HTCondor-CE: Configuration

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

Brian Lin
University of Wisconsin — Madison

HTCondor Configuration Refresher

- HTCondor-CE configuration lives in /etc/condor-ce/
    - Main configuration file in `/etc/condor-ce/condor_config`
    - Add local configuration to `/etc/condor-ce/config.d/` (files processed in lexicographic order)
- `condor_ce_config_val` to inspect config values
- `condor_ce_reconfig` to apply new configuration
- ...or restart the condor-ce service just to be sure

Authentication and Authorization

- Authentication can be configured via the HTCondor-CE unified mapfile `/etc/condor-ce/condor_mapfile`
    - One mapping per line with the following format:
    `<AUTH METHOD> <AUTH NAME> <HTCONDOR PRINCIPLE>`
    - Auth names supports perl-compatible regular expressions
    - Selected mapping is determined by first-match
- HTCondor principles (`<USERNAME>@<DOMAIN>`) determine authorization level
    - <hostname>@daemon.htcondor.org: authorized as a daemon
    - .*@users.htcondor.org: authorized to submit jobs
    - GSS_ASSIST_GRIDMAP: a special value telling HTCondor-CE to call out to another service for user mapping, e.g. LCMAPS, Argus
- https://htcondor-ce.readthedocs.io/en/latest/installation/htcondor-ce/#configuring-authentication

Authentication and Authorization

An example from our HTCondor-CE, lhcb-ce.chtc.wisc.edu:

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

Authentication and Authorization

Authentication method:

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

GSI is the default auth method for remote clients

Authentication and Authorization

The "authentication name":

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

 In this case, this is the subject distinguished name (DN) of the host certificate:

EC-JRC - HTCondor-CE: Configuration

Authentication and Authorization

The HTCondor principle:

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

Authentication and Authorization

This line allows all the HTCondor-CE daemons to authenticate with each other.

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

NOTE: HTCondor CE 4 will use filesystem authN, eliminating the need for these types of lines!

Authentication and Authorization

Mapping via VOMS FQANs are possible in the authenticated name

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

- `<SUBJECT DN>,<VOMS FQAN 1>,...,<VOMS FQAN N>`
- This line maps all X.509 credentials with an LHCb primary VOMS FQAN to the `nu_lhcb` user

Authentication and Authorization

Explicit mapping for a single user:

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

Authentication and Authorization

Callout to external service:

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

via `/etc/grid-security/gsi-authz.conf`, containing the following for LCMAPS:

```
globus_mapping liblcas_lcmaps_gt4_mapping.so lcmaps_callout
```

Or for Argus:

```
globus_mapping /usr/lib64/libgsi_pep_callout.so argus_pep_callout
```

Authentication and Authorization

Unauthenticated fallbacks

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

Authentication and Authorization

Finally, map local accounts to themselves.

```
GSI "/DC=org/DC=incommon/<snip>/CN=lhcb-ce.chtc.wisc.edu" lhcb-ce.chtc.wisc.edu@daemon.htcondor.org
GSI ".*,/lhcb/Role=pilot/Capability=.*" nu_lhcb@users.htcondor.org
GSI "/DC=org/DC=cilogon/<snip>/CN=Brian Lin A106521" blin@users.htcondor.org
GSI (.*) GSS_ASSIST_GRIDMAP
GSI "(/CN=[-.A-Za-z0-9/= ]+)" \1@unmapped.htcondor.org
CLAIMTOBE .* anonymous@claimtobe
FS (.*) \1
```

The UID_DOMAIN (@users.htcondor.org) is automatically appended, i.e. they have submit privileges

HTCondor Configuration

- For HTCondor batch systems, you must set
    - `JOB_ROUTER_SCHEDD2_POOL` - the <HOST>:<PORT> of the local HTCondor central manager
- If you've changed the default location of the batch HTCondor's SPOOL directory
    - `JOB_ROUTER_SCHEDD2_SPOOL` - location of the local SPOOL directory (`condor_config_val SPOOL`)

https://htcondor-ce.readthedocs.io/en/latest/installation/htcondor-ce/#configuring-the-batch-system

Non-HTCondor Configuration

Configure the Batch GAHP (a.k.a. BLAHP) via
`/usr/libexec/condor/glite/etc/batch_gahp.config`

1. Disable Batch GAHP delegation of proxy certificates:
   `blah_disable_wn_proxy_renewal=yes`
   `blah_delegate_renewed_proxies=no`
   `blah_disable_limited_proxy=yes`
2. If your batch system tools exist outside of `/usr/bin`, also edit `*_binpath` (e.g., `slurm_binpath=/opt/slurm/bin`). **NOTE** no spaces around the '='!

https://htcondor-ce.readthedocs.io/en/latest/installation/htcondor-ce/#configuring-the-batch-system

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 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`

