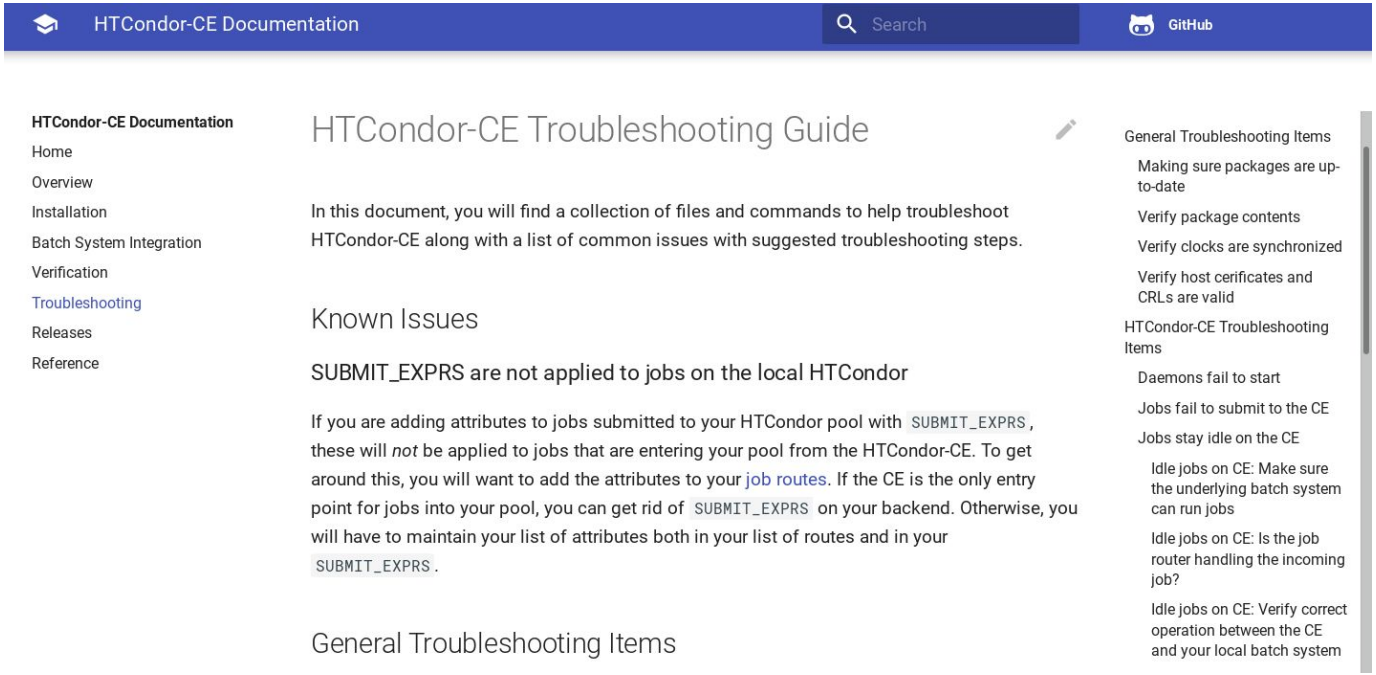
HTCondor-CE: Troubleshooting

HTCondor Workshop 2019 - EU Joint Research Centre

Brian Lin
University of Wisconsin — Madison

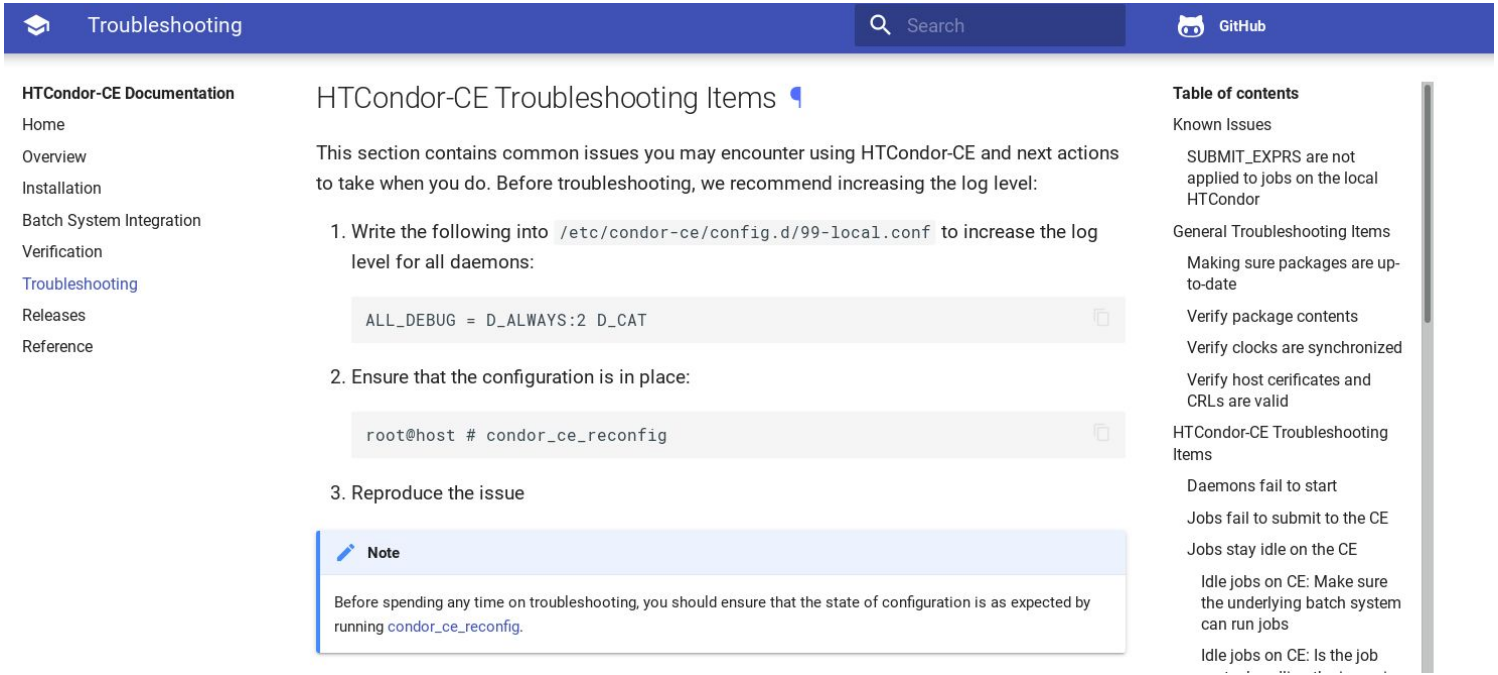
Troubleshooting HTCondor-CE

