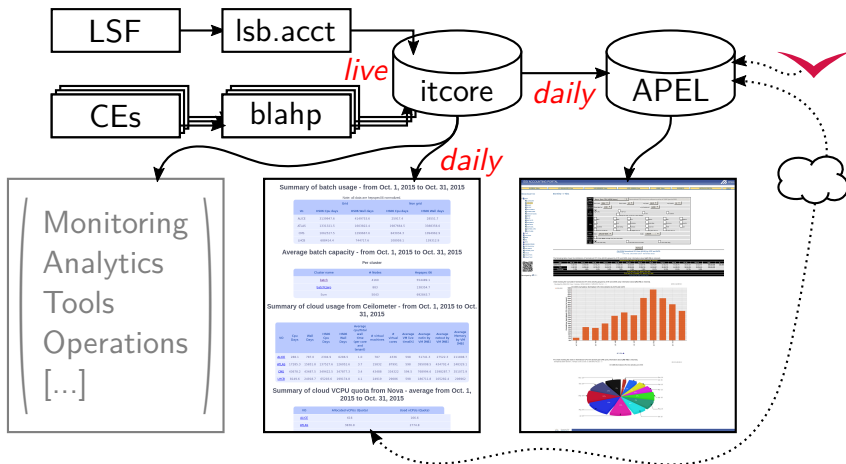




Accounting Data Streams

Jérôme Belleman

Streams



4



Site	H026: Cope days		H026: Wall days	
	mean	sd	mean	sd
AGCE	21.29997.8	4.943713.9	2317.8	28002.7
ATLAS	1305.320.8	360.8822.6	198786.8	3088336.6
CRS	1982507.5	339988.7	643056.3	10884683.9
LACB	889054.5	584717.6	358858.3	129333.8

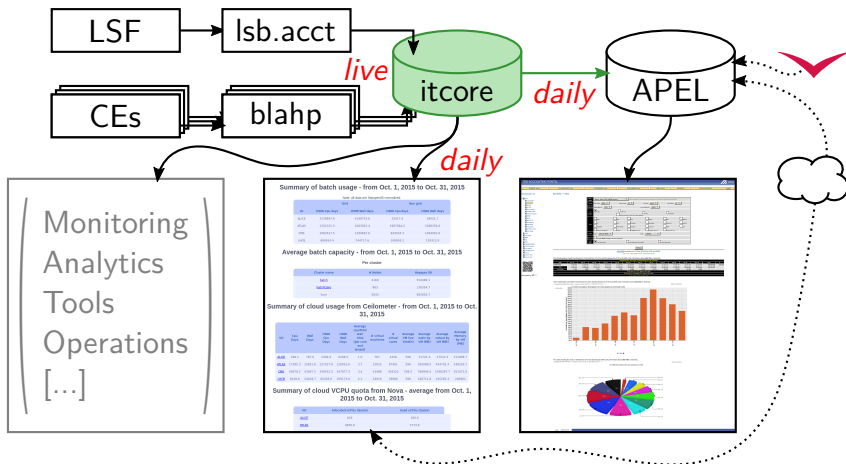
	Per cluster	
Cluster name	# nodes	logper MB
hadoop	1180	0111100.1
hadoop2	910	1307554.7
Sum	5540	0013143.7

V9	Cov	Dys	Udd	Shd	Average					if virtual	if not used	Average	Average	Average
					MSD	MSD	MSD	MSD	MSD					
BASE	190.5	381.0	2084.9	4388.5	1.8	767	4356	596	35740.3	37512.5	24588.8			
ELITE	1735.3	2083.8	12707.8	12081.9	2.7	13482	87891	596	38080.8	43495.8	24038.8			
ELITE	43676.2	43673.7	186820.5	31777.3	3.6	60888	334122	596.8	138804.6	219484.7	30273.8			
ELITE	43676.2	43673.7	43702.8	43702.8	4.5	108888	334122	596.8	138804.6	30273.8	30273.8			

VD	Allocated vCPU (physical)	Used vCPU (physical)
AGE	434	366.0
ELAS	500.0	2714.0

Monthly Accounting,
<https://batchmon.cern.ch/acct>

Joining Data in itcore



LSF Records in itcore (I)

EVENTTYPE	Always JOB_FINISH
VERSION	E.g. 7.06
EVENTTIME	End time
JOBID	E.g. 123897609
USERID	Numeric
USERNAME	UNIX user name
OPTIONS	Bit flags for job processing, I'm not familiar with it
NUMPROCESSORS	Number of cores
JSTATUS	64 for success, 32 for failure
SUBMITTIME	Submission time
BEGINTIME	Unused, unclear, always EPOCH
TERMTIME	Unused, unclear, always EPOCH
STARTTIME	Job start/dispatch time
QUEUE	One of the 60.
RESREQ	E.g. type==SLC5_64
FROMHOST	Submission host, e.g. lxplus302
NUMASKEDHOSTS	Seems to be the max number of worker nodes
ASKEDHOSTS	Probably a result of bsub -m

LSF Records in itcore (II)

