

What's new in the HTCondor Software Suite (HTCSS) ? What's coming up?

European HTCondor Workshop 2022

Todd Tannenbaum
Center for High Throughput Computing
University of Wisconsin-Madison

Some New Terminology

› Releases

- Long Term Support (LTS) Releases Channel
- Feature Releases Channel

› HTCondor Software Suite (**HTCSS**)

- Access Point (AP)
- Execution Point (EP)
- HTCondor Pool =
Central Manager +
Execution Point(s) (EP)
- HTCondor Compute Entrypoint
(HTCondor-CE)

Release Channels

- › **Long-Term Support (LTS) Releases** (*formerly 'stable series'*)
 - Only bug fixes
 - vMajor.0.Update (e.g. 9.0.1, 9.0.2, ... 9.0.15, 10.0.0, 10.0.1, 10.0.2, ...)
 - Last year: v9.0 just released with 71 documented enhancements
 - Today: v9.0.17
- › **Feature Releases** (*formerly 'developer series'*)
 - Bug fixes plus new features
 - vMajor.Minor.Update (e.g. 9.1.0, 9.2.0, 9.3.0, 9.3.1, 9.4.0, ...)
 - Today: v9.12.0 ***This is the v10.0.0 release candidate!***

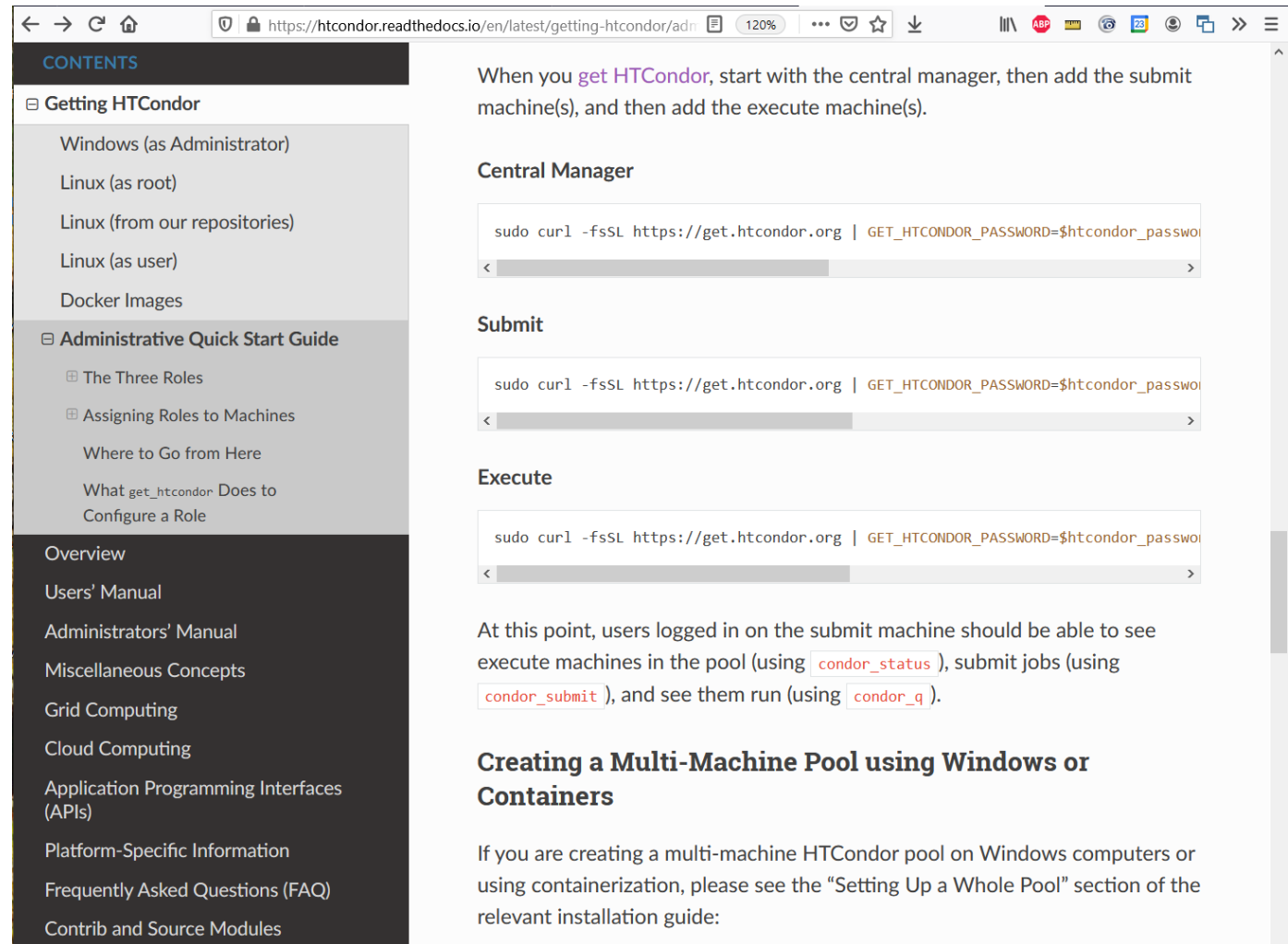
Update since last year

- › Not an exhaustive list!
- › For a complete list of "what has changed"
 - HTCondor
 - Highlights at:
<https://htcondor.org/htcondor/release-highlights/>
 - Details at:
<https://htcondor.readthedocs.io/en/latest/version-history/development-release-series-91.html>
 - HTCondor CE
 - Highlights at:
<https://htcondor.org/htcondor-ce/release-highlights/>
 - Details at:
<https://htcondor.com/htcondor-ce/v5/releases/#htcondor-ce-5-version-history>

What was introduced in HTCSS LTS v9.0.x?

Fast "one-line" install & configure

- › Secure by default using ID tokens
- › Demo and IDTokens explanation later this morning



The screenshot shows a web browser displaying the HTCondor documentation. The left sidebar contains a table of contents with the following items:

