

What's new in HTCondor? What's coming?

European HTCondor Workshop Sept 2021

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

What's new since last year?

Outline

- › PAtH
- › Change in Version number scheme
- › HTCondor version 9.x released
- › What is cooking in the CHTC kitchens...

PATH_h

**PARTNERSHIP to ADVANCE
THROUGHPUT
COMPUTING**

<https://path-cc.io/>

A Partnership Launched!

On October 1, 2020 we started the 5 years, \$4.5M annual budget NSF “**Partnership for Advanced Throughput Computing (PATH)**” project

“The Partnership to Advance Throughput Computing (PATH) project will expand Distributed High Throughput Computing (dHTC) technologies and methodologies through innovation, translational effort, and large-scale adoption to advance the Science & Engineering goals of the broader community.”

Aligned with NSF Cyberinfrastructure blueprint

An organic partnership

- Partnership between the UW-Madison **CHTC** and the **OSG** Consortium
- Builds on decades of collaboration, common vision and shared principals
- Two main elements of PATH are the **HTCondor Software Suite (HTCSS)** and the **Fabric of Capacity Services (FoCaS)** offered by the OSG
- Involves 40 individuals at seven institutions
- Committed to community building

Some upcoming changes resulting from PATH...

1. HTCondor → **HTCSS**
HTCondor Software Suite



Some upcoming changes resulting from PATH...

1. HTCondor → **HTCSS**
HTCondor Software Suite
2. Need for DevOps in OSG → More updates
→ **New Version Number Scheme**

HTCondor
High Throughput Computing



HTCondor
Software Suite



Version Number Changes

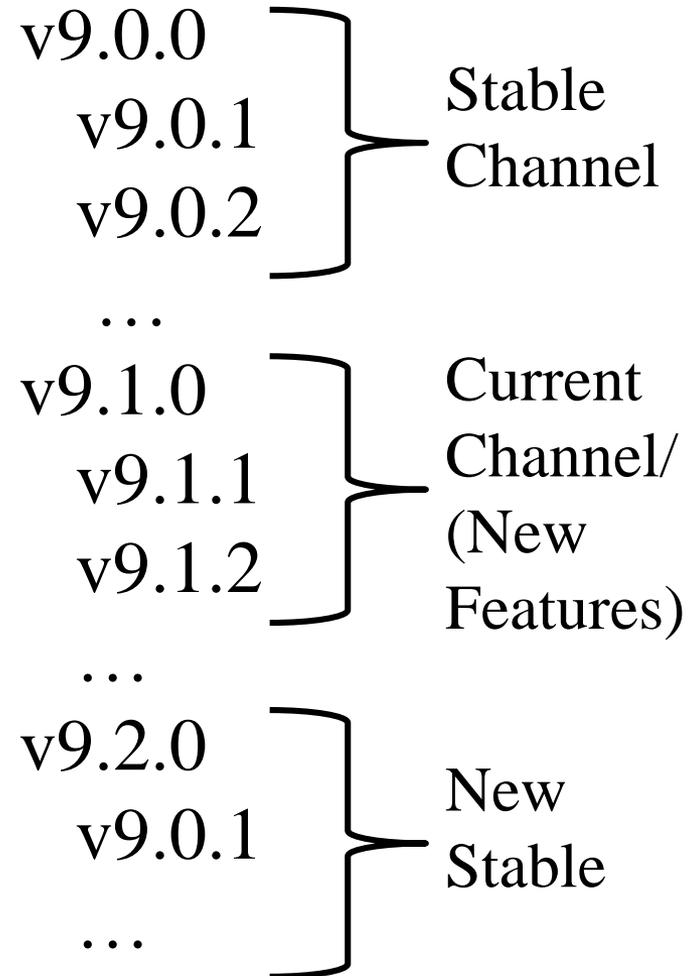
> Historical Version Scheme

<EPIC.MAJOR.MINOR>

- First Digit = "Epic" (incremented pseudo-arbitrarily every several years)
- Second Digit = Major Release (incremented at each new series number every ~12-18 months)
 - *Stable Channel (bug-fixes only series):* second digit is even
 - *Current Channel (new features series):* second digit is odd
- Third Digit = New release in the series

> What about update patches?

v9.1.2.&del><patch #> ???

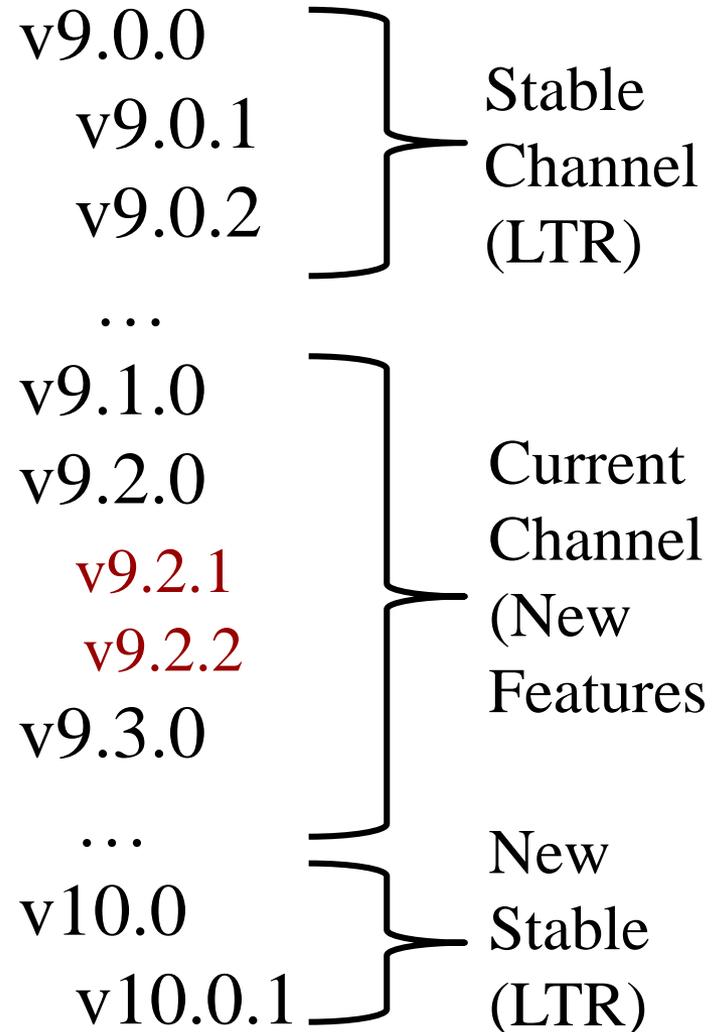


Version Number Changes

> New Version Scheme

<MAJOR.MINOR.PATCH>

- Adopts some ideas from semantic versioning (<https://semver.org/>)
- First Digit = Major Release (incremented at each new series number every ~12-18 months)
- Second Digit = Minor Release
 - Stable Channel (*bug-fixes only*): second digit = 0
 - Current Channel (*new features*): second digit > 0
- Third Digit = Update patch



Big News: HTCondor v9!

- › HTCondor v9.0.x – first introduced Apr 2021
- › Contains 71 documented enhancements introduced over the past 14 v8.9.x releases.
- › Detailed Version History in the Manual
<https://htcondor.readthedocs.io/en/latest/version-history/>

Typical Installation procedure before HTCondor v9.0... a bit daunting...

- › Figure out what OS and distribution using
- › Add proper repositories and signing keys with system packaging commands (yum / apt commands...)
- › Install binaries
- › *If you want a security beyond host-based authentication*
 - *Read Security section of the manual*
 - *Configure security settings*
- › Figure out if your system is using systemd
 - Hint : likely yes if on bare metal, likely no if in a container...
- › Start up HTCondor services, configure to start on reboot

L I V E
D E M O



You may have questions...

- › *"I don't want to pipe something off the internet into a root shell on my production machines..."*
 - You don't have to... `get_htcondor` has a non-root mode that just displays the commands it would use to do the install (without doing anything).
- › *"Is this configuration secure by default?"*
 - Yes. HTCondor services will authenticate across servers via tokens and use encryption on the network.
- › *"How about upgrading an existing pool to v9.0?"*
 - **Tune in to Todd Miller's talk "Upgrading to HTCondor 9.0" on Weds (second presentation)**

So what's new in v9.0.x?



New functionality 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 functionality for end-users

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

Topic of Christina's talk tomorrow (Tuesday) + other talks on Thursday

New functionality for sys admins

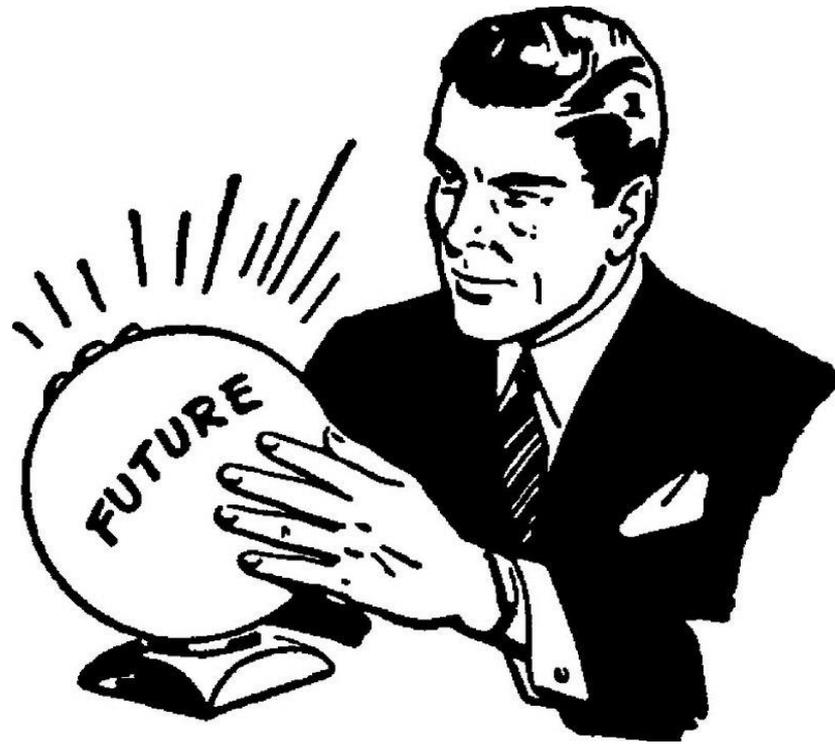
- Fundamental security model changes:
 - Secure by default: authentication, integrity check, encryption
 - Authorization via Tokens: IDTOKENS, SciTokens
 - Oauth2 workflow to interoperate with other services
- New packaging
 - Native packaging has releases, release candidates, updates, and daily builds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

New functionality for sys admins

- Fundamental security model changes:
 - Secure by default
 - Authorization via Tokens: IDTOKENS
 - Oauth2 workflow to interoperate
- New packaging
 - Native packaging has
 - Official container
 - Condor
- Elastic Search
- Head transform language for HTCondor-CE, schedd
- Queue 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...

**Topic of Greg's talk on Weds
(talk #1)**

What's coming to v9.1.x and beyond?



REST API

- Python (Flask) webapp for querying HTCondor jobs, machines, and config
- Runs alongside an HTCondor pool
- Listens to HTTP queries, responds with JSON
 - Built ontop of Python API
 - other cool tools coming courtesy Python API...
....like condor_watch_q !



https://htcondor.readthedocs.io/en/latest/man-pages/condor_watch_q.html

Organize Jobs into Job Sets

- › Today job "clusters" mostly behave as expected
 - Can remove all jobs in a cluster
 - Can edit all jobs in a cluster
- › But some operations are missing
 - Append jobs to a set (in a subsequent submission)
 - Move an entire set of jobs from one schedd to another
 - Job set **aggregates** (for use in polices?)

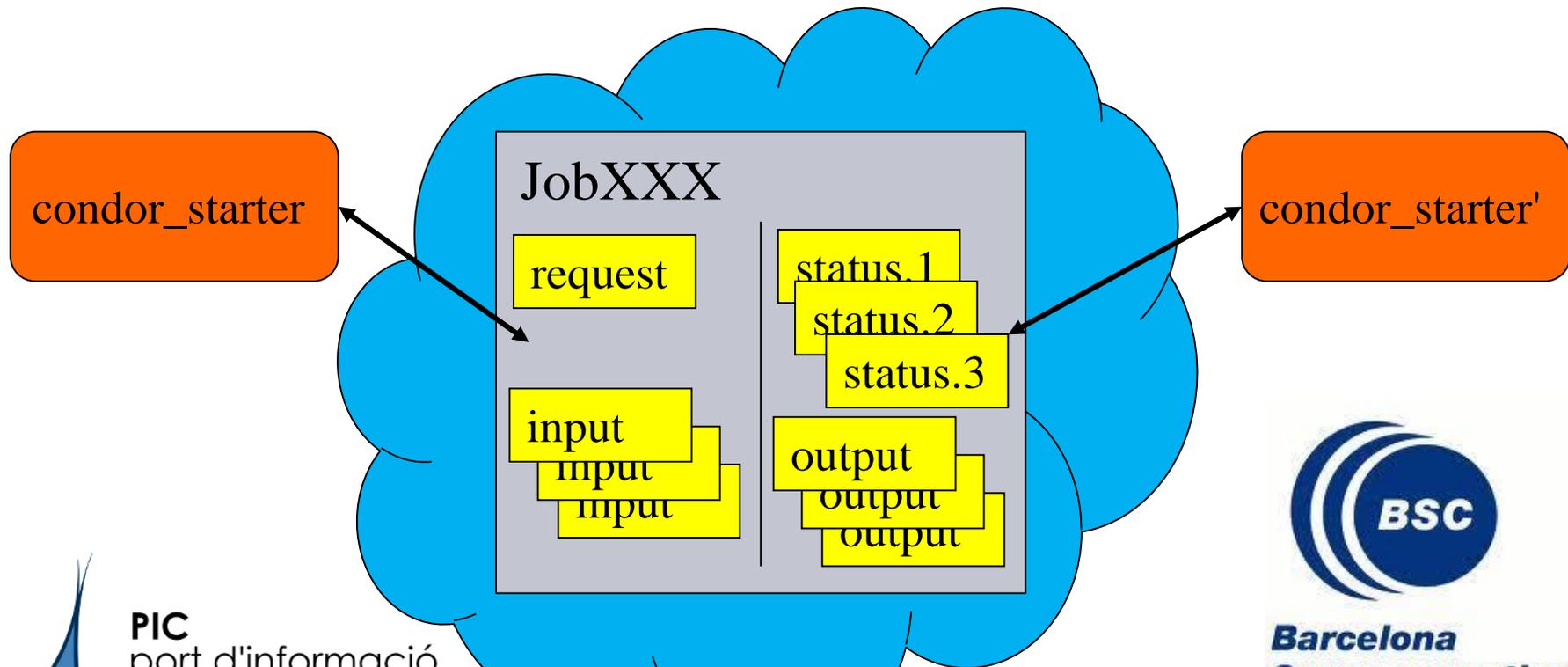
Job Sets, cont

- › Users want to think about a set of jobs as it relates to their mental model (and NOT based upon when they submit). For instance:
 - Set of jobs analyzing genome 52
 - Set of jobs analyzing images captured on date xxx
- › Experimenting with new command line user interface
 - *condor* <noun> <verb>
 - "*condor set submit*", "*condor set status*", ...
 - "*condor job submit*", "*condor job status*", ...
 - "*condor dag submit*", "*condor dag status*", ...
 - Legacy tools (*condor_q*, *condor_submit*, ...) not going anywhere...

HTCondor Annex

- › Instantiate an HTCondor Annex to dynamically add additional execute slots into your HTCondor environment
- › Continue work on mechanisms to enable end-users to provision an Annex on
 - **Local Clusters** (e.g. campus Slurm HPC clusters)
 - **Clouds** (e.g. 1st talk tomorrow from Google)
 - **HPC Centers / Supercomputers**
- › Run a job set as a campaign at an HPC site

No internet access to/from HPC nodes? File-based communication between execute nodes



PIC
port d'informació
científica

Read more about our current
approach at <http://tiny.cc/f158cz>



**Barcelona
Supercomputing
Center**
Centro Nacional
de Supercomputación

Improve GPU Management

- › HTCondor has long been able to detect GPU devices and schedule GPU jobs (CUDA/OpenCL)
- › More recently also:
 - Monitor/report job GPU processor utilization
 - Monitor/report job GPU memory utilization
- › ***In the works:***
 - Nvidia Multi-Instance GPU (MIG) support
 - Concurrent jobs on one GPU device.

Questions:

- Specify GPU memory for scheduling ?
- Limit to jobs from the same user ? Same job set ?

Improve GPU Management

- › HTCondor has long been able to detect GPU devices and schedule GPU jobs (GPUs)
- › More recently also:
 - Monitor/report job GPU utilization
 - Monitor/report node GPU utilization

- › *In the future* support multi-instance GPU (MIG) support concurrent jobs on one GPU device.

Questions:

- Specify GPU memory for scheduling ?
- Limit to jobs from the same user ? Same job set ?

Topic of John's (TJ) talk later today (after our group photo)

Unified Container Runtime Support

- › From "Docker Universe" to "Container Universe" ... just jobs with a container image specified regardless of runtime
 - Allow jobs to specify specific containers, allow admins to specify default containers
- › Podman support
- › Private images from docker hub compliant repositories
 - Cache images locally, but do not reuse private images across users

Kubernetes Integration

- › HTCondor-CE ability to provision resources from a k8s cluster.
- › Officially support easy "one command" installation of an HTCondor pool onto k8s.
 - Note: we have an initial prototype of this work already available

Yet More Security Work...

- › Federal Information Processing Standard (FIPS) compliance for mainstream HTCSS product
- › Implement plan to replace GSI and Grid Community Toolkit functionality. See timeline at: <http://wiki.htcondor.org/index.cgi/wiki?p=PlanToReplaceGridCommunityToolkit>
 - Note: HTCSS v9.1.0 already released with ARC-CE REST interface support
- › Improve some rough edges: allow an IDTOKEN alone to work well with remote administrative tools like condor_off, condor_drain, condor_reconfig

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.

PATH PARTNERSHIP to ADVANCE
THROUGHPUT
COMPUTING