The Use of OpenSource in the Grid Computing and Data Acquisition System in High Energy and Particle Physics Research Projects

R.Randriatoamanana
(F.Hernandez)

CNRS/IN2P3

Octobre 27-28, 2009 @ GCOS, Jakarta - Indonesia

1. Who we are and where we come from
   - Speakers
   - CNRS
   - IN2P3
2. High Energy and Particle Physics (HEP)
   - Large Hardon Collider/Grid Computing
   - AUGER/CDAS
3. The most common tools used by HEP community
   - www
   - Scientific Linux
   - ROOT
   - OpenAFS

Richard Randriatoamanana

Background

- Bachelor's Degree in Computer Science and Applied Mathematics (University of Bordeaux I)
- Master's Degree in Computer Science Engineering (Institut d'Ingénierie Informatique de Limoges)

Positions held

- Computer Science Research Engineer in a CNRS research laboratory (since 2002) and Technical expert leader of the data acquisition system development for the Pierre AUGER's international Cosmic Ray Observatory.
- CNRS Information Security System Officier of LPNHE, since 2004.
- IT Unit Head at LPNHE since 2007.

Fabio Hernandez

- Working in the field of computing for high-energy physics research since 1992
  - Software development for scientific data management (data transfer over high throughput networks, mass storage and retrieval, cataloguing, etc.) and operations of IT services for scientific research
- Involved in **grid computing projects since 2000** and in particular in the deployment of computing infrastructure for the LHC since 2004
  - Technical leader of the French contribution to this infrastructure (1 tier-1, 4 tier-2s, 3 tier-3s)
  - Member of the Management Board and grid Deployment Board of the Worldwide LHC Computing Grid collaboration
- **Deputy director IN2P3/CNRS computing centre**, which hosts and operates the French WLCG tier-1

History and Missions

The **C**entre **N**ational de la **R**echerche **S**cientifique is a government-funded research organization, under the administrative authority of France's Ministry of Research. It was founded in 1939 by governmental decree, CNRS has the following missions:

- To evaluate and carry out all research capable of advancing knowledge and bringing social, cultural, and economic benefits for society.
- To contribute to the application and promotion of research results.
- To develop scientific information.
- To support research training.
- To participate in the analysis of the national and international scientific climate and its potential for evolution in order to develop a national policy.

History and Missions

The **C**entre **N**ational de la **R**echerche **S**cientifique is a government-funded research organization, under the administrative authority of France's Ministry of Research. It was founded in 1939 by governmental decree, CNRS has the following missions:

- To evaluate and carry out all research capable of advancing knowledge and bringing social, cultural, and economic benefits for society.
- To contribute to the application and promotion of research results.
- To develop scientific information.
- To support research training.
- To participate in the analysis of the national and international scientific climate and its potential for evolution in order to develop a national policy.

History and Missions

The **C**entre **N**ational de la **R**echerche **S**cientifique is a government-funded research organization, under the administrative authority of France's Ministry of Research. It was founded in 1939 by governmental decree, CNRS has the following missions:

- To evaluate and carry out all research capable of advancing knowledge and bringing social, cultural, and economic benefits for society.
- To contribute to the application and promotion of research results.
- To develop scientific information.
- To support research training.
- To participate in the analysis of the national and international scientific climate and its potential for evolution in order to develop a national policy.

History and Missions

The **C**entre **N**ational de la **R**echerche **S**cientifique is a government-funded research organization, under the administrative authority of France's Ministry of Research. It was founded in 1939 by governmental decree, CNRS has the following missions:

- To evaluate and carry out all research capable of advancing knowledge and bringing social, cultural, and economic benefits for society.
- To contribute to the application and promotion of research results.
- To develop scientific information.
- To support research training.
- To participate in the analysis of the national and international scientific climate and its potential for evolution in order to develop a national policy.

History and Missions

The **C**entre **N**ational de la **R**echerche **S**cientifique is a government-funded research organization, under the administrative authority of France's Ministry of Research. It was founded in 1939 by governmental decree, CNRS has the following missions:

- To evaluate and carry out all research capable of advancing knowledge and bringing social, cultural, and economic benefits for society.
- To contribute to the application and promotion of research results.
- To develop scientific information.
- To support research training.
- To participate in the analysis of the national and international scientific climate and its potential for evolution in order to develop a national policy.

