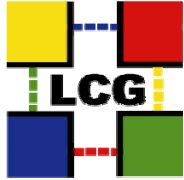




CERN site report

Operational aspects of Grid Services at the Tier-0

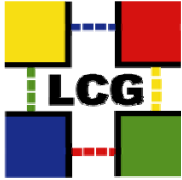




- **General infrastructure**
 - Computer Center refurbishment
 - LCG backbone
 - CPU, Disk, Tape
- **Running Services at CERN**
 - Operational 'Standard Model'
- **Grid Services – service manager's perspective**
 - Workload Management
 - Data Management
 - Authentication, Authorization, Information, Monitoring, VO boxes
- **Conclusions**

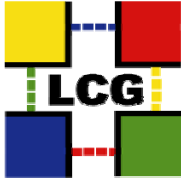


General Infrastructure



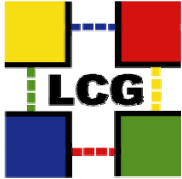
Computer Center refurbishment

- **What still needs to be done in 2007? Very little!**
 - Electrical power: full 2.5 MW UPS coverage for physics from Feb 5th.
 - Cooling: Major refurbishment and upgrade in 2006
 - Little work outstanding: just connection of six ducts to the cold air plenum to improve underfloor airflow on the “barn” side of the machine room.
- **However:**
 - Demands for “critical power” (i.e. with indefinite diesel backup) significantly exceed our 2001/2 estimates
 - The major growth is in power required for networking and databases.
 - There is not much scope for flexibility. Consider replicating elsewhere servers/services needed to maintain operations at other sites in the event of a power cut at CERN
 - Overall power demand is likely to exceed 2.5MW capacity by 2010 at the latest; Planning for a future upgrade has started.
 - Foresee load of 10MW in 2016 rising to over 20MW in 2020 even at rather low predictions of the increase in year on year computing requirements.



Network status

- **LCG backbone in the Computer Center fully deployed**
 - Bandwidth between routers has been upgraded to 80 Gb/s this week
 - Additional router may have to be added for the new hardware
- **In 2006, all physics IP-services were moved to LCG backbone**
 - General services (incl Ixplus) remain on the GPN
- **In general, the deployment went smoothly**
- **To be done:**
 - Connect experimental zones to LCG network
 - Alice have been done already
 - Replace 63 HP switches in the Computer Center
 - Impact on services in Q1

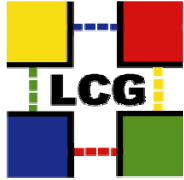


Hardware arrival

- **2007: CPU, disk, tape requirements from LHC experiments (*c-RRB, Oct-06*) are met already:**
 - CPU: 7.6 MSI2K (~3000 boxes)
 - Disk: 1.5 PB (300 diskservers)
 - Tape: 5 PB tape space (10000 cartridges, IBM, STK)
- **Hardware is under acceptance, expect deployment in coming weeks**
- **No more hardware purchases foreseen for this year**
- **2008: large jump in capacity:**
 - CPU: 21.1 MSI2K (1000+ boxes)
 - Disk: 4.2 PB
 - Tape: increment by 8 PB
- **Need to decide soon on what configurations to buy to meet CERN procurement deadlines**
 - Tender adjudication Oct 2007
 - Installation Dec/Jan