HOSTFACTOR	Normalisation factor
NUMEXHOSTS	Actual number of worker nodes, can be < NUMASKEDHOSTS
EXECHOSTS	Actual worker node names, truncated, sadly
JOBNAME	Job name set by the user
COMMAND	Parent process command line
RU_ETIME	User CPU time in s
RU_STIME	System CPU time in s
RU_MINFLT	Number of page reclaims
RU_MAJFLT	Number of page faults
RU_NSWAP	Number of times the process was swapped out
DEPENDCOND	Job dependency condition specified by the user
MAILUSER	Name of the user to whom job related mail was sent
PROJECTNAME	Probably a result of bsub -P
EXITSTATUS	UNIX exit status of the job
MAXNUMPROCESSORS	Seems to equal NUMPROCESSORS most of the time
LOGINSHELL	Sometimes NULL, sometimes /bin/bash, ...
IDX	Job array index
MAXRMEM	Total resident memory usage of all processes in the job
MAXRSWAP	Total virtual memory usage of all processes in the job

LSF Records in itcore (III)

EXITINFO	Job termination reason
CHARGEDSAAP	Charged group/user, e.g. /u_ATLAS/grid_ATLASPRD/atlprd80
APP	Probably always NULL
RUNTIMEESTIMATION	Usually -1, rarely filled in
PUBLISHED	APEL message generation time

CE Records in itcore

USERDN	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=atlpilo2/CN=5...
TIMESTAMP	Which turns out to be the time of submission LSF
CEID	ce206.cern.ch:8443/cream-lsf-grid_ops
LRMSID	LSF job ID, e.g. 123897609
USERFQAN	/alice/Role=NULL/Capability=NULL /alice/alarm/Role=NULL/Ca...

Individual APEL Records (I)

EVENTTIME	End time
JOBID	E.g. 123897609
USERNAME	UNIX user name
NUMPROCS	Number of cores
STARTTIME	Job start/dispatch time
HOSTFACTOR	Normalisation factor
EXECHOSTS	Actual worker node names, truncated, sadly
RU_ETIME	User CPU time in s
RU_STIME	System CPU time in s
IDX	Job array index
MAXMEM	Total resident memory usage of all processes in the job
MAXSWAP	Total virtual memory usage of all processes in the job
CHARGEDSAAP	Charged group/user, e.g. /u_ATLAS/grid_ATLASPRD/atlprd80
PUBLISHED	APEL message generation time
USERDN	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=atlpilo2/CN=5...
TIMESTAMP	Which turns out to be the time of submission LSF
CEID	ce206.cern.ch:8443/cream-lsf-grid_ops
LRMSID	LSF job ID, e.g. 123897609
USERFQAN	/alice/Role=NULL/Capability=NULL /alice/alarm/Role=NULL/Ca...

Individual APEL Records (II)

APEL Message	Our DB	Use
Site	From config file	E.g. CERN-PROD
SubmitHost	CEID	E.g. local-batch
LocalJobId	JOBID IDX	
FQAN	CHARGEDSAAP USERFQAN	VO, VOGroup, VORole
WallDuration	EVENTTIME STARTTIME	
CpuDuration	RU_UTILITY RU_STIME	
Processors	NUMPROCESSORS	Needed for multicore
NodeCount	EXECHOSTS	Needed for multicore
StartTime	STARTTIME	
EndTime	EVENTTIME	
MemoryReal	MAXRMEM	Not needed
MemoryVirtual	MAXRSWAP	Not needed
ServiceLevelType	From config file	E.g. Si2k
ServiceLevel	HOSTFACTOR	E.g. 1000
InfrastructureType	USERFQAN	Local/grid
SubmitHostType	USERFQAN	Local/grid, LSF/Condor, phys/VM
GlobalUserName	USERNAME USERDN	Grid-wide usage/users

https://wiki.egi.eu/wiki/APEL/MessageFormat#Job_Records

APEL Summaries? (I)

2 types:

1. Sent by the APEL client. CPU, wall totals + benchmark. Normalisation done by APEL.
2. https://wiki.egi.eu/wiki/APEL/MessageFormat#Summary_Job_Records, CPU, wall + normalisation of each.

Checking with APEL which one to go for

APEL Summaries? (II)

Site	GOCDB sitename
Month	Month of summary
Year	Year of summary
WallDuration	Sum of wall clock times for all jobs in the month
CpuDuration	Sum of CPU time for all jobs in the month
NormalisedWallDuration	Sum of normalised wall clock time for all jobs
NormalisedCpuDuration	Sum of normalised CPU times for all jobs
NumberOfJobs	Total number of jobs
GlobalUserName	Number of nodes
VO	User's VO
VOGroup	User's VOMS group
VORole	User's VOMS role
SubmitHost	The CE-ID or LRMS hostname
Infrastructure	Grid OR local
Processors	Number of processors
NodeCount	Number of nodes

https://wiki.egi.eu/wiki/APEL/MessageFormat#Summary_Job_Records

APEL Summaries? (III)

Combinations:

- 13 of them
- “All combinations are not usually present”
- How do we choose?

APEL Summaries? (IV)

```
mysql> describe Summaries;
```

Field	Key
SiteID	PRI
Month	PRI
Year	PRI
GlobalUserNameID	PRI
VOID	PRI
VOGroupID	PRI
VORoleID	PRI
SubmitHostId	PRI
InfrastructureType	PRI
ServiceLevelType	PRI
ServiceLevel	PRI
NodeCount	PRI
Processors	PRI

APEL Summaries? (V)

Normal frequency practice:

- Updated summaries each day
- Replacing previous day

Anyway, do we want to?

Links

- <https://twiki.cern.ch/twiki/bin/view/Batch/AccountingDoc>
- <http://accounting.egi.eu/tier1.php>
- <https://batchmon.cern.ch/acct>
- <https://wiki.egi.eu/wiki/APEL/MessageFormat>
- <https://twiki.cern.ch/twiki/bin/viewauth/Batch/APELMessageFormats>



www.cern.ch