- CONTENTS
- Getting HTCondor
 - Windows (as Administrator)
 - Linux (as root)
 - Linux (from our repositories)
 - Linux (as user)
 - Docker Images
- Administrative Quick Start Guide
 - The Three Roles
 - Assigning Roles to Machines
 - Where to Go from Here
 - What get_htcondor Does to Configure a Role
- Overview
- Users' Manual
- Administrators' Manual
- Miscellaneous Concepts
- Grid Computing
- Cloud Computing
- Application Programming Interfaces (APIs)
- Platform-Specific Information
- Frequently Asked Questions (FAQ)
- Contrib and Source Modules

The main content area shows the 'Getting HTCondor' section. It includes a paragraph: "When you get HTCondor, start with the central manager, then add the submit machine(s), and then add the execute machine(s)."

Below this, there are three sections, each with a terminal command:

- Central Manager**

```
sudo curl -fsSL https://get.htcondor.org | GET_HTCONDOR_PASSWORD=$htcondor_password
```
- Submit**

```
sudo curl -fsSL https://get.htcondor.org | GET_HTCONDOR_PASSWORD=$htcondor_password
```
- Execute**

```
sudo curl -fsSL https://get.htcondor.org | GET_HTCONDOR_PASSWORD=$htcondor_password
```

Below the commands, a paragraph states: "At this point, users logged in on the submit machine should be able to see execute machines in the pool (using `condor_status`), submit jobs (using `condor_submit`), and see them run (using `condor_q`)."

Finally, there is a section titled "Creating a Multi-Machine Pool using Windows or Containers" with the text: "If you are creating a multi-machine HTCondor pool on Windows computers or using containerization, please see the 'Setting Up a Whole Pool' section of the relevant installation guide:"