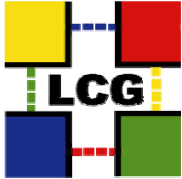


Running Services at CERN

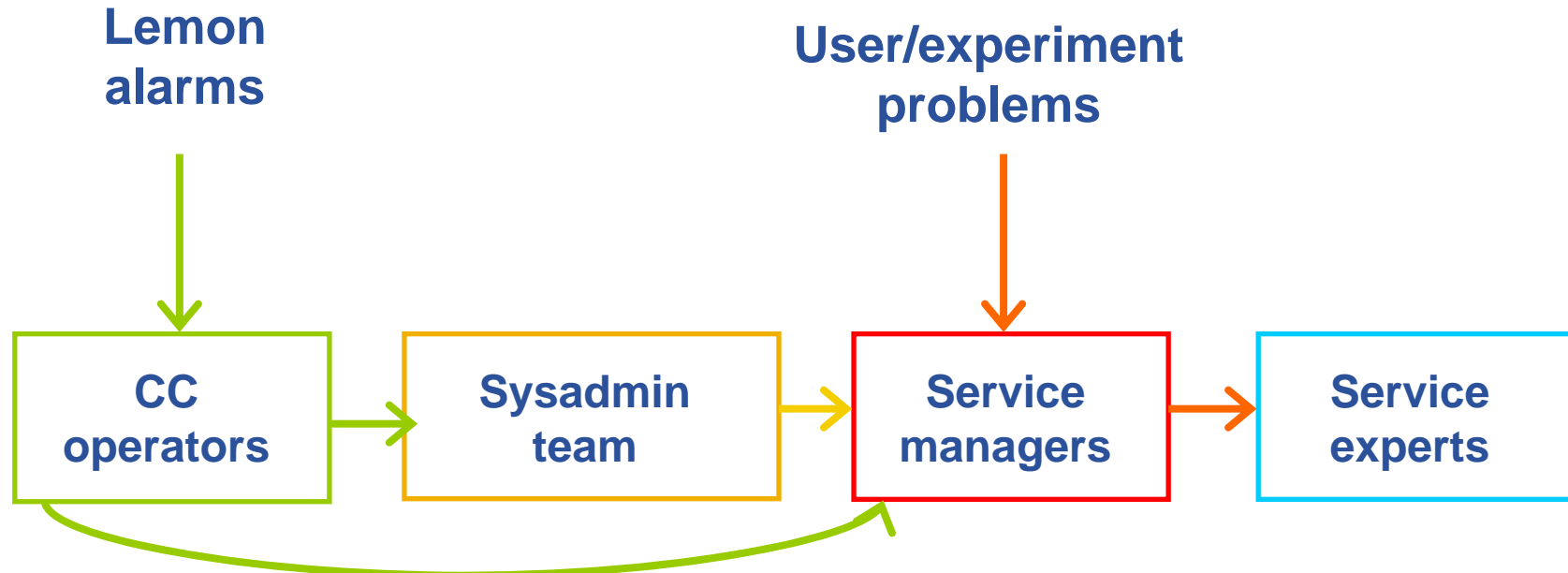


Grand Unification

- **In 2006, we have extended our operational ‘Standard Model’ to cover Grid services:**
 - Use ELFms toolset for installing, configuring, updating services
 - And automate what can be automated
 - Monitor hardware, system, applications
 - Apply recovery actions, raise alarms, plot performance
 - Define workflows involving the different operations teams
 - With an emphasis on documentation!
- **This applies to ~15 services, and should scale to 5000 servers!**
- **WLCG Service Coordination Meeting**
 - Weekly CERN internal meeting for WLCG service providers focusing on issues on a timescale of ~1 week to ~1 month
 - Includes Experiment Integration Support team
- **Operations Morning Meeting**
 - SysAdmins, Service managers, service experts discussing punctual operational problems (yesterday’s alarms, ongoing Grid activities)
 - Experiments call in during service challenges
- **Weekly review of tickets**



Grid services support flow



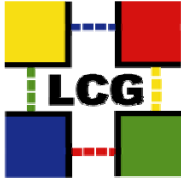
• 24 x 7 coverage
 • 1st level alarm handling
 • Driven by procedures

• 2nd level alarms
 • System installation
 • Manage Hardware configuration

• Service responsible
 • Applies s/w upgrades,
 • Configuration changes, provides

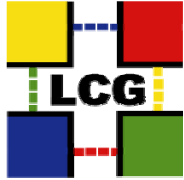
• Entry point for support lines
 • SMod, GMod – person on rota

Problem reports come to the Service Managers via many different flows, using many different tools, directly and indirectly. May end up in bug to developers
 This still needs some tuning
 Piquet Working Group looking into the possibility to establish piquet services for critical services



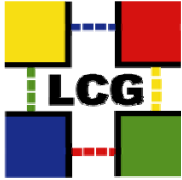
Examples

- **Very simple example: daemon dies regularly**
 - Lemon actuator tries to restart it.
 - An operator alarm is raised in case of failure to restart
- **Simple**
 - Hardware problem generates Lemon alarm.
 - Operator escalates to SysAdmins, who open a Vendor Call for repair
 - If necessary, the Service Manager is asked to drain the server.
- **Not so simple: New problem, for which no alarm exists yet ☺**
 - User complains “connection times out”
 - Service Manager investigates, and fixes or escalates
 - An alarm is added
- **Problem to be tuned:**
 - Different reporting channels start different flows. LFC.Support != GGUS ticket...



