GRID Activities at LIP

Portuguese CMS Group Meeting

Jorge Gomes - jorge@lip.pt
(LIP Computer Centre)

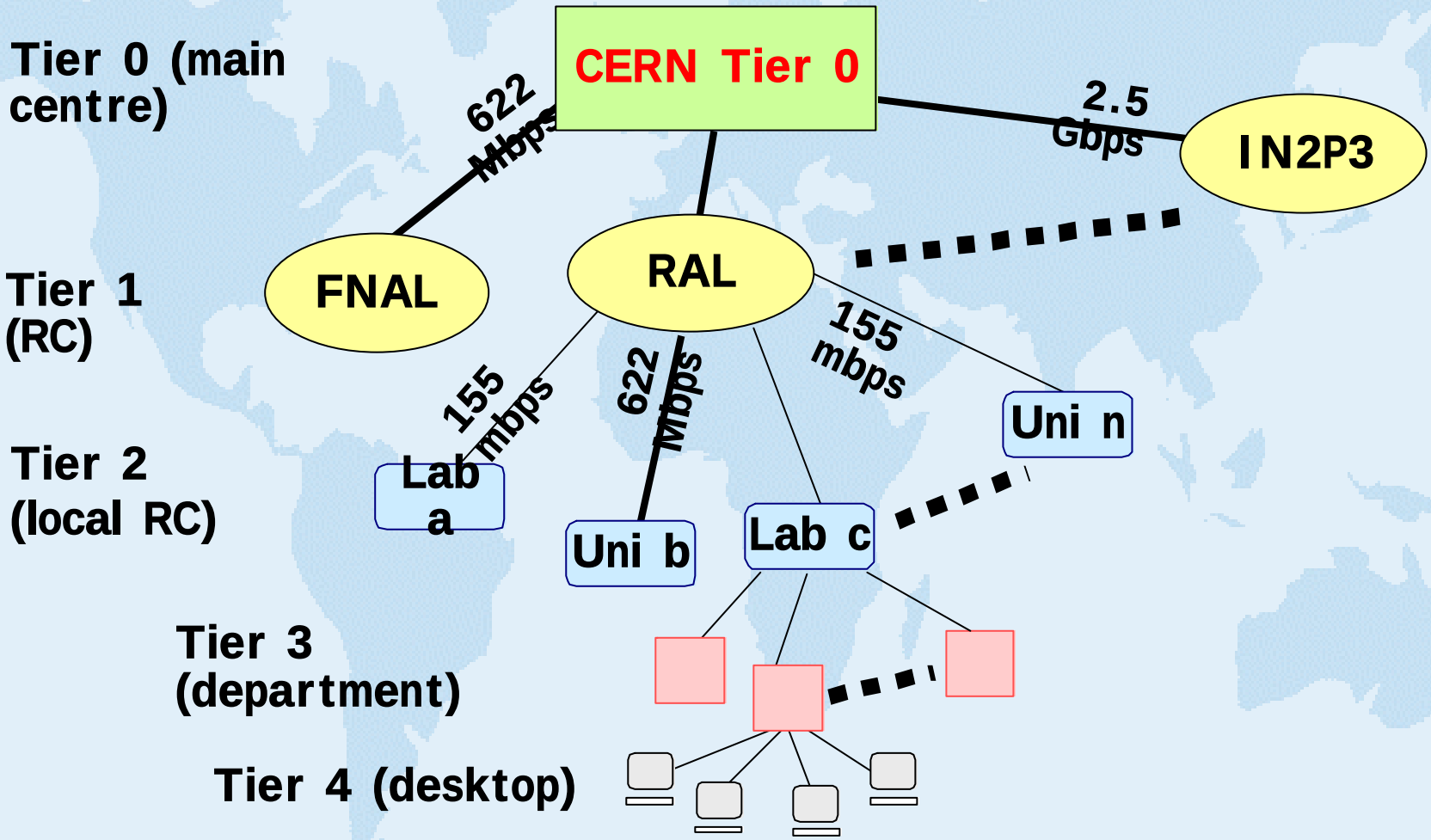
Lisbon 16-Nov-2000

Laboratório de Instrumentação e Física Experimental de Partículas

LHC Computing

- LHC experiments will produce after filtering a recording rate of 100 MegaByte/sec.
- Each LHC experiment foresees a recorded raw rate of one PetaByte/year at the start of LHC operation.
- The demands for processor power, storage and networking are at least 2 to 3 orders of magnitude more than we known how to handle today.
- The complexity of accessing and processing this data is increased substantially by:
  - the size and global span of the experiments;
  - limited wide area network bandwidth.

Regional Centre Hierarchical Model



Regional Centre Model Motivations

- Maximise the intellectual contribution of physicists all over the world.
- Bring computing facilities geographically closer to the home institutes will enable cheaper and higher bandwidth.
- A hierarchy of centres with data storage will ensure that network problems will not interfere with physics analysis.
- Allow the usage of the expertise and resources residing in computer centres throughout the world.
- Maximise contributions from each member country.

The Grid Metaphor

- A computational grid provides a metaphor for a coherent set of computing resources physically distributed across a number of geographically separated sites.
- A computational grid exhibits a uniform interface to its resources providing a dependable and pervasive service.
- Computational grid works like a electrical power grid in which resources are provided to users through a simplified interface hiding the underlying complexity of the network.
- LHC major distributed computing requirements match perfectly the grid concept making it the ideal test-bed.
- The grid is seen as a fundamental step towards the solution of the LHC computing problem.

The DATAGRID Project

