Operations in the biomed VO

Tristan Glatard
CNRS-Creatis, Lyon, France

EGI technical forum
14/09/2010

Creatis

biomed overview

- **Users**
  - 300+ DNs from ~20 different countries
- **Applications fields**
  - Bioinformatics
  - Drug discovery
  - Medical imaging
- **Resources**
  - 238 CEs from 129 sites
  - 74 SEs
- **Heterogeneous tooling**
  - Job submission: custom scripts, pilot-job systems, workflows, ...
  - Data management: LFC, AMGA, SRB/iRODS
  - No central monitoring service such as Dashboard

Usage stats

- **Activity (Sep 09 – Aug 10), according to CESGA**
  - 186 kjobs / month
  - 480,225 CPU hours / month (speed-up = 657)
  - CPU / data transfer = 37.1%
- **Data transfers is the main source of known errors**
- **For instance in** [Li et al, HealthGrid 2010]
  - Cannot download input (76.7%) ; cannot upload output (9.9%)
  - Pilot/master com. Problem (8.8%)
  - Application error (4.8%)

VO administration

- **User management**
  - Registration, group creation
  - Organization of LFC directory structure
- **Technical teams**
  - 6 teams of 1 or 2 persons
  - On shift for 1 or 2 weeks
  - Operational procedures
    - Problem detection, monitoring
    - Problem resolution
    - Long-term fixes
- **Targets**
  - LFC, SE availability, full SEs, SE decommissioning, VOMS

Monitoring system

- **Probes included in Hudson servers (CNRS-I3S)**

Hudson

search | log in

Hudson » Biomed VO Storage Elements

ENABLE AUTO REFRESH

Build History

Build Queue

No builds in the queue.

Build Executor Status

#	Status	
1	Idle	
2	Idle	
3	Idle	
4	Idle	
5	Idle	
6	Idle	
7	Idle	
8	Idle	
9	Idle	
10	Idle	
11	Idle	
12	Idle	
13	Idle	
14	Idle	
15	Idle	
16	Idle	
17	Idle	
18	Idle	
19	Idle	
20	Idle	

Egee Biomed VO storage elements tests

All | **Biomed VO Storage Elements** | Biomed VO general tests

S	W ↓	Job	Last Success	Last Failure	Last Duration
		SE griditse01.na.infn.it	N/A	5 hr 40 min (#7)	1 min 2 sec
		SE gridsrm.pi.infn.it	N/A	5 hr 40 min (#7)	23 sec
		SE pcncp22.ncp.edu.pk	1 mo 29 days (#16)	1 mo 27 days (#23)	1.4 sec
		SE se1-egee.srce.hr	1 mo 1 day (#23)	29 days (#29)	0.43 sec
		SE serv02.hep.phy.cam.ac.uk	N/A	5 hr 36 min (#7)	19 sec
		SE t2-se-00.to.infn.it	N/A	5 hr 35 min (#7)	19 sec
		SE srm.glite.ecdf.ed.ac.uk	N/A	29 days (#29)	25 sec
		SE se2.egee.cesga.es	2 mo 7 days (#18)	1 mo 29 days (#31)	45 min
		SE lcgse.psn.ru	N/A	1 mo 27 days (#23)	13 sec
		SE fornax-se.itwm.fhg.de	1 mo 16 days (#19)	1 mo 13 days (#29)	20 sec
		SE se01.marie.hellasgrid.gr	1 day 23 hr (#29)	1 day 17 hr (#30)	2 min 28 sec
		SE agh3.atlas.unimelb.edu.au	29 days (#28)	29 days (#29)	2 min 42 sec
		SE dpm.cyf-kr.edu.pl	2 mo 11 days (#29)	N/A	20 sec
		SE egee-ce1.gup.uni-linz.ac.at	4 mo 21 days (#16)	N/A	4.6 sec
		SE egee-se.grid.niif.hu	1 mo 6 days (#29)	1 mo 10 days (#15)	17 sec
		SE egee2.irb.hr	5 hr 40 min (#7)	N/A	2 min 25 sec
		SE epgse1.ph.bham.ac.uk	1 mo 6 days (#29)	N/A	19 sec
		SE eymir.grid.metu.edu.tr	1 day 17 hr (#30)	5 days 5 hr (#15)	5 sec
		SE fal-pygrid-30.lancs.ac.uk	5 hr 40 min (#7)	N/A	4.7 sec
		SE plethon.grid.ucy.ac.cy	1 mo 6 days (#29)	N/A	21 sec
		SE fornax-se2.itwm.fhg.de	2 mo 25 days (#29)	N/A	13 sec
		SE gfe02.grid.hep.ph.ic.ac.uk	5 hr 40 min (#7)	N/A	2 min 30 sec

Creatis

LFC

- **Tested functionality**
  - `lfc-ls /grid` response time is less than 30s.
- **Problem detection, monitoring**
  - Detected only once the problem happens
  - Top-priority issue (stops production)
- **Problem resolution (LFC admins @ CC-IN2P3)**
  - Increase the number of LFC threads (up to 99)
  - Add servers behind DNS alias (currently 2)
  - Sword of Damocles in peak production periods
- **Long-term fixes**
  - VO-level: split users / groups in different LFCs
  - Infrastructure-level: master-slave LFC
  - Avoid single point of failure with a distributed catalog?