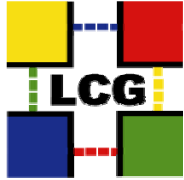
2006: Hardening the services

- **Hardware refresh**
 - basic grid services run on more reliable hardware:
 - dual power supplies, hot-swappable
 - mirrored disks (130 – 220 GB), hot-swappable
 - 2GB or 4GB of RAM
 - Currently reviewing VO box deployments
- **Redundant services**
 - Load-balanced services with servers spread over multiple network switches (LFC, FTS, BDII, SRM, ...)
 - Use of HA Linux (MyProxy, VOMS)
- **Optimized RAID, filesystem options (RB's)**



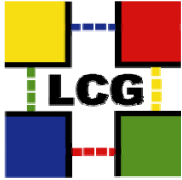
2006: deployment improvements

- **Improved deployment of gLite point releases**
 - Avoid drift in versions of deployed packages
 - Avoid expiry of CA's
- **Deployed many new Quattor components**
 - Control software configurations
- **We use YAIM to configure Grid services:**
 - Configuration information stored inside Quattor
 - A Quattor component generates siteinfo.def, and runs YAIM
 - Now also on SE_castor nodetype 😊
- **Result: more consistent deployments, easier to manage, easier to debug, improved stability**
 - But OS upgrades occur 'at random' and have caused problems on occasion...



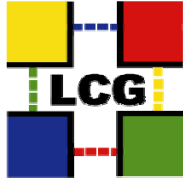
Future software deployment plans

- **We plan to only apply software updates between LHC runs**
- This means that we will prepare updates during LHC runs
- Aim to validate the complete software stack (OS + middleware + applications). Need experiment help!
- To be done in 2007: make a test cluster available to validate upcoming releases

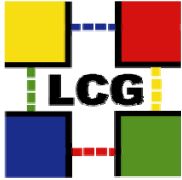


Physics Database services

- **Physics Database services fully based on Oracle RAC/Linux**
 - 7 production RACs clusters, incl WLCG (FTS, LCG, VOMS, ...)
 - A validation service (2-node RACs) is also provided for testing and validating new applications and new s/w and o/s upgrades
 - ORACLE 10gR2 as main platform
- **Currently doubling resources and services**
 - Upgrade production RACs from 4-node to 6- or 8-nodes
 - 110 mid-range servers and 110 disk arrays (1100 disks)
- **Service procedures are in place**
 - On-call team for 24x7 coverage: 6 DBAs and 5 developers
 - Backups on tape and on disk, recovery procedures validated
 - Default backup retention policy and frequency has been agreed with experiments/project
 - Monitoring: Oracle Enterprise Manager for DBAs



- **CERN Certificate Authority reimplemented**
 - **It is now an online CA**
 - Certificates can be issued in 2 clicks through a Web interface
 - In production since Nov 1
- **Host Certificates issuing can be scripted**
 - SOAP Web service
- **Simplifies certificate handling dramatically!**
- **We have integrated automatic host certificate management (including alerts) in the ELFms tools**
- ***And we need it: we have 460 nodes with host certificates, ie 2 expire every working day...***



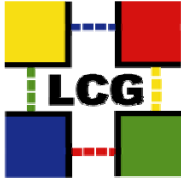
2006: LAN DB sets

- **‘Landb sets’: production machines with common access patterns**
 - ‘Ixplus nodes to be accessed through SSH’
 - ‘diskservers accessed through gridFTP’
 - ‘Castor instance dedicated to a VO’
- **These sets are used to configure**
 - the site firewall
 - **Eliminates the need for security scans on production machines!**
 - the Technical Network between experimental zones and the CC
- **We have automated the updates of the Landb sets**

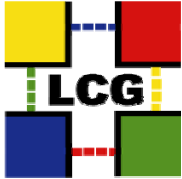


Grid Services

Workload Management



- **We currently run 5 LCG-CE**
 - More CE's have been added over time to cope with increased load
 - But this configuration caused problems:
 - CEs reported different snapshots of the same batch system
published information inconsistent among the CEs
 - Information providers independently querying job status from each CE overloaded LSF master, that started to drop requests
 - Solutions: production CE's share cached status information
 - We also run a CE dedicated to 64-bit SLC4 WN's in pre-production
- **... And 17 LCG-RB's**
 - 7 LCG-RB 3.0.6, partially quattorized, new hardware
 - 10 LCG-RB 2.7.0, not quattor-managed, old hardware
- **Phase-out by Mid-2007?**