History and Missions

The **C**entre **N**ational de la **R**echerche **S**cientifique is a government-funded research organization, under the administrative authority of France's Ministry of Research. It was founded in 1939 by governmental decree, CNRS has the following missions:

- To evaluate and carry out all research capable of advancing knowledge and bringing social, cultural, and economic benefits for society.
- To contribute to the application and promotion of research results.
- To develop scientific information.
- To support research training.
- To participate in the analysis of the national and international scientific climate and its potential for evolution in order to develop a national policy.

Research fields

As the largest fundamental research organization in Europe, CNRS carried out research in all fields of knowledge, through its 9 institutes (2 of which have the status of national institutes):

- Institute of Chemistry (INC)
- Institute of Ecology and Environment (INEE)
- Institute of Physics (INP)
- Institute of Biological Sciences (INSB)
- Institute for Humanities and Social Sciences (INSHS)
- Institute for Mathematical Sciences (INSMI)
- Institute of Information and Engineering Sciences and Technologies (INST2I)
- National Institute of Nuclear and Particle Physics (IN2P3)
- National Institute for Earth Sciences and Astronomy (INSU)

Key figures (february 2009)

- **33,600** employees of which 26,000 are CNRS tenured employees (11,600 researchers and 14,400 engineers and support staff)
- **1,100** research units (90% are joint research laboratories with universities and industry)
- **5,000** foreign visiting scientists (PhD students, post-docs and visiting researchers)
- **18** International Joint Units (UMI)

Budget for 2009

3.367 billion Euros of which 607 million come from revenues generated by CNRS contracts.

National Institute of Nuclear and Particle Physics

- **A national CNRS institut**, created in 1971, it's mission is to unite and promote research activities in the fields of nuclear physics, particle physics and astroparticle, with common programs on behalf of CNRS and universities, in partnership with CEA;
- Very large collaborations (or projects) for research conducted by IN2P3 and organized around increasingly sophisticated and expensive instruments (accelerators and detectors) shared by a worldwide community in laboratories. These collaborations occur particularly with accelerators located at **CERN (Geneva), GANIL (France), at SLAC (Stanford, USA), FNAL (USA) and DESY (Germany)**.
- 20 labs (LPNHE, etc.) and 1 computing centre.

Key figures

42 billion euros annual budget and ~2453 employees (10% in ICT).

National Institute of Nuclear and Particle Physics

- **A national CNRS institut**, created in 1971, it's mission is to unite and promote research activities in the fields of nuclear physics, particle physics and astroparticle, with common programs on behalf of CNRS and universities, in partnership with CEA;
- Very large collaborations (or projects) for research conducted by IN2P3 and organized around increasingly sophisticated and expensive instruments (accelerators and detectors) shared by a worldwide community in laboratories. These collaborations occur particularly with accelerators located at **CERN (Geneva), GANIL (France), at SLAC (Stanford, USA), FNAL (USA) and DESY (Germany)**.
- 20 labs (LPNHE, etc.) and 1 computing centre.

Key figures

42 billion euros annual budget and ∼2453 employees (10% in ICT).

National Institute of Nuclear and Particle Physics

- **A national CNRS institut**, created in 1971, it's mission is to unite and promote research activities in the fields of nuclear physics, particle physics and astroparticle, with common programs on behalf of CNRS and universities, in partnership with CEA;
- Very large collaborations (or projects) for research conducted by IN2P3 and organized around increasingly sophisticated and expensive instruments (accelerators and detectors) shared by a worldwide community in laboratories. These collaborations occur particularly with accelerators located at **CERN (Geneva), GANIL (France), at SLAC (Stanford, USA), FNAL (USA) and DESY (Germany)**.
- 20 labs (LPNHE, etc.) and 1 computing centre.

Key figures

42 billion euros annual budget and ~2453 employees (10% in ICT).

National Institute of Nuclear and Particle Physics

- **A national CNRS institut**, created in 1971, it's mission is to unite and promote research activities in the fields of nuclear physics, particle physics and astroparticle, with common programs on behalf of CNRS and universities, in partnership with CEA;
- Very large collaborations (or projects) for research conducted by IN2P3 and organized around increasingly sophisticated and expensive instruments (accelerators and detectors) shared by a worldwide community in laboratories. These collaborations occur particularly with accelerators located at **CERN (Geneva), GANIL (France), at SLAC (Stanford, USA), FNAL (USA) and DESY (Germany)**.
- 20 labs (LPNHE, etc.) and 1 computing centre.

Key figures

42 billion euros annual budget and ∼2453 employees (10% in ICT).

