
CERN

Evolution of Batch Services at CERN

Jérôme Belleman, Ulrich Schwickerath, Iain Steers – Batch Service, IT-PES-PS

Pledges 2014 – WLCG

HEP-SPEC06	LHCb	All LHC VOs
CERN (Tier-0)	34 000	356 000
All	218 000	2 178 000

Evolution of Requests

Significantly higher data rates for Run 2. LHCb requests:

HEP-SPEC06	2014 pledge	2015 req.	Prel. 2016 req.
CERN (Tier-0)	34 000	36 000	51 000
All	218 000	240 000	315 000

Current Situation – Batch Jobs

- $\approx$ 4 000 nodes, $\approx$ 60 000 cores, 400 000 jobs/day, some 45 000 concurrently running jobs
- 20% to 40% of total usage is grid jobs (via CEs)

Current Situation – Transitions

- Quattor to Puppet (done)
- SLC5 to SLC6 (done)
- Physical to virtual worker nodes within CERN's Agile Infrastructure (90% done)

Workload Management

- Since the late 1990s, CERN use a commercial product: Platform Inc.'s Load Sharing Facility (LSF)
- IBM acquired Platform Inc. in 2011/2012
- CERN's licence is perpetual, maintenance currently covered until November 2017
- We run 7.0.6

Pain Points with LSF (I)

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
Licence-free system	Licensed product

Pain Points with LSF (II)

Worker node scaling:

- Resource requirements to grow by $> 100\%$
- Unclear what future distribution of batch vs. cloud resources will be
- Limit already constrains us to use unnaturally large VMs (whole hypervisor)
- Limit unchanged with LSF 8/9, although can set up multiple instances that can submit to each other

Pain Points with LSF (III)

Cluster dynamism:

- LSF reconfigurations are expensive – some 10 minutes of unresponsiveness
- We are running it once per day
- Sometimes reconfiguration fails, leading to loss of queues, etc.
- Some operations require two reconfigurations, hence up to 48 hours before becoming effective

Pain Points with LSF (IV)

Query rate:

- LSF is not (cannot be) protected against users hammering the system with expensive queries
- Number of cases in the past where submissions and job dispatch were seriously affected by query activity

Alternatives to LSF 7

LSF 8 or 9:

- Not really addressing any of our pain points

PBS offsprings:

- Too much trouble reported by other LCG sites

SLURM:

- Scalability with many cores/parallel computing
- Scaling limits on worker nodes and job slots

Grid Engine:

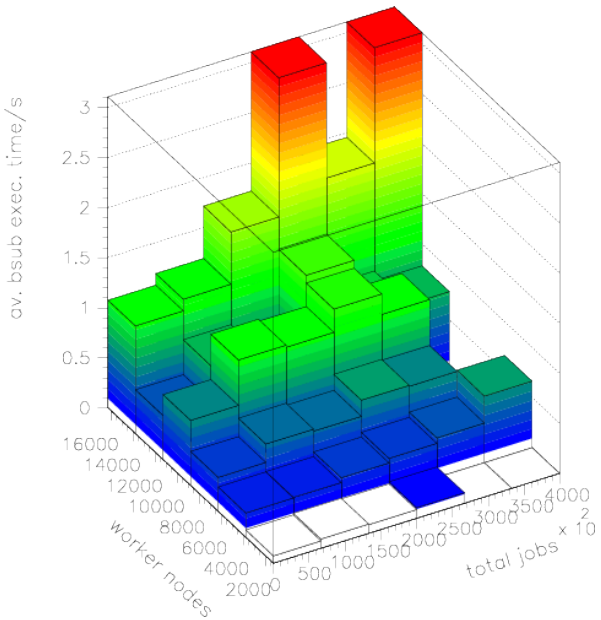
- Univa Grid Engine only serious contender
- Commercial, similar architecture to LSF

Alternatives to LSF 7: HTCondor

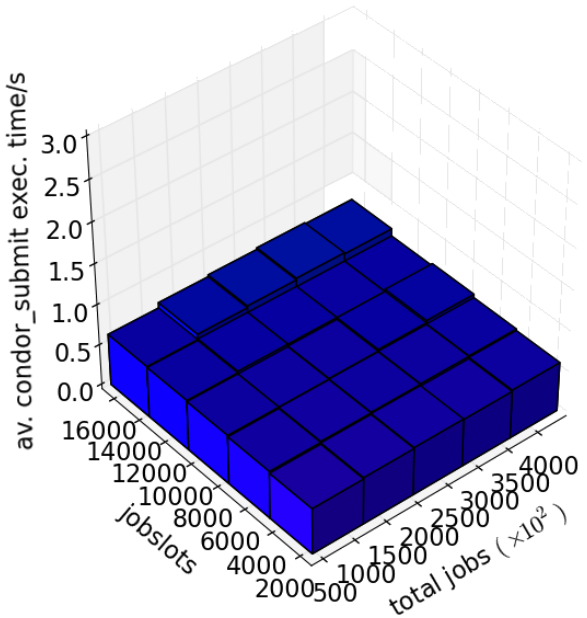
- Open-source, academic environment
- Already in widespread use in WLCG, e.g. FNAL, BNL, RAL – good experience
- CERN's requirements are different
  - CERN cluster already largest and growing
  - CERN need to also support local job submission with AFS token passing/extension
- Scaling test (shadows on LSF worker nodes) looked promising
  - 2 central managers, 20 schedulers/submission nodes, 1 300 worker nodes with 62 500 job slots
  - Architecture promises to support further scale-out (unlike LSF, GE, SLURM, etc.)

HTCondor Scaling Behaviour

Job submission time as a function of the number of worker nodes and total number of jobs



LSF



HTCondor

HTCondor Observations

HTCondor in experiment frameworks (even as CE)

- Can be used as cloud scheduler
- Potential for future further integration

Contact with enthusiastic developer community

- Excellent two-way discussions
- Not all is there, but confident that it will be

Tests so far very successful

- Scaling, adding/removing worker nodes
- Failing central manager/scheduler unproblematic
- Query scaling an issue, worked on by developers

HTCondor Deployment Plan (I)

Opened pilot service for Grid submission only

- Mostly transparent to users
- No need for AFS token passing and extension
- Looking for friendly user community to try out

Grow that service

- Up to taking all Grid submissions
- Overflowing into LSF part via `condor_glidein` possible

HTCondor Deployment Plan (II)

Once ready, small service for local jobs

- Condor submission and queries to look like LSF?
- User support (documentation, tutorials, etc.) will be integral part of deployment
- Again, looking for friendly user community to try out

Grow to full size

- Reducing LSF capacity
- Close interaction with user community

HTCondor Deployment Timescale

- Very small prototype (10 workers) exists
  - Reviewed by computer security team
  - Now working on CE integration, machine/job features, GLUE 2, accounting
- Grid submission with real payload starting in 1H2015
- Timescale for local submission service to be defined:
  - Hoping for pilot by end 2015, but...
  - Priority on full-scale and production-quality service for Grid submissions
- Target: terminate LSF service by end of Run 2

Conclusion

- Move off LSF is inevitable
- HTCondor is the only viable – and a very attractive – option
- A lot of work still to be done
- We'll do everything reasonable to ensure a smooth transition
- Lot of fun, collaborating with HTCondor development team and other sites very rewarding
- We need, and we count on, your patience and your support

www.cern.ch