

LIP-Lisbon: operational report

Mário David
(*david@lip.pt*)

LIP Lisboa

- **LIP-Lisbon: Tier2 (associated Tier 1 PIC): ATLAS and CMS.**
- **WN's/Cluster:**
 - Shared between local and grid users.
 - OS: SLC4 since a long time.
 - Problems trying to install SLC3 in new HW, in particular for new sites.
 - To much dependency of gLite distro on SLC3 (or RHEL3 like distros).
 - Batch system: SGE (being certified EGEE → Official support):
 - Powerful and simple management and configuration: Nice admin GUI also for WN monitoring.
 - What you configure is what you get: policies, priorities, etc..
 - Some WN's with XEN:
 - At least 1 core/1GB mem per virtual machine, new machines with 2GB per core.
 - Shared with other projects: EELA, I2G, NGI. Plus support for MPI.
- **SE dCache:**
 - Local users: pnfs mounted on WN's and Workstations, dcap.
 - Grid users: gsidcap, gsiftp
 - Backup with AMANDA: special “tar” compiled with dcap lib.

- **Local:**

- Installation and configuration:
 - *All nodes: YUM and YAIM, except:*
 - *WN's: Quattor, and in near future the dcache pool nodes.*
- Monitoring with Nagios: still need more sensors for grid MW specific services/deamons.
- Local network monitoring with Cacti.
- Iperf to debug network throughput issues and bottlenecks.

- **Global:**

- SAM, Gstat: SWE-ROC implemented mail notification when a site fails: very useful before we get CIC/GGUS tickets.
- SAM Admin page: very useful helps in debugging the problems.
- GGUS and LCG-Rollout:
 - *When there are problems that site admins are not able to solve:*
 - *LCG-Rollout is faster and wider “audience” but in general solutions are not gathered for the future.*
 - *GGUS in principle more reliable, but sometimes slow, some SU's almost unresponsive:*
 - *“Does anyone uses the search engine in GGUS?”*
- GOCDDB: for (un)schedule downtimes/maintenance periods.
- CIC portal:
 - *RC reports.*
 - *VO configuration parameters, although the VO managers should take a real effort to put everything there.*

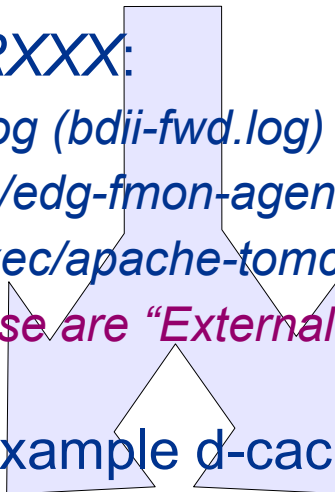
- **Start on documentation:**

- Striped from my mail to Alessandra, Jeff and Daniele:
 - “As far as I know there is nowhere documented the full set of services which should be running for each node type, what are the relevant configuration files for each daemon and what/where are their log files.”
 - “The guide should contain a kind of “ChangeLog” and all known issues of the current and previous releases. Of course if a given issue was solved in the meantime, only the Chagelog will mention it.”
 - If I miss 1 or 2 (or more releases) I don't know anymore which are the issues I have to worry about, and if the issues of the previous ones are still there or not.

- **SAM level/tests:**

- Not yet clear when failures are due to “local” services (sites) or “core” services. I know it is being worked on as we speak.
- A lot of improvement needed in the interaction CIC-CIC Operator-SAM-GOCDB-GGUS.
 - *Describe example GGUS tickets #22222 and #22250: PPS-LIP in downtime when the tickets were opened.*

- **General:**
 - Clear log messages: Documentation of errors, possible causes.
 - **Huge effort** going through the LCG Rollout, but I think we would “catch” a lot of “fish” here and....
 - All logs in `/var/log/DIRXXX`:
 - `/opt/bdii/var/bdii.log (bdii-fwd.log)`
 - `/opt/edg/var/fmon/edg-fmon-agent.log`
 - `/opt/d-cache/libexec/apache-tomcat-5.5.20/logs/catalina.out`
 - **OK so a lot of these are “External” so “External” puts wherever they want.**
 - `logrotate` everything: example d-cache logs.
- **“Again documentation”: A proper “admin manual”:**
 - What services should be running in each node type, are there “optional” services?
 - What is the “behavior” of the service, it's config files, it's log files, ...
 - Daemons in “debug/foreground” mode.
 - **Huge effort:** Troubleshooting/FAQ. (Presently completely dispersed when it exists).



- **Scheduled interventions:**
 - *Electric power maintenance.*
 - *Major movement of machines due to construction work in Comp. Center.*
 - *Major upgrades in SE dCache, LFC.*
- **Unscheduled interventions:**
 - *Unwanted electric power cut's:*
 - They occur in general on **weekends** and **holidays** followed sometimes by the failure in the UPS's.
 - *Network failures (sometimes in those cases not even “we” can reach GOCDB to put the downtime).*
 - *“Bad” releases: bugs/issues appearing after an upgrade of the MW, those “guys” pass unnoticed by the certification, pre-prod,...*
 - *From time to time one or more dCache components gives problems, although in many cases a restart of the d-cache daemons or pools is sufficient to put the system up and running. (Example of CIC/GGUS ticket #23122 from last Monday).*

- ???
- Really don't know what to say here.
-
- Raise awareness of site admins to “**certain types**” of “Subjects” in the EGEE broadcasts?

- **Since many months we update “on the fly” the following:**
 - lcg-CE and WN's: with restart of services when needed.
 - Recent glite 3.1 WN compiled in SLC4, we will do several bunches of WN's at a time:
 1. *Take out some WN's.*
 2. *Re-install.*
 3. *Put back into production.*
 4. *See how they behave, and back to 1.*
 - top-BDII, lcg-RB: with restart of services when needed.
 - MON box.
 - MyProxy (never changes!).
- **Scheduled downtime for:**
 - SE dCache, we take as much care as possible even for “minor” upgrades of the dcache-server.
 - LFC take special care for the MySQL database, specially if it implies a change in the DB schema.

- We will not “upgrade” either the lcg-CE or the lcg-RB into glite-CE and glite-WMS, we will deploy those services in addition, and will let all run in parallel until we have confidence in the glite flavors.

- **Updates to production every 1 or 2 weeks:**
 - Major effort for small sites/small # of human resources.
 - I understand the philosophy of “continues updates without breaking the production” but... We still have “breaks” in some services due to bugs/issues entering a production infrastructure.
 - Wouldn't it be preferable of have less number of updates “per/month” and more bug fixes per update?
 - OK, so this would have to be a big effort in the improvement of the quality of the MW reaching production, in order that a new “major” release is “Production Quality” and doesn't break anything.
 - Does it make sense to have “very buggy” services in production?
 - *glite-CE (and glite-WMS), where site admins spend a lot of their time trying to understand why it does not work properly, or when they have it running, it's again a headache after the last upgrade??*
- I didn't install any glite-CE or glite-WMS in PPS, because I don't want to spend most of my time debugging it, or asking why it does not work properly anymore after the last upgrade.

- **Communication with users:**
 - Local users problems: personal or phone contact and e-mail.
 - External users: GGUS, sometimes e-mail.
- **Site problems detection by COD:**
 - The percentage of problem detection before COD, is very high, in general COD get's at it first early Monday morning, or the day after a holiday, i.e., when the problem occurs at weekends, holidays or during th night.