New in v9.0 for end-users

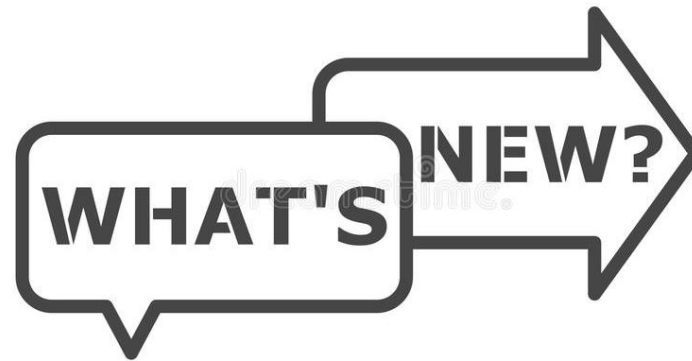
- Transfer job data to/from web servers, Box.com, Amazon S3, Google Drive, MS OneDrive via file transfer plugins now supporting uploads, authentication, and improved error handling
- Improved support for GPUs, including GPU utilization info and support for using GPUs in containers
- New `condor_watch_q` tool that efficiently provides live job status updates
- New tools and mechanisms to support jobs that checkpoint
- Fixed interactive submission of Docker jobs
- Many improvements to the Python API, including Python 3 support and new bindings for DAGMan and chirp
- DAGMan now provides a method for inline jobs to share submit descriptions
- You may now change some DAGMan throttles while the DAG is running
- You can assign priorities to DAGMan (scheduler universe) jobs
- And more...

New in v9.0 for sys admins

- Fundamental security model changes:
 - Secure by default
 - Authorization via Tokens: IDTOKENS, SciTokens
 - Oauth2 workflow to interoperate with other services
- New packaging
 - Native packaging has releases, release candidates, and daily builds
 - Official containers on Docker Hub designed to work well with k8s
 - Conda
- Push job information into Elastic Search
- More powerful classad transform language for HTCondor-CE, schedd
- Dataflow mode for jobs
- Container support improvements
- Cgroup limit improvements
- Cap the number of cores allocated to individual users
- New platforms (e.g. EL8, Ubuntu 20)
- And more...



What is coming in HTCSS v10.0.x ?



Updates to Web and Manual to reflect new terminology and the Suite idea

The screenshot shows the HTCondor Software Suite website. The top navigation bar includes links for Software, Help & Support, Community, and About. A search bar is located on the right. The main content area features a grid of links for HTCondor, HTCondor-CE, and General sections. A prominent announcement for 'HTCondor Now Hiring Software Engineer' is dated May 17, 2022. Below this is a banner for 'Week' with a 'View Event Page' button. A text block describes the HTCondor Software Suite (HTCSS) as providing services for automating and managing batch workloads. The bottom section contains a 'Tweets by @HTCondor' widget, a 'News' section with recent updates, and a 'Latest Releases' section listing versions for HTCondor (9.8.1, 9.0.12) and HTCondor-CE (5.1.3).

HTCondor
Overview
Download
Documentation
Release Highlights
Release Schedule

HTCondor-CE
Overview
Download
Documentation
Release Highlights

General
News
Security
Documentation
Download

HTCondor Now Hiring Software Engineer
May 17, 2022
[View All Software News](#)

and users the chance to exchange ideas and experiences, to learn about the latest research, to experience live demos, and to influence our short and long term research and development directions
[View Event Page](#)

Week

HTCondor Software Suite (HTCSS) provides an array of services for automating and managing batch workloads and computing resources.

Tweets by @HTCondor

HTCondor @HTCondor
Next up is Marco Mascheroni from the University of California, San Diego, presenting "Managing a dynamic pool for CMS."
Click the link below to view the slides bit.ly/3IGU5A6

Managing a dynamic pool for CMS
Marco Mascheroni, Antonio Perez-Caleiro Yzquierdo, Saeed Talemi, Edita Kozmavic
On behalf of the CMS Submission Infrastructure team
HTCondor Workshop Spring 2021

News

HTCondor Now Hiring Software Engineer
May 17, 2022

High-throughput computing as an enabler of black hole science
May 12, 2022

Speaker line-up for HTCondor Week 2022 announced
May 9, 2022

Expanding, uniting, and enhancing CLAS12 computing with OSG's fabric of services
May 2, 2022

