HTCondor-CE: Troubleshooting

HTCondor Workshop 2019 - EU Joint Research Centre
Brian Lin
University of Wisconsin — Madison
Troubleshooting HTCondor-CE

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting

HTCondor-CE Troubleshooting Guide

In this document, you will find a collection of files and commands to help troubleshoot HTCondor-CE along with a list of common issues with suggested troubleshooting steps.

Known Issues

SUBMIT_EXPRS are not applied to jobs on the local HTCondor

If you are adding attributes to jobs submitted to your HTCondor pool with `SUBMIT_EXPRS`, these will not be applied to jobs that are entering your pool from the HTCondor-CE. To get around this, you will want to add the attributes to your `job routes`. If the CE is the only entry point for jobs into your pool, you can get rid of `SUBMIT_EXPRS` on your backend. Otherwise, you will have to maintain your list of attributes both in your list of routes and in your `SUBMIT_EXPRS`.

General Troubleshooting Items

- Making sure packages are up-to-date
- Verify package contents
- Verify clocks are synchronized
- Verify host certificates and CRLs are valid
- HTCondor-CE Troubleshooting Items
  - Daemons fail to start
  - Jobs fail to submit to the CE
  - Jobs stay idle on the CE
  - Idle jobs on CE: Make sure the underlying batch system can run jobs
  - Idle jobs on CE: Is the job router handling the incoming job?
  - Idle jobs on CE: Verify correct operation between the CE and your local batch system
Troubleshooting HTCondor-CE

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#htcondor-ce-troubleshooting-items

HTCondor-CE Troubleshooting Items

This section contains common issues you may encounter using HTCondor-CE and next actions to take when you do. Before troubleshooting, we recommend increasing the log level:

1. Write the following into `/etc/condor-ce/config.d/99-local.conf` to increase the log level for all daemons:

   ```
   ALL_DEBUG = D_ALWAYS:2 D_CAT
   ```

2. Ensure that the configuration is in place:

   ```
   root@host # condor_ce_reconfig
   ```

3. Reproduce the issue

   **Note**

   Before spending any time on troubleshooting, you should ensure that the state of configuration is as expected by running `condor_ce_reconfig`.

Table of contents

Known Issues
SUBMIT_EXPRS are not applied to jobs on the local HTCondor

General Troubleshooting Items
Making sure packages are up-to-date
Verify package contents
Verify clocks are synchronized
Verify host certificates and CRLs are valid

HTCondor-CE Troubleshooting Items
Daemons fail to start
Jobs fail to submit to the CE
Jobs stay idle on the CE
Idle jobs on CE: Make sure the underlying batch system can run jobs
Idle jobs on CE: Is the job...
Troubleshooting HTCondor-CE

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#htcondor-ce-troubleshooting-tools

HTCondor-CE Troubleshooting Tools

HTCondor-CE has its own separate set of of the HTCondor tools with ce in the name (i.e., `condor_ce_submit` vs `condor_submit`). Some of the the commands are only for the CE (e.g., `condor_ce_run` and `condor_ce_trace`) but many of them are just HTCondor commands configured to interact with the CE (e.g., `condor_ce_q`, `condor_ce_status`). It is important to differentiate the two: `condor_ce_config_val` will provide configuration values for your HTCondor-CE while `condor_config_val` will provide configuration values for your HTCondor batch system. If you are not running an HTCondor batch system, the non-CE commands will return errors.

`condor_ce_trace`

Usage

Jobs go on hold

- Held jobs: no matching routes, route job limit, or route failure threshold
- Held jobs: Missing/expired user proxy
- Held jobs: Invalid job universe

Identifying the corresponding job ID on the local batch system

- HTCondor batch systems
- Non-HTCondor batch systems
- Jobs removed from the local HTCondor pool become resubmitted (HTCondor batch systems only)
Troubleshooting HTCondor-CE


HTCondor-CE Troubleshooting Data

The following files are located on the CE host.

MasterLog

The HTCondor-CE master log tracks status of all of the other HTCondor daemons and thus contains valuable information if they fail to start.

