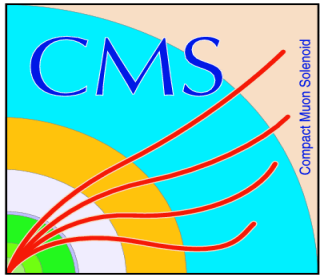
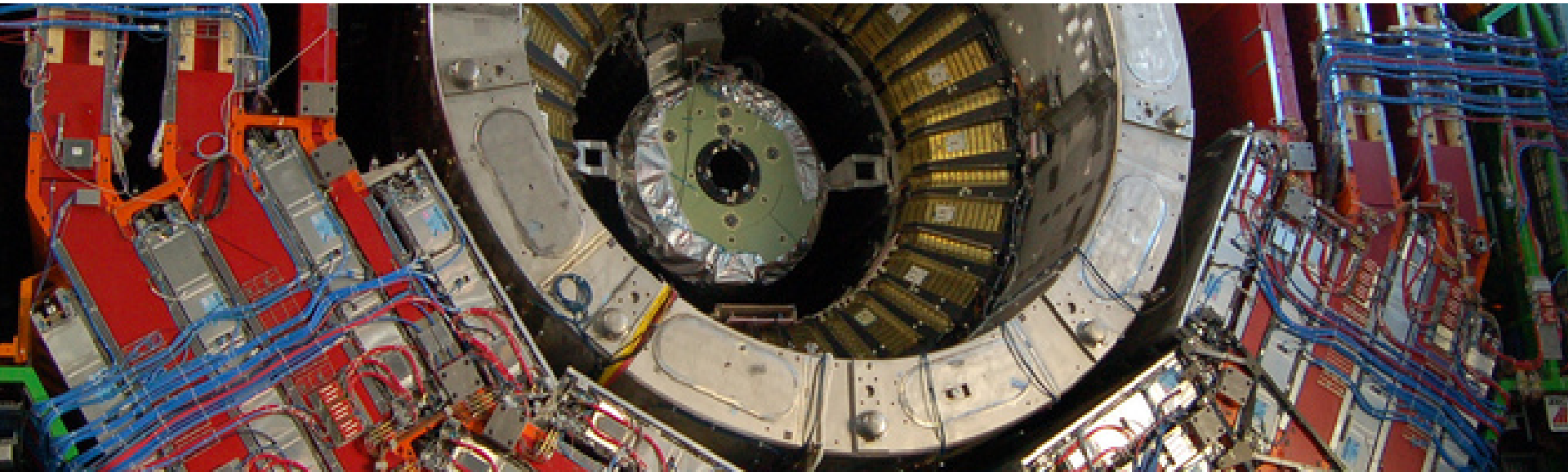


Open Science Grid

HTCondor as the job gateway: HTCondor-CE

Brian Lin, Marian Zvada

HTCondor-CE Support Training
Indianapolis, October 20–21, 2014



Introducing the CE concept

- At the heart of the OSG Compute Element* (CE) is the job gateway software. The job gateway software anchors three core pieces of functionality:

*Compute Element: job gateway to a cluster, also called gatekeeper.

Introducing the CE concept

- At the heart of the OSG Compute Element* (CE) is the job gateway software. The job gateway software anchors three core pieces of functionality:
- Remote access
  - The job gateway provides a network service that remote clients can contact and interact with
- Authentication and authorization
  - The job gateway is responsible for authenticating the client and deciding on what actions it is authorized to perform.
- Resource allocation
  - The job gateway accepts an abstract description of a resource to allocate and actualizes the resource request within the local environment.

*Compute Element: job gateway to a cluster, also called gatekeeper.

HTCondor-CE introduction

- Currently, Globus GRAM provides the abstraction, sandbox movement, and remote submission layers for the OSG-CE.

HTCondor-CE introduction

- Currently, Globus GRAM provides the abstraction, sandbox movement, and remote submission layers for the OSG-CE.

- Why HTCondor-CE?