Latest Releases

HTCondor

Feature	9.8.1
April 26, 2022	
Long Term Support	9.0.12
April 19, 2022	

HTCondor-CE

Latest	5.1.3
January 20, 2022	

Upcoming OS and Architecture Support



Enterprise Linux 8 / Stream 8

- Available Now
- Based on Alma Linux

ARM and PowerPC architectures

- Available with v10.0.0
- Based on Alma Linux

Enterprise Linux 9 / Stream 9

- Challenges due to shifts in several libraries (e.g. OpenSSL, cgroups)
- Interoperability with HTCSS v8.x may be limited (upgrade!)
- ETA: End of the year

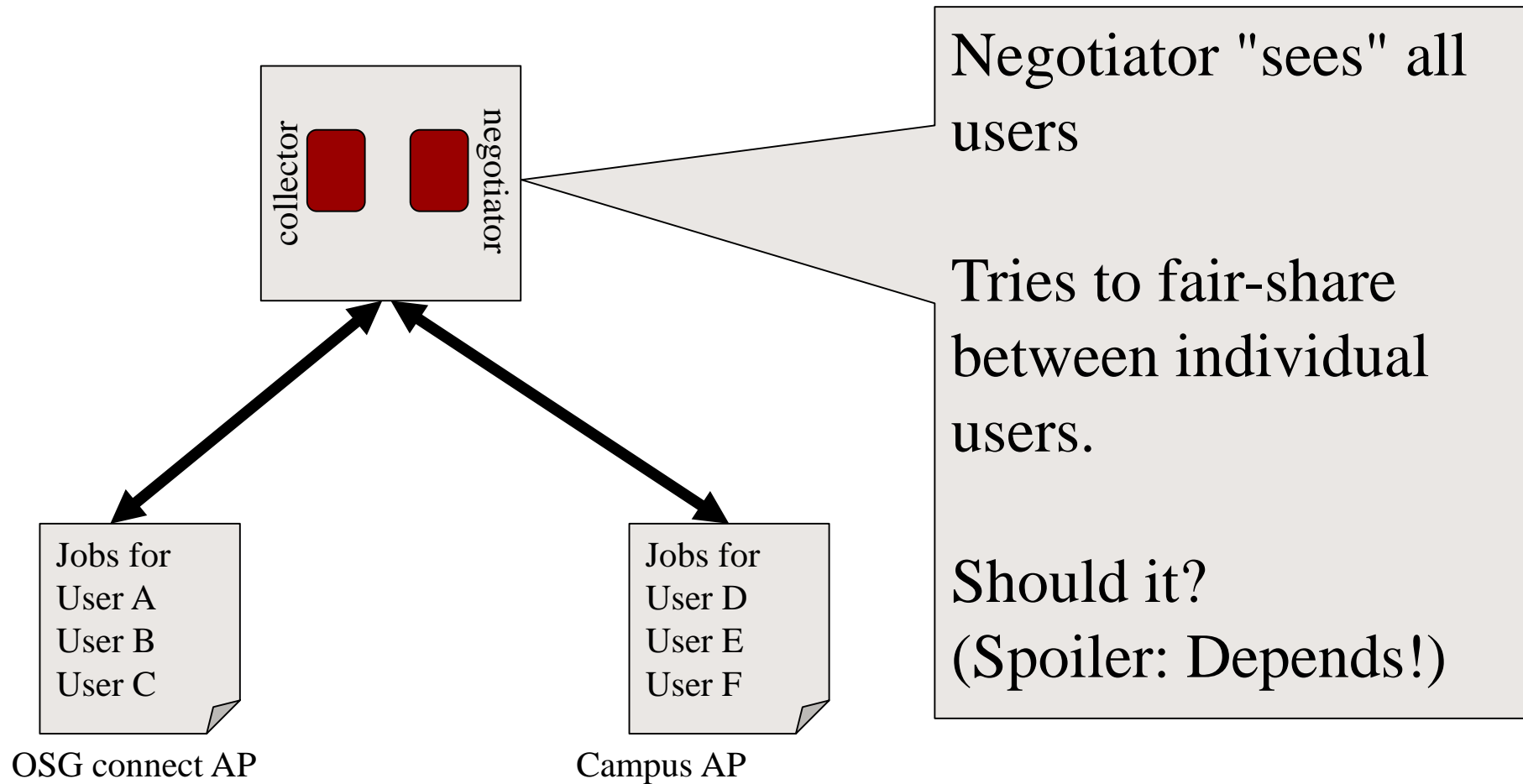
Some Access Point Enhancements

Stop Runaway Trains Jobs!

