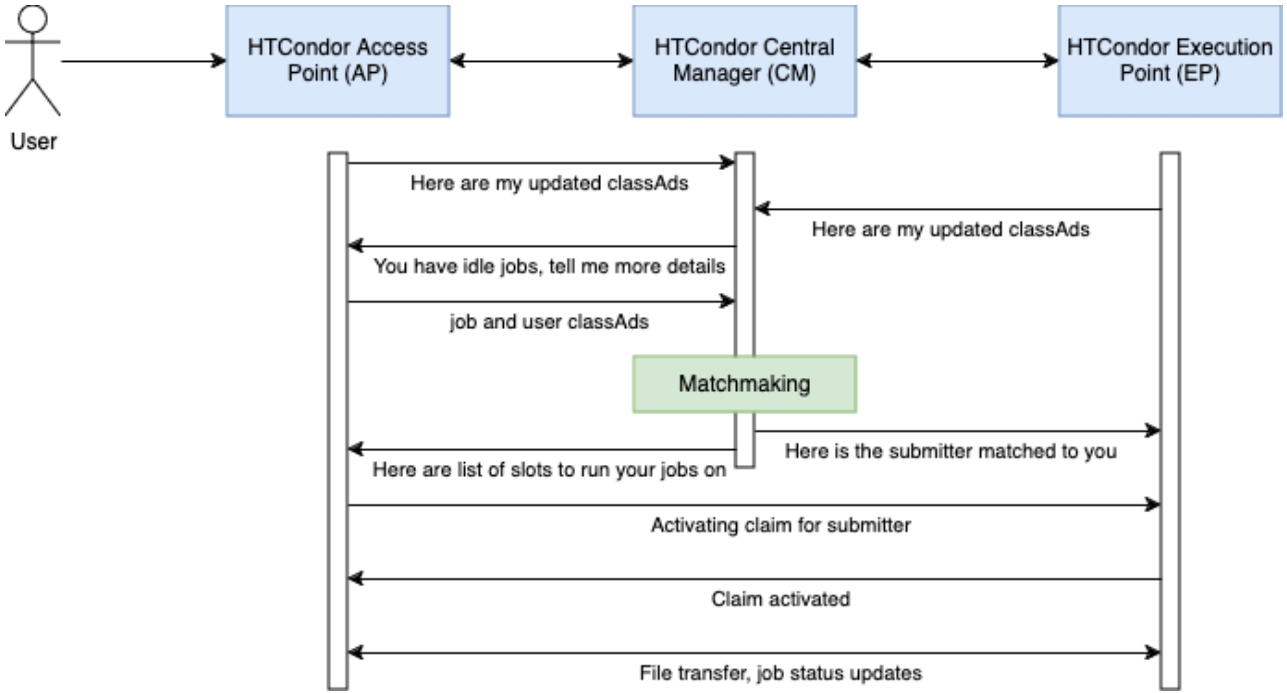
HTCondor Software Suite (HTCSS) – A simple pool

- A simple HTCondor pool setup looks like this:



HTCondor Software Suite (HTCSS) – A simple pool

- A simple HTCondor pool setup looks like this:



Fermilab

HTCondor Software Suite (HTCSS) – Installation

- You can install HTCondor on your provided containers as follows:

```
[root@fkhan ~]# cd /packages/
[root@fkhan packages]# ls
condor-23.0.10-1.el9.x86_64.rpm  htop-3.3.0-1.el9.x86_64.rpm
[root@fkhan packages]# dnf install -y ./condor-23.0.10-1.el9.x86_64.rpm
Last metadata expiration check: 0:35:28 ago on Thu May 16 17:46:41 2024.
Dependencies resolved.
================================================================================
 Package              Architecture          Version                  Repository                  Size
================================================================================
Installing:
 condor               x86_64                23.0.10-1.el9            @commandline               7.9 M

Transaction Summary
================================================================================
Install  1 Package

Total size: 7.9 M
Installed size: 25 M
Downloading Packages:
Running transaction check
```