Computing Centre

- High-throughput data processing facility
  - not co-located with an experimental site
- Missions & competencies
  - mass storage & computing
  - infrastructure services for scientific distributed international collaborations : web hosting, webcast, mail, etc.
  - round-the-clock service
- Users
  - ∼70 collaborations: nuclear physics, particle physics and astro-particle physics
  - More recently, bio-medical applications, data repositories for social sciences

Key figures

12 million euros annual budget and ∼70 FTE in human resources.

Computing Centre

- High-throughput data processing facility
  - not co-located with an experimental site
- Missions & competencies
  - mass storage & computing
  - infrastructure services for scientific distributed international collaborations : web hosting, webcast, mail, etc.
  - round-the-clock service
- Users
  - ~70 collaborations: nuclear physics, particle physics and astro-particle physics
  - More recently, bio-medical applications, data repositories for social sciences

Key figures

12 million euros annual budget and ~70 FTE in human resources.

Computing Centre

- High-throughput data processing facility
    - not co-located with an experimental site
- Missions & competencies
    - mass storage & computing
    - infrastructure services for scientific distributed international collaborations : web hosting, webcast, mail, etc.
    - round-the-clock service
- Users
    - ~70 collaborations: nuclear physics, particle physics and astro-particle physics
    - More recently, bio-medical applications, data repositories for social sciences

Key figures

12 million euros annual budget and ~70 FTE in human resources.

Computing Centre

- High-throughput data processing facility
  - not co-located with an experimental site
- Missions & competencies
  - mass storage & computing
  - infrastructure services for scientific distributed international collaborations : web hosting, webcast, mail, etc.
  - round-the-clock service
- Users
  - ∼70 collaborations: nuclear physics, particle physics and astro-particle physics
  - More recently, bio-medical applications, data repositories for social sciences

Key figures

12 million euros annual budget and ∼70 FTE in human resources.

IN2P3 Computing DataCenter at Lyon, France

Laboratory of Physics Nuclear and High Energy, LPNHE

- A *unité mixte de recherche* of IN2P3, CNRS and Universities Paris 6/7 located in the UPMC's campus in Jussieu/Paris.
- Employs 12 research groups, 3 technical and 2 support services.
- Is engaged in several large experimental research programs pursued in the context of international collaborations with very large research facilities around the world, centers of particle accelerators and observatories (AUGER, ATLAS, HESS, etc.)
- Hosts an ICT unit that ensures the administration of information systems and is involved in developments by coordinating experiments in acquisition systems, software and databases management.

Key figures

A local node of computing grid, a part of global grids that are LCG and EGEE: grid-based infrastructure, composed of hundreds of computing sites, linked by high-speed networks around the world.

Laboratory of Physics Nuclear and High Energy, LPNHE

- A *unité mixte de recherche* of IN2P3, CNRS and Universities Paris 6/7 located in the UPMC's campus in Jussieu/Paris.
- Employs 12 research groups, 3 technical and 2 support services.
- Is engaged in several large experimental research programs pursued in the context of international collaborations with very large research facilities around the world, centers of particle accelerators and observatories (AUGER, ATLAS, HESS, etc.)
- Hosts an ICT unit that ensures the administration of information systems and is involved in developments by coordinating experiments in acquisition systems, software and databases management.

Key figures

A local node of computing grid, a part of global grids that are LCG and EGEE: grid-based infrastructure, composed of hundreds of computing sites, linked by high-speed networks around the world.

Laboratory of Physics Nuclear and High Energy, LPNHE

- A *unité mixte de recherche* of IN2P3, CNRS and Universities Paris 6/7 located in the UPMC's campus in Jussieu/Paris.
- Employs 12 research groups, 3 technical and 2 support services.
- Is engaged in several large experimental research programs pursued in the context of international collaborations with very large research facilities around the world, centers of particle accelerators and observatories (AUGER, ATLAS, HESS, etc.)
- Hosts an ICT unit that ensures the administration of information systems and is involved in developments by coordinating experiments in acquisition systems, software and databases management.

Key figures

A local node of computing grid, a part of global grids that are LCG and EGEE: grid-based infrastructure, composed of hundreds of computing sites, linked by high-speed networks around the world.

Laboratory of Physics Nuclear and High Energy, LPNHE

- A *unité mixte de recherche* of IN2P3, CNRS and Universities Paris 6/7 located in the UPMC's campus in Jussieu/Paris.
- Employs 12 research groups, 3 technical and 2 support services.
- Is engaged in several large experimental research programs pursued in the context of international collaborations with very large research facilities around the world, centers of particle accelerators and observatories (AUGER, ATLAS, HESS, etc.)
- Hosts an ICT unit that ensures the administration of information systems and is involved in developments by coordinating experiments in acquisition systems, software and databases management.