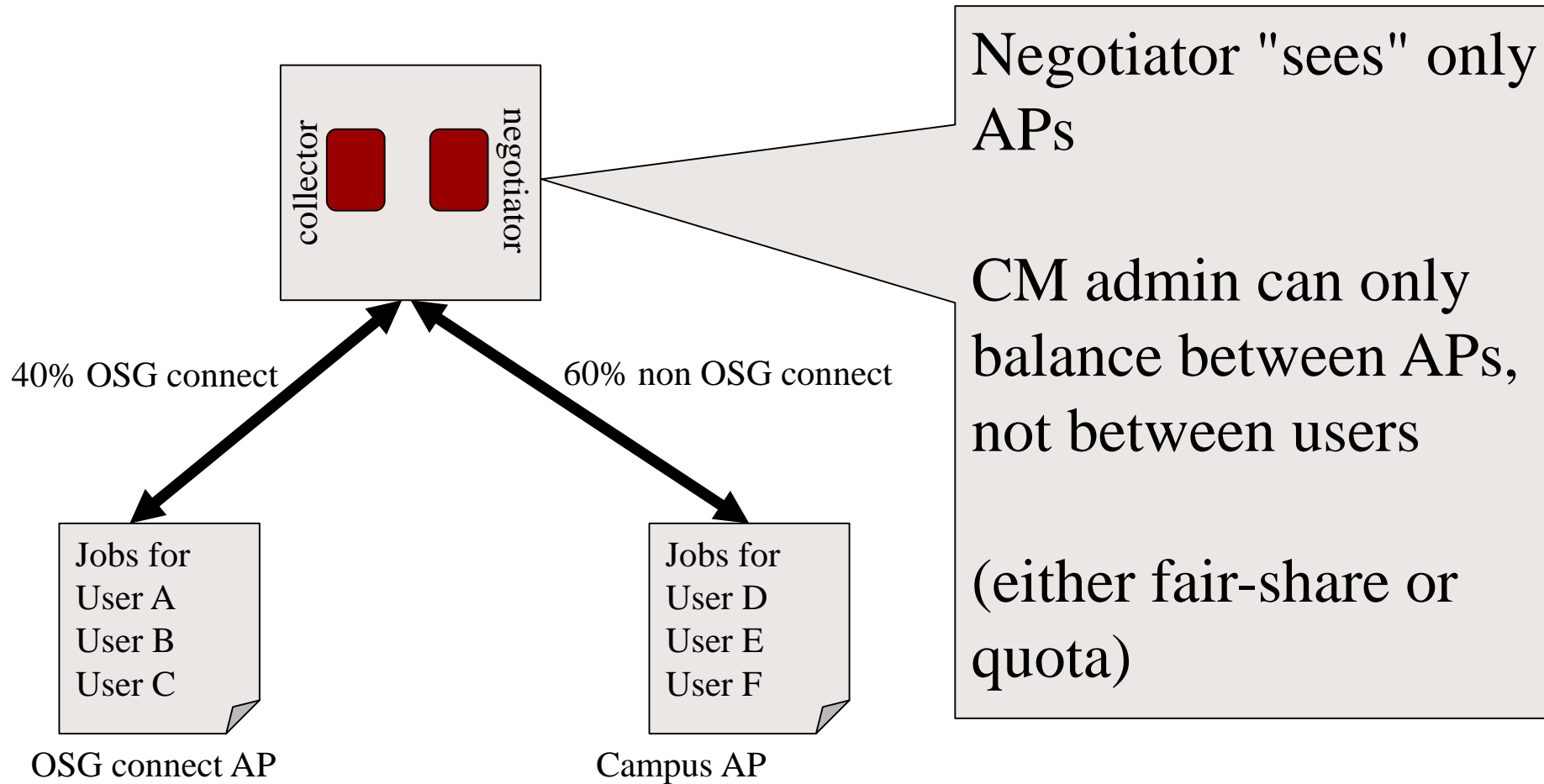
- › Stop runaway jobs!
- › Following can go in your submit file:
 - `allowed_execute_duration = <seconds>`
(from exec to exit)
 - `allowed_job_duration = <seconds>`
(include job transfers)