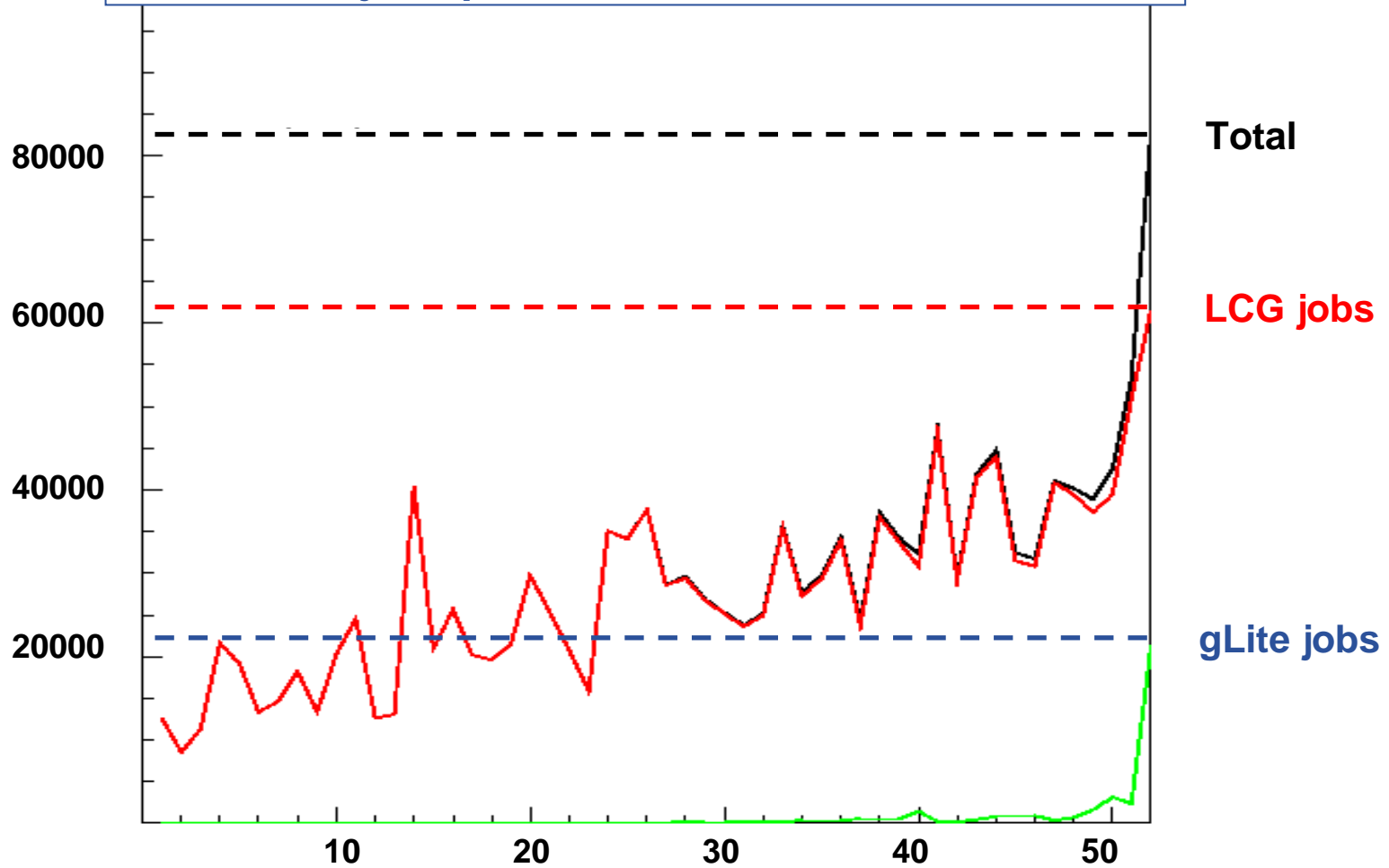


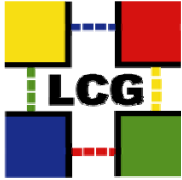
Workload Management System

-
-
-
-

Grid jobs per week on Ixbatch in 2006

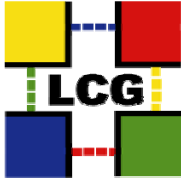
16





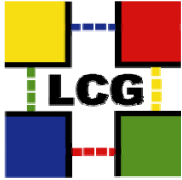
Worker nodes

- **Early 2006: doubling number of WN in Ixbatch showed scaling problems**
 - Increase in LSF reconfiguration times, up to 2 hours. Problematic to debug and difficult to fix, but service running smoothly since July
- **Today we have 2000 WNs in Ixbatch, 800 are being added (4 →8 MSI2K)**
- **Currently running LSF 6.1. We will evaluate the scalability improvements in LSF 7.0. Platform are using our cluster to demonstrate scalability to 10,000 nodes**
- **SLC3 → SLC4 transition**
 - We now run 100 SLC4 WN's, still pre-production
 - Once green light given from PPS, we will:
 - Make CE that submits to SLC4 available
 - Agree with expts on when to reinstall their batch share with SLC4



