CERN Site Report
HTASC, Pisa, 13 June 2003

Excerpts from reports to Focus last week by
F. Hemmer and L. Robertson

Thanks to all contributors from IT groups

Internet Services & Desktops

- Mail and collaborative Services
  - Deployment of the new mail infrastructure, 50% of active users migrated
  - Spam Fight improvements
  - New services: Calendar, Webmail interface, Encrypted sessions
- Web Services
  - PHP service investigation
  - Evaluation of WebDav service to access DFS file system from any platforms (Linux, Mac, …) – See picture next slide
- Windows Services
  - About 5500 W2k/XP managed PCs
  - Windows XP Workstation fully supported.
    Now the default for new machines.
  - VPN pilot service in production, ACB service may be restructured
  - Windows Terminal Service prototype will soon be proposed to allow access to selected Windows applications (office) from any platforms (Linux, Mac, …) – See picture slide after next
  - New server hardware will be deployed, Windows server 2003 being evaluated
- Pilot PC rental scheme
  - http://cern.ch/pcdesktop/html/PCRental/PCrental.htm
- Mac Support
  - Support information issued at last Desktop Forum

Databases

- POOL 1.0 has been released (together with EP-SFT / SPI)
  - RedHat 7.3 and gcc 3.2
  - Release notes and further information
    - http://lcgapp.cern.ch/project/persist/relnotes.html
    - http://lcgapp.cern.ch/project/persist
- POOL file catalogue based on EDG-WP2 (rlstest.cern.ch)
  - Uses Oracle 9iAS and Oracle 9i database
  - Will be stress-tested by cataloguing ~1.5M files
  - Selected by POOL_CATALOG env. variable
  - More details on this pre-production service:
    - http://cern.ch/edg-wp2/replication/rls-for-pool.html
- POOL 1.1 will come with experiment specific production catalogues (e.g. rlscms.cern.ch)
  - 2nd half of June
  - Other catalogue implementations still transparently supported

Databases (2)

- Oracle databases for physics (PDB) services
  - general purpose Oracle9i server (2 node Sun cluster) ready for production
    - Based on COCOTIME requests
  - CMS tracker database upgraded to Oracle 9i R2
  - Crystal II database in preparation
  - New backup service based on TSM 5.2 being established by FIO-DS
    - needed for hot incremental backups directly from Oracle9i
  - Using RedHat ES 2.1 operating system on Intel servers
- Objectivity/DB end of service
  - COMPASS data migration (300TB) has been finished
    - migrated from Objectivity to Oracle + DATE files
    - original tapes are being gradually released for reuse
  - HARP raw data (30TB) has been migrated
    - metadata migration sill ongoing
  - Plan to protect Objy s/w tree in AFS from 1 July
    - Access by experiment group code

Architecture and Data Challenges

- The design of the new CASTOR stager has started (ADC)
  - A proposal will be circulated for feedback/comments in coming weeks.
  - A CASTOR users meeting, where the new design will be discussed, is foreseen by end of June (date still to be decided)
- Next CERN Linux release
  - Change of Red Hat support strategy (12 months support for "Desktop release") affects CERN
  - CERN Linux user group meeting on 27.05.03 to discuss impact on next CERN release
- 1GB/s-to-tape data challenge (ADC)
  - simple CDR setup (filter farm → disk → tape)
  - 920 MB/s average over a period of 3 days with a 8 hour period of 1.1 GB/s and peaks of 1.2 GB/s
  - 40 CPU servers, 60 disk servers, on average 45 tape servers
  - Heavy use of the OpenLab equipment (10 Gbit Enterasys router, 20 2xIA64 HP)

Security: Proposed Short-Term Actions (next 3-6 months)

Following serious break-ins the following actions are planned:

- AFS password expiry enforced
  - Passwords must be changed at least once per year
  - Users informed by email 30 days before expiry
- Hardware address registration enforced for portables
  - CERN users register at http://cern.ch/register
  - Visitors request temporary access via a web form
- Off-site ftp closure
  - ftp replaced by SSH (Scp/Sftp) with some exceptions
  - On-site ftp will be available
  - No centrally supported graphical tool on NICE. Free versions (e.g. Winscp, iXplorer) available; license negotiations with ssh.com failed, can be bought individually
- Rules for systems connected to CERN network
  - Will define acceptable network and security practice

Hardware address registration for portables using DHCP

CERN Users

- Register hardware address(es) for portables
  - http://cern.ch/register and select "Register Portable"
  - Requires a NICE account and internal network access

