Debugging Common Problems in HTCondor

Jaime Frey
Center for High-Throughput Computing

Typical User Problems

- Administrators should also understand these problems and solutions.

- User problems become the administrator's problem, and being able to explain to the user what is happening with their jobs will be necessary.

Typical User Problems

- Can’t submit jobs
- Jobs never start
- Jobs start but go on hold
- Jobs start but go back to idle unexpectedly

From the User's Perspective

- Basics
  - Is HTCondor installed?
  - Are the tools in the path?
- If the administrator has done a typical install, the path and environment should be fine.
- Run 'condor_version' to verify it works.

Can’t Submit Jobs

- When submitting, HTCondor checks the locations specified for your output files to make sure they are writable after the job completes
  - UNIX file permissions
  - Typo in a pathname
- Same for the job’s log file

Can’t Submit Jobs

- When submitting, HTCondor also checks your input files to make sure they are readable.
  - UNIX file permissions
  - Typo in a pathname
- HTCondor also checks that the job’s log can be written to.

Can't Submit Jobs

- Unable to contact the condor_schedd
- Are you logged into a submit machine? Or is this an execute machine or central manager?
- You can use 'ps' to see if any HTCondor daemons are running
- Is the condor_schedd overwhelmed or system load very high?
  - Not necessarily a user problem

Can’t Submit Jobs

- Unable to authenticate to the condor_schedd.
  - Shouldn’t be an issue if you are submitting on the same machine where the schedd is running
  - Can be an issue if you do “remote submits” since those authentication mechanisms require special configuration by the administrator

Can’t Submit Jobs

- Not authorized
- SUBMIT_REQUIREMENTS check not met
  - For example, to restrict which executable is run
  - To enforce which Accounting_Group a user claims to be part of
  - Controlled by your HTCondor administrator

Jobs Never Start

- So, you were successful at submitting the job, but now when you run ‘condor_q’ you see it stay in the “Idle” state forever.

- First, the Matchmaking process is NOT instantaneous, so some patience is required. We are a High-Throughput system.

Jobs Never Start

- Depends a lot on the pool policy
- Will another user's job get evicted or do you need to wait for a free slot?
- Are your job requirements reasonable?
  - Are you asking for an amount of CPU, Disk, Memory, or other resource that doesn't exist in your pool?
  - Or even if it's rare, you may have to wait quite a while to get access that resource

Jobs Never Start

- Is there some attribute in your job that is not satisfying the StartD requirements?
- Is there some attribute in your job that is making it “unattractive” to the StartD rank?
- Remember that each StartD might have a different configuration for Requirements and Rank (like the Owners of machines)

Jobs Never Start

- Helpful tools:
  - condor_q –analyze
  - condor_q –better-analyze
  - condor_q –better-analyze –reverse

- Will check and analyze the requirements expression of the job (or machine) to see if it matches
- Offers suggestions when it doesn't match

Jobs Go On Hold

- Many reasons jobs could go on hold:
- Job's own periodic_hold expression
- The adminstrator's "SYSTEM_PERIODIC_HOLD" expression

- These are typically used to hold the job when it violates some condition (using too much RAM, Disk, or CPU)

Jobs Go On Hold

- When file transfer fails
- Unable to write the input files into the Job Sandbox (rare)
- Unable to find an output file that was specified in the submit file (common)
- Unable to write the output back to the submit machine (rare)

Jobs Go On Hold

- You can run ‘condor_q –held’ to see which jobs are held and also the reason why.
- You can edit already-queued jobs using ‘condor_qedit’ to change the command line arguments or the name of an output file (among many other things).
- After editing, you can run ‘condor_release’ to let the job run again.

Jobs Run but then Become Idle

- This doesn't necessarily indicate a problem!
- Your job may have been evicted due to user priority and is simply waiting to be rescheduled by the system
- The machine's "PREEMPT" or "KILL" policy may have stopped your job for using too many resources
  - In this case, you should edit your Request_Cpus / Request_Memory / Etc.

Jobs Run but then Become Idle

- Remember you can always look in your job's log file for hints
- You are specifying a log file for your job, right?

- If you see excessive "Shadow Exception" messages, that may indicate a mis-configuration of the system by the administrator.

My Job Doesn't Run Correctly!

- Does it work correctly outside HTCondor?
  - ARE YOU SURE?!?!?
- Check that the environment for the job is the same as when it is running from the command line.

My Job Doesn't Run Correctly!

- Use 'condor_ssh_to_job' while it is running and you can check on it in real-time.
  - Check memory footprint, disk usage, load.
  - Output files being written correctly?
  - Attach to it with gdb to inspect the stack.
- Also, 'condor_submit –interactive'
  - Sets up the job environment and input files
  - Gives you a command prompt where you can then start job manually to see what happens

From the Admin's View

- Each running HTCondor daemon keeps a log file:
  - MasterLog
  - SchedLog
  - ShadowLog
  - etc.
- These logs can contain an enormous amount of information. The level of verbosity is configurable per-daemon.