- **Location:** `/var/log/condor-ce/MasterLog`
- **Key contents:** Start-up, shut-down, and communication with other HTCondor daemons
- **Increasing the debug level:**
  
  a. Set the following value in `/etc/condor-ce/config.d/99-local.conf` on the CE host:

  ```
  MASTER_DEBUG = D_ALWAYS:2 D_CAT
  ```

Log Levels

- Useful for temporary debugging
- Log level can be adjusted per daemon (e.g., SCHEDD_DEBUG) or across all daemons (ALL_DEBUG)
- Most common, helpful log levels for HTCondor-CE:
  - D_CAT D_ALL :2 - shows the log level for each line (helpful for debugging HTCondor bugs!) and increases the log level of general messages
  - D_SECURITY - show verbose authentication messages
  - D_NETWORK - show messages for TCP/UDP connections
- Warning, this makes logs very chatty! Adjust the log sizes and number of logs kept:
  - MAX_<SUBSYS>_LOG - Max size of each log file, e.g. MAX_JOB_ROUTER_LOG
  - MAX_NUM_<SUBSYS>_LOG - Max number of logs kept, e.g. MAX_NUM_JOB_ROUTER_LOG
HTCondor-CE Startup

systemctl start condor-ce

service condor start

condor_ce_on

Legend:
- Startup
- Authorization
- Command/Logs

/var/log/condor-ce/SchedLog
/var/log/condor-ce/CollectorLog
/var/log/condor-ce/JobRouterLog

/var/log/condor-ce/MasterLog
Troubleshooting Startup

If all goes well, command-line queries should show the following daemons:

# condor_ce_status -any
MyType       TargetType       Name
Collector   None               My Pool - fermicloud068.fnal.gov@fermicloud068.fnal.gov
Scheduler   None               fermicloud068.fnal.gov
DaemonMaster None               fermicloud068.fnal.gov
Job_Router   None               htcondor-ce@fermicloud068.fnal.gov

If not...

# condor_ce_status -any
Error: communication error
CEDAR:6001:Failed to connect to <131.225.155.18:9619>
Troubleshooting Startup

systemctl start condor-ce
service condor start
condor_ce_on

Legend:
Startup
Failed AuthZ
Command/Logs

/var/log/condor-ce/MasterLog
/var/log/condor-ce/SchedLog
/var/log/condor-ce/CollectorLog
/var/log/condor-ce/JobRouterLog

03/20/19 16:05:58 ERROR: AUTHENTICATE:1003:Failed to authenticate with any method
Troubleshooting Startup

03/20/19 16:05:58 ERROR: AUTHENTICATE:1003:Failed to authenticate with any method

1. Update CA certificates and CRLs
2. Verify host cert validity
   ```bash
   openssl x509 -in /etc/grid-security/hostcert.pem -dates
   ```
3. Verify CE host system clock
4. Verify unified mapfile. Do you have a GSI line mapping the CE host certificate to “<hostname>@daemon.htcondor.org”?
5. Run `condor_ce_host_network_check`

Startup issues should be reduced in HTCondor-CE 4 thanks to filesystem authN!
Verifying HTCondor-CE Job Submission

After daemons have started up, verify job submission from the CE host:

1. Verify that local job submissions complete successfully from the CE host, e.g. `sbatch`, `condor_submit`, `qsub`, etc.
2. Verify the CE’s network configuration with `condor_ce_host_network_check`
3. Verify end-to-end job submission with `condor_ce_trace`
   a. First, from the CE host
   b. Next, from a remote host with the `htcondor-ce-client` tools

https://htcondor-ce.readthedocs.io/en/latest/verification/
Troubleshooting CE Jobs: HTCondor

1. Grid Job
2. Routed Job
3. HTCondor Negotiation

/var/log/condor-ce/SchedLog
/var/log/condor-ce/JobRouterLog
/var/log/condor/SchedLog
Troubleshooting the CE Schedd

1. No authentication errors in the SchedLog? Make sure that the firewall is open

2. Authentication errors?
   a. Set `SCHEDD_LOG = $(SCHEDD_LOG) D_SECURITY`
   b. Check for issues with `/etc/condor-ce/condor_mapfile`
      i. If using a callout to an authentication service via `GSS_ASSIST_GRIDMAP`, check LCMAPS/Argus error logs
      ii. If not, do you have a GSI line mapping the job’s DN or VOMS attribute to the proper “<user>@htcondor.org”?
   c. Make sure that mapped users exist
   d. Ensure CAs, CRLs, and VO information is up-to-date
   e. Verify CE host system clock
Troubleshooting CE Jobs

For a regular HTCondor pool:

