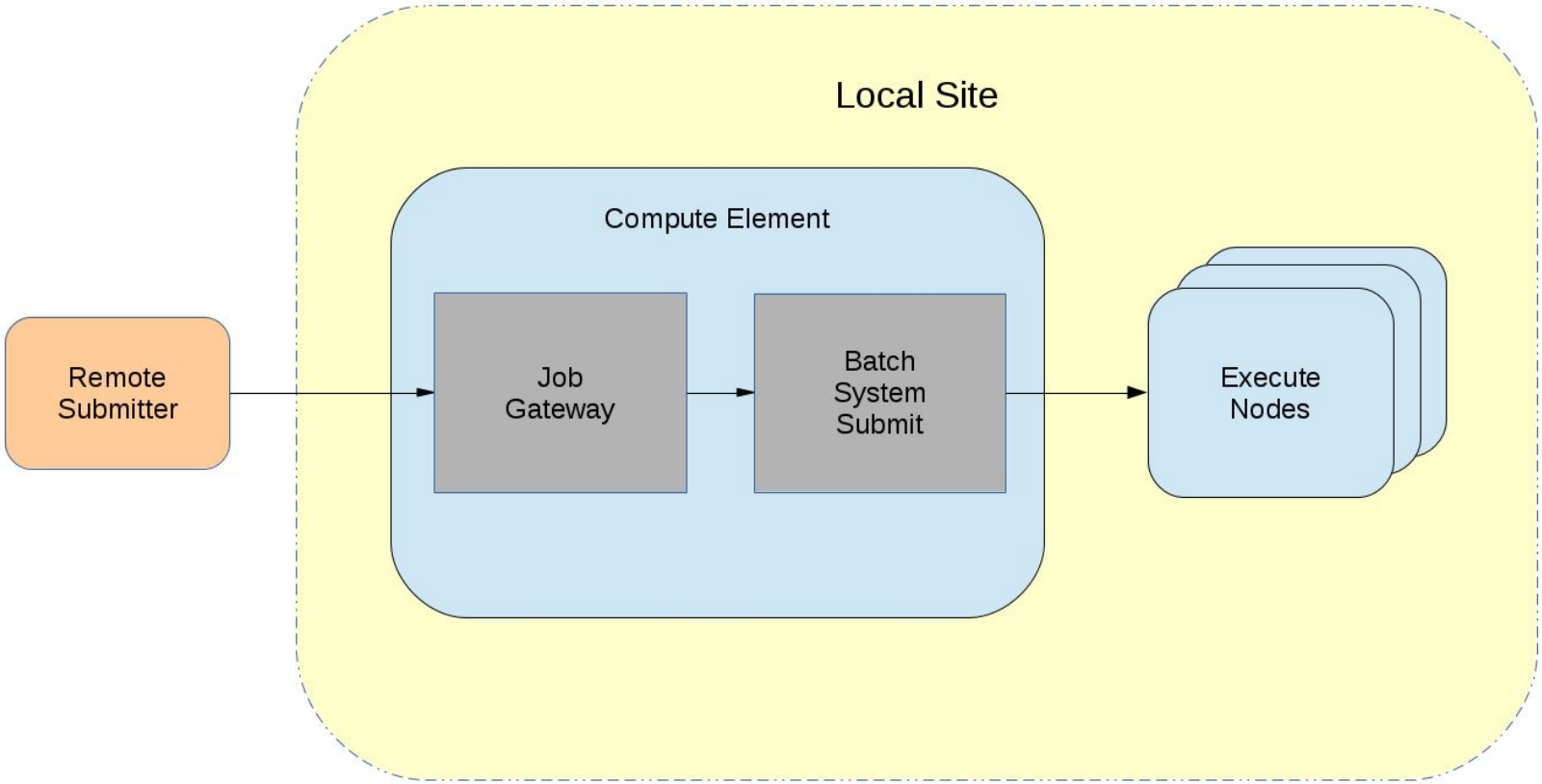
HTCondor-CE:
Managing the Grid With HTCondor

Brian Lin
OSG Software Team
HEPix Oct 2015
blin@cs.wisc.edu

Anatomy of a Compute Element (CE)