Key figures

A local node of computing grid, a part of global grids that are LCG and EGEE: grid-based infrastructure, composed of hundreds of computing sites, linked by high-speed networks around the world.

Laboratory of Physics Nuclear and High Energy, LPNHE

- A *unité mixte de recherche* of IN2P3, CNRS and Universities Paris 6/7 located in the UPMC's campus in Jussieu/Paris.
- Employs 12 research groups, 3 technical and 2 support services.
- Is engaged in several large experimental research programs pursued in the context of international collaborations with very large research facilities around the world, centers of particle accelerators and observatories (AUGER, ATLAS, HESS, etc.)
- Hosts an ICT unit that ensures the administration of information systems and is involved in developments by coordinating experiments in acquisition systems, software and databases management.

Key figures

A local node of computing grid, a part of global grids that are LCG and EGEE: grid-based infrastructure, composed of hundreds of computing sites, linked by high-speed networks around the world.

Sommaire

CERN

- the European Organization for Nuclear Research, is one of the world's largest and most respected centres for scientific research.
- founded in 1954, the CERN Laboratory sits astride the Franco-Swiss border near Geneva.
- its business is fundamental physics, finding out what the Universe is made of and how it works.
- the world's largest and most complex scientific instruments are used to study the basic constituents of matter - the fundamental particles. By studying what happens when these particles collide, physicists learn about the laws of Nature.
- the instruments used are particle accelerators and detectors. Accelerators boost beams of particles to high energies before they are made to collide with each other or with stationary targets. Detectors observe and record the results of these collisions.

CERN

- the European Organization for Nuclear Research, is one of the world's largest and most respected centres for scientific research.
- founded in 1954, the CERN Laboratory sits astride the Franco-Swiss border near Geneva.
- its business is fundamental physics, finding out what the Universe is made of and how it works.
- the world's largest and most complex scientific instruments are used to study the basic constituents of matter - the fundamental particles. By studying what happens when these particles collide, physicists learn about the laws of Nature.
- the instruments used are particle accelerators and detectors. Accelerators boost beams of particles to high energies before they are made to collide with each other or with stationary targets. Detectors observe and record the results of these collisions.

CERN

- the European Organization for Nuclear Research, is one of the world's largest and most respected centres for scientific research.
- founded in 1954, the CERN Laboratory sits astride the Franco-Swiss border near Geneva.
- its business is fundamental physics, finding out what the Universe is made of and how it works.
- the world's largest and most complex scientific instruments are used to study the basic constituents of matter - the fundamental particles. By studying what happens when these particles collide, physicists learn about the laws of Nature.
- the instruments used are particle accelerators and detectors. Accelerators boost beams of particles to high energies before they are made to collide with each other or with stationary targets. Detectors observe and record the results of these collisions.

CERN

- the European Organization for Nuclear Research, is one of the world's largest and most respected centres for scientific research.
- founded in 1954, the CERN Laboratory sits astride the Franco-Swiss border near Geneva.
- its business is fundamental physics, finding out what the Universe is made of and how it works.
- the world's largest and most complex scientific instruments are used to study the basic constituents of matter - the fundamental particles. By studying what happens when these particles collide, physicists learn about the laws of Nature.
- the instruments used are particle accelerators and detectors. Accelerators boost beams of particles to high energies before they are made to collide with each other or with stationary targets. Detectors observe and record the results of these collisions.

CERN

- the European Organization for Nuclear Research, is one of the world's largest and most respected centres for scientific research.
- founded in 1954, the CERN Laboratory sits astride the Franco-Swiss border near Geneva.
- its business is fundamental physics, finding out what the Universe is made of and how it works.
- the world's largest and most complex scientific instruments are used to study the basic constituents of matter - the fundamental particles. By studying what happens when these particles collide, physicists learn about the laws of Nature.
- the instruments used are particle accelerators and detectors. Accelerators boost beams of particles to high energies before they are made to collide with each other or with stationary targets. Detectors observe and record the results of these collisions.

The experiment

- the Large Hadron Collider (LHC) is a gigantic scientific instrument that spans the border between Switzerland and France about 100 m underground.
- the world's largest and most powerful particle accelerator consists of a 27 km ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles.
- six experiments are all run by international collaborations, bringing together scientists from institutes all over the world. Each experiment is distinct, characterised by its unique particle detector. **ATLAS, CMS, ALICE, LHCb, TOTEM and LHCf**

Key figures

Chilling magnets (-271°C) colder than outer space! Inside the accelerator, two beams of particles travel at close to the speed of light with very high energies before colliding with one another.