Submit file → Idle → Held → Running → Complete

Submit file → Suspend → History file

Submit file → Complete → History file
Troubleshooting CE Jobs

For HTCondor-CEs (i.e. Condor-G submission)
# Troubleshooting CE Jobs

```bash
# condor_ce_q -nobatch

-- Schedd: lhcb-ce.cttc.wisc.edu : <128.104.100.65:9618>... @ 03/20/19 21:31:19

<table>
<thead>
<tr>
<th>ID</th>
<th>OWNER</th>
<th>SUBMITTED</th>
<th>RUN_TIME</th>
<th>ST</th>
<th>PRI</th>
<th>SIZE</th>
<th>CMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>153501.0</td>
<td>nu_lhcb</td>
<td>3/18 13:30</td>
<td>2+07:56:31</td>
<td>R</td>
<td>0</td>
<td>733.0</td>
<td>DIRAC_clpM0A_pilotwrapper.py</td>
</tr>
<tr>
<td>154043.0</td>
<td>nu_lhcb</td>
<td>3/19 13:43</td>
<td>1+07:41:29</td>
<td>R</td>
<td>0</td>
<td>1709.0</td>
<td>DIRAC_RpJK9Q_pilotwrapper.py</td>
</tr>
<tr>
<td>154066.0</td>
<td>nu_lhcb</td>
<td>3/19 13:43</td>
<td>1+07:41:31</td>
<td>R</td>
<td>0</td>
<td>1465.0</td>
<td>DIRAC_RpJK9Q_pilotwrapper.py</td>
</tr>
<tr>
<td>154088.0</td>
<td>nu_lhcb</td>
<td>3/19 14:09</td>
<td>1+07:14:33</td>
<td>R</td>
<td>0</td>
<td>1709.0</td>
<td>DIRAC_ekQezG_pilotwrapper.py</td>
</tr>
<tr>
<td>154091.0</td>
<td>nu_lhcb</td>
<td>3/19 14:09</td>
<td>1+07:14:32</td>
<td>R</td>
<td>0</td>
<td>1709.0</td>
<td>DIRAC_ekQezG_pilotwrapper.py</td>
</tr>
<tr>
<td>154258.0</td>
<td>nu_lhcb</td>
<td>3/19 17:36</td>
<td>1+03:37:18</td>
<td>R</td>
<td>0</td>
<td>1221.0</td>
<td>DIRAC_lIr4FB_pilotwrapper.py</td>
</tr>
</tbody>
</table>
```
Troubleshooting CE Jobs

```
# condor_ce_q -help status
[...]

JobStatus codes:
1 I IDLE
2 R RUNNING
3 X REMOVED
4 C COMPLETED
5 H HELD
6 > TRANSFERRING_OUTPUT
7 S SUSPENDED
```

See hold reasons with `condor_ce_q -held`
Common Hold Reasons

- **Spooling input data files**: the remote client is sending input files, should clear up after the transfer is complete

- **HTCondor-CE held job due to...**
  - **missing/expired user proxy**: job X.509 proxy was removed or expired. If the number of these failures is small, it’s safe to remove the job (pilots are cheap!)
  - **invalid job universe**: HTCondor-CE only routes vanilla, local, scheduler, and standard universe
  - **no matching routes, route job limit, or route failure threshold**: see 'HTCondor-CE Troubleshooting Guide': job sat in the queue for > 30 min without being picked up by the job router
    - No routes match the job:
      \[ \text{condor\_ce\_q <JOB ID> | condor\_ce\_job\_router\_info -match-jobs -ignore-prior-routing -jobads -} \]
    - All routes are full: \text{condor\_ce\_router\_q}
    - Route failure threshold: check the JobRouterLog or GridmanagerLog for local batch system submission failures. Fix the underlying issue and restart the HTCondor-CE services.
Troubleshooting Idle CE Jobs

- Jobs are indefinitely idle? Check for job router matching issues
  - For jobs still in the queue:
    
    ```bash
    # condor_ce_q -l <JOB-ID> | condor_ce_job_router_info -match-jobs
    -ignore-prior-routing -jobads -
    ```
  - For jobs that have left the queue:
    

- Wrap ClassAd expressions with the `debug()` function, evaluation details will appear in the JobRouterLog
- Ensure that you can submit jobs to your local batch system from the CE host
- Errors will appear in the JobRouterLog and the local SchedLog if there are communication issues between HTCondor-CE and the local HTCondor
Tracking HTCondor Batch System Jobs

1. Grid Job
2. Routed Job
3. HTCondor Negotiation

/var/log/condor-ce/SchedLog
/var/log/condor-ce/JobRouterLog
/var/log/condor/SchedLog
Tracking HTCondor Batch System Jobs

- Find the chain of job IDs using one of the following methods:
  - Query the CE schedd: `condor_ce_q -af RoutedToJobId <ORIGINAL JOB ID>`
  - Find relevant lines in the JobRouterLog
    09/17/14 15:00:57 JobRouter (src=86.0, dest=205.0, route=Local_Condor): claimed job
  - Query the local schedd: `condor_q -af RoutedFromJobId`
Tracking Non-HTCondor Batch System Jobs

1. Grid Job
2. Routed Job
3. Auth
4. qsub, sbatch, etc.

/var/log/condor-ce/GridmanagerLog.<user>
Tracking Non-HTCondor Batch System Jobs

- For non-HTCondor batch systems, find the batch system job ID:
  - Query the CE schedd routed job:
    $ condor-ce_q <ROUTED JOB ID> -af GridJobId
    <snip> lsf/20141206/482046
  - If the batch system jobs has completed, find relevant lines in the GridmanagerLog. Look for <BATCH SYSTEM>/<DATE>/<JOB ID>
    lsf/20141206/482046
  - Alternatively, for completed jobs with HTCondor >= 8.9.4:
    $ condor-ce_history <ROUTED JOB ID> -match 1 -af LastGridJobId
Troubleshooting the Gridmanager

If you see failures during the GM_SUBMIT phase, this means that the Batch GAHP/BLAHP is having issues submitting jobs to the local batch system

1. Verify that local job submission to the batch system works
2. Set the following in /usr/libexec/condor/glite/etc/batch_gahp.config:
   `blah_debug_save_submit_info=<DIR_NAME>`
   This saves generated submit files that HTCondor-CE uses for submission to<br/>&lt;DIR_NAME&gt;
3. Test local job submission with the generated file
4. Please report any issues found this way!

Troubleshooting the Gridmanager

A successful query of the local LSF batch system by the Gridmanager daemon

09/17/14 15:07:24 [25543] (87.0) gm state change: GM_SUBMITTED -> GM_POLL_ACTIVE
09/17/14 15:07:25 [25543] GAHP[25563] -> 'R'
09/17/14 15:07:25 [25543] GAHP[25563] -> 'S' '1'
09/17/14 15:07:25 [25543] GAHP[25563] -> '3' '0' 'No Error' '4' ' [ BatchjobId = "482046"; JobStatus = 4; ExitCode = 0; WorkerNode = "atl-prod08" ]'

Troubleshooting the Gridmanager

Routed job ID

09/17/14 15:07:24 [25543] (87.0) gm state change: GM_SUBMITTED -> GM_POLL_ACTIVE
09/17/14 15:07:25 [25543] GAHP[25563] -> 'R'
09/17/14 15:07:25 [25543] GAHP[25563] -> 'S' '1'
09/17/14 15:07:25 [25543] GAHP[25563] -> '3' '0' 'No Error' '4' '[ BatchjobId = "482046";
JobStatus = 4; ExitCode = 0; WorkerNode = "atl-prod08" ]'

Troubleshooting the Gridmanager

LSF job ID

09/17/14 15:07:24 [25543] (87.0) gm state change: GM_SUBMITTED -> GM_POLL_ACTIVE
09/17/14 15:07:25 [25543] GAHP[25563] -> 'R'
09/17/14 15:07:25 [25543] GAHP[25563] -> 'S' '1'
09/17/14 15:07:25 [25543] GAHP[25563] -> '3' '0' 'No Error' '4' '[ BatchjobId = "482046";
JobStatus = 4; ExitCode = 0; WorkerNode = "atl-prod08" ]'

Troubleshooting the Gridmanager

If there are issues, errors should show up here. If the messages do not provide enough information, run the BLAHP commands by hand:

```
/usr/libexec/condor/glite/bin/lsf_status.sh lsf/20140917/482046
```

09/17/14 15:07:24 [25543] (87.0) gm state change: GM_SUBMITTED -> GM_POLL_ACTIVE
09/17/14 15:07:25 [25543] GAHP[25563] -> 'R'
09/17/14 15:07:25 [25543] GAHP[25563] -> 'S' '1'
09/17/14 15:07:25 [25543] GAHP[25563] -> '3' '0' 'No Error' '4' "[ BatchjobId = "482046"; JobStatus = 4; ExitCode = 0; WorkerNode = "atl-prod08" ]"

Please report any issues found this way!
Additional Resources

- Troubleshooting Guide

- Still have question, issues, or comments?
  htcondor-users@cs.wisc.edu