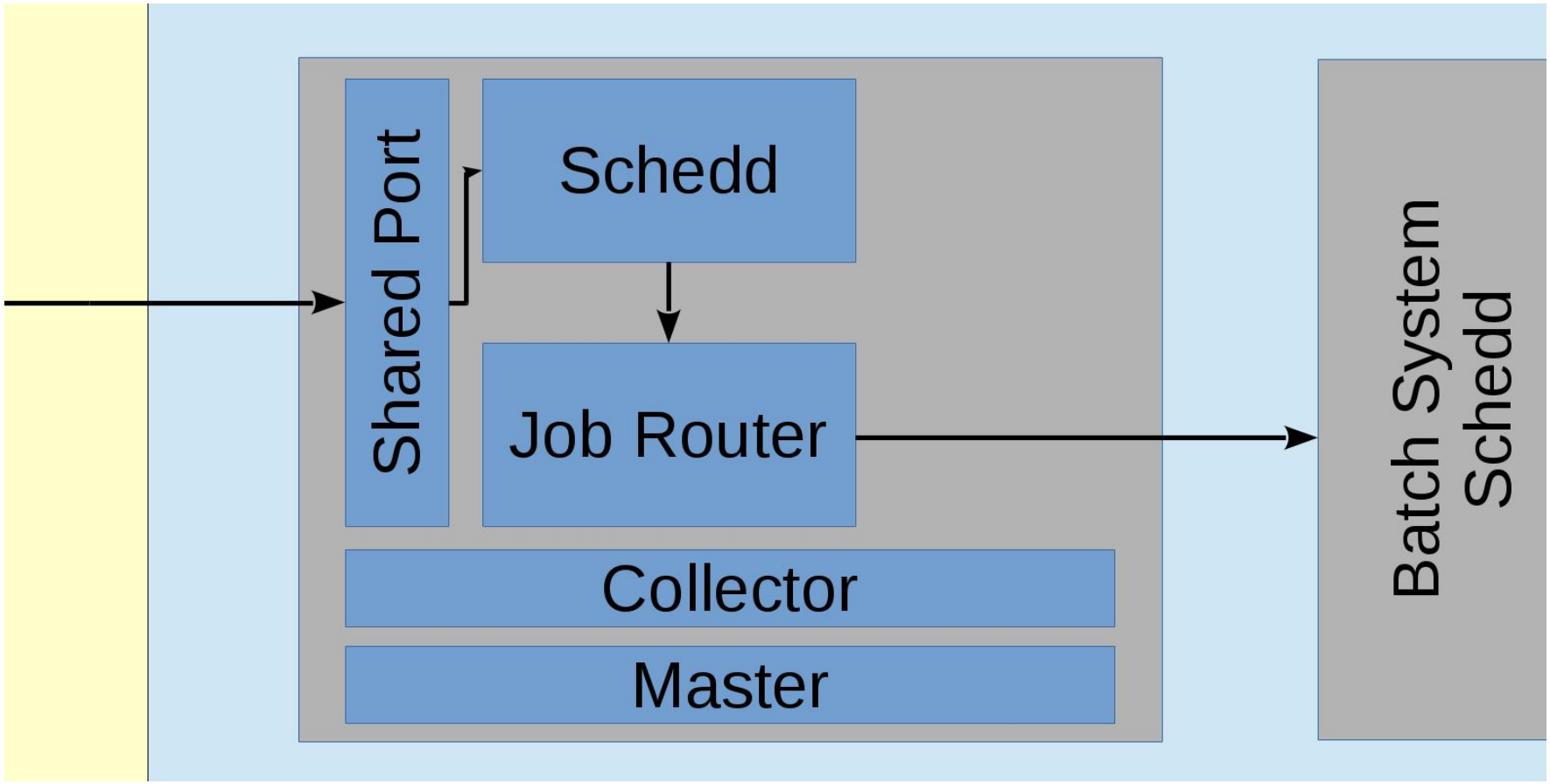


HTCondor-CE: Job Gateway Software

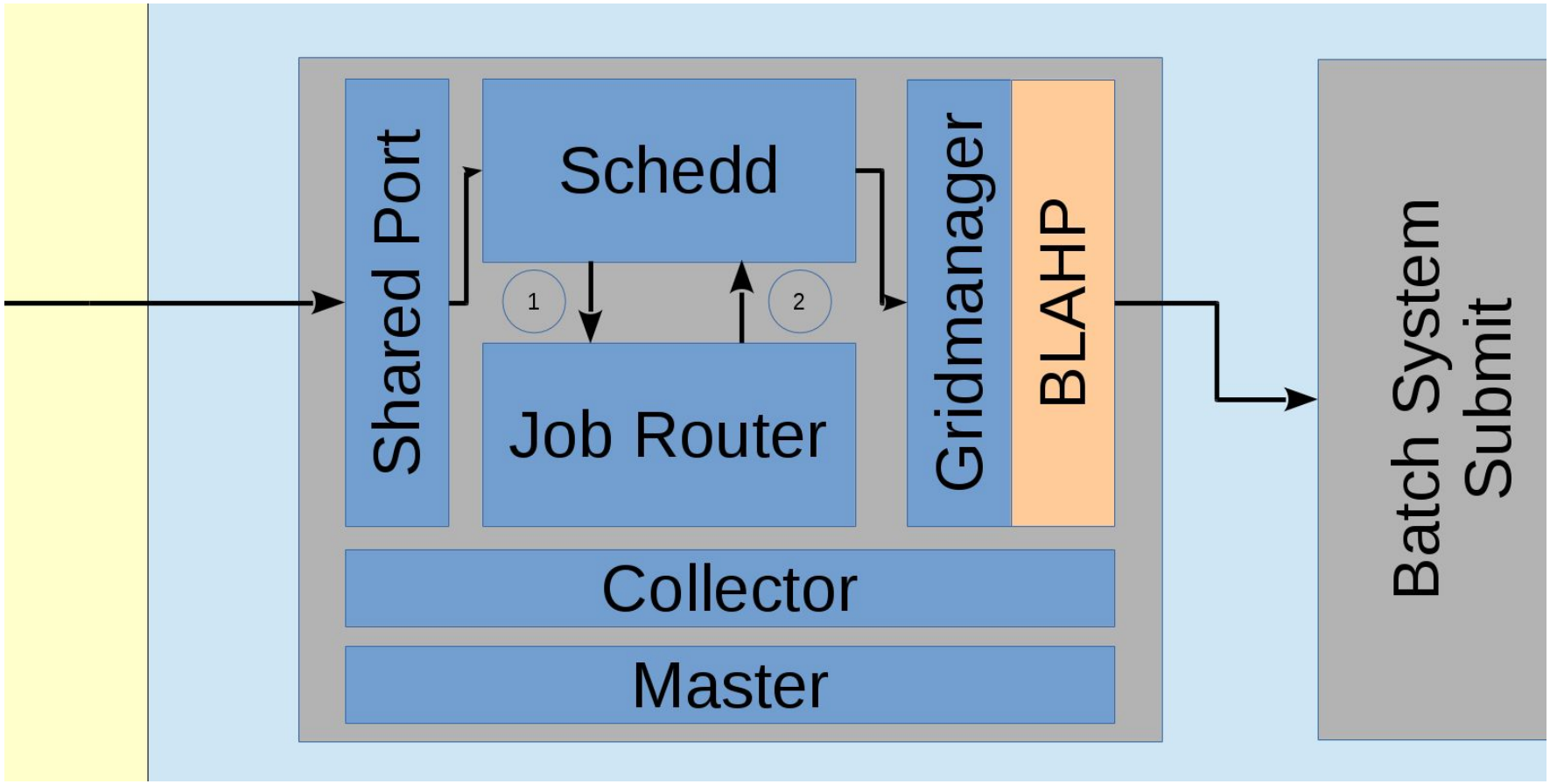
HTCondor-CE is just a special configuration of HTCondor distributed as an RPM.