The experiment

- the Large Hadron Collider (LHC) is a gigantic scientific instrument that spans the border between Switzerland and France about 100 m underground.
- the world's largest and most powerful particle accelerator consists of a 27 km ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles.
- six experiments are all run by international collaborations, bringing together scientists from institutes all over the world. Each experiment is distinct, characterised by its unique particle detector. **ATLAS, CMS, ALICE, LHCb, TOTEM and LHCf**

Key figures

Chilling magnets (-271°C) colder than outer space! Inside the accelerator, two beams of particles travel at close to the speed of light with very high energies before colliding with one another.

The experiment

- the Large Hadron Collider (LHC) is a gigantic scientific instrument that spans the border between Switzerland and France about 100 m underground.
- the world's largest and most powerful particle accelerator consists of a 27 km ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles.
- six experiments are all run by international collaborations, bringing together scientists from institutes all over the world. Each experiment is distinct, characterised by its unique particle detector. **ATLAS, CMS, ALICE, LHCb, TOTEM and LHCf**

Key figures

Chilling magnets (-271°C) colder than outer space! Inside the accelerator, two beams of particles travel at close to the speed of light with very high energies before colliding with one another.

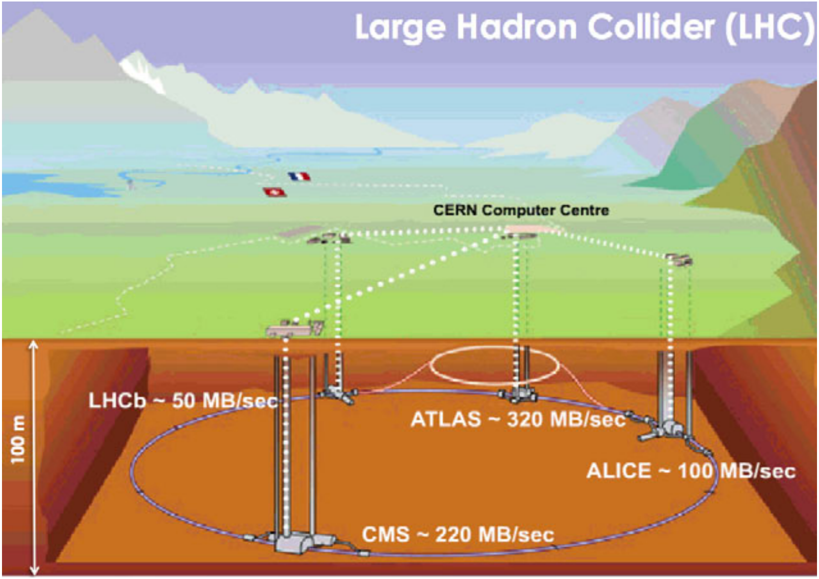
The experiment

- the Large Hadron Collider (LHC) is a gigantic scientific instrument that spans the border between Switzerland and France about 100 m underground.
- the world's largest and most powerful particle accelerator consists of a 27 km ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles.
- six experiments are all run by international collaborations, bringing together scientists from institutes all over the world. Each experiment is distinct, characterised by its unique particle detector. **ATLAS, CMS, ALICE, LHCb, TOTEM and LHCf**