- CERN proposal with CNRS, ESA, INFN, NIKEF and PPARC partially funded by the EU, will start in January 2001.
- Objectives:
  - Establish a Research Network that will enable the development of the technology components essential for the implementation of a large scale GRID.
  - Demonstrate the effectiveness of this technology through the deployment of experiment applications involving real users.
  - Demonstrate the ability to build, connect and manage distributed clusters based on commodity components capable of scale to large computing requirements.
- LIP is involved in grid test-bed activities in collaboration with the CERN DATAGRID project.

Networking for The Grid

- Two elements are indispensable to implement a grid:
  - high performance networks;
  - network quality of service.
- High performance depends on the economic capacity.
- QoS depends on network technology.
- LIP will work with FCCN in order to help the Portuguese research network meet QoS requirements.
- A new Portuguese funding for projects on high bandwidth applications will soon open, The funding will finance the required bandwidth both at national and international level.

Computing Requirements

Minimum computing resources required at LIP for 2005

PC99 CPU	STORAGE	FARM Bandwidth	WAN Bandwidth
190	46 TB	440 Mbits/s	20 Mbits/s

Test-bed computing resources required at LIP for 2001

PC99 CPU	STORAGE	FARM Bandwidth	WAN Bandwidth
36	2.3 TB	100 Mbits/s	3 - 6 Mbits/s
19%	5%	25%	15 % - 30 %

Current Grid Activities at LIP

- Hardware
  - Test of cpu and I/O solutions
  - Test of rack mount solutions
- Software
  - Evaluation of scheduling systems (pbs, lsf, condor)
  - Implementation of a LIP certification authority
  - Installation and test of the Globus grid toolkit
    - Security and X509 certificates
    - Information services and LDAP
    - Job submission
    - Resource management
    - Remote file management
    - I/O
- Co-ordination with the DATAGRID project