- Accepts remote jobs via Condor-C
- GSI authentication; LCMAPS authorization
- Interface with local batch systems via job router daemon (plus the grid manager daemon for Torque, PBS Pro, SLURM, LSF and SGE batch systems)

Anatomy of HTCondor-CE: HTCondor Batch System



Anatomy of HTCondor-CE: Non-HTCondor Batch System



Job Router Configuration

- Job router filters and transforms incoming grid jobs into “routed” jobs
- Configured using declarative ClassAds with the JOB_ROUTER_ENTRIES variable
- Each entry in JOB_ROUTER_ENTRIES is combined with the JOB_ROUTER_DEFAULTS configuration variable to create each job route

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example HTCondor Job Route

Cameron has an HTCondor pool and she wants CMS jobs submitted to her CE to be forwarded to her pool and requesting x86_64 Linux machines and setting the attribute “foo” on her routed job to “bar”. All other jobs should be submitted to the pool without any changes.

```
JOB_ROUTER_ENTRIES = [ \
     name = "condor_pool_cms"; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
     set_requirements = (Arch == "X86_64") && (TARGET.OpSys == "LINUX"); \
     set_foo = “bar”; \
] \
[ \
     name = “condor_pool_other”; \
     TargetUniverse = 5; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Example PBS Job Route

Cameron has a PBS pool and she wants CMS jobs submitted to her CE to be forwarded to her pool. All other jobs should be submitted to her pool without any changes

```
JOB_ROUTER_ENTRIES = [ \
     name = "pbs_pool_cms"; \
     TargetUniverse = 9; \
     GridResource = "batch pbs"; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
] \
[ \
     name = "pbs_pool_other"; \
     TargetUniverse = 9; \
     GridResource = "batch pbs"; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