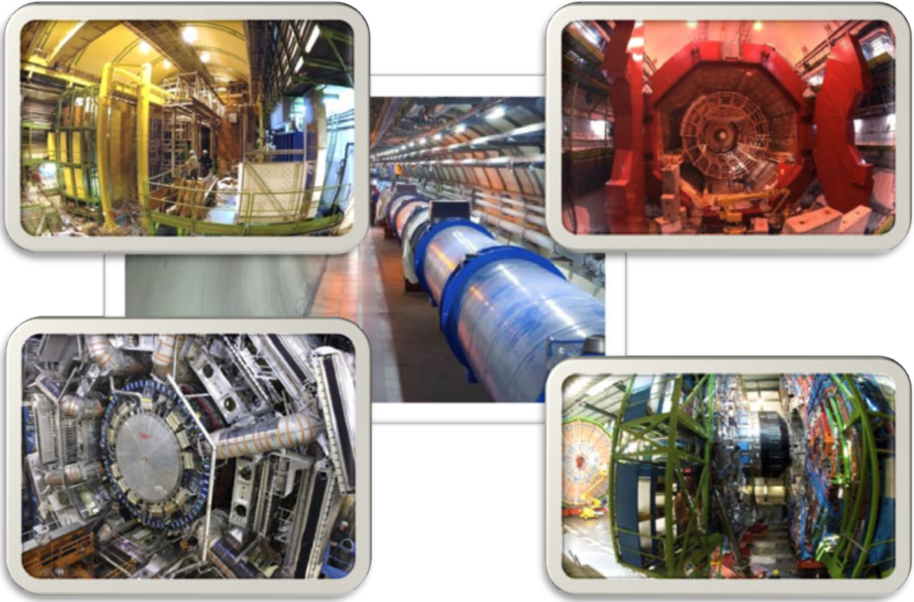
Key figures

Chilling magnets (-271°C) colder than outer space! Inside the accelerator, two beams of particles travel at close to the speed of light with very high energies before colliding with one another.

LHC topology



LHC in work...



LHC Computing Challenge

Data Volume

High rate * large number of channels * 4 experiments
→ 15 PetaBytes of new data each year

Compute power

Event complexity * Nb. events * thousands users
→ 100 k of (today's) fastest CPUs and 140 PB of storage.

Worldwide analysis & funding

Computing funding locally in major regions & countries (33)
Efficient analysis everywhere
→ GRID technology

Behind a multi grid-infrastructure

European multi-science grid Enabling Grids for E-SciencE (EGEE).

LHC Computing Challenge

Data Volume

High rate * large number of channels * 4 experiments
→ 15 PetaBytes of new data each year

Compute power

Event complexity * Nb. events * thousands users
→ 100 k of (today's) fastest CPUs and 140 PB of storage.

Worldwide analysis & funding

Computing funding locally in major regions & countries (33)
Efficient analysis everywhere
→ GRID technology

Behind a multi grid-infrastructure

European multi-science grid Enabling Grids for E-SciencE (EGEE).

LHC Computing Challenge

Data Volume

High rate * large number of channels * 4 experiments
→ 15 PetaBytes of new data each year

Compute power

Event complexity * Nb. events * thousands users
→ 100 k of (today's) fastest CPUs and 140 PB of storage.

Worldwide analysis & funding

Computing funding locally in major regions & countries (33)
Efficient analysis everywhere
→ GRID technology

Behind a multi grid-infrastructure

European multi-science grid Enabling Grids for E-SciencE (EGEE).

LHC Computing Challenge

Data Volume

High rate * large number of channels * 4 experiments
→ 15 PetaBytes of new data each year

Compute power

Event complexity * Nb. events * thousands users
→ 100 k of (today's) fastest CPUs and 140 PB of storage.

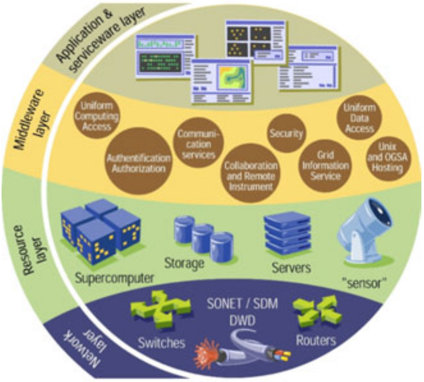
Worldwide analysis & funding

Computing funding locally in major regions & countries (33)
Efficient analysis everywhere
→ GRID technology

Behind a multi grid-infrastructure

European multi-science grid Enabling Grids for E-SciencE (EGEE).

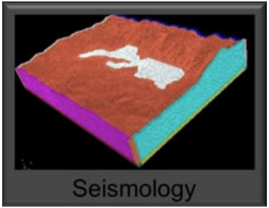
How does the Grid works ?

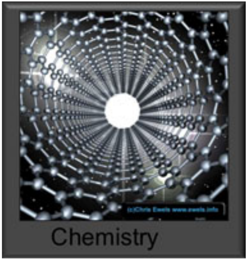


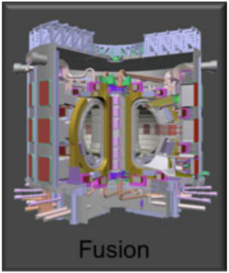
Use the Grid to unite computing resources of particle physics institutes around the world

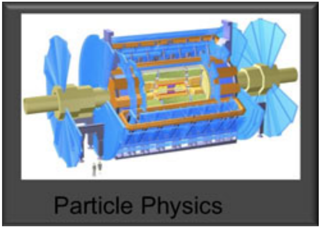
- It relies on special software, called middleware.
- Middleware automatically finds the data the scientist needs, and the computing power to analyse it.
- Middleware balances the load on different resources. It also handles security, accounting, monitoring and much more.