External Visitors

- Fill out a web form to request visitor access
  - Contact data, Hardware Address(es), Start date, Duration (1-4 weeks), CERN contact
- CERN contact receives an email
  - Can accept/refuse responsibility for the visitor
  - Signature requires a NICE account and internal network access
  - Unsigned requests are cancelled after 5 days
- Hardware address(es) accepted by DHCP servers
  - Goal is approx 5 mins after request accepted by CERN contact

External Networking

- DataTAG provides GLUE middleware:
  - For LCG-0
  - To be incorporated in EDG2.0
- INTERNET2 landspeed data transfer record:
  - 2.38 Gbit/s between CERN and Sunnyvale (CA)
- 10 Gbit/s upgrade of shared CERN-SWITCH connection to GEANT at identical financial conditions

Fabric Infrastructure and Operations

- Public interactive/batch RH6 services terminated as agreed 31st May
  - RH7 installation and management using well advanced pilots of EDG automation tools.
  - Minimal & restricted RH6 service for Objectivity users until end-2003.
- STK tape complex migrated to B513 basement
  - Silos installed on loan from STK, then two silos, 24,000 cartridges, 25 tape drives and backup servers moved in 10-day period with minimal perturbation to users.
  - 9940A drives all removed from service at end-May.
- Recruitment underway for insourced SysAdmin
  - Initial service to take over from Serco from September
- Plans for new B513 substation approved
  - Civil engineering to start in August
  - Machine room upgrade in parallel now STK silos in vault.

LCG

from L. Robertson's slides

Goal of the LHC Computing Grid Project - LCG

To prepare and deploy the LHC computing environment
to help the experiments' analyse the data coming from the detectors

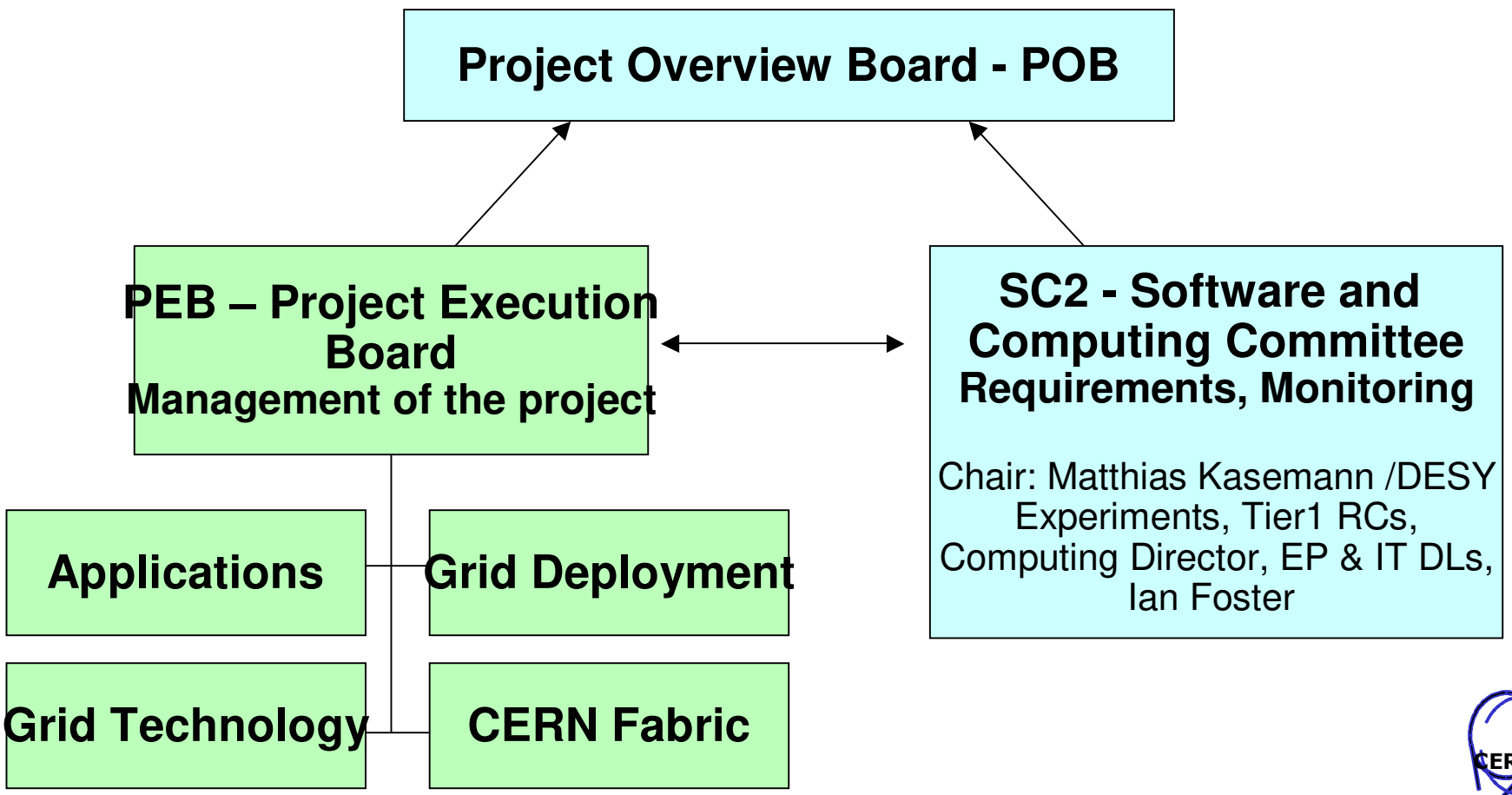
Phase 1 – 2002-05
development of common applications, libraries, frameworks, prototyping of the environment, operation of a pilot computing service

