Future of Batch Processing at CERN: a Condor Pilot Service

HEPiX Fall 2014

Jérôme Belleman, Daniel Pek, Ulrich Schwickerath
CERN IT

October 2014

Outline

1 Context

2 Minimum Viable Service

3 Next Features Towards Production

Section 1

Context

Current Setup

- LSF 7.0.6
- $\approx 4\,000$ nodes
- SLC5 $\xrightarrow{95\%}$ SLC6
- Physical $\xrightarrow{85\%}$ Virtual machines
- Quattor $\xrightarrow{99\%}$ Puppet
- $> 65\,000$ cores
- $400\,000$ jobs/day
- $\pm 70\,000$ running jobs

Goals and Concerns

Goals	Concerns with LSF
30 000 to 50 000 nodes	6 500 nodes max
Cluster dynamism	Adding/Removing nodes requires reconfiguration
10 to 100 Hz dispatch rate	Transient dispatch problems
100 Hz query scaling	Slow query/submission response times

Evaluating Alternatives to LSF

After HEPiX Fall 2013 – Ann Arbor:

- LSF 8/9 advertised to only marginally scale higher.
- SLURM showed scalability problems too.
- Son of Grid Engine only briefly reviewed, because...
- ... HTCondor looked promising.

Settling on Condor

After HEPiX Spring 2014 – Annecy:

- Condor scaled encouragingly
- Focus on functions (Grid, fairshare, authentication, AFS)
- Pleasant experience

→ Now setting up a pilot service

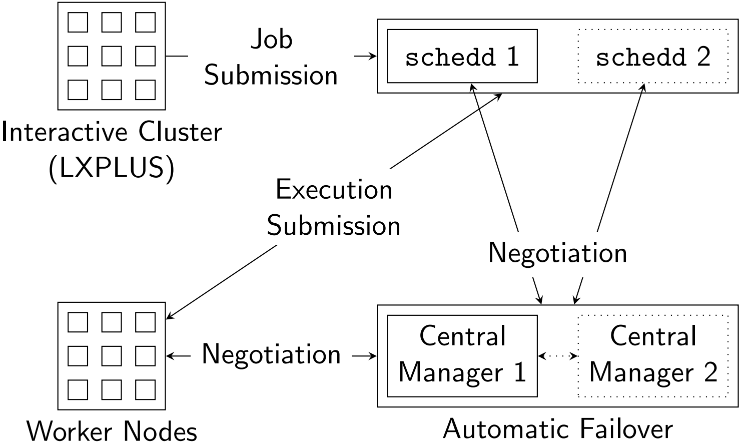
Section 2

Minimum Viable Service

Phase 1: Grid Submissions Only

- No AFS
- Submission authorisation for free
- Naturally sticky submission hosts

Cluster Setup



Interactive Cluster (LXPLUS)

Job Submission

schedd 1

schedd 2

Execution Submission

Negotiation

Worker Nodes

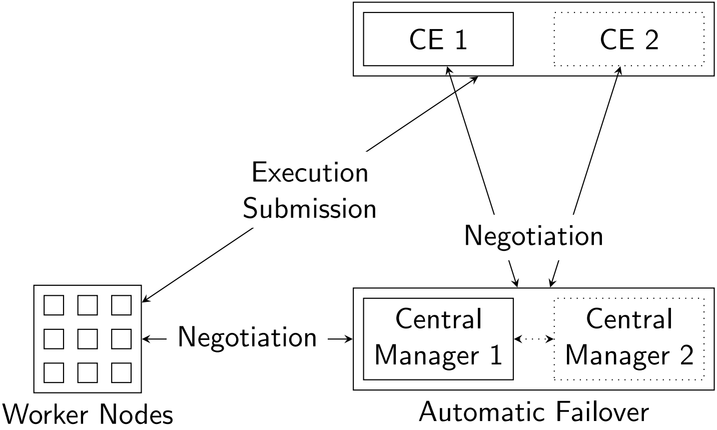
Negotiation

Central Manager 1

Central Manager 2

Automatic Failover

Cluster Setup



Node Flavours

Virtual machines	
1 central manager	4 cores, 8 GB
1 `schedd` (CE)	8 cores, 16 GB
< 10 worker nodes	8 cores, 16 GB

Queues

- Grid: same as LSF queues, as agreed with experiments
- Local: opportunity to review setup
- Plan to set restrictions on resources

Accounting Groups and Fairshare

LSF hierarchical dynamic priority:

$$\begin{aligned}
\text{Priority} &= \frac{\text{number of shares}}{t_{\text{CPU}} \times \text{CPU time factor} + t_{\text{wall}} \times \text{wall time factor}} \\
&= \frac{\text{number of shares}}{t_{\text{CPU}} \times 0.7 + t_{\text{wall}} \times 0.0} \\
&= \frac{\text{number of shares}}{t_{\text{CPU}} \times 0.7}
\end{aligned}$$

Condor:

- No hierarchical dynamic priority, flat by user
- → Trying hierarchical group quotas

CREAM CE

- Currently 1 CE submitting to Condor
- Contributing to HEP-Puppet HTCondor module
- GIP publishes static info about our `grid_*` queues

TODO:

- Dynamic info
- Machine/job features
- GLUE 2
- Accounting

Security

- Close everything in the cluster, then open as needed
- Worker node authorisation: machine lists
- Daemon-to-daemon authentication: GSI
- User authentication: Kerberos
- Local authentication on CEs: filesystem
- Pending for review from our security team
- We've made several bug reports

Tests and Monitoring

- Sample jobs successfully run from CE
- gLExec tests pass
- Basic Ganglia setup

Section 3

Next Features Towards Production

Phase 2: Local Submissions Too

- AFS
- Local job submissions
- Authorisation
- Scalable & available setup
- . . .

Requirement Specification

- Been sharing/updating a document for last few months
- Understanding our LSF setup better
- Learning about our real needs

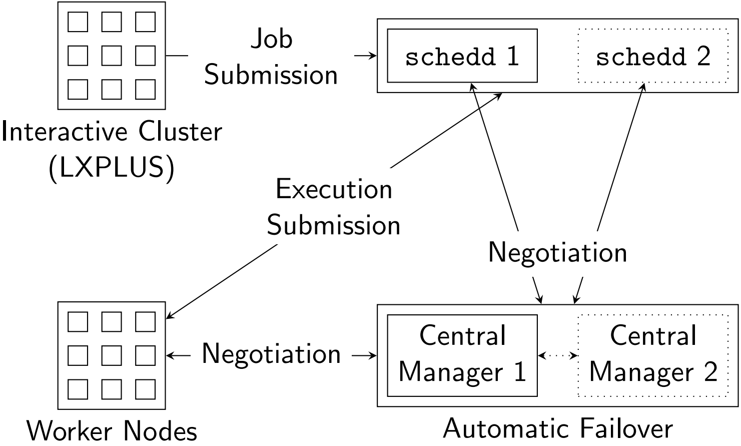
→ A useful process

AFS Token Management

Ongoing work:

- Managed to pass a Kerberos ticket, no extension yet
- May choose to only manage Kerberos tickets
- Long-lived Kerberos credentials. . .
- . . . or dated ones?
- SSH keys to authenticate users?
- Need to transfer a tracker at submission

Local Job Submissions



- Costly global queries, locals only to host submitted from
- Land on host you submitted from?

→ Sticky submission hosts

Job Submission Wrapper

- Get credentials? Set off process renewing them?
- Enforce group ownership?
- Deployment: keep `condor_submit` out of reach?

Later. . .

- High availability: schedulers and central managers
- Host normalisation
- Local job accounting

Conclusions

- Open Grid submissions to pioneers once past security
- Couple of bugs
- Always good collaboration with developers
- Now looking into AFS

Thanks!

Questions?