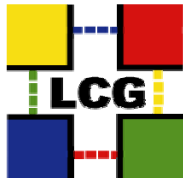
- **Installed on AFS**
 - To be available from lxplus, desktops, laptops, etc
 - No more interference with non-grid lxplus users
 - Preferred by users

- **Problems porting UI to SLC4**
 - Impacts lxplus alias change SLC3 → SLC4 (on Jan 29...)



Grid Services

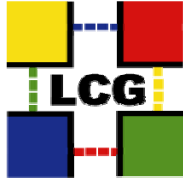
Data Management



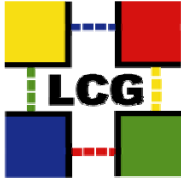
- **CERN tier-0 export FTS deployment**
 - 10 mid-range servers
 - Database on separate RAC cluster
- **Stable 24 x 7 running**
 - Some stability issues encountered and solved in S/W
 - Fabric monitoring and operational procedures in place to catch problems
 - Integrated into experiment software, who are submitting stably to the service since Summer 06
- **Over 6 PB successfully exported from CERN since May 2006**
 - Though much of it “dteam” test data...
- **New monitoring tools available for the “overall service”**
 - Looking at aggregate failure rates on the different export channels
- **Plans:**
 - Better monitoring of the “distributed service”
 - Integrated into the grid operations infrastructure
 - It’ll send an alarm if it detects a problem on your site...
 - FTS 2.0 pilot being deployed



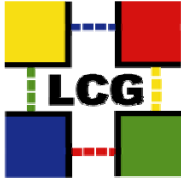
- **Deployed at CERN for all 4 LHC experiments, and dteam/ops VO's**
- **Hardware upgraded last November**
 - 2 load-balanced nodes per experiment service
 - Database on separate RAC
- **Generally stable 24 x 7 running**
 - Monitoring, alarms, procedures in place
 - Significant operational experience gained
 - Monitoring tools added at CERN for LFC
- **No major changes foreseen**



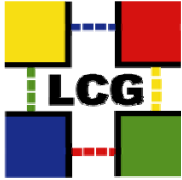
- **SRM v1 endpoints**
 - srm.cern.ch: shared between all VO's
 - srm-durable-`{atlas,lhcb}`.cern.ch: to access durable Disk-1 service classes
- **Service stability reflects underlying MSS...**
 - Shared endpoint can lead to operational problems ☹️
- **Work in progress: deploy Castor SRM v2.2 service**
 - Pre-production endpoint since November
 - Software implementation is proceeding, with regular updates
 - Service is being defined: Hardware configuration, monitoring metrics, alarms, procedures.



- **Castor is the MSS backend for the SE**
- **In spring 2006, we deployed dedicated Castor-2 instances for all LHC VO's**
 - A public instance for all non-LHC was added in September
- **Instances configured and sized specifically for each VO**
 - ~250 TB per VO, with ~6 disk pools
- **Service is still maturing while software is stabilizing**
 - Successfully used in data challenges
 - Long list of operational problems (being) addressed
 - We have deployed many upgrades since Spring 2006
- **Main operational problems:**
 - Limitations in LSF plugin limit the number of jobs in queue
→ limit is being removed, deployment end of Q1
 - Disk-1...

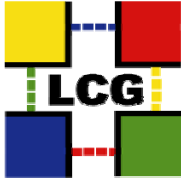


- **Disk-1...**
 - what is the problem?
 - Disk-1 assumes that user responsibility for managing the contents of the diskpool
 - Castor does not provide the means to do this
 - Proposed solutions
 - Reject write request to full diskpool
 - Access control to diskpools
 - NB: We backup Disk-1 data to tape purely for operational reasons!
- **Plans:**
 - Upgrading the instances:
 - Add disk space to setups 1.1 → 1.6 PB
 - Upgrade firmware on ~4000 WD disks... ☹️
 - Move databases to Oracle certified hardware 😊
 - Operational enhancements:
 - Service now stable enough for sysAdmins to handle routine tasks
 - More monitoring, alarms, procedures
 - Improved LSF plugin → end of Q1
 - Migrate non-LHC users to CASTOR-2