Phase 2 – 2006-08
acquire, build and operate the LHC computing service

Project Organisation

LHCC
Technical Review

Computing RRB
Resources



Project Execution Board

Project Management
project leader
area managers
SC2, GDB chairs
resource & planning officers

Experiment Delegates

External Projects
EDG
VDT
GridPP
INFN Grid

Other Resource Suppliers
IN2P3
Germany
CERN-IT

Applications Area Projects

- Software Process and Infrastructure (SPI) *(Alberto Aimar)*
  - Librarian, QA, testing, developer tools, documentation, training, …
- Persistency Framework (POOL) *(Dirk Duellmann)*
  - Hybrid data store (ROOT+relational database manaement system)
- Core Tools and Services (SEAL) *(Pere Mato)*
  - Foundation and utility libraries, basic framework services, system services, object dictionary and whiteboard, grid enabled services, mathematical libraries
- Physics Interfaces (PI) *(Vincenzo Innocente)*
  - Interfaces and tools by which physicists directly use the software. Interactive (distributed) analysis, visualization, grid portals
- Simulation
  - simulation framework, Geant4, FLUKA, generator services, physics validation

Initial set of Middleware

- VDT - includes components from several different US projects, including Globus and Condor
- European DataGrid – EDG – includes replica catalogue, resource broker, replica management tools, gatekeeper
- Lower functionality than was expected to be available at this time
- So additional components will be added as they become available, and demonstrate sufficient reliability

Where are we with Grid Technology?

- We now have practical experience of deploying grid software from a number of projects
  - in 2002 - successful Monte Carlo productions by experiments using grid software in US (VDT) and Europe (EDG, Nordugrid)
  - demonstration of integrated VDT-EDG grid at Super-Computing 2002 Exhibition
- Much more difficult to achieve production quality than expected
  - expectations were probably too high
  - reliability, scalability, manageability still to be tackled

Where could the middleware come from for LHC startup?

- New architecture defined by Globus –
  - OGSA – Open Grid Systems Architecture
  - based on Web Services
  - backed by several industrial companies
  - waiting for first experience with first implementation of OGSA from Globus
- Not yet clear if industrial products will be useful on LHC timescale
- Proposal made to EU 6th Framework programme for re-engineering of EDG middleware - EGEE
- Proposal for extension of US NSF Middleware Initiative (NMI) to "harden", support US base components

If EGEE is funded .. (we should know by end June)

- EGEE –
  - 2 year project – 2004-05
  - Could be extended for a second two years
  - Middleware re-engineering
  - General European Science Grid Infrastructure
  - Applications support – extending grid experience to other sciences
- Middleware timeline –
  - Second half of 2003 – pre-project planning and design, establish implementation teams
  - 2004 – implement & deliver first set of re-engineered tools
- and if EGEE is not funded…..