HTCondor Software Suite (HTCSS) – Start up

- Start HTCondor inside the container as follows:

```
[root@fkhan packages]# condor_master &
[1] 168
[root@fkhan packages]# ps faux
USER         PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
root          31  0.0  0.0   5616  4716 pts/1    Ss   18:20   0:00 /bin/bash
root         200  0.0  0.0   7560  3384 pts/1    R+   18:26   0:00  \_ ps faux
root           1  0.0  0.0   4996  4024 pts/0    Ss+  18:20   0:00 /bin/bash
condor       169  0.0  0.0  24620 13136 ?        Ss   18:26   0:00 condor_master
root         198  0.0  0.0   8096  3136 ?        S    18:26   0:00  \_ condor_procd -A /var/run/condor/procd_pipe
condor       199  0.1  0.0  19948 14436 ?        Ss   18:26   0:00  \_ condor_shared_port
[1]+  Done                    condor_master
[root@fkhan packages]#
```

- condor_master is the parent process that spawns and manages the rest
- condor_procd manages the inter process communication between HTCondor daemons
- condor_shared_port provides a common interface to reach all HTCondor daemons over a network

HTCondor Software Suite (HTCSS) – Directory structure

- HTCondor places all important files under /etc/condor
  - /etc/condor/config.d for configuration files
  - /etc/condor/tokens.d for tokens
  - /etc/condor/passwords.d for pool passwords
  - /etc/condor/ganglia.d for monitoring metrics

HTCondor Software Suite (HTCSS) – Configuration files

- HTCondor configuration files can be found under /etc/condor/config.d
- Create a new file for your configuration tweaks. For example, I created a 99-local.conf as shown below and added a configuration knob DAEMON_LIST:

```
[root@fkhan config.d]# ls
00-htcondor-9.0.config  10-stash-plugin.conf
[root@fkhan config.d]# touch 99-local.conf
[root@fkhan config.d]# vim 99-local.conf
[root@fkhan config.d]# cat 99-local.conf
DAEMON_LIST = SCHEDD
[root@fkhan config.d]#
```

- What does DAEMON_LIST mean? HTCondor manual has all the answers: https://htcondor.readthedocs.io/en/23.0/admin-manual/configuration-macros.html

HTCondor Software Suite (HTCSS) – Making changes

- condor_reconfig is used to make HTCondor re-read the configuration files i.e. in other words, to make changes take effect
- I added SCHEDD to the daemon list, after reconfig you can see an additional HTCondor daemon running:

```
[root@fkhan config.d]# cat 99-local.conf
DAEMON_LIST = SCHEDD
[root@fkhan config.d]# condor_reconfig
Sent "Reconfig" command to local master
[root@fkhan config.d]# ps faux
USER         PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
root          31  0.0  0.0   5616  4716 pts/1    Ss   18:20   0:00 /bin/bash
root         209  0.0  0.0   7560  3432 pts/1    R+   18:34   0:00  \_ ps faux
root           1  0.0  0.0   4996  4024 pts/0    Ss+  18:20   0:00 /bin/bash
condor       169  0.0  0.0  24744 14344 ?        Ss   18:26   0:00 condor_master
root         198  0.0  0.0   8096  3136 ?        S    18:26   0:00  \_ condor_procd -A /var/run/condor/procd
condor       199  0.0  0.0  20060 14620 ?        Ss   18:26   0:00  \_ condor_shared_port
condor       206  0.2  0.0  21868 16724 ?        Ss   18:34   0:00  \_ condor_schedd
```

HTCondor Software Suite (HTCSS) – Troubleshooting

- /var/log/condor has log file for troubleshooting issues

```
[root@fkhan config.d]# ls /var/log/condor/
KernelTuning.log  MasterLog  ProcLog  SchedLog  ScheddRestartReport  SharedPortLog
[root@fkhan config.d]#
```

- HTCondor provides a diverse set of commands to interact with daemons. The commands can be found here: https://htcondor.readthedocs.io/en/23.0/man-pages/index.html
- For the exercise this week, the following commands will come in handy:
  - condor_q
  - condor_status
  - condor_config_val
  - condor_token_* (all its variants)

HTCondor Software Suite (HTCSS) – Hands-on exercise

- Tomorrow in the first session you will be setting up a HTCondor pool of your own!
- You are a system manager now so this will not be a guided exercise
  - I will be around to answer questions and help you
- We will do this in teams:
  - Each team will have 5 to 6 people
  - You will work together to setup a HTCondor pool with 2 APs, 1 CM and 2 EPs. Teams of six can setup an extra AP or EP (your choice)
  - These components should be individually functional and adequately connected to each other
- Good luck!