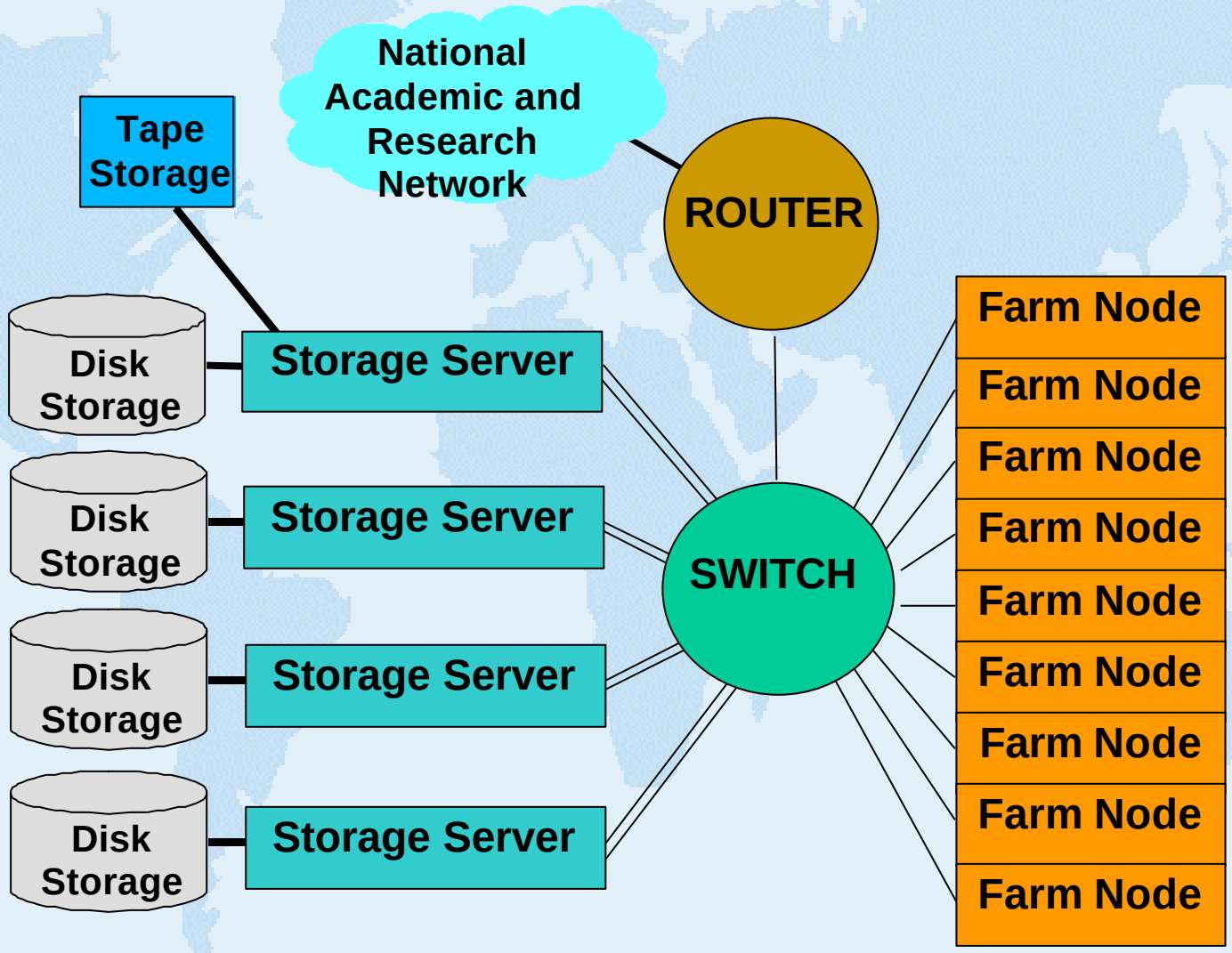
Grid Test Facility

- Cluster of Linux systems with RH 6.1
  - lflip01 - PIII 600MHz 256MB
  - lflip02 - AMD K7 600MHz 256MB
  - lflip03 - PII 400MHz 256MB
  - lflip04 - PIII 600MHz 256MB SCSI
    (cluster server/globus gatekeeper)
  - lfna5001 - PPro 200MHz 32MB
  - lfdel01 - PIII 700MHz 256MB
  - pcjorge - PII 300MHz 128MB
  - lfna5002 - PIII 800MHz 256MB
  - lfdel02 - PIII 866MHz 256MB
- Interconnection through fast ethernet using a Catalyst 5000
- The cluster is being used for grid testing and production (mostly by delphi and medica) which makes grid related testing difficult.
- System console on serial port