Centres taking part in the LCG prototype service – 2003-05

Tier 0

- CERN

Tier 1 Centres

- Brookhaven National Lab
- CNAF Bologna
- Fermilab
- FZK Karlsruhe
- IN2P3 Lyon
- Rutherford Appleton Lab
- University of Tokyo
- CERN

Other Centres

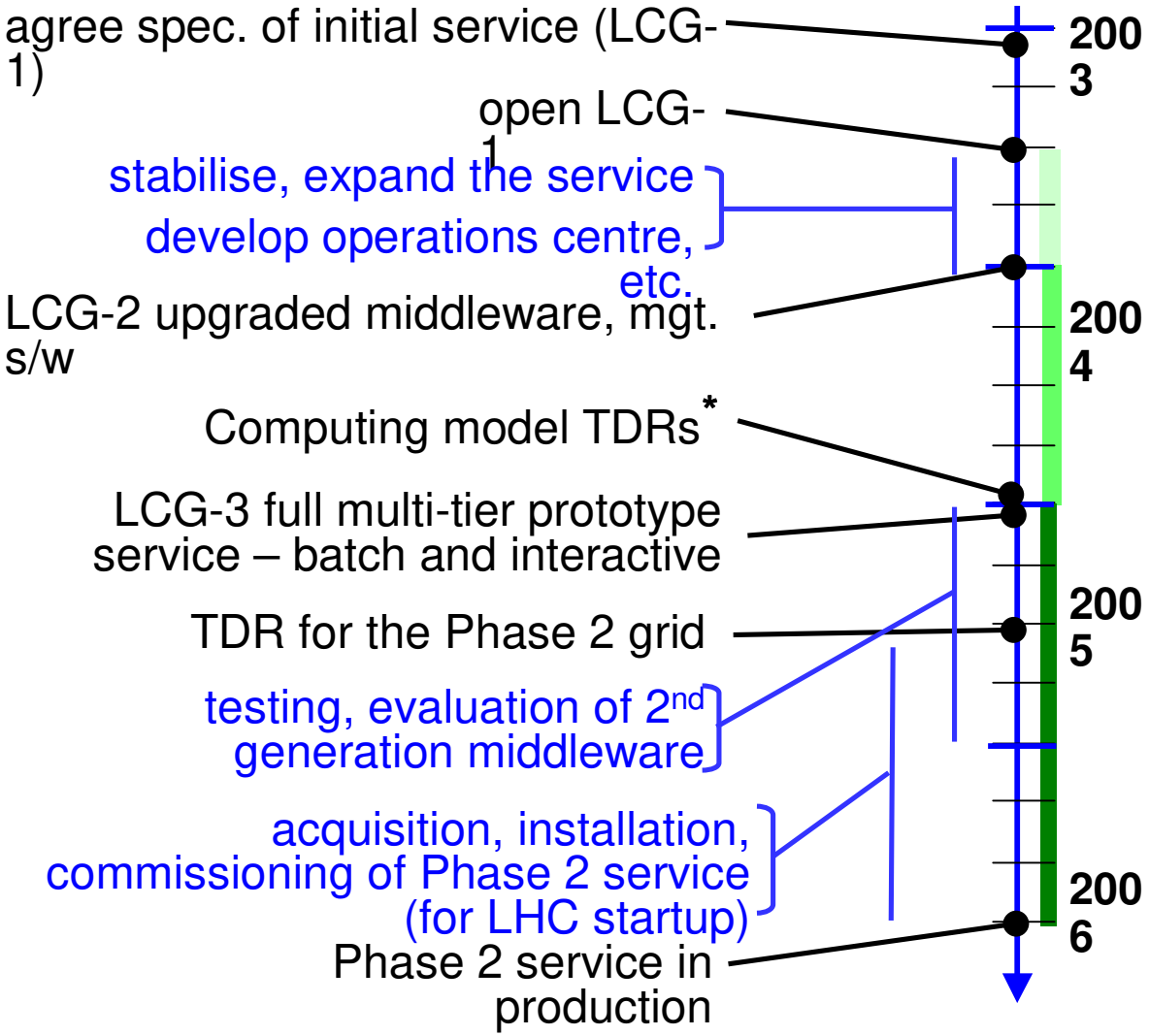
- Academia Sinica (Taipei)
- Barcelona
- Caltech
- GSI Darmstadt
- Italian Tier 2s(Torino, Milano, Legnaro)
- Manno (Switzerland)
- Moscow State University
- NIKHEF Amsterdam
- Ohio Supercomputing Centre
- Sweden (NorduGrid)
- Tata Institute (India)
- UCSD
- UK Tier 2s
- University of Florida- Gainesville
- University of Prague
- ……

Grid Deployment Board

Coordination of the LHC Grid Operation & Management

- Membership –
  - chair – Mirco Mazzucato (Padova)
  - one member for each country with an active Regional Centre
  - two members from each LHC experiment
  - area managers, SC2 chair, project leader
- Monthly meetings since October 2002
  - CERN + Milano (October), Taipei (March), Oxford (July)
- Working groups established to agree on operational issues for LCG-1
  - middleware, resources and scheduling, security and access, operations, user support
- Security group
  - chair – David Kelsey