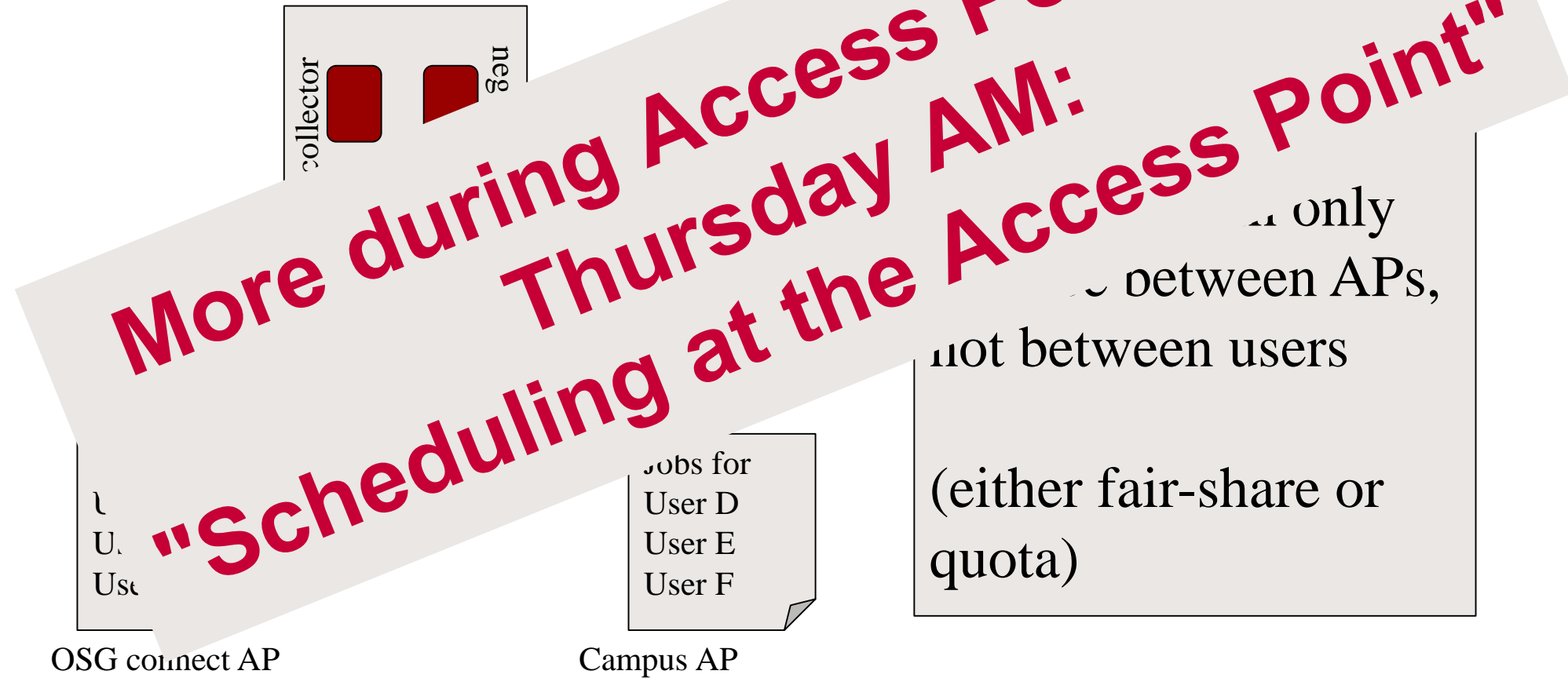
AP/CM interactions: As it is Today



A new mode for AP/CM interactions



A new mode for AP/CM interactions



Submit File Enhancements

- › Can make "+CustomAttributes" first class with knobs **EXTENDED_SUBMIT_[COMMANDS|HELPPFILE]:**

Instead of:

```
+SiteJobType = "analysis"  
queue
```

Just:

```
SiteJobType = analysis  
queue
```

- › Define job submit templates:

```
use template: Matlab( mymatrix.m )  
queue
```

Experimenting with new command line user interface

› *htcondor* <noun> <verb>

- "*htcondor job submit*", "*htcondor job status*", ...