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



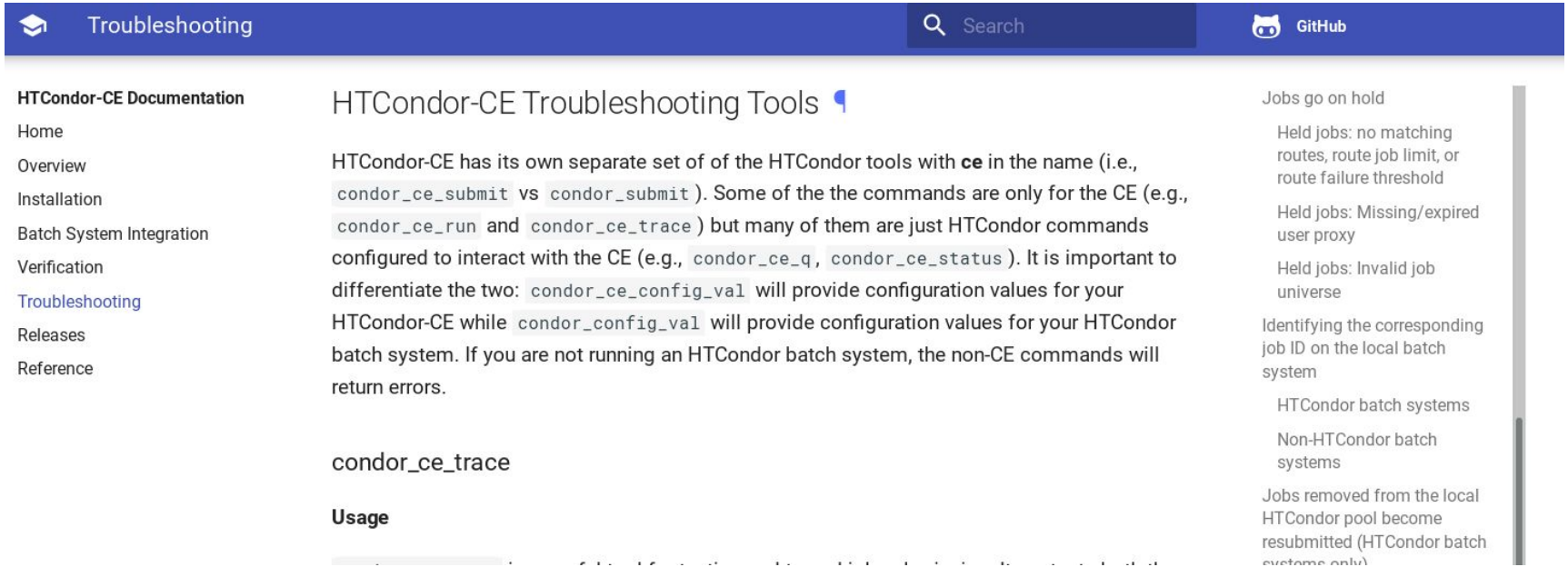
Troubleshooting HTCondor-CE

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



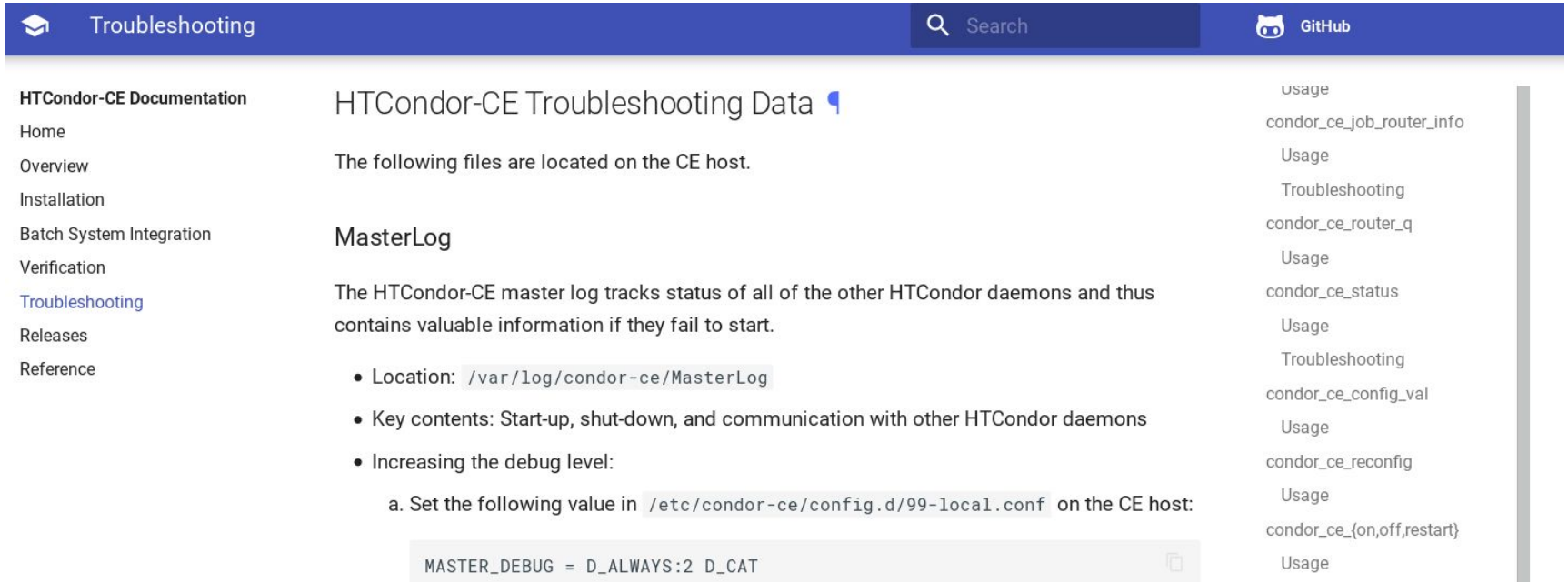
