



HTCondor-CE our Experience

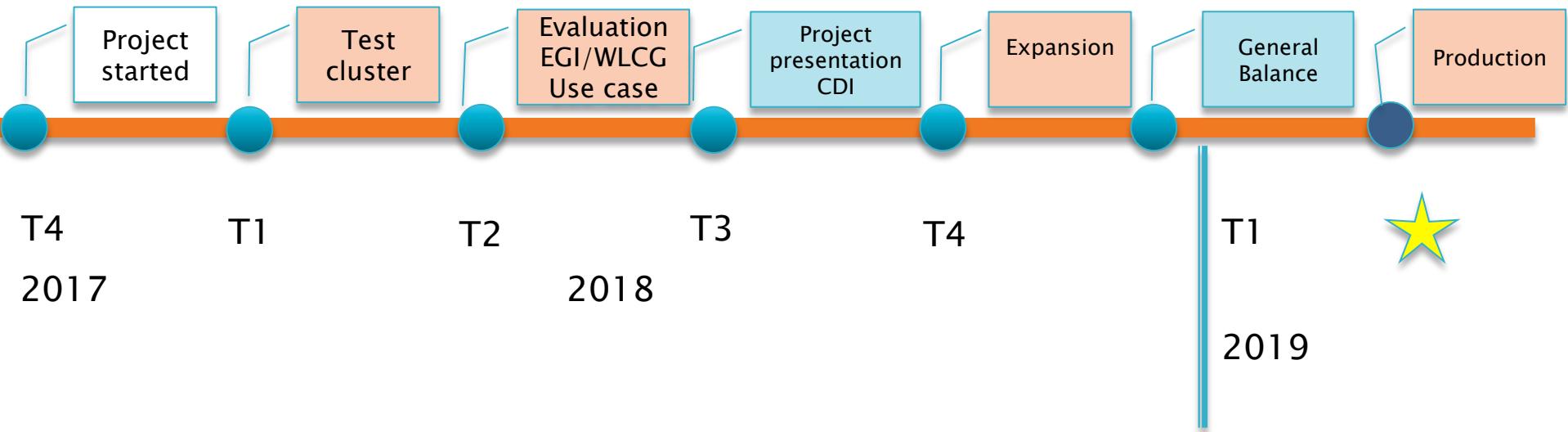
C. Eloto, N. Fournials, E. Vamvakopoulos, V. Hamar

- ▶ HTCondor project
- ▶ HTCondor-CE
 - Architecture
 - Deployment
 - Configuration
 - External grid nodes configuration
 - Monitoring
- ▶ HTCondor
 - Configuration at CC
- ▶ Future Work
- ▶ Conclusions

« POC - Comment se positionne HTCondor, à minima pour couvrir le périmètre de l'activité calcul (HTC, HPC, GPGPU) ? »

- ▶ Trying to:
 - Find an efficient and simple configuration
 - Be open to new ideas

Project roadmap

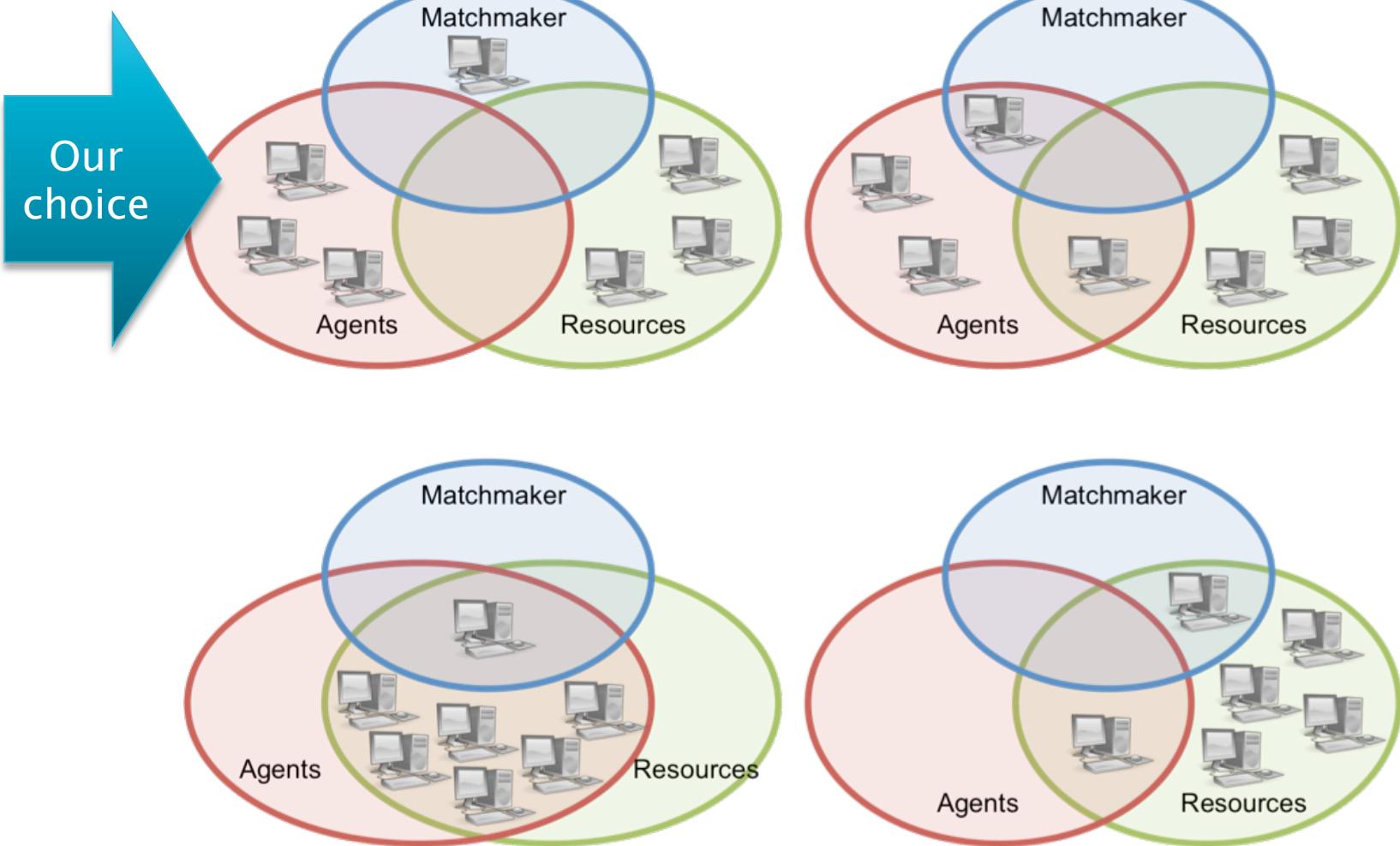


[EGI BROADCAST] End of Support for CREAM-CE
Fri, 1 Feb 2019 13:18:36 +0100 (CET)

Activities

Deliverable

Architecture



<http://condorpy.readthedocs.io/en/latest/htcondor.html>

Added Grid Universe support for Azure, SLURM, Cobalt

- ▶ Speak to Microsoft Azure
- ▶ Speak native SLURM protocol
- ▶ Speak to Cobalt Scheduler
 - Argonne Leadership Computing Facilities

Jaime:
Grid
Jedi



Also HTCondor-CE "native" package

- › HTCondor-CE started as an OSG package
- › IN2P3 wanted HTCondor-CE without all the OSG dependencies....
- › Now HTCondor-CE available stand-alone in HTCondor repositories

Deployment

- ▶ HTCondor_CE cernops module
 - Customized to fit our needs
 - A few parameters changed
 - Deploy a HTCondor-CE is easy and reproducibly.

https://github.com/cernops/puppet-htcondor_ce

- ▶ Used in conjunction with HTCondor HEP-Puppet module
 - Very useful deployment tool for HTCondor

<https://github.com/HEP-Puppet/htcondor>

HTCondor configuration with puppet

Oliver Freyermuth, Luke Kreczko, Kashif Mohammad, Frederic Schaer, Peter Wienemann

LCG-FRANCE

HTCondor Puppet module

- <https://github.com/HEP-Puppet/htcondor>
- Almost 10,000 downloads ([Puppet forge](#))
- Covers the most common use cases
 - Setup managers (w/o high-availability), schedulers, and workers
 - Kerberos authentication
 - Singularity configuration
 - Fully configurable via [hiera](#) (YAML files)
- Since 2.1.0: no (condor) knob left behind
 - Big thanks to @ccninfo

Table of Contents

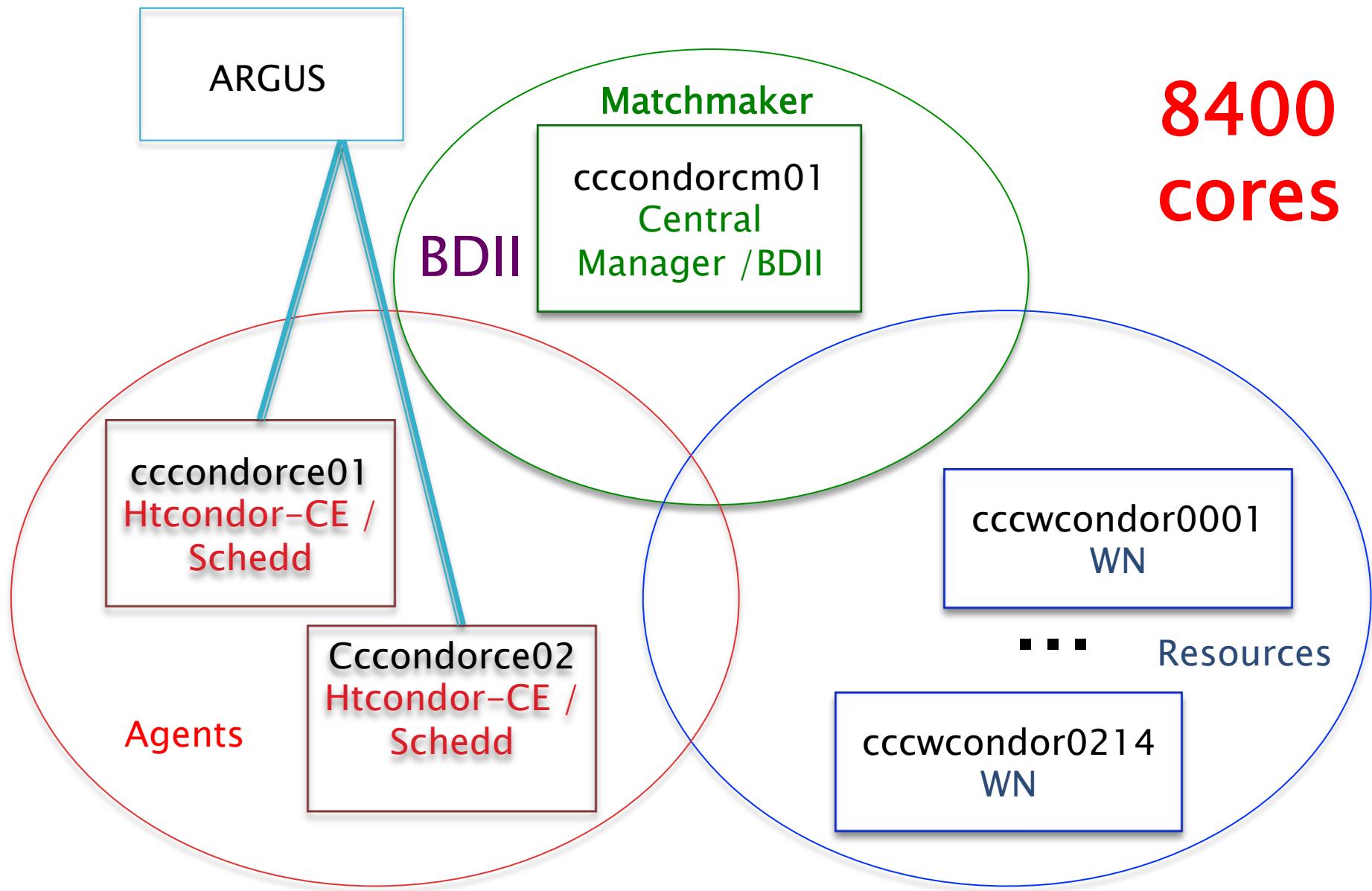
1. Overview - What is the htcondor module?
2. Module Description - What does the module do?
3. Setup - The basics of getting started with htcondor
4. Singularity container support
5. Kerberos authentication support
6. Additional logging parameters
7. Additional custom parameters
8. Limitations - OS compatibility, etc.
9. Development - Guide for contributing to the module
 - Contributing to the htcondor module
 - Running tests - A quick guide

Deployment - Hieradata

```
htcondor_ce::pool_collectors:
  - 'cccondorcm01.in2p3.fr'
htcondor_ce::condor_view_hosts: []
htcondor_ce::ce_version: '3.3.0-1.el7'
htcondor_ce::lrms_version: '8.6.12-0.445603.el7'
htcondor_ce::uid_domain: 'in2p3.fr'
htcondor_ce::gsi_regex: '^VO\=GRID-FR\&C\=FR\O\=CNRSVOU\=CC-IN2P3\CN\=([A-Za-z0-9.\-]*$)'
htcondor_ce::gsi_backend: 'argus'
htcondor_ce::argus_server: 'cctbargus01.in2p3.fr'
htcondor_ce::argus_port: 8154
htcondor_ce::argus_resourceid: 'http://cc.in2p3.fr/ce'
htcondor_ce::use_static_shadow: false
htcondor_ce::job_router_entries: >-
  [
    eval_set_environment = debug(strcat("HOME=/tmp CONDORCE_COLLECTOR_HOST=", CondorCECollectorHost, " ", \
TargetUniverse = 5; \
  name = "Local_Condor"; \
  set_VOName = ifThenElse(isUndefined(X509UserProxyVOName),"LOCAL",X509UserProxyVOName); \
  set_AcctSubGroup = toUpper(\
  ....
  eval_set_RequestMemory = ifThenElse(WantWholeNode is true, !isUndefined(TotalMemory) ? TotalMemory*95/100 : JobMemory, OriginalMemory); \
  ]

site_htcondor_ce::prerelease_repo_enabled: false
# BDII
htcondor_ce::install_bdii: true
htcondor_ce::supported_vos:
  - atlas
  - .....
htcondor_ce::goc_site_name: 'IN2P3-CC'
htcondor_ce::benchmark_result: '10.26-HEP-SPEC06'
htcondor_ce::execution_env_cores: 40
```

Until now ...



- ▶ htcondor-ce-bdii RPM
- ▶ LDAP user is authenticated using GSI authorization

```
55 * * * * /ldap voms-proxy-init --cert /var/lib/ldap/hostcert.pem --key /var/lib/ldap/hostkey.pem –valid  
02:00
```

- ▶ To check the publication

```
$ ldapsearch -LLL -x -H ldap://cccondorcm01.in2p3.fr:2170 -b o=glue
```

- ▶ Note: The resource must be added into site BDII configuration to be published

```
$ vi /etc/bdii/gip/site-urls-CC.conf
```

....

```
CONDOR01 ldap://cccondorcm01.in2p3.fr:2170/mds-vo-name=resource,o=grid
```

- ▶ GGUS Ticket:
 - cccondorcm01.in2p3.fr 128334
- In particular, I would like to know if the number provided by cccondorcm01 is correct to take into account for the total cores provided by your site.
- ▶ Changes in the /var/lib/bdii/gip/provider/htcondor-ce-provider file:

vcpus = 0 (line 140)

vcpus=sum(total_cores.values()) (line 163)

cpu=tup[5] => cpu=vcpus (line 179)

arch=tup[0].lower()

os=tup[1].lower(),

```
$ ldapsearch -LLL -x -H ldap://cccondorcm01.in2p3.fr:2170 -b o=glue | grep CPU
GLUE2ComputingManagerTotalLogicalCPUs: 2680
GLUE2ExecutionEnvironmentLogicalCPUs: 2680
GLUE2ExecutionEnvironmentCPUMultiplicity: singlecpu-multicore
```

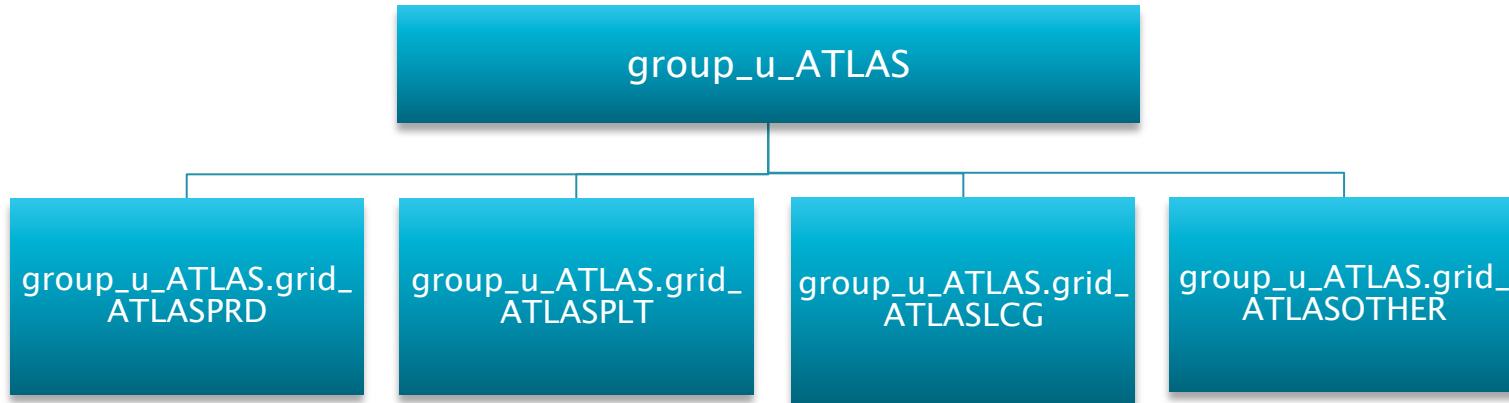
- ▶ Add the resource to your argus policies

```
resource "http://cc.in2p3.fr/ce" {  
    obligation "http://glite.org/xacml/obligation/local-environment-map" {  
    }  
  
    action ".*" {  
        rule permit { pfqan="/atlas/Role=pilot" }  
        rule permit { pfqan="/atlas/Role=lcgadmin" }  
        rule permit { pfqan="/atlas/Role=production" }  
        rule permit { pfqan="/atlas/Role=software" }  
        ...  
    }  
}
```

- ▶ The grid-mapfile is static in argus server, for historical reasons specific roles could be matched only to one account or to an account of the VO pool:

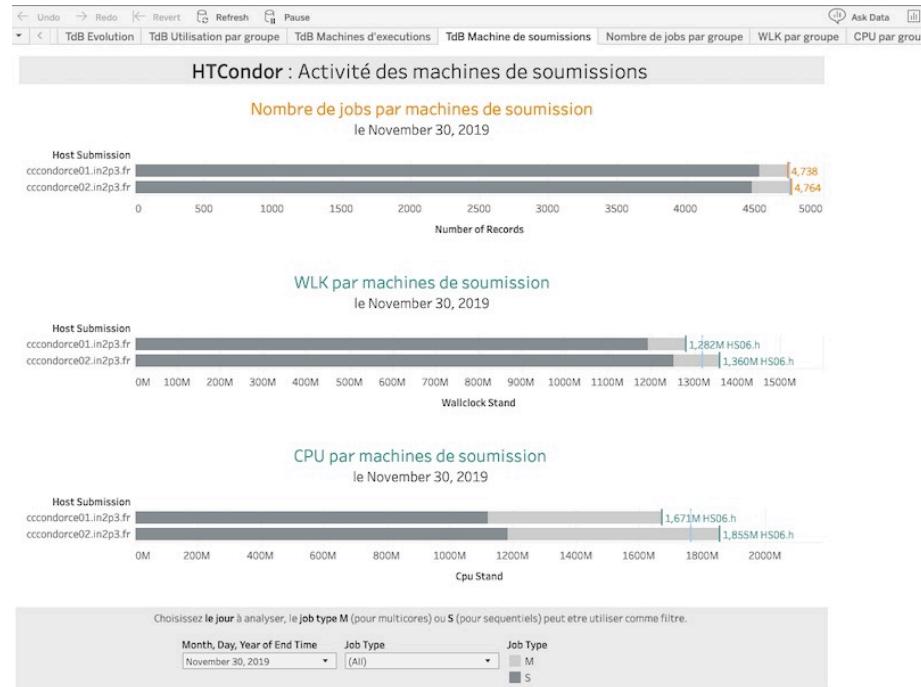
```
"/cms/Role=lcgadmin" cmsgrid  
"/cms/Role=pilot" .cms
```

- ▶ Job routers in local htcondor and htcondor-CE
- ▶ Hierarchy tree is based on VO names and user proxy role.
 - group_u_VONAME
 - group_u_VONAME.grid_{VONAME}{ROLE}



- ▶ Assuming that each proxy role represents a different activity, but a VO can use the same role for different activities.