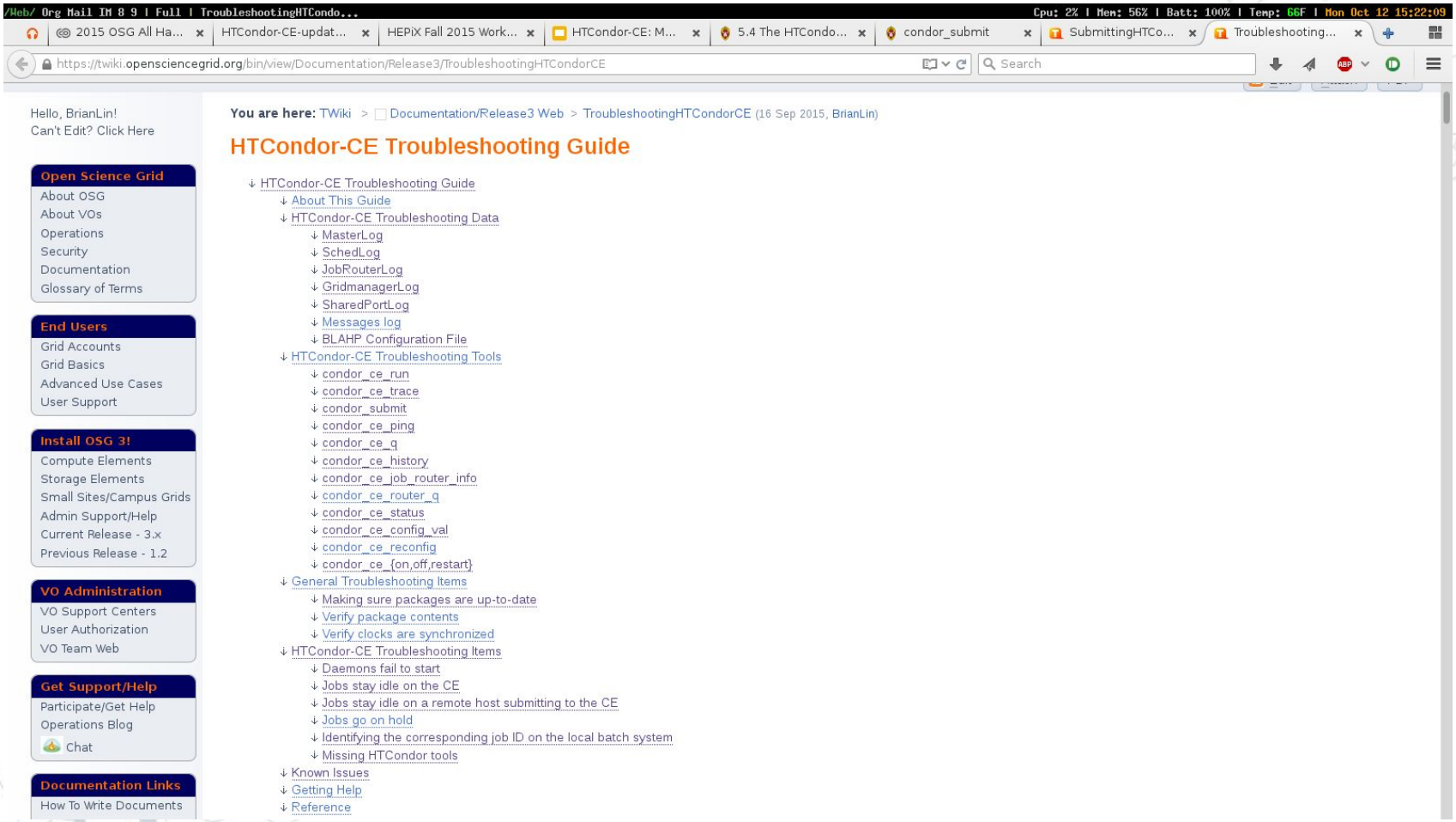
Example PBS Job Route

Cameron has a PBS pool and she wants CMS jobs submitted to her CE to be forwarded to her pool. All other jobs should be submitted to her pool without any changes

```
JOB_ROUTER_ENTRIES = [ \
     name = "pbs_pool_cms"; \
     TargetUniverse = 9; \
     GridResource = "batch pbs"; \
     Requirements = target.x509UserProxyVOName =?= "cms"; \
] \
[ \
     name = "pbs_pool_other"; \
     TargetUniverse = 9; \
     GridResource = "batch pbs"; \
     Requirements = target.x509UserProxyVOName =!= "cms"; \
]
```

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

Abbreviated Troubleshooting Tips



Abbreviated Troubleshooting Tips

- Increase log level in your configuration e.g. ALL_DEBUG = D_FULLDEBUG
- Logs + troubleshooting tools
  - condor_ce_trace: End-to-end job testing
  - condor_ce_host_network_check: DNS issues required expertise in HTCondor
  - condor_ce_job_router_info: Why aren't my jobs being routed?
- If all else fails, contact us! goc@opensciencegrid.org

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

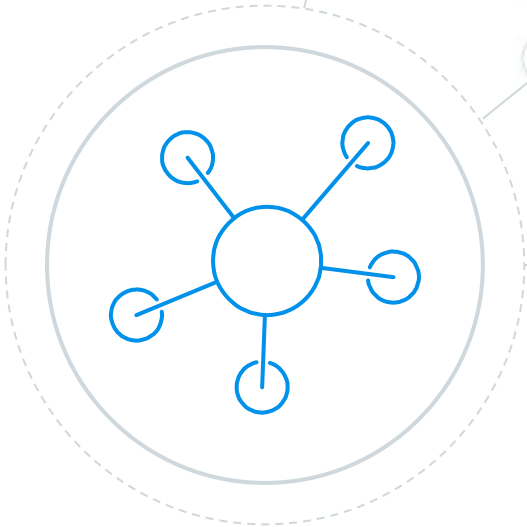
Why Switch to HTCondor-CE?

- If you're using HTCondor as your batch system
  - One less software provider: just a special HTCondor configuration
  - Take full advantage of HTCondor features e.g. Docker universe
- If not, there are still some advantages
  - Scalability: tests achieved 16k max jobs on one CE
  - Declarative ClassAd language: flexible routing policy
  - Job traceability: via debugging tools + logs
  - Fewer open ports: 1 for HTCondor > 8.3.2, otherwise 2

Why Switch to HTCondor-CE?

