Oracle Grid Engine at CC-IN2P3 Report after One year

Philippe.Olivero @cc.in2p3.fr

Hepix Fall meeting 2012 - Beijing

dapnia

cea

saclay

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Overview

- Reminder
- Current situation
- Evolution since 1 year
- Positive points
- Negative points
- Oracle Support
- Futur

Reminder

- June 2010 Start of the migration to OGE (from BQS)
- December 2011 End of the migration to OGE
- Since June 2011/06 Oracle Support

- Work done before starting :
  - AFS token support and token renewal
  - Access control to Services (NFS, GPFS mounts)
  - Prolog, Epilog, JSV, load-sensors (used disk space and memory per machine)
  - Interface to GE in CEs (collaboration with CESGA)
  - Accounting adaptation (local and grid)

Current situation (1/2)

- OGE 6.2 Update 6_17
- master : bi-proc - quad cores – 48GB of RAM (Dell R 610) Solaris 10
- ~ 750 machines in the farm
- only one instance of GE for all our needs :
  - Sequential jobs (~640 machines ~15K cores)
  - multicores jobs (~40 machines ~600 cores)
  - parallels jobs (64 machines 1024 cores)
  - interactive jobs (4 machines, 32 cores)

- >16K execution threads or jobs
- ~100 000 ended jobs/day

Current situation (2/2)

- No more failover server (shadow server) but
  - automatic restart procedure in case of service breakdown (20 mns)
  - backup server ready to start

- Spooling files (system internal status) using Berkeley DB with NFS V4
- Stability only since 3 months (patches, bugs fixed, configuration tuning)
- MultiCores and parallel jobs on separate machine groups (workarounds)
- Acceptable scheduler round time now (~30s, peaks to 200s max)

- People ~1 FTE (against 3 FTE with last system BQS (inc 2.2 for Devs))
  - 1 sysadmin (instals, hooks) , 1 GE expert (advices, tests)
  - 2 GE admins, 1 GE operator , 1 GE support expert

Summary of the OGE Farm

					Total			
					Nombre de machines	Nombre de thread d'exécution	Puissance machine (HS06)	%
GE	prod	INTER	Pwg 1950b	8	4	32	302	0 %
			Total		4	32	302	0 %
		SEQ	Pwg C6100	24	315	7 560	69 556	43 %
			Pwg C6220	32	136	4 352	47 785	30 %
			Pwg M610	16	187	2 992	27 706	17 %
			Total		638	14 904	145 047	90 %
		MULTICORES	Pwg C6100	24	1	24	221	0 %
			Pwg M610	16	33	528	4 889	3 %
			Virtual KVM	16	3	48	444	0 %
			Total		37	600	5 555	3 %
		PARA	Pwg M610	16	64	1 024	9 482	6 %
			Total		64	1 024	9 482	6 %
	Total				743	16 560	160 386	100 %
Total					743	16 560	160 386	100 %

Evolution since one year

- Non stable period during months :
  - Frequent switches between master and shadow servers
  - Berkeley DB corrupted with Shadow and NFS
    - (shadow given up now)
  - Scheduling run too often more than 5 mns (up to 1 hour , peaks to 3hours !)
    - Separate multi-cores jobs now
  - Several tries of Berkeley db between NFS and Local, with shadow or not
  - Timeouts with end-users commands (qstat, qlogin, qsub)
  - Adding machines is always generating a heavy load on master
- Different upgrades during this period : U6.12 -> U6.14 -> U6.15 -> U6.17
- Direct phone calls to Oracle developpers needed

Positive points

- Batch System FTEs decreased from 3 to less than 1 FTE (now!)
- One single farm for all types of jobs (but dedicated groups of machines)
- Ease of administration and configuration per user/group/machine, …
  - Useful concepts (hostGroups, UserSets, Projects, RQS, Complexs …)
  - Hierarchic objects with inheritance
  - Fine scheduling regulation (3 classes of fairshare policies)
  - X11 GUI to provide control over all components

- Powerfull RQS (Ressource Quota Sets) used to limit accesses to Storage Services
- Parallel jobs integration easier (Paral. Env. objects) - GE integration in MPI, Mpich2
- Good documentation (Administration and user guides)

Negative points

- Difficulties to get jobs information (different command for running and ended jobs)
- Lost information when job is ended (required resources)
- Ended jobs information in a flat file : tedious to extract it, needing shared FS
- Not easy to monitor jobs or to analyse post mortem
- No smooth spawn of jobs (« Distribution rate »)
- No native system to rotate log and accounting files

- No satisfying service stability yet
- Still a memory leak => preemptive restart

- No native support to interface a cloud :

  Service Delivery Manager and its Cloud Adapter has been given up

Oracle Support

- It has been difficult to access Oracle developpers in first months
  - We can now join them by phone call when needed
  - We have got now a dedicated support for us

- No roadmap for GE (out of their policy)
- Version 6.2 since more than one year now
- Only mportant bugs fixed
  - loss of AFS tokens,
  - delete of ended jobs repository, …

- Several submitted points raised as RFEs, but no date to implement them
  - Currently, 6 RFEs and 2 bugs in queue for next releases

Future

- Upgrade to 6,2 Update 8 scheduled on december (next general outage)
- Suppress multicores dedicated nodes to face LHC new needs
- Implement an integration to openStack (IaaS)
- Improve end-users oriented monitoring
-
- Stay in touch with Univa, and see more deeply their flavor
- Enforce backups for GE admin and GE System-administrator

Questions ?