- ▶ Using condor_history command to generate json files by hour and by schedd and to save into a database.
 - Testing Apache Spark
 - Tableau – Data visualization tool
- ▶ Adapting our local scripts to generate APEL accounting files.

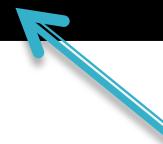


Jobs submitted through HTCondor-CE

- ▶ Fifemon looked like the way to go
 - But CC-IN2P3 doesn't have graphite in its monitoring stack
 - Adapting it to Collectd + Elastic Search was time consuming and quite unsuccessful.
- ▶ Collectd + dedicated plugin
 - Currently writing a Python plugin for HTCondor
 - Using Python Bindings to gather classads
 - Though the Collector
 - Negotiator
 - Collector
 - Defrag
 - Schedd
 - Startd
 - Through the Schedd
 - Schedd ads (jobads of all jobs, regardless of their state)

- ▶ Collectd + plugin for HTCondor-CE as well
 - Query the HTCondor-CE collector with a simple trick

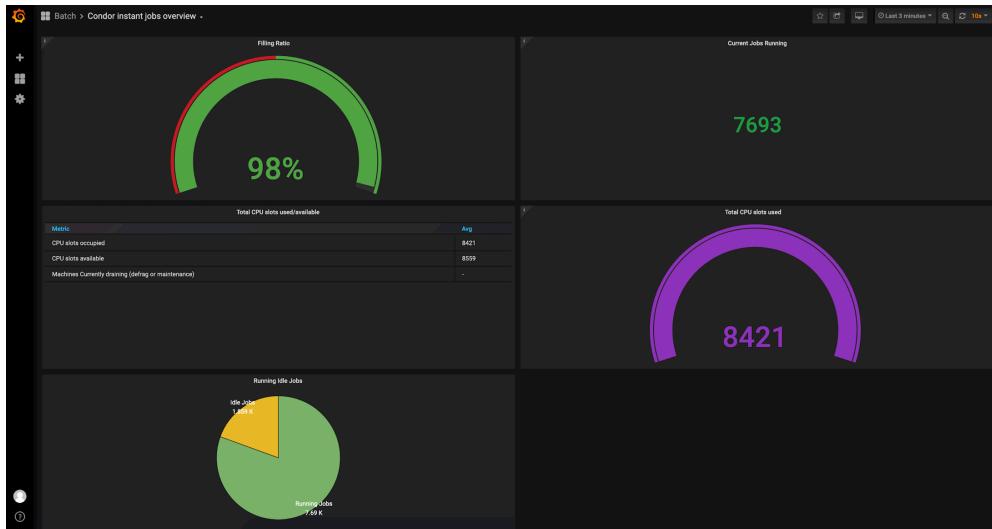
```
Import htcondor
ce_coll = htcondor.Collector('localhost:9619')
ce_coll.query()
```



Changing the port you get connected to CE

- We gather
 - Job general stats
 - Schedd daemon metrics
- Currently putting efforts in writing a proper reusable code for public release

▶ Grafana to plot the metrics



Cluster usage

Jobs by group



- ▶ A lot of things to do!!
- ▶ A short term:
 - Migrate all grid resources before February 2020.
 - Publish Collectd python plugin for HTCondor-CE and HTCondor

A big thank you !!!

- ▶ To HTCondor developers, specially to:
 - Miron Livny
 - Todd Tannenbaum
 - John (TJ) Knoeller
 - Brian Lin
 - Zachary Miller
- ▶ CC-IN2P3 colleagues
 - Christelle Eloto
 - Cécile Evesque
 - Nicolas Fournials
 - Ghita Rahal
 - Emmanouil Vamvakopoulos