  - With the HTCondor team, the OSG has been working to provide an alternate job gateway implementation, the HTCondor-CE.
  - The HTCondor-CE is a special configuration of the HTCondor software which provides the three core pieces of functionality described previously.

HTCondor-CE overview (I)

- Special configuration of HTCondor
- Sits on CE* of each cluster (schedd)
- Allows:
  - Remote job submission and management
  - Strong authentication (GSI/VOMS)
  - Logging and monitoring
  - Scalability
  - Work with (most) existing batch systems

HTCondor-CE overview (II)

- HTCondor-CE provides (based on three fundamentals of the CE concept):
- Remote access
  - Based on the internal CEDAR protocol.
  - CEDAR provides a RPC and messaging mechanism over UDP or TCP, and can provide various levels of integrity or encryption based upon the session parameters.
- Authentication and authorization
  - Based on Globus libraries for GSI and authorization callout.
- Resource allocation
  - Grid jobs are taken and transformed to local jobs using the JobRouter component.
    - Any software HTCondor can interact with is a potential backend. This includes EC2, OpenStack, or even another HTCondor-CE!

HTCondor-CE Building Blocks

- HTCondor-C
  - Submit jobs from one HTCondor scheduler to another (submit machine to CE)
- Job Router
  - Transform jobs (localize jobs at CE)
- BLAHP
  - Submit jobs to non-HTCondor batch systems (PBS, SGE, SLURM, etc.)
  - blahp is the executable which then calls, for example, qstat / qsub / qdel.
  - blahp has another layer of customization if, for example, you need to tweak qsub arguments. Most useful things can be done via the JobRouter transform.