Timeline for the LCG computing service



used for simulated event productions

principal service for LHC data challenges – batch analysis and simulation

validation of computing models

**** TDR – technical design report***

LCG-1&2 Deployment Goals

- Production service for Data Challenges in 2H03 & 2004
  - Initially focused on batch production work
- Experience in close collaboration between the Regional Centres
  - Must have wide enough participation to understand the issues,
- Learn how to maintain and operate a global grid
- Focus on a production-quality service
  - Robustness, fault-tolerance, predictability, and supportability take precedence; additional functionality gets prioritized
- LCG should be integrated into the sites' physics computing services – should not be something apart
  - This requires coordination between participating sites in:
    - Policies and collaborative agreements
    - Resource planning and scheduling
    - Operations and Support

Progress

- Certification process defined (January)
  - This has been done – agreed common process with EDG
  - Have agreed joint project with VDT (US):
    - VDT provide basic level (Globus, Condor) testing suites
    - We provide higher level testing
  - but need much more effort on devising & writing basic and application-level tests
- Packaging/configuration mechanism defined
  - Group (EDG, LCG, VDT) have documented an agreed common approach
  - Now will proceed with a staged implementation
    - Basic for LCG-1 in July, and more developed later
- Roll-out of the pilot service on schedule
  - tests the process, helps to define the procedures
  - then we will have to formalise agreements on efforts

Progress

- Certification process
  - Process defined in January
  - Certification testbed has been implemented at CERN, will be extended to Wisconsin, Moscow, FNAL (and an INFN site)
  - Joint projects with US groups (Globus, VDT) and Russia to collaborate on tests
- Packaging and configuration process
  - Basic process agreed - implementation proposal under way for longer term
  - For July will use LCFGng for service nodes, optional tools for clusters
- Operations centre
  - RAL will set up the first operations centre, IN2P3-Lyon will collaborate on tools
- Call centre
  - FZK have agreed to lead the activity – building a prototype support portal with problem tracking system for July

LCG-0 Deployment Status

validating the packaging & distribution process

	Site	Scheduled	Status
Tier 1			
0	CERN	15/2/03	Done
1	CNAF	28/2/03	Done
2	RAL	28/2/03	Done
3	FNAL	30/3/03	Done
4	Taipei	15/4/03	Done
5	FZK	30/4/03	In progress
6	IN2P3	7/5/03	In prep.
7	BNL	15/5/03	In prep.
8	Russia (Moscow)	21/5/03	In prep.
9	Tokyo	21/5/03	Done
Tier 2			
10	Legnaro (INFN)	After CNAF	Done

Events during the last few months

High Performance Data Recording

- December 2002 - ALICE-IT Computing Data Challenge
  - reached ~300 MB/s sustained (for 7 days) from an emulated DAQ system into the CASTOR mass storage system (disk and tape)
  - peak date rate ~350 MB/s
  - the goal was 200 MB/s.

Cluster Scaling

- January 2004
  - ATLAS used 230 testbed nodes for an Online Computing Data Challenge to test event building and run control issues

Technology tracking

- PASTA III technology survey complete
  - many institutes involved
  - conclusions presented at CHEP

GigaByte/second data recording challenge

2003

- **Migration of LXBATCH systems to LCG-1 grid service**
- **Migration of equipment into the vault in Building 513**
  - selected servers from computer room, all new systems, tape robots
  - clearing current computer room for power upgrade in 2005
- **Civil engineering for new high voltage equipment**
- **1 GByte/s IT Computing Data Challenge in April**
  - takes advantage of a period of overlap of old and new tape drives
  - 'CDR' system at → 50 cpu server + 50 disk server + 50 tape drives
  - target → 1 GByte/s through CASTOR to tape
- **Further steps in monitoring, installation automation programme of Computer Centre clusters**

Priorities for 2003

- Establish the LHC Grid as an operational service
  - important to keep the pressure/focus on the July release of the LCG service
  - reliability → platform for Monte Carlo production
- Integrate LXBATCH in the service
- Establish a plan for the middleware required for Phase 2
  - base functionality & performance
  - identify suppliers and maintainers
- Initial object persistency release – mid-year
  - POOL and ROOT I/O
- Distributed production environment
  - integrating experiment specific code with common applications services
- GEANT 4
  - validation activity → development priorities for LHC

Summary

- Applications area *scope* close to completion
- Good collaboration with & participation of experiments
- POOL release at end of June on schedule
- Grid service – good collaboration with regional centres
- Middleware going slower than hoped
- EGEE decision crucial for future strategy
- In July we shall see the real problems of operating a grid as a service

Thanks, Francesco!