Creatis

Data transfers

- **Tested functionality**
  - lcg-cr, lcg-cr, lcg-del from a UI to all SEs
- **Problem detection, monitoring**
  - Two Hudson servers (different DNs and UIs)
  - SEs showing in the top BDII are tested every 6 hours
- **Problem resolution**
  - Submit and follow-up on team ticket if none is open
  - Several DN-specific issues
- **Long-term fixes**
  - SEs in maintenance should be removed from BDII
  - Broadcast fixes to SEs (e.g. change in CA name)

Creatis

Full SEs

- **Tested functionality**
  - lcg-cr does not produce “No space left on disk” message
  - Critical issue when jobs write on site's default SEs
- **Problem detection**
  - Same as before (i.e. once the problem happened...)
- **Problem resolution**
  - Open team ticket ; assign it to VOSupport
  - Check the **consistency of lcg-infosites** w.r.t available space
  - Ask list of LFNs/DNs having SURLs on SE **to the LFC admin**
  - Ask **the users** to move or delete their data
  - Send **reminders** every week until at least 50% (or at least 500 Go for big SEs) of the SE space for biomed is free
- **Long-term fixes**
  - Raise alarms when x% of the SE is full
  - Provide data migration service (can crash LFC)

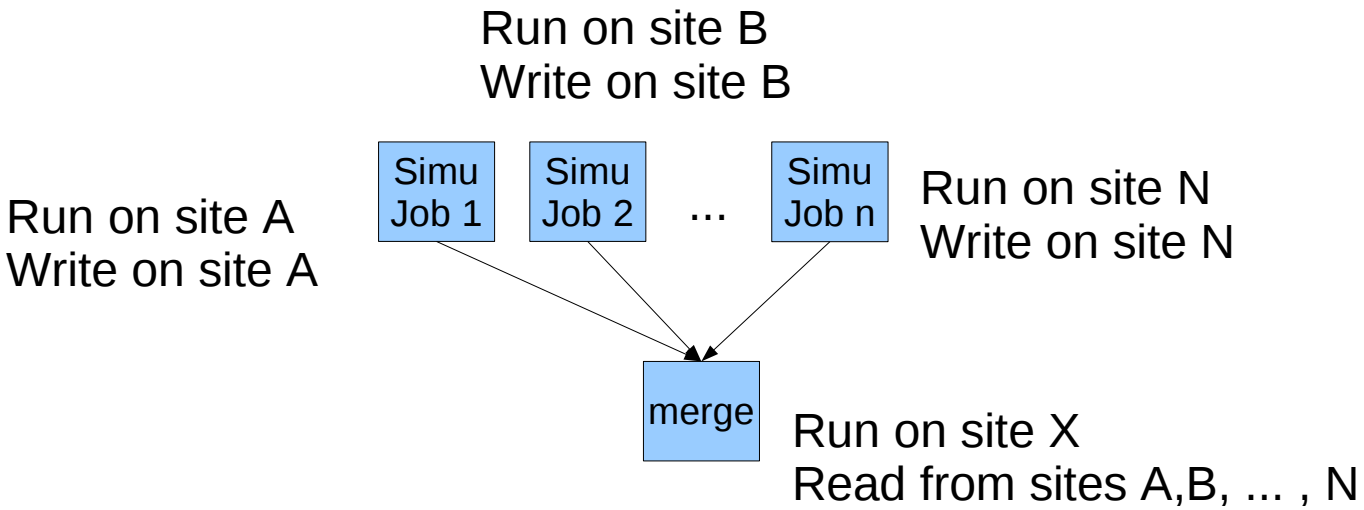
Creatis

Decommissioned SEs

- **Problem detection**
  - Lucky day: get information from broadcast emails
  - Otherwise: when the SE is decommissioned
- **Problem resolution**
  - Open a GGUS team ticket including the expected decommissioning date
  - Ask the site admin to forbid lcg-cr writes from now on
  - **while** (decommissioning deadline is not passed)
    - Ask list of LFNs/DNs having SURLs on SE
    - Notify the concerned users
  - **end while**
  - Send decommissioning green light to SE.
- **Long-term fixes**
  - Improve decommissioning notifications
  - Provide data migration service

Monitoring SE/CE pair: use-case

- **Example on a merge scenario**



- **Alternates**
  - SE A, B, ... N are accessible from site X (not likely when N increases)
  - Simulation jobs adjust their destination SE based on SE/CE monitoring

Creatis

Conclusions / lessons learned

- **Results**
  - May-August: ~17 team tickets per month (used to be ~35)
- **Missing / flawed information**
  - Available storage space on SEs (for monitoring)
  - Decommissioning announcements
  - Remove site in maintenance from BDII
- **Monitoring, error detection**
  - How to monitor transfers from SEs to CEs?
  - Job submission (test WMS/CE pairs)

Creatis

Conclusions / lessons learned

- **Error resolution: site-level**
  - Integrate tests in Nagios box (with French NGI) ; site operators detect/solve them autonomously
  - Better broadcast / share incident fixes among sites (redundant tickets / solutions)
- **Error resolution: VO-level**
  - Missing critical data management services
    - Data migration
    - Improved LFC clients for VO mgt (list of SURLs/SE on an SE, change in SE name)
  - LFC = single-point of failure
  - SE tags (OK vs faulty)
- **These doesn't seem biomed-specific**
  - Actions from EMI?
  - Technical recommendations from EGI to VOs?