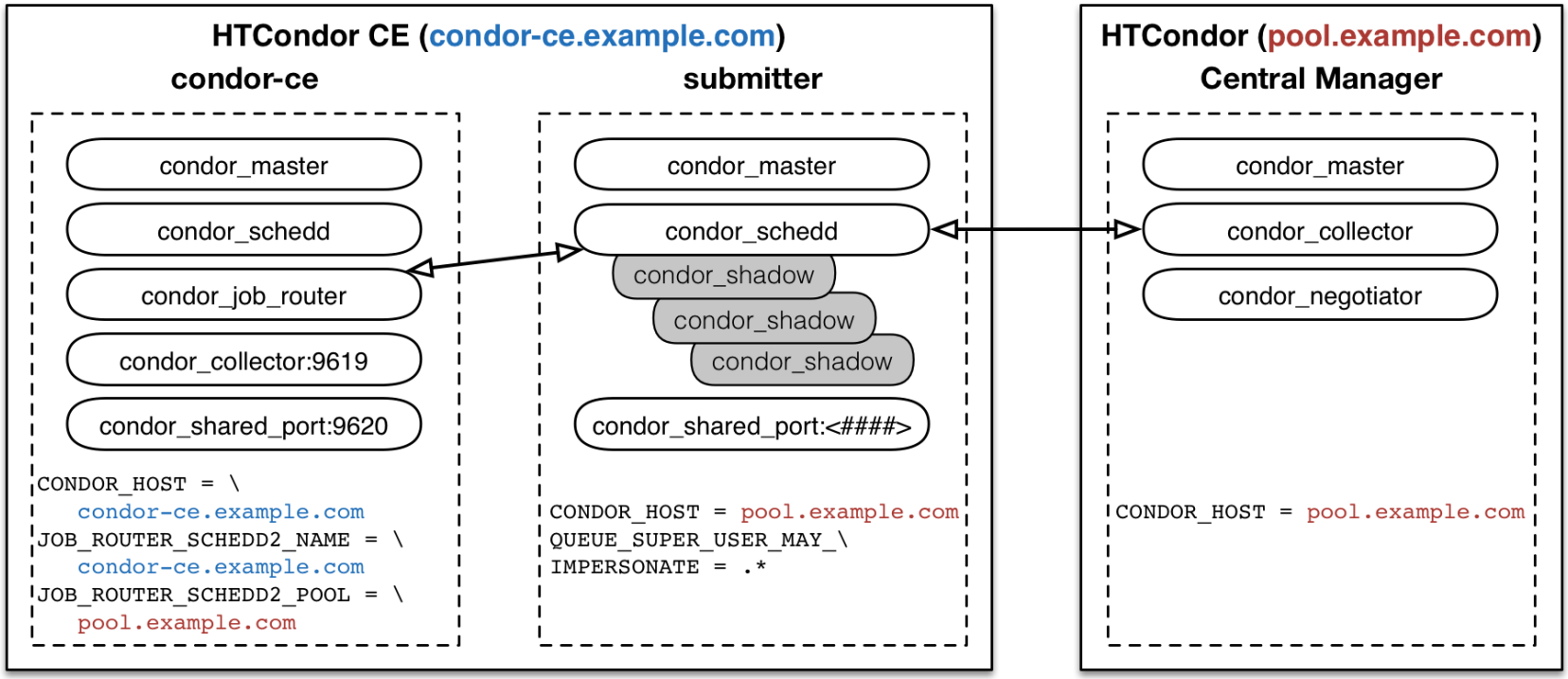
HTCondor-CE Building Blocks

- HTCondor-C
  - Submit jobs from one HTCondor scheduler to another (submit machine to CE)
- Job Router
  - Transform jobs (localize jobs at CE)
- BLAHP
  - Submit jobs to non-HTCondor batch systems (PBS, SGE, SLURM, etc.)
  - blahp is the executable which then calls, for example, qstat / qsub / qdel.
  - blahp has another layer of customization if, for example, you need to tweak qsub arguments. Most useful things can be done via the JobRouter transform.

- HOW IT WORKS?
  - Submit workflow for the HTCondor-CE running on the site with the:

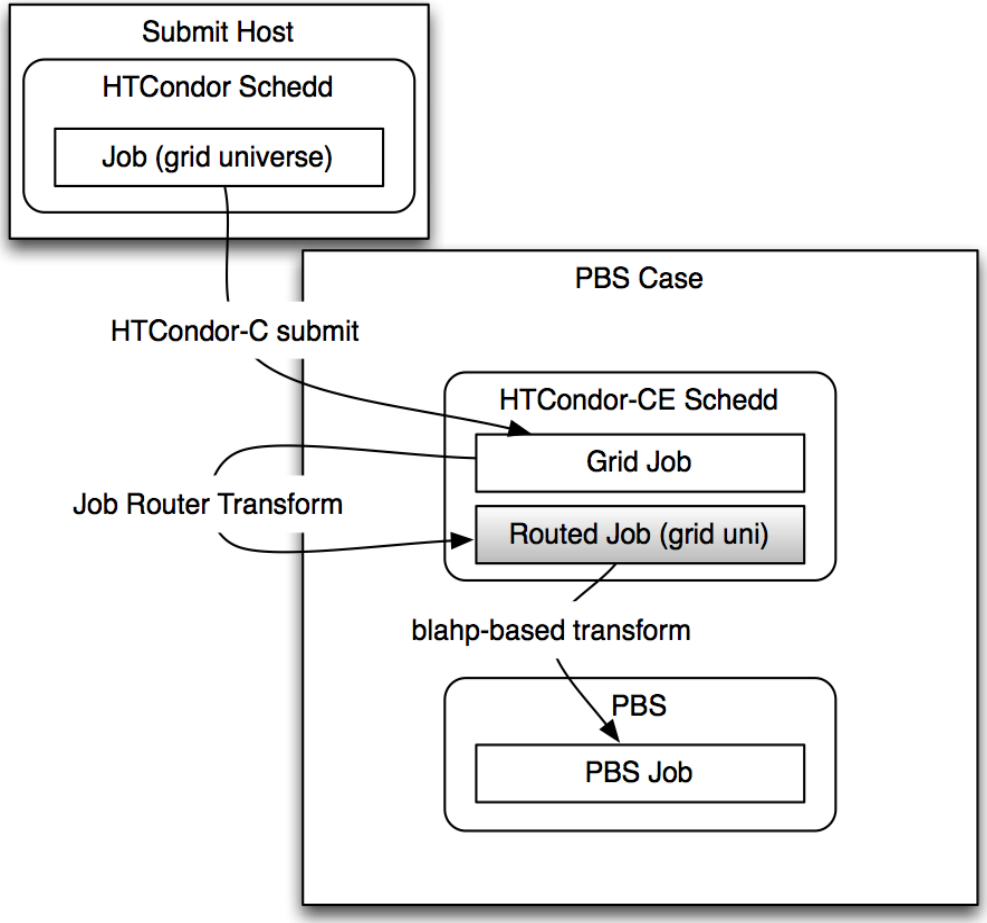
HTCondor-CE: How it works?