Troubleshooting HTCondor-CE

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



Troubleshooting HTCondor-CE

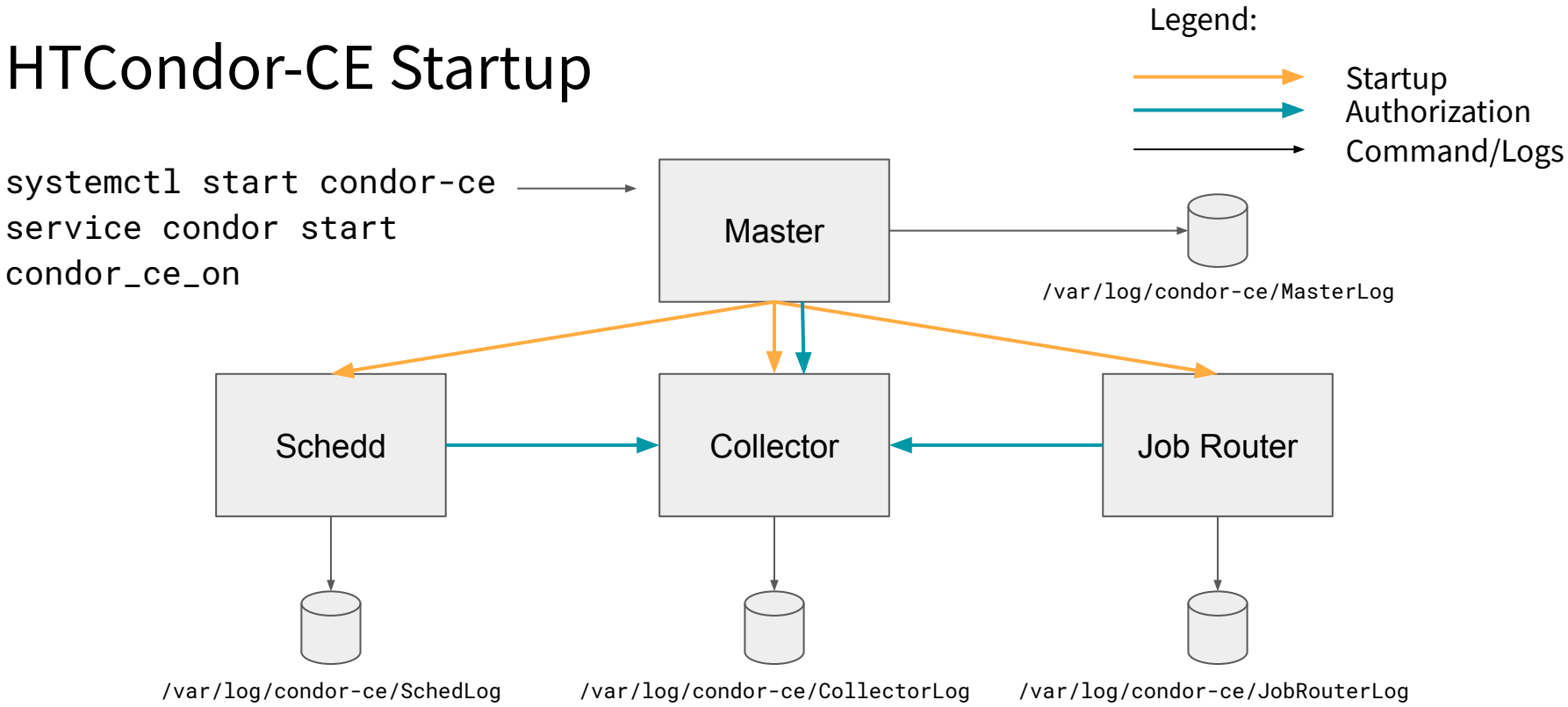
https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#htcondor-ce-troubleshooting-data



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



Troubleshooting Startup

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

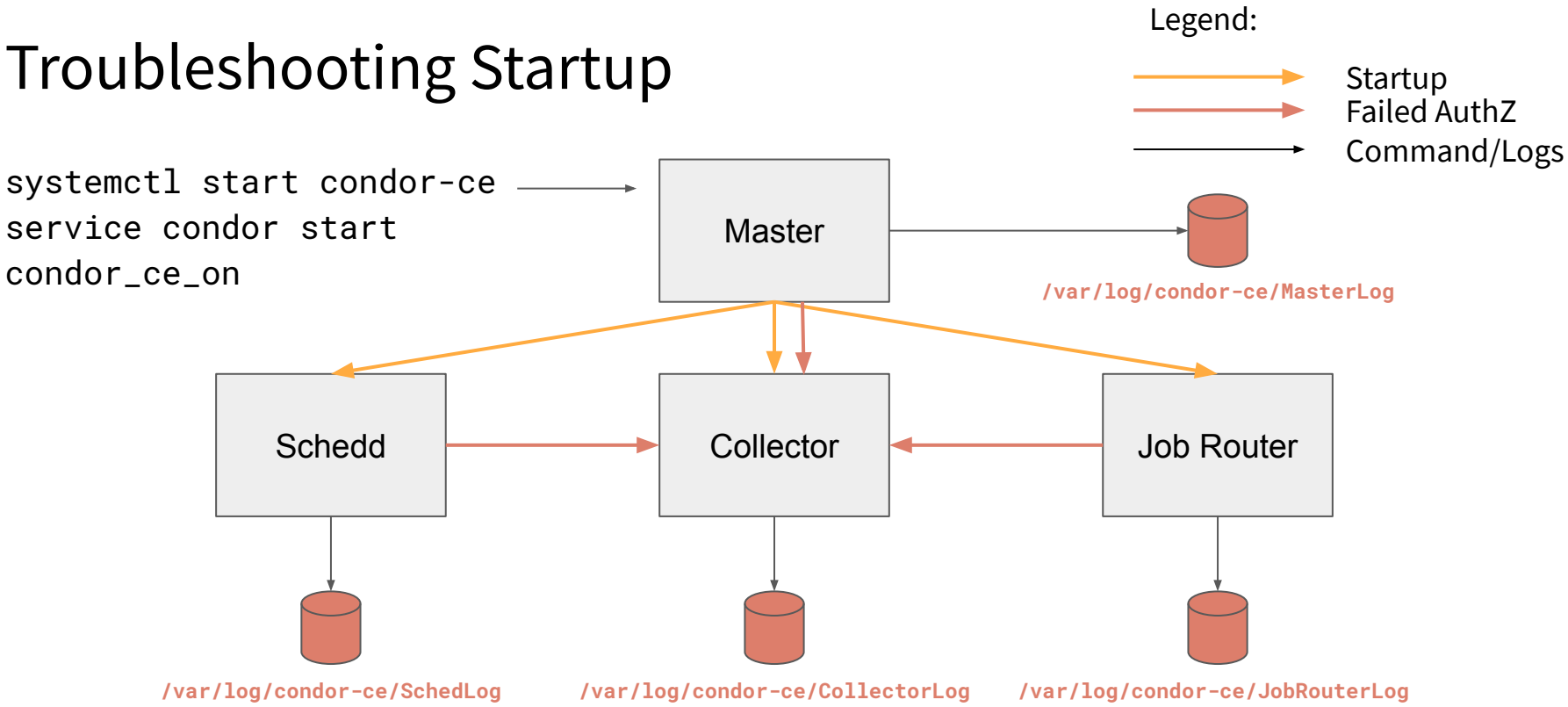
```
# condor_ce_status -any
MyType              TargetType          Name

Collector           None                My Pool - fermicloud068.fnal.gov@fermiclo