Grid Services

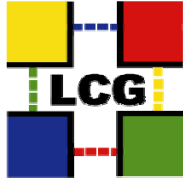
*Authentication, Authorization,
Information, Monitoring, VO boxes*



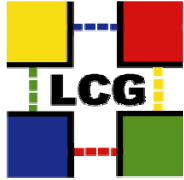
- **Two services myproxy.cern.ch and myproxy-fts.cern.ch**
 - myproxy.cern.ch is the default for sites to configure if they do not have their own MyProxy server nearby (e.g. at a T1)
 - myproxy.cern.ch trusts many remote services to renew proxies, in particular all the VOBOX nodes
- **Both services consist of 2 midrange servers**
 - LinuxHA managed takeover and replication
- **MyProxy Serving around 2,000 requests per hour, very comfortably**
 - For the FTS Myproxy the load is much less, because the same proxy is used for many transfers
- **Service stability: MyProxy not an area of concern**



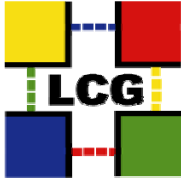
- **Two DNS aliases: voms.cern.ch, lcg-voms.cern.ch**
 - voms used for grid-mapfile, lcg-voms used for registration (VOMRS)
 - Same database, same config, both give voms-proxies for redundancy
 - [siteinfo.def](#) regulates the above
 - Database: 4 listeners for load distributions and automatic failover, implemented in Oracle RAC
- **Hardware**
 - Voms: single mid-range server
 - Lcg-voms: HA linux, 2 midrange servers
- **Lcg-voms: 20 requests per hour, response time 0.2 s**
- **Biggest stability issue: [tomcat memory problems](#)**
 - A configuration option introduced during Xmas seems to have solved the problem...
- **Outlook for next ~6 months**
 - Oracle optimisation
 - change of Java version
 - [new vomrs design](#) (as being discussed in parallel this week)



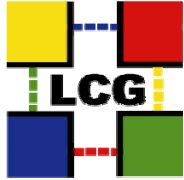
- **We run 3 different BDII's:**
 - Global LCG BDII
 - 6 midrange servers
 - Load-balanced alias: lcg-bdii
 - CERN Site LCG BDII
 - 2 midrange servers
 - Load-balanced alias: prod-bdii
 - SAM BDII
 - 2 midrange servers
 - Load-balanced alias: sam-bdii
- **Biggest operational problem: high loads.
Fixed in July 2006 by adding more servers.**
- **Now, stable running.**



- **SAM submission infrastructure (UI, WMS, RB) on new hardware**
- **Largely quattor managed, Lemon monitored**
- **2 UIs**
 - SAM Client software installed by APT, configured and monitored manually
 - Quattor managed cron jobs with time-distribution: load balancing and fail-over
 - configured to provide load balancing and fail-over of RBs and WMSes
- **2 gLite WMS**
- **2 LCG RBs**
 - RB middleware managed manually by APT
- **SAM server upgrade in progress**
 - 2 redundant machines
 - will be managed by Quattor and monitored by Lemon
 - should be ready by the end of January 2007

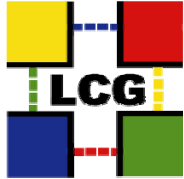


- **Two machines:**
 - OS managed by Quattor and monitored by Lemon (SLC3)
 - application logs monitored by separate process that send email alerts to admins
 - cron jobs to restart services in case of failures
- **Backend machine (grvw002)**
 - R-GMA archiver (Tomcat) and summarization scrips (cron jobs)
 - middleware managed manually
- **Frontend machine (grvw001)**
 - GridView Portal (Apache) and web services (Tomcat)
 - web servers managed manually



VO boxes

- **By construction, VO boxes run VO applications...**
- **... but we try to manage them all in the same way**
 - Quattorized installation/configuration, standard Lemon alarms, etc.
 - Experiments manage their applications, and can provide application-specific monitoring + associated procedures
- **Currently reviewing the setup and the configurations**
 - Hardware refresh: mid-range servers, diskservers



Conclusions

- **General infrastructure largely in place**
- **We have extended our operational ‘Standard Model’ to manage the services in the Computer Center**
- **Hardware refresh for Grid Services nearly finished**
- **Hugely improved Grid operations in 2006**
- **Focus 2007:**
 - SLC3 → SLC4
 - SRM 2.2
 - Final preparations for LHC running (eg. Piquet service)
- **We are confident that the fabric and operational infrastructure to run reliable services is in place**