- HTCondor batch system



HTCondor-CE: How it works?

- PBS batch system



HTCondor-CE: Work List (I)

- Harden
- Scale
- Audit
- Transform via local policy
- Troubleshooting tools
- Package

HTCondor-CE: Work List (II)

- Harden and Scale
  - BLAHP
    - Improved file cleanup
    - Better error messages on failure
    - Handle errors more gracefully
  - HTCondor-C

HTCondor-CE: Work List (II)

- Harden and Scale
  - BLAHP
    - Improved file cleanup
    - Better error messages on failure
    - Handle errors more gracefully
  - HTCondor-C

- Security Audit
  - Record actions by the user that affect the job queue
    - Submission, removal, modification
  - Record how the user was authenticated
  - Record job credential files
  - Time-based rotation

HTCondor-CE: Job Router (I)

- A key technology is the Job Router, which creates a copy of the job and transforms it according to a set of rules.

HTCondor-CE: Job Router (I)

- A key technology is the Job Router, which creates a copy of the job and transforms it according to a set of rules. In other words:
  - we use the condor_jobrouter daemon for transforming the job for the local site.
  - this daemon creates a copy of the job and applies a set of admin-prescribed transformations aka routes.
    - These can either be done via a ClassAd policy (declarative way) or a script callout.
    - The site customizations will no longer be overwritten by RPM upgrades!
  - The JobRouter can create the job copy directly in a site schedd, doing the site batch system submission for HTCondor sites.

HTCondor-CE: JobRouter ClassAd Policy

```
JOB_ROUTER_ENTRIES = \
  [ \
    GridResource = "batch pbs"; \
    TargetUniverse = 9; \
    name = "Local_PBS_cms"; \
    set_remote_queue = "cms"; \
    Requirements = target.x509UserProxyVOName =?= "cms"; \
  ] \
  [ \
    GridResource = "batch pbs"; \
    TargetUniverse = 9; \
    name = "Local_PBS_other"; \
    set_remote_queue = "other"; \
    Requirements = target.x509UserProxyVOName =!= "cms"; \
  ]
```

- More details/recipes for the routes:
  https://twiki.grid.iu.edu/bin/view/Documentation/Release3/JobRouterRecipes

HTCondor-CE: Job Router (II)

- Previously (GRAM), job transformations were specified in an imperative language (perl). The JobRouter includes a "hook" which allows the sysadmin to specify a script in any language.
  - e.g. JobRouter script JOB_ROUTER_DEFAULTS (python)

- NEW PHILOSHOPHY
  - The pilot describes the resources it needs and the site implementation details are hidden by the JobRouter.
  - Sites have the option of exposing internal configurations, but we'd like to encourage VOs to get to "site-independent pilot submission" - only the endpoint name is different!

HTCondor-CE: Troubleshooting Tools

- Diagnose communication problems
- Detailed diagnosis of failures
  - Can you connect to the server?
  - Can you authenticate with the server?
  - Are you authorized by the server?
- Troubleshooting data and list of tools

  https://twiki.opensciencegrid.org/bin/view/Documentation/Release3/TroubleshootingHTCondorCE

HTCondor-CE: Troubleshooting Tools

- Diagnose communication problems
- Detailed diagnosis of failures
  - Can you connect to the server?
  - Can you authenticate with the server?
  - Are you authorized by the server?
- Troubleshooting data and list of tools

  https://twiki.opensciencegrid.org/bin/view/Documentation/Release3/TroubleshootingHTCondorCE

Questions?