Hardware

- Evaluation of a 2U rack mount system
  - Motherboard ASUS CUR-DLS dual CPU
  - 2x PIII 800MHz
  - 512MB SDRAM
  - On-board Fast ethernet, Graphics adapter, Ultra160 SCSI
- Evaluation of a 1U rack system
  - Motherboard ASUS CUS-L2 with on-board graphics adapter
  - 1x PIII 866MHz
  - 256MB SDRAM
- Test of IDE Ultra DMA 100 controllers from promise
- Test of large (60GB) IDE UDMA100 disks from IBM
- Objective: define the systems and configurations to be bought next year for the grid test-bed

Future Grid Test-bed Infra-structure



Software - Scheduling Systems

- Currently running PBS (Portable Batch System)
  - Free version (commercial version available but extremely expensive)
  - Easy to use and configure for simple setups
  - Implementation of policies through recompilation
  - Many bugs and security problems
  - Poor documentation
- Condor
  - Free version but no source code available (maintenance must be paid)
  - Complex to setup
  - Some reliability problems
  - Allows complex configurations and policies
  - Allows checkpoint and migration
  - Documentation out of date
- LSF
  - Expensive (80K Esc per CPU)
  - Simple to configure
  - Allows complex configurations and policies
  - Reliable
  - Good documentation

Software - Certificates

- Globus uses X509 certificates for user and system authentication.
- Certificates must be signed by Certification Authorities (CAs).
- A LIP CA has been established using the SSLeay toolkit.
- Upon deployment Globus generates a system certificate for each system to be deployed.
- The public certificate must then be signed by the certification authority and installed in the Globus system.
- User certificates are generated through a globus command.
- The user must then send the certificate to the CA which will return it after being signed.
- The LIP CA public certificate must be distributed to all Globus systems in which LIP certificates are expected to be used.
- To cancel a certificate a CRL (revogation list) must be issued by the CA.

Software - Certificates

- Problems and open questions:
  - How to distribute CA certificates ?
  - Should we have a HEP CA hierarchy ?
  - How to handle mapping between usernames and CERTs ?
  - Which CERT to use ? Who should issue the CERTs ?
    - A single CERT per user recognized everywhere
    - A CERT per user and per institute
  - How to cancel certificates in an efficient way ?
    - CRL distribution and installation
    - CRLs can grow and become huge
  - Common HEP security policy needed ?
    - Can all CAs be trusted ?
  - HEP wide security coordination is needed ?

Software - Globus

- Installation of the Globus 1.1.3 toolkit in the LIP farm.
- Each node has a fork service and lflip04 has a pbs service.
- A distributed Globus information service (MDS) with GIIS has been established.
  - Institutional MDS service "dc=lip, dc=pt, o=Grid" (lflip04:3890)
  - Local MDS service in each system
- GSI authentication through the LIP CA.
- Problems:
  - Documentation dispersed, confusing and out of date, almost no docs on MDS.
  - Some odd behaviors of certain tools not fully understood.
  - Several bugs found specially on GASS (remote cache system).
  - Debugging Globus is not an easy task.
- Two small documents on Globus at LIP have been produced.
  - Configuring and managing certificates and the LIP CA.
  - Installing and deploying Globus at LIP.

Software – Things to do (short term)

- Explore the Grid Information System more deeply.
- Evaluation of the new 2.4 Linux kernel features.
- Evaluation of LINUX raid capabilities.
- Evaluation of LINUX log file systems.
- Evaluation of OpenAFS.
- Installation of grid enabled high performance file transfer tools.
- Monitoring and accounting software for the farm.

END