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



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
   `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!

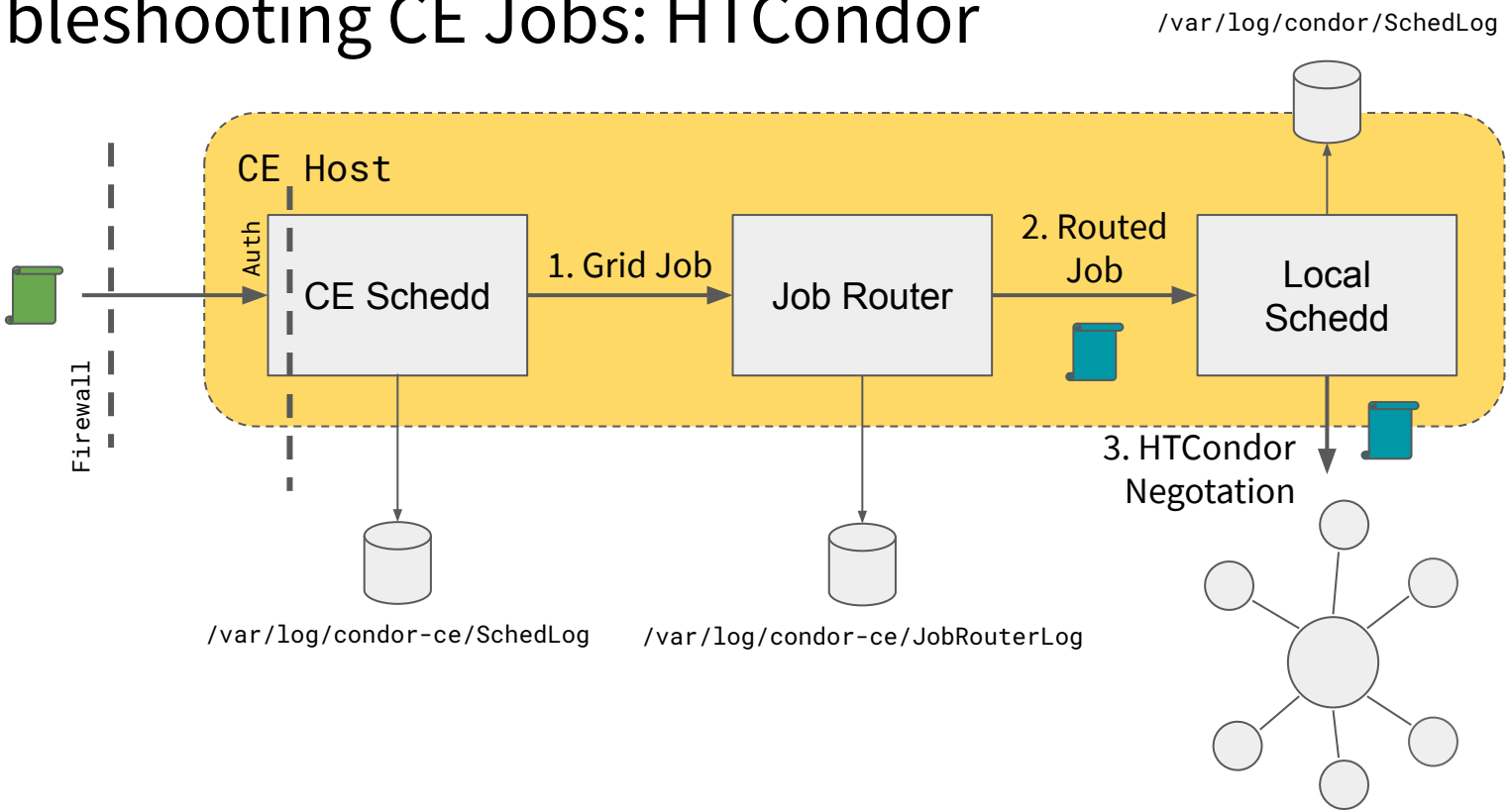
Verifying HTCondor-CE Job Submission

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

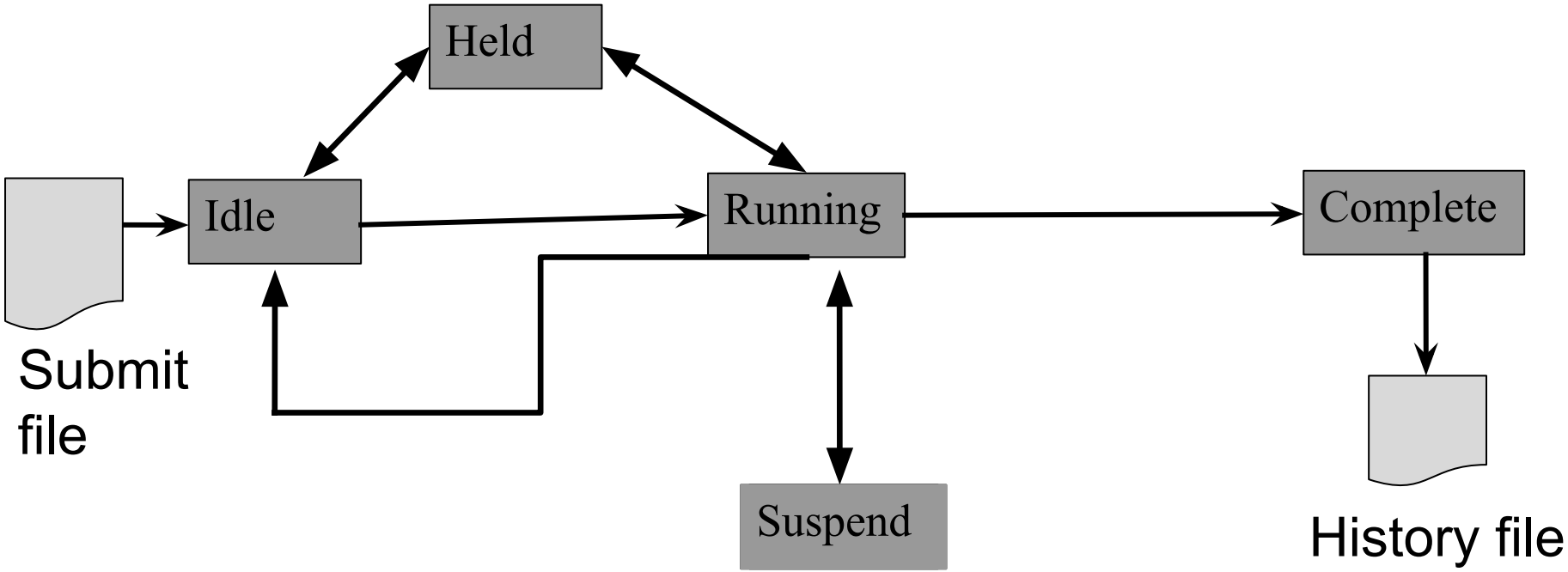


Troubleshooting the CE Schedd

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

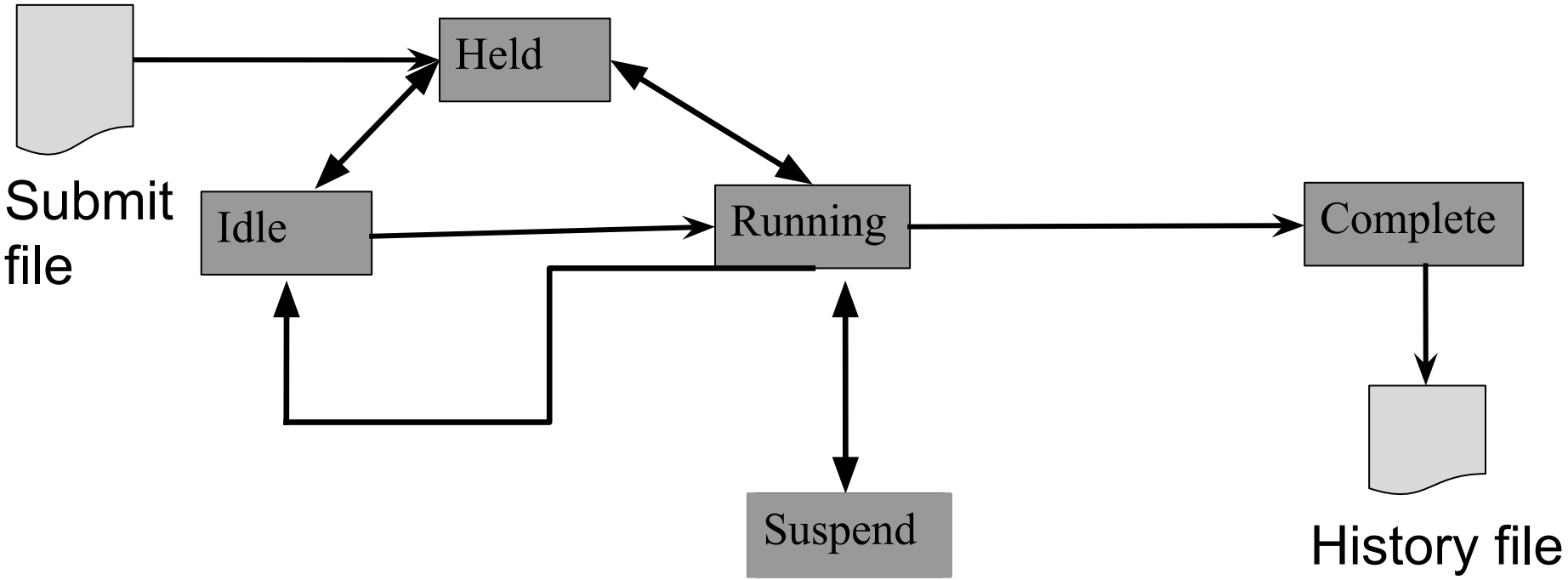
Troubleshooting CE Jobs

For a regular HTCondor pool:



Troubleshooting CE Jobs

For HTCondor-CEs (i.e. Condor-G submission)



Troubleshooting CE Jobs

```
# condor_ce_q -nobatch

-- Schedd: lhcb-ce.chtc.wisc.edu : <128.104.100.65:9618?... @ 03/20/19 21:31:19
 ID       OWNER          SUBMITTED      RUN_TIME ST PRI SIZE   CMD
153501.0   nu_lhcb       3/18 13:30   2+07:56:31 R  0  733.0 DIRAC_clpM0A_pilotwrapper.py
154043.0   nu_lhcb       3/19 13:43   1+07:41:29 R  0   1709.0 DIRAC_RpJK9Q_pilotwrapper.py
154066.0   nu_lhcb       3/19 13:43   1+07:41:31 R  0   1465.0 DIRAC_RpJK9Q_pilotwrapper.py
154088.0   nu_lhcb       3/19 14:09   1+07:14:33 R  0   1709.0 DIRAC_ekQezG_pilotwrapper.py
154091.0   nu_lhcb       3/19 14:09   1+07:14:32 R  0   1709.0 DIRAC_ekQezG_pilotwrapper.py
154258.0   nu_lhcb       3/19 17:36   1+03:37:18 R  0   1221.0 DIRAC_lIr4FB_pilotwrapper.py
```

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`

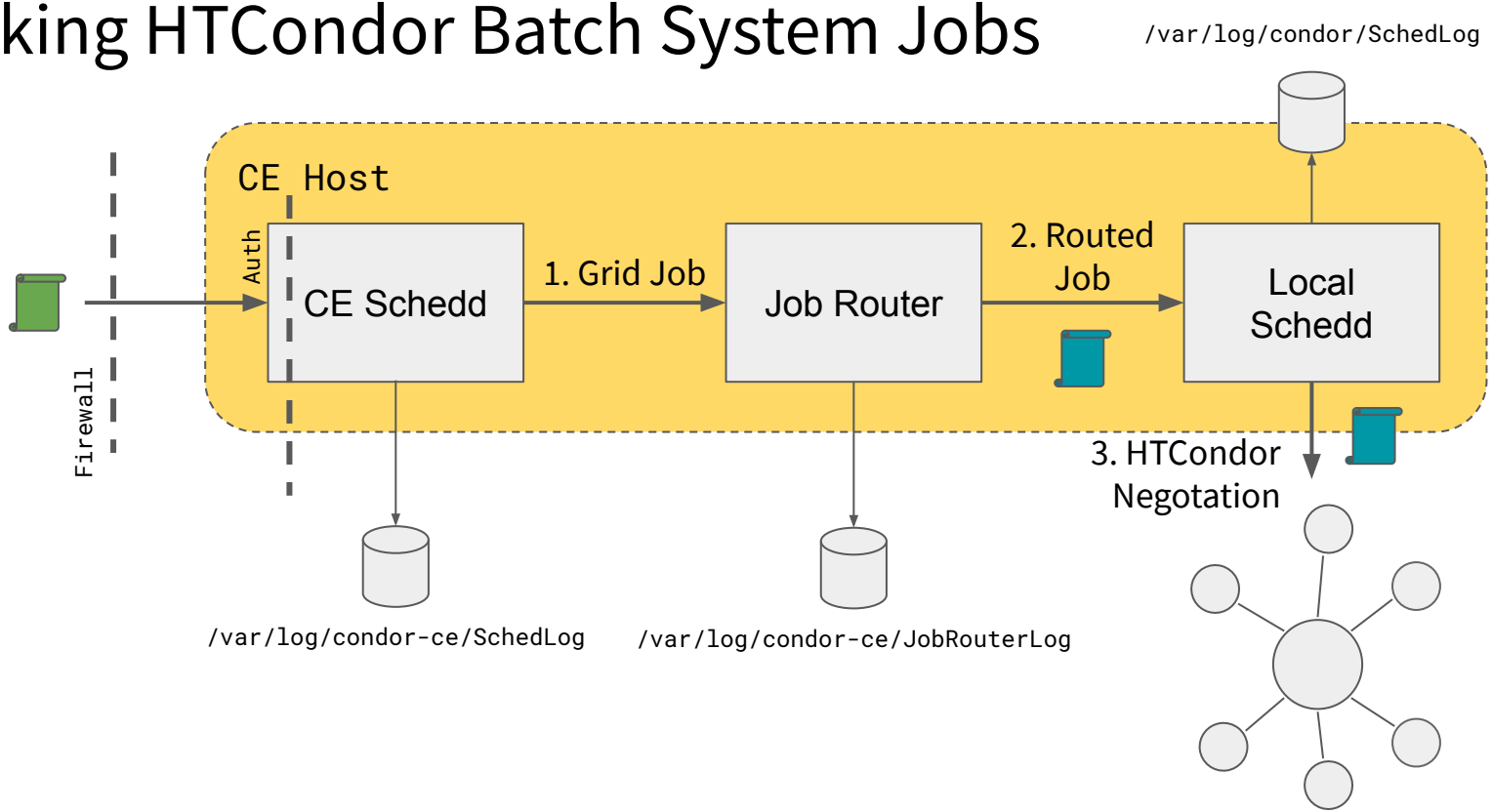
Common Hold Reasons

- **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:
      `condor_ce_q <JOB ID> | condor_ce_job_router_info -match-jobs \ -ignore-prior-routing -jobads -`
    - All routes are full: `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:
    `# condor_ce_q -l <JOB-ID> | condor_ce_job_router_info -match-jobs -ignore-prior-routing -jobads -`
  - For jobs that have left the queue:
    https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#condor_ce_job_router_info
- 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

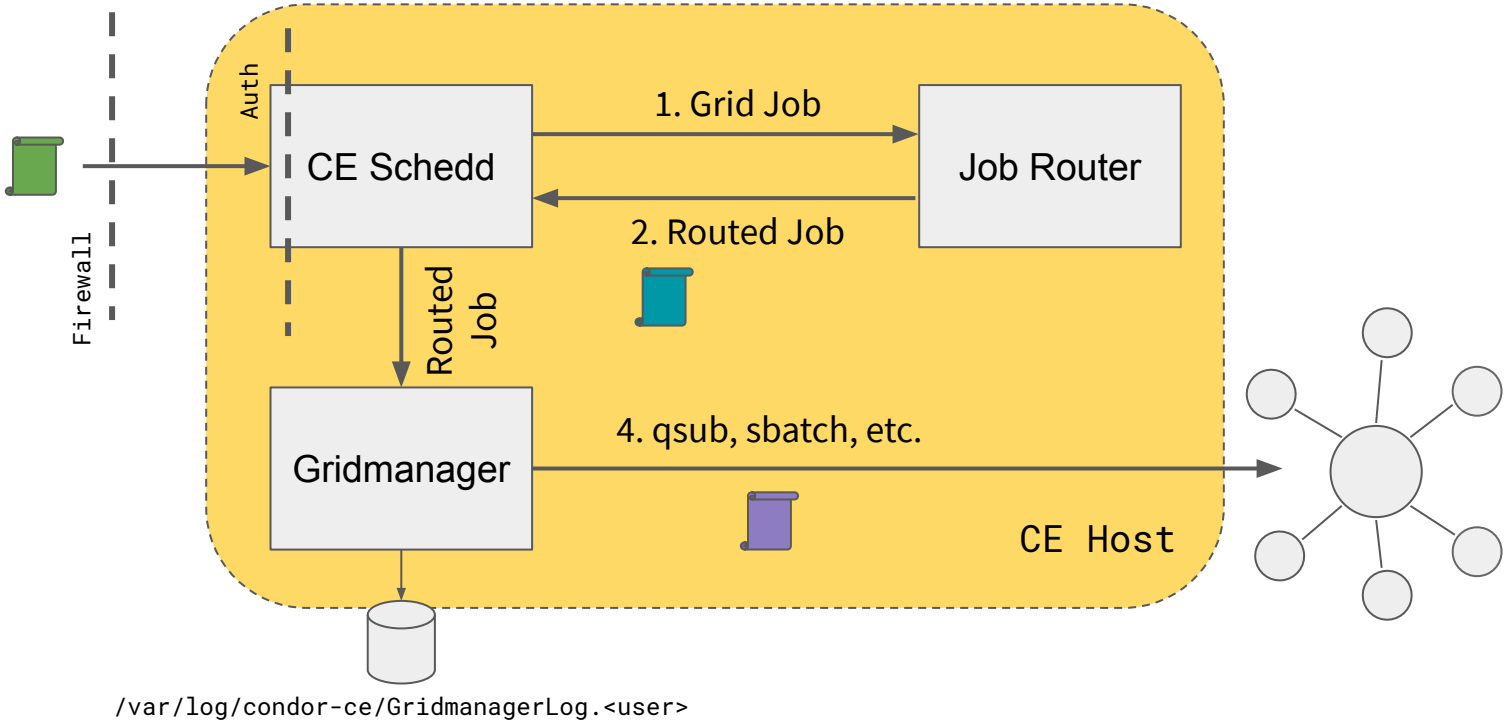
Tracking HTCondor Batch System Jobs



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`
- https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#htcondor-batch-systems

Tracking Non-HTCondor Batch System Jobs



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
    ```
- https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#non-htcondor-batch-systems