- If you're using HTCondor as your batch system
  - One less software provider: just a special HTCondor configuration
  - Take full advantage of HTCondor features e.g. Docker universe
- If not, there are still some advantages
  - Scalability: tests achieved 16k max jobs on one CE
  - Declarative ClassAd language: flexible routing policy
  - Job traceability: via debugging tools + logs
  - Fewer open ports: 1 for HTCondor > 8.3.2, otherwise 2
- And some disadvantages…
  - Declarative ClassAd language: can get complicated
  - BLAHP layer is not feature complete

Deployment in the OSG

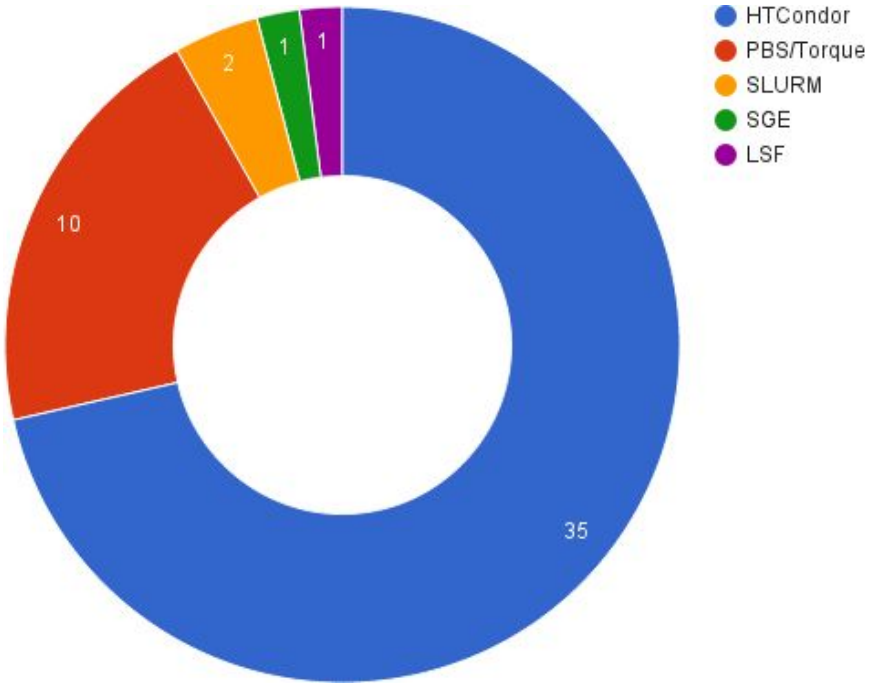


HTCondor-CE and OSG infrastructure

New Installations

- Last year, 10-15 installations of HTCondor-CE in the OSG. Today, we have 49!

```
condor_status -schedd -pool collector.opensciencegrid.org:9619 -af
              OSG_BatchSystems | sort | uniq -c
```



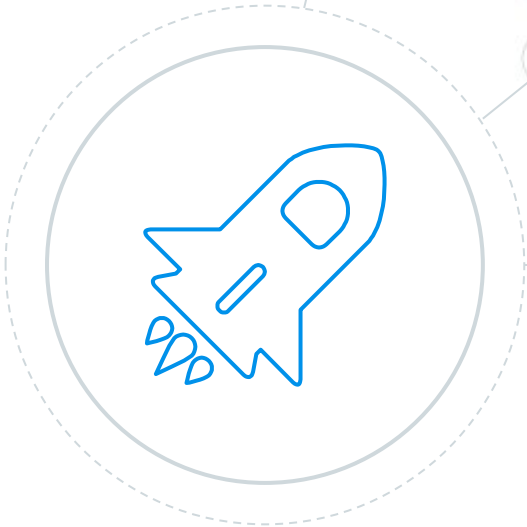
HTCondor-CE Central Collector

- Each site HTCondor-CE advertises itself and its resources to the central HTCondor collector in the OSG
- Next step in information services (resource availability in the grid) after BDII
  - Consolidates software providers: just another special HTCondor config
  - Only resource provisioning: no site reporting
  - Extendable: we can add more attributes easily
- Also accepts StashCache ClassAds

HTCondor-CE Central Collector

```
$ condor_ce_info_status --name='GLOW CE'
[
      OSG_BatchSystems = "Condor";
      MaxWallTime = 1440;
      CPUs = 8;
      Name = "GLOW CE";
      Memory = 16030;
      OSG_Resource = "GLOW-OSG";
      Transform =
            [
                  set_MaxMemory = RequestMemory;
                  set_xcount = RequestCPUs
            ];
      grid_resource = "condor osggrid01.hep.wisc.edu osggrid01.hep.wisc.edu:9619";
      Requirements = TARGET.RequestCPUs <= CPUs && TARGET.RequestMemory <= Memory;
      OSG_ResourceGroup = "GLOW"
]
```