Grid Applications



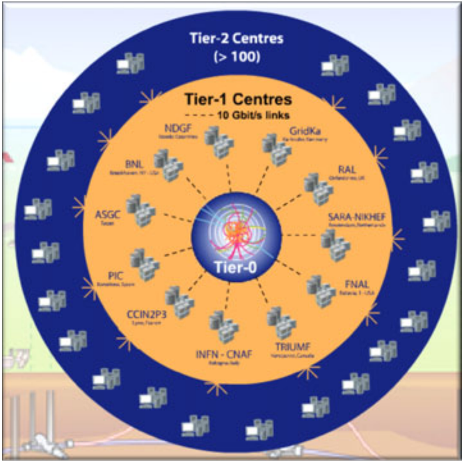






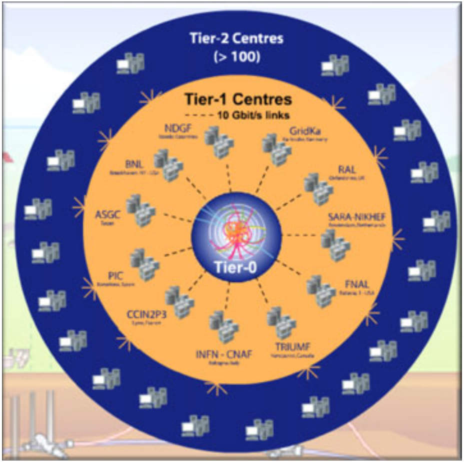
BUILDING THE INFORMATION SOCIETY
ITU
International Telecommunication Union

Tier-0-1-2



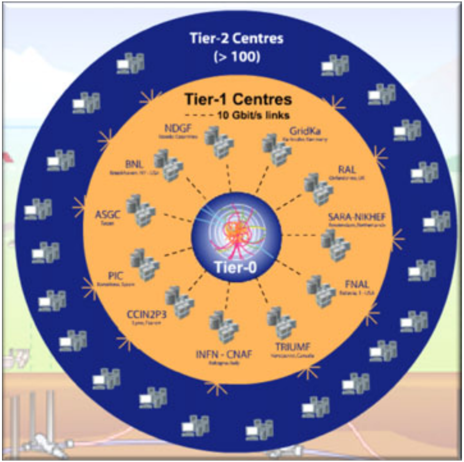
- Tier-0 (CERN)
  Data recording
  First-pass reconstruction Data distribution
- Tier-1 (11 centres)
  Permanent storage
  Re-processing Analysis
- Tier-2 (>200 centres)
  Simulation
  End-user analysis

Tier-0-1-2



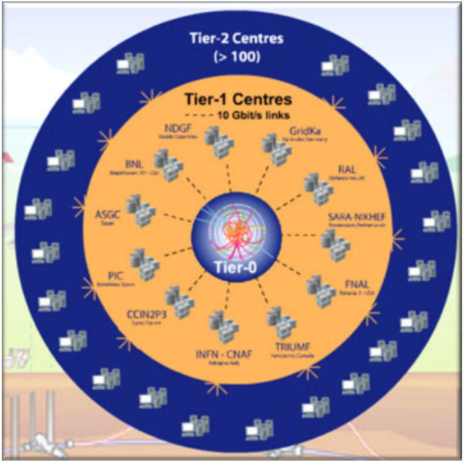
- Tier-0 (CERN)
  Data recording
  First-pass reconstruction Data distribution
- Tier-1 (11 centres)
  Permanent storage
  Re-processing Analysis
- Tier-2 (>200 centres)
  Simulation
  End-user analysis

Tier-0-1-2



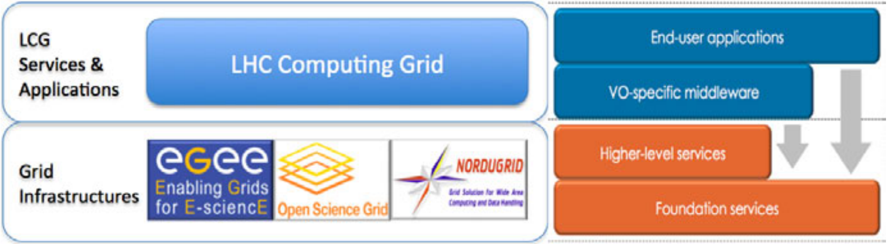
- Tier-0 (CERN)
  Data recording
  First-pass reconstruction Data distribution
- Tier-1 (11 centres)
  Permanent storage
  Re-processing Analysis
- Tier-2 (>200 centres)
  Simulation
  End-user analysis