Information Services

- An HTCondor-CE Central Collector requires no extra configuration, central grid operations can just install the `htcondor-ce-collector` package and start the `condor-ce-collector` service!
- To report to a central collector, specify the hostname and port. For example, `CONDOR_VIEW_HOST = collector.opensciencegrid.org:9619`
- Advertise Schedd ads to the Central Collector (the default in HTCondor-CE 4): `CONDOR_VIEW_CLASSAD_TYPES = Scheduler`
- Add arbitrary attributes to the Schedd ad reported to the Central Collector:
  `FOO = "Bar"`
  `SCHEDD_ATTRS = $(SCHEDD_ATTRS) FOO`

Configuring Job Routes

Job Router Configuration

- Declare your site policy
- Each route is described with ClassAds
- Job routes are constructed by combining each entry in JOB_ROUTER_ENTRIES with the JOB_ROUTER_DEFAULTS
- https://htcondor-ce.readthedocs.io/en/latest/batch-system-integration/

```
$ condor_ce_job_router_info
-config
Route 1
Name        : "Local_Condor"
Universe    : 5
MaxJobs     : 10000
MaxIdleJobs : 2000
GridResource :
Requirements : true
ClassAd     :
    [
[...]
```

Job Router Matching

- Each job that matches is compared to each job route's requirements expression (`Requirements = True` by default) in order
    - Currently, this is hash table order (i.e. "random" but fixed) instead of the order in which the entries are configured
    - Aiming for a fix in HTCondor 8.8.6:
      http://condor-wiki.cs.wisc.edu/index.cgi/tktview?tn=7284
- To use round-robin matching behavior, set the following in your configuration (not within the routes):
  `JOB_ROUTER_ROUND_ROBIN_SELECTION = True`

EC-JRC - HTCondor-CE: Configuration

Job Router ClassAds

Special job route functions are used to transform jobs, evaluated in the following order.

1. Copy an attribute from the original job ad to the routed job ad:
    ```
    copy_foo = "original_foo";
    ```
2. Delete an attribute from the original job ad from the routed job ad:
    ```
    delete_foo = True;
    ```
3. Set an attribute in the routed job ad to a value or expression
    ```
    set_requirements = (OpSys == "LINUX");
    ```
4. Set an attribute in the routed job ad to value that is evaluated in the context of the original job ad.
    ```
    eval_set_Experiment = strcat("cms.", Owner);
    ```

set_ vs eval_set_

In a job route, `set_Experiment = strcat("cms.", Owner);`
    →`Experiment = strcat("cms.", Owner)` in the routed job

Use `eval_set_*` if you just want the value of an expression in the routed job.
In a job route, `Owner = blin; eval_set_Experiment = strcat("cms.", Owner);`
    →`Experiment = cms.blin` in the routed job

set_ vs eval_set_

Use `set_*` for any of the following scenarios:

-   To use ClassAd expressions in the routed job on your HTCondor batch system
-   For non-HTCondor batch systems to `set_remote_cerequirements` or default resource requests
-   To set intermediate variables for multiple `eval_set_*` statements

EC-JRC - HTCondor-CE: Configuration

set_ vs eval_set_

```
set_VONameExpr =
    ifThenElse(isUndefined(X509UserProxyVOName),"LOCAL",X509UserProxyVOName);
set_AcctSubGroup = toUpper(
  ifThenElse(isUndefined(X509UserProxyFirstFQAN),
    "noproxy",
    ifThenElse(regexp("production",X509UserProxyFirstFQAN),
      strcat(VOName,"prd"),
      ifThenElse(regexp("pilot",X509UserProxyFirstFQAN),
        strcat(VOName,"plt"),
        ifThenElse(regexp("lcgadmin",X509UserProxyFirstFQAN),
          strcat(VOName,"lcg"),
          "other")))));
eval_set_AccountingGroup =
  strcat("group_u_",VONameExpr,".","grid_",AcctSubGroup,".",OWNER);
eval_set_VOName = VONameExpr;
```

Setting Default Resource Requests

Use `set_*` for the following resource requests

- `default_xcount` to set the default number of cores
- `default_maxMemory` to set the default maximum memory (in MB)
- `default_maxWalltime` to set the default maximum walltime (in minutes)
- `default_queue` to set the default batch system queue (non-HTCondor only)

EC-JRC - HTCondor-CE: Configuration

Job Router Defaults

HTCondor-CE automatically generates JOB_ROUTER_DEFAULTS. Modify existing attributes at your own risk but feel free to add to it!

For HTCondor:

```
JOB_ROUTER_DEFAULTS @=jrd
$(JOB_ROUTER_DEFAULTS)
[
    # Route jobs to an HTCondor batch system
    TargetUniverse = 5;
]
@jrd
```

For Slurm:

```
JOB_ROUTER_DEFAULTS @=jrd
$(JOB_ROUTER_DEFAULTS)
[
    # Route jobs to a Slurm batch system
    GridResource = "batch slurm"
]
@jrd
```