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 <DIR_NAME>
3. Test local job submission with the generated file
4. Please report any issues found this way!

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#for-non-htcondor-batch-systems

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 15:07:24 [25543] GAHP[25563] <- 'BLAH_JOB_STATUS 3 lsf/20140917/482046'
09/17/14 15:07:24 [25543] GAHP[25563] -> 'S'
09/17/14 15:07:25 [25543] GAHP[25563] <- 'RESULTS'
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" ]'
```

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#for-non-htcondor-batch-systems

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 15:07:24 [25543] GAHP[25563] <- 'BLAH_JOB_STATUS 3 lsf/20140917/482046'
09/17/14 15:07:24 [25543] GAHP[25563] -> 'S'
09/17/14 15:07:25 [25543] GAHP[25563] <- 'RESULTS'
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" ]'
```

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#for-non-htcondor-batch-systems

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 15:07:24 [25543] GAHP[25563] <- 'BLAH_JOB_STATUS 3 lsf/20140917/482046'
09/17/14 15:07:24 [25543] GAHP[25563] -> 'S'
09/17/14 15:07:25 [25543] GAHP[25563] <- 'RESULTS'
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" ]'
```

https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/#for-non-htcondor-batch-systems

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 15:07:24 [25543] GAHP[25563] <- 'BLAH_JOB_STATUS 3 lsf/20140917/482046'
09/17/14 15:07:24 [25543] GAHP[25563] -> 'S'
09/17/14 15:07:25 [25543] GAHP[25563] <- 'RESULTS'
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
  https://htcondor-ce.readthedocs.io/en/latest/troubleshooting/troubleshooting/
- Still have question, issues, or comments?
  htcondor-users@cs.wisc.edu