Tier-0-1-2



- Tier-0 (CERN)
  Data recording
  First-pass reconstruction Data distribution
- Tier-1 (11 centres)
  Permanent storage
  Re-processing Analysis
- Tier-2 (>200 centres)
  Simulation
  End-user analysis

LCG infrastructure topology



Foundation services

- Authentication
- Authorization
- Virtual organization membership service
- Computing element
- Remote job submission to the site's batch system
- Storage element
- Information system
- Accounting of resources

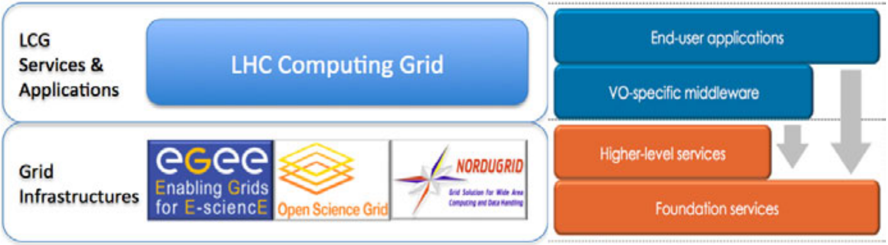
Higher-level services

- Workload management
- Data management
- VO software installation

VO-specific middlewares

- Metadata catalogues
- Data placement services
- Monitoring & alerting services
- Database replication service

LCG infrastructure topology



Foundation services

- Authentication
- Authorization
- Virtual organization membership service
- Computing element
- Remote job submission to the site's batch system
- Storage element
- Information system
- Accounting of resources

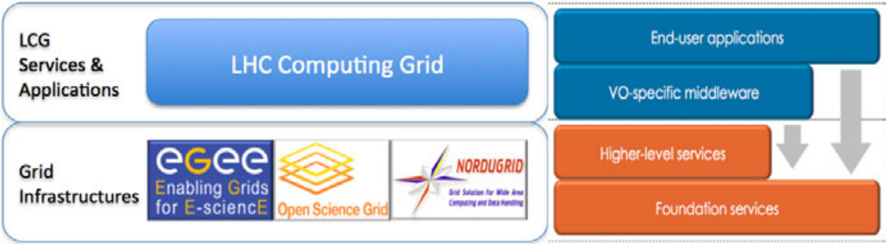
Higher-level services

- Workload management
- Data management
- VO software installation

VO-specific middlewares

- Metadata catalogues
- Data placement services
- Monitoring & alerting services
- Database replication service

LCG infrastructure topology



Foundation services

- Authentication
- Authorization
- Virtual organization membership service
- Computing element
- Remote job submission to the site's batch system
- Storage element
- Information system
- Accounting of resources

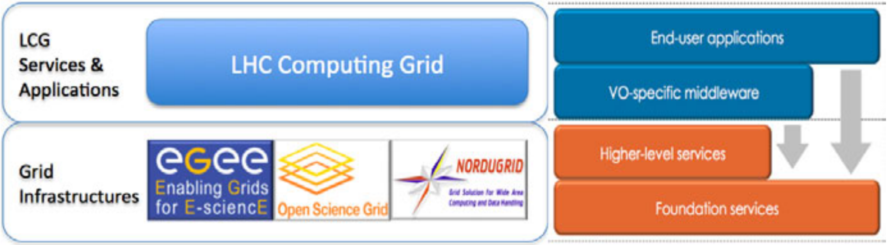
Higher-level services

- Workload management
- Data management
- VO software installation

VO-specific middlewares

- Metadata catalogues
- Data placement services
- Monitoring & alerting services
- Database replication service

LCG infrastructure topology



Foundation services

- Authentication
- Authorization
- Virtual organization membership service
- Computing element
- Remote job submission to the site's batch system
- Storage element
- Information system
- Accounting of resources

Higher-level services

- Workload management
- Data management
- VO software installation

VO-specific middlewares

- Metadata catalogues
- Data placement services
- Monitoring & alerting services
- Database replication service

Middleware explained

Security

- Virtual Organization Management (VOMS)
- MyProxy

Data Management

- File catalogue (LFC)
- File transfer service (FTS)
- Storage Element (SE)
- Storage Resource Management (SRM)

Extracted slide Jurgