

Batch Systems Review

Michel Jouvin

LAL, Orsay

jouvin@lal.in2p3.fr

HEPiX, Annecy, May 2014

Introduction

- Most of the information in this presentation called from a WLCG pre-GDB devoted to batch systems
 - > March 2014
 - > Agenda: <https://indico.cern.ch/event/272785/>
 - > Part of an ongoing work to review the batch system situation
 - > European-centric review
- Most (European) “well known experts” of batch systems present
 - > CESGA (Grid Engine) apologized not being able to join
 - > Covering Torque/MAUI, Grid Engine, LSF, HTCondor, SLURM

Meeting Goals

- Share experience about the different batch systems
 - > First part of the meeting was a batch system review by sites with a concrete experience
- Identify strengths and weaknesses
 - > Base features of a batch system
 - > Multi-core job support
 - > Handling of dynamic WNs
- Review missing bits for EMI MW integration
 - > Job submission and management
 - > Accounting
 - > Monitoring

Torque/MAUI

- Used by most sites, including T1s
 - Torque reasonably maintained but we are still running very old (unmaintained) versions
 - Still used for Moab, the commercial replacement for MAUI
 - No known showstopper for migration to recent versions but some validation/configuration work to be done (e.g. munge)
 - MAUI is a requirement and has been unmaintained for years
 - MAUI is feature rich when Torque has very basic scheduling capabilities
 - Running unmaintained SW is a potential concern, even though every security vulnerability has been fixed by the community
- PIC and NIKHEF reported a successful experience with Torque/MAUI at the 3K job slot scale
 - Not yet convinced of the benefit of moving to something else
 - No major problem so far with MAUI, take in charge its development remains an option...

Univa Grid Engine

- ◉ All the features of major batch systems
 - Fair share, back filling, multi-core job support...
 - Several fair share strategies
- ◉ Several big sites (T1s + large T2s) migrated to Grid Engine
 - UNIVA seems the only alive variant
 - Commercial variant with very good support: sites happy
 - Son of GE (open-source) still alive but not used as far as we know
 - Good feedback: presentations given by KIT and CCIN2P3
 - No scalability issues at the 15-20K job slot scale
 - Well integrated with the MW
 - CCIN2P3 using its site specific integration
- ◉ Multi-core job support without dedicated resources successfully experimented at KIT
 - Using dynamic reservations: 0.5% of CPU usage loss

LSF

- Robust, feature rich, commercial batch system
- Used successfully at CNAF and at several INFN sites
 - National license for INFN
 - CNAF: 1400 WNs, 18K job slots, 100K jobs/day
 - Also used at CERN but no report during the meeting
- Lots of tools developed by CNAF to help with LSF monitoring and to integrate it with the dynamic WN infrastructure (WNoDeS)
 - Local development to control packing of jobs on nodes
 - Development in progress for helping with multi-core job placement optimization
- No plan to move to something else
 - But technical feasibility of moving has been assessed recently

HTCondor

- RAL adopted it 6 months ago for its production cluster as a replacement for Torque/MAUI
 - Already used at most OSG sites
 - No major issue migrating: simple configuration, simple to administer, reliable
 - Scalability tests done at a very large scale
 - During test reached 30K simultaneous jobs without problems, 10K in prod
 - Dynamic cluster membership: no predefined list of WN
 - cgroups support may help to prevent resource exhaustion by jobs
- Integrated both with ARC CE and CREAM CE (and OSG!)
 - RAL running 3 ARC and 3 CREAM
- Multi-core job support enabled: several features helping with it
 - See detailed presentation at the Multi-core job TF
- Already a couple of other sites in UK, with ARC CE

SLURM

- Modern, highly scalable, open source batch system
 - Easy to configure
 - Good multi-core job support
 - Good community support + commercial support
 - Successfully tested at the scale of 10K jobs, limit probably higher
- Widely adopted in Nordic countries
 - All Finnish scientific computing centers, Sweden moving towards
 - Also adopted by Swiss CSCS: an HPC center and a WLCG T2
- Working with both ARC CE and CREAM CE
 - EMI-3 required for APEL accounting
- Some weak points also...
 - Release quality, preference for a share file system, identical configuration file on every node at any time...

MW Support and Accounting

- MW support now available for all 5 batch systems in EMI
 - Job submission and management for CREAM: BLAH
 - BDII publication: recent fixes released to fix all known issues
- CREAM Accounting: solutions available for the 5 batch systems
 - No problem with ARC accounting (JURA): no parser involved
 - HTCCondor: currently based on a script converting to Torque format, need to be enhanced as a real parser.
 - No objection/difficulty to do it but no interest expressed when EMI-3 parsers were written

Multi-Core Jobs

- Most of the work happening in the WLCG Ops Coord TF dedicated to multi-core job deployment
 - Fulfill demand of experiments to have ~30% of multicore slots next fall
- Pragmatic work to evaluate technical possibilities of each implementation and find appropriate solutions
 - Hold dedicated workshops on each implementation
 - Avoid starting partitionning of the resources
- Entropy (mix of job types) hardly achieved with WLCG jobs
 - Multi-core jobs increase the need for an efficient back filling strategy to avoid wasting resources
 - But back filling requires short single core jobs advertised as such: not currently the case in WLCG
 - Despite many short jobs, e.g. in Atlas
 - Need to discuss more with VOs this need for a mix of job type

Multi-Core Jobs (UGE)

- Most advanced experience by KIT
 - Described in details during pre-GDb by M. Alef
- UGE scheduler seems very good to allow concurrent scheduling of single core and multi-core jobs
 - Minimal impact on global usage demonstrated at KIT: ~0.5%
 - Parameter to balance the number of multi-core jobs considered at each scheduling pass against the global usage loss
 - At KIT, optimal number is 10 (max_reservation)
- Based on job reservations
 - No pre-defined number of cores per reservation: each job requests the number of cores needed through the JDL
 - At each sched pass, max_reservation multi-core jobs considered
 - Scheduler collects the appropriate number of core for each job with potential backfilling
 - No static partitioning, no max number of multi-core jobs

Multi-Core Jobs (MAUI)

- ◉ Torque/MAUI situation not so bad compared to initial feedback
 - > Credit to Jeff Templon for the real work
- ◉ Similar approach as UGE implemented using MAUI partitions managed by an external script
 - > 2 partitions of nodes: single core and multicore
 - > Standing reservations to allocate block of cores (8)
 - > A cron job dynamically moving nodes from one partition to another according to the load: NIKHEF ready to share it/
 - > NIKHEF observed very good results in term of farm occupancy (98%)
- ◉ See presentations
 - > <https://indico.cern.ch/event/298050/contribution/3/material/slides/1.pdf>
 - > <https://indico.cern.ch/event/305625/contribution/0/material/slides/1.pdf>




Multi-Core Jobs (HTCondor)

- RAL has a very positive experience: enabled multi-core job since the beginning of their move to HTCondor (last Fall)
 - See dedicated talk by I. Collier
- Some features helping with dynamic support of multi-core jobs
 - Partitionable resources: ability to partition a node to run several “small jobs” (compared to node resources)
 - Not only for cores: also memory and disks
 - condor_defrag daemon: allows to do partial drain of WNs to help collecting cores for multi-core jobs
 - Recover from resource partitioning
 - Several configuration parameters allowing to implement different policies

Comparison Table...

- ⦿ A concrete outcome from the meeting...
- ⦿ A summary table produced in Twiki to help sites wanted to review their batch system choice
 - > <https://twiki.cern.ch/twiki/bin/view/LCG/BatchSystemComparison>
 - > Weaknesses, not only strengths/features...
 - > Scale at which problems where observed
 - > Contact of reference sites
- ⦿ Why not in HEPiX web site?
 - > Happened in the WLCG context because of the Torque/MAUI concerns and the work on multicore job support
 - > Recognized as a typical HEPiX topic: no desire to fight against/ignore HEPiX
 - > Difficult to move the page as it has been already advertize but no problem to refer to it and **contribute to it**

... Comparison Table

Functionality	Torque/Maui	SLURM	HTCondor	USGE/SoGE	LSF
Number of sites 	101	10	10	14	7
Developer Support	maui not supported	yes	yes	yes	yes
Community Support	yes for torque	yes through slurm-dev list 	yes through the HTCondor-users list	no/?	no
Dcumentation	Good	Good	Good	less than satisfying/?	Good
Cream integration	partial	partial	partial	full	full
ARC-CE integration	full	full	full	full	no
APEL support	Cream		no	yes	yes
Info System support 	Cream	Cream	no	Cream	Cream
Getting arguments from CE	ARC-CE	ARC-CE	ARC-CE	ARC-CE/Cream	Cream
Licenses/Costs	free	free	free	Univa/free	IBM
Scalability	no	Debated. Probably depending on the hardware type (HPC vs HTC) as much as the configuration of the DB and plug-ins.	yes	yes	6500 nodes
IPv6 support	no	no	yes, with limitations	no	yes
Distribution format	source, rpm	source	source, rpm, tarball	tarball	tarball
High Availability	no	head node failover	central manager & job queue failover	head node failover	head node failover
Stability	low	high	high	high	high
Installation ease	yes	without DB yes	yes	yes	yes

Conclusions

- Very good discussions based on actual experiences
 - A lot of valuable information
- The summary table is a live material to help sharing experience and findings
 - Please, contribute to it!
- A lot of work in progress, in particular for multi-core job support
 - The number one challenge for the future
- Some topics not discussed due to lack of time
 - Dynamic WN handling
- An area for future collaboration between HEPiX and WLCG, as it happened for IPv6?