From the Admin's View

- Find the location of the log directory:
  - condor_config_val LOG
- Look at the debug levels for each daemon:
  - condor_config_val –dump _DEBUG

From the Admin's View

- Let's consider the SCHEDD_DEBUG setting in the condor_config.
- Controls the verbosity of the SchedLog
- Individual subsystems can be added:
  - D_NETWORK
  - D_SECURITY
  - D_COMMAND
  - etc.
- D_ALL:2 is the most verbose level

From the Admin's View

- Because log files can be huge, they have a certain maximum size and are rotated as needed.

- See Section 3.3.4 in the manual for full debugging subsystem configuration.

From the Admin's View

- You can remotely fetch a log:
- condor_fetchlog <machine> <subsys>
  - condor_fetchlog abc.wisc.edu SCHEDD
- By default, you can only fetch logs from an "administrator" authorized machine (like the Central Manager).
  - Like everything, this is configurable

condor_master Won't Start

- It is possible that the condor_master cannot write to its own log file. In this case, it will refuse to start and exit with status 44.

- The condor_master also checks to see if another instance of HTCondor is already running. In this case it does not start a new instance and instead prints a message in the MasterLog file.

condor_master Won't Start

- Possible error in the configuration file that made it unparsable
- Specified a condor_config file that doesn't exist or has permissions that make it unreadable.
- Almost all other situations should result in at least something being written to log file.

From the Admin's View

- Okay, now that we have the logs, we have access to the information that we will need to debug problems.

- Let's move on to some common problems and how they are identified.

From the Admin's View

- When I run condor_status, I don't see any output!
- This means that the condor_startd is unable to advertise the slots to the collector
  - Is the condor_startd running? (Use 'ps')
  - Network connectivity issue? (Firewall?)
  - Authorization issue?
  - Start by looking at the StartLog of an execute machine that should be reporting

From the Admin's View

- Obvious errors in the StartLog:
  - Is the right collector specified?
  - Do you see messages about "Can't connect"?
  - Error sending data?
  - Timing out?
  - Update was denied?

From the Admin's View

- You should also check the CollectorLog on the central manager to see if the information is coming in correctly
  - Do you see "Command received"?
  - Error reading data?
  - Timing out?
  - Update was denied?

From the Admin's View

- Authorization issue
  - You will see "PERMISSION DENIED" in the CollectorLog on the Central Manager
- It generally means that the ALLOW_WRITE or ALLOW_DAEMON setting on the Central Manager is not permitting the other machines to send updates
- Run 'condor_config_val –dump ALLOW_' on the Central Manager

From the Admin's View

- Check the list of authorized IP addresses
- Wildcards and netmasks are permitted:
  - 10.0.0.*
  - *.wisc.edu
  - 192.168.0.0/24
- Make sure to condor_reconfig the Central Manager after making any changes.

From the Admin's View

- The entire pool is "Idle" even though there are jobs in the queues!

- Any Ideas?

From the Admin's View

- The entire pool is "Idle" even though there are jobs in the queues!

- Negotiator is not making matches…

From the Admin's View

- The entire pool is "Idle" even though there are jobs in the queues!

- Negotiator is not making matches…
  - Is it running?
  - What are the Machines' "START" expressions?
  - Would you expect jobs to match?

From the Admin's View

- Negotiator *is* making matches, but somehow the SchedD is failing to finalize the match when claiming the StartD
- Examine the SchedD, StartD logs
- Look for "ERROR", "WARNING", "FAILED"
- Look at the preceding lines of the log to try to determine what led to the failure
- If needed, increase the verbosity level to get more information in the log.

From the Admin's View

- When examining logs, also pay attention to the time stamps.
  - Long gaps could indicate a problem where HTCondor was forced to block while waiting for something to happen
  - Example: Your DNS server is down or very slow, and HTCondor can't resolve hostnames
- Number of open file descriptors can be seen as well. See if you are perhaps bumping against the 'limits'.

The Wrong Jobs Are Running!

- Double check the user priorities using ‘condor_userprio’

- A handy way to see what’s happening:
  - condor_q –allusers –global –run
  - condor_status –run

From the Admin's View

- Suppose some user has submitted "too many" jobs
- The SchedD may become unresponsive, and you'll be unable to examine or modify the job queue.
- Similarly, too many simultaneous updates to the Collector can cause it to slow down
- Examine the logs to see if it is excessively busy, or possible hung or blocked.

From the Admin's View

- Use the condor_sos command!
  - condor_sos condor_q
  - condor_sos condor_status
- This sends the command in such a way that it moves to "the front of the line" and is serviced first.
- Useful for admins to diagnose and fix system problems.

Still Stuck?

- Send email to htcondor-users@cs.wisc.edu
  - Community mailing list which is very responsive
  - Always include OS and distro, version of HTCondor, specific error messages or problematic behavior
- Email htcondor-admin@cs.wisc.edu
  - Best-effort support from HTCondor developers
  - Include the same information