Deployment at CERN

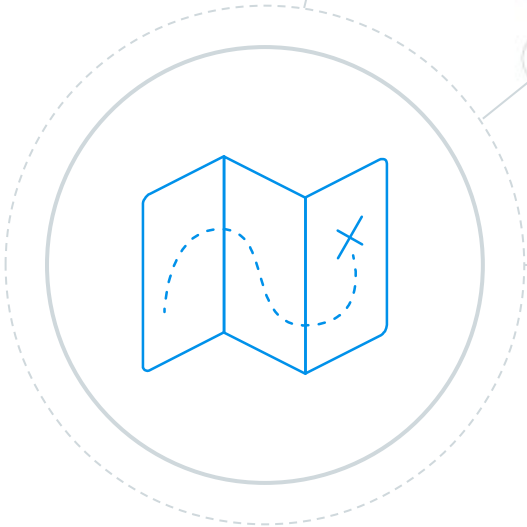


HTCondor-CE at a non-OSG site

HTCondor-CE at CERN

- Running HTCondor-CE in front of an HTCondor pool
  - Rebuilt the RPM for their repos
  - Required changes to OSG-specific configuration
  - Main blocker was due to an upstream bug in HTCondor
- Quickly ramping up, they hope to add more nodes to their pool and more CEs to scale
- Central collector combined with a BDII publisher for Virtual Organizations (VOs) that haven't made the transition

HTCondor-CE: Days of Future Past



Development since HEPix 2014 and upcoming work

A Look Back: Docker Universe

- Available in HTCondor 8.3.6, requires an HTCondor pool
- An example job route (WARNING: untested, try at your own risk!)

```
[ TargetUniverse = docker; \
name = “sl7_HEP”; \
set_docker_image = sl7_and_HEP_stack; \
]
```

- More information can be found
  - Todd Tannenbaum’s talk on Friday
  - http://research.cs.wisc.edu/htcondor/manual/v8.4/2_12Docker_Universe.html
  - http://research.cs.wisc.edu/htcondor/manual/v8.4/3_14Setting_Up.html

A Look Back

- Local collector accepts pilot payload ads
- BLAHP improvements: Basic PBS Pro support and LSF fixes
- Added condor_ce_network_host_check, condor_ce_info_status. Made improvements to existing tools
- EL7 support

Future Work

- Improvements to job router configuration
  - Easily add defaults to your job routes

```
MERGE_JOB_ROUTER_DEFAULT_ADS = True

JOB_ROUTER_DEFAULTS = $(JOB_ROUTER_DEFAULTS_GENERATED) [TargetUniverse = 5; \
set_foo = “bar”;]
```

  - Flexible AccountingGroup assignments

```
set_AccountingGroup = ifThenElse(regexp(“.*Brian Lin.*”, \
x509UserProxySubject), “naughty users”, AccountingGroupOSG)
```

- Improvements to the BLAHP
  - PBS Pro MPI support
  - Integration into HTCondor source

Future Work

- Expanding the central collector
  - Long-lived resource ads
  - More non-CE resource types
  - ATLAS Global Information System (AGIS) adapter

Resources

- HTCondor manual

  http://research.cs.wisc.edu/htcondor/manual/v8.4/ref.html

- HTCondor-CE Overview

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

- HTCondor-CE Install Guide

  https://twiki.grid.iu.edu/bin/view/Documentation/Release3/InstallHTCondorCE

- HTCondor-CE Troubleshooting Guide

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

- HTCondor-CE Job Router Configuration Examples

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

- Submitting Test Jobs to HTCondor-CE

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

Credits

I'd like to thank:

- The HTCondor team
- Brian Bockelman at UNL
- Iain Steers at CERN
- All the sites that helped test HTCondor-CE

Special thanks to all the people who made and released these awesome resources for free:

- Presentation template by SlidesCarnival
- Photographs by Unsplash & Death to the Stock Photo (license)

Thanks!

Any questions?

You can find me at:

blin@cs.wisc.edu

osg-software@opensciencegrid.org

Support requests should go to:

goc@opensciencegrid.org