EC-JRC - HTCondor-CE: Configuration

Job Router Entries

- Use the multiline config syntax
- Each route is enclosed by [ ]
- Each route requires a `Name` attr
- Since we're using ClassAds, we can use ClassAd functions!
- TARGET ensures that the job attribute is used to match the route

```
JOB_ROUTER_ENTRIES @=jre
[
  Name = "atlas_mcore";
  TargetUniverse = 5;
  Requirements = regexp("^usatlas",
TARGET.Owner);
  set_default_xcount = 8;
]
[
  Name = "everything_else";
  TargetUniverse = 5;
  set_default_xcount = 1;
]
@jre
```

Job Router Entries

- Use the multiline config syntax
- Each route is enclosed by [ ]
- Each route requires a Name attr
- Since we're using ClassAds, we can use ClassAd functions!
- TARGET ensures that the job attribute is used to match the route

```
JOB_ROUTER_ENTRIES @=jre
[
  Name = "atlas_mcore";
  TargetUniverse = 5;
  Requirements = regexp("^usatlas",
TARGET.Owner);
  set_default_xcount = 8;
]
[

  Name = "everything_else";
  TargetUniverse = 5;
  set_default_xcount = 1;
]
@jre
```

Job Router Entries

- Use the multiline config syntax
- Each route is enclosed by [ ]
- Each route requires a `Name` attr
- Since we're using ClassAds, we can use ClassAd functions!
- TARGET ensures that the job attribute is used to match the route

```
JOB_ROUTER_ENTRIES @=jre
[
  Name = "atlas_mcore";
  TargetUniverse = 5;
  Requirements = regexp("^usatlas",
TARGET.Owner);
  set_default_xcount = 8;
]
[
  Name = "everything_else";
  TargetUniverse = 5;
  set_default_xcount = 1;
]
@jre
```

Job Router Entries

- Use the multiline config syntax
- Each route is enclosed by [ ]
- Each route requires a `Name` attr
- Since we're using ClassAds, we can use ClassAd functions!
- TARGET ensures that the job attribute is used to match the route

```
JOB_ROUTER_ENTRIES @=jre
[
  Name = "atlas_mcore";
  TargetUniverse = 5;
  Requirements = regexp("^usatlas",
TARGET.Owner);
  set_default_xcount = 8;
]
[
  Name = "everything_else";
  TargetUniverse = 5;
  set_default_xcount = 1;
]
@jre
```

EC-JRC - HTCondor-CE: Configuration

Job Router Entries

- Use the multiline config syntax
- Each route is enclosed by [ ]
- Each route requires a `Name` attr
- Since we're using ClassAds, we can use ClassAd functions!
- TARGET ensures that the job attribute is used to match the route

```
JOB_ROUTER_ENTRIES @=jre
[
  Name = "atlas_mcore";
  TargetUniverse = 5;
  Requirements = regexp("^usatlas",
TARGET.Owner);
  set_default_xcount = 8;
]
[
  Name = "everything_else";
  TargetUniverse = 5;
  set_default_xcount = 1;
]
@jre
```

EC-JRC - HTCondor-CE: Configuration

HTCondor-Specific Route Configuration

A common use of set_* for HTCondor batch systems is to set periodic expressions or Requirements on the routed job

```
JOB_ROUTER_ENTRIES @=jre
[
  name = "Setting HTCondor expressions";
  # Puts the routed job on hold if the job has tried to start more than once
  set_Periodic_Hold = (NumJobStarts > 1);
  # Remove routed jobs if their walltime is longer than 3 days and 5 minutes
  set_Periodic_Remove = (RemoteWallClockTime > (3*24*60*60 + 5*60));
  # Release routed jobs if the condor_starter couldn't start the executable
  set_Periodic_Release = (HoldReasonCode == 6);
  # Routed jobs require worker nodes that have Gluster
  set_Requirements = (HasGluster == True)
]
@jre
```

Non-HTCondor Route Configuration

For non-HTCondor batch system directives not covered, there's default_remote_cerequirements:

```
set_default_remote_cerequirements = strcat("Walltime == 3600 && AccountingGroup =="",
x509UserProxyFirstFQAN, "\"");
```

Results in $Walltime and $AccountingGroup shell variables that can be used in the relevant /usr/libexec/condor/glite/bin/*_local_submit_attributes.sh for your batch system. An example PBS scrip, whose output is appended to the PBS JDL:

```
#!/bin/bash

echo "#PBS -l walltime=$Walltime"
echo "#PBS -A $AccountingGroup"
```

EC-JRC - HTCondor-CE: Configuration

Non-HTCondor Route Configuration

In HTCondor-CE 4, this becomes much simpler!

```
[
  Name = "New CERequirements format";
  set_WallTime = 3600;
  set_AccountingGroup = x509UserProxyFirstFQAN;
  set_default_CERequirements = "Walltime,AccountingGroup";
]
```

Can be used with the same `pbs_local_submit_attributes.sh` script

Additional Resources

- Documentation
    - https://htcondor-ce.readthedocs.io/en/latest/installation/htcondor-ce/#configuring-htcondor-ce
    - https://htcondor-ce.readthedocs.io/en/latest/batch-system-integration/
- Questions or comments? htcondor-users@cs.wisc.edu

EC-JRC - HTCondor-CE: Configuration