New Command Line User Interface, cont

\$ htcondor job status 123.45

Job 123.45 is currently running.

It started running 2.1 hours ago.

It was submitted 3.6 hours ago.

Its current memory usage is 2.5 GB out of 4.0 GB requested.

Its current disk usage is 3.8 GB out of 5.5 GB requested.

It has restarted 2 times.

Experimenting with new commands

More Weds AM:

"Submitting Sets of Jobs"

- › *htcondor* <...>
 - "*htcondor job submit*", "*htcondor job status*", ...
 - "*htcondor dag submit*", "*htcondor dag status*", ...
 - "*htcondor jobset submit*", "*htcondor jobset status*", ...
 - "*htcondor annex create*", "*htcondor annex status*", ...
- › Legacy tools (*condor_q*, *condor_submit*, ...) not going anywhere...

Experimenting with new commands

More Thurs AM:

"Bring Your Own Capacity Demo"

- › *htcondor* <...>
 - *"htcondor jobset submit"*, *"htcondor jobset status"*, ...
 - *"htcondor dag submit"*, *"htcondor dag status"*, ...
 - *"htcondor jobset submit"*, *"htcondor jobset status"*, ...
 - *"htcondor annex create"*, *"htcondor annex status"*, ...
- › Legacy tools (*condor_q*, *condor_submit*, ...) not going anywhere...

Some Execution Point Enhancements

New "Container Universe"

- › EP advertises container runtimes available, and uses whichever one can get the job done

- › New world order:

```
universe = container
```

```
container_image = /cvmfs/my/image/dir/
```

```
# Or container_image = docker://Debian
```

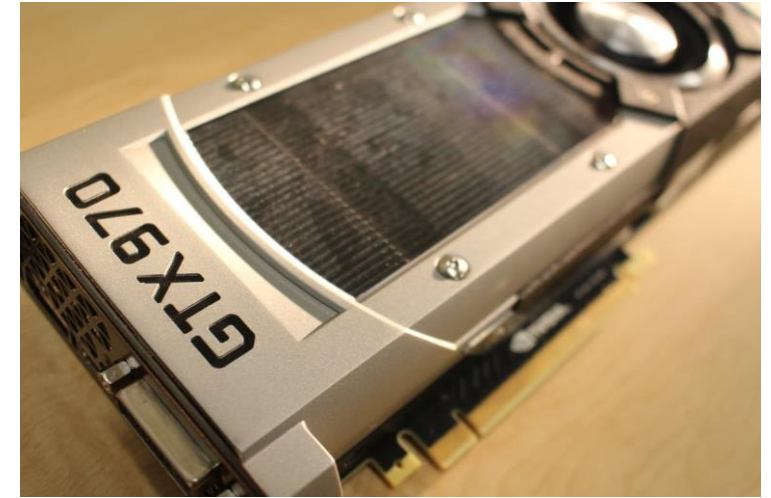
```
# Or container_image = myImage.sif
```

```
# Or container_image = http://xxx/image.sif
```



GPU Support Improvements

- › HTCondor has long been able to detect GPU devices and schedule GPU jobs
- › More recently also:
 - Monitor/report job GPU processor utilization
 - Monitor/report job GPU memory utilization
- › ***New in v9.8.0: Support for heterogeneous GPUs in one server***
 - E.g. a server with two different models of GPU cards
 - NVIDIA Multi-Instance GPU (MIG) partitioning



Submit File Example:

```
Executable = foo.exe  
RequestGPUs = 1  
RequireGPUs =  
    Capability > 7.0  
Queue
```


GPU Support Improvement

- › HTCondor has long been able to detect GPU devices and schedule GPU jobs

- › More recently also:

- Monitor/report job GPU usage
- Monitor/report job GPU error

- › **New in HTCondor 8.6.0:**
GPU-aware scheduling

- E.g. a job can request multiple GPU cards
- NVIDIA (and AMD) MIG partitioning

More during Execution Point
talks Friday AM:
"Worker nodes with GPUs"

Submit File Example:

```
Executable = foo.exe  
RequestGPUs = 1  
RequireGPUs =  
    Capability > 7.0  
Queue
```

Some Data Management Enhancements

File Transfer Error Propagation

› Hold Reason and Reason Codes now useful

12 [TransferOutputError]	An error occurred while transferring job output files or checkpoint files.	The Unix errno number.
13 [TransferInputError]	An error occurred while transferring job input files.	The Unix errno number.

Example of "condor_q -hold" (output directory missing on the Access Point)

Old Message:

Error from slot1@TODDS480S: STARTER at 127.0.0.1 failed to send file(s) to <127.0.0.1:50288>; SHADOW at 127.0.0.1 failed to write to file C:\condor\test\not_there\blah: (errno 2) No such file or directory

New Message:

Transfer output files failure at access point SUBMIT1 while receiving files from execution point slot3@NODE5. Details: writing to file C:\condor\test\not_there\blah: (errno 2) No such file or directory

Todd's Favorite File Transfer Enhancements

- › Submit macro **preserve_relative_paths = True**
`transfer_output_files = result_data/x`
ends up creating 'result_data/x' on EP job sandbox, not 'x'.
- › Config knob **STARTD_ENFORCE_DISK_LIMITS = True**
 - Create an ephemeral filesystem for the job sandbox on the EP.
 - Uses LVM or makes a loop-back file system
 - Reserves space on disk for the job
 - Enables EP to perform fast usage queries and cleanup

Todd's Favorite File Transfer Enhancements

- › Submit macro `preserve_relative_transfer_output` ends up creating
- › Config `HTCSS = True`
 - Create job sandbox on the EP.
 - Use local file system
 - Reserve disk for the job
 - Enable to perform fast usage queries and cleanup

More during Weds AM:
"The HTCSS Data Story"

Security Enhancements



- › Host-based security is no longer the default. 😊
- › Turn-key installation with "get_htcondor" tool sets up all the security knobs for you
- › For IDTOKENS authentication, no longer need a signing key on every EP host in your pool
 - EP belongs to a pool and trusts the Central Manager
 - Central Manager can give capabilities for admin tools
- › TOFU ! "Trust on First Use" ability for SSL
- › Improvements against replay attacks



Security Enhancements

- › Host-based security is no longer
- › Turn-key installation with security knobs for the
- › For IDTC need a signing key
 - EP is the Central Manager
 - Central give capabilities for admin tools
- › TOFU ! "on First Use" ability for SSL
- › Improvements against replay attacks

**More during Security focus talks
Weds PM**



HTCondor-CE Work

- › New syntax for writing job routes (classad transform syntax) has nice benefits including:
 - Statements evaluated in the order they are written
 - Use of variables that are not included in resultant job ad
 - Use of simple case statements
 - See <https://htcondor.com/htcondor-ce/v5/configuration/writing-job-routes/>
- › Mechanism for a (hosted) CE to access capacity needing 2fa with ssh instead of just ssh alone
- › Ability for a job route to insert an IDTOKEN

GSI to Tokens Transition

- › Grid Community Toolkit support is going away, so HTCSS use of GSI authentication phased out
 - Use tokens instead (SCITOKENS, IDTOKENS)
 - Full details here:
<https://htcondor-wiki.cs.wisc.edu/index.cgi/wiki?p=PlanToReplaceGridCommunityToolkit>
- › NOTE: HTCSS will still delegate and manage an X.509 proxy with the job, but the GSI authentication method is gone from v9.4 +

Thank you!

Follow us on Twitter!
<https://twitter.com/HTCondor>



This work is supported by NSF under Cooperative Agreement OAC-2030508 as part of the PATh Project. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF

PATh PARTNERSHIP to ADVANCE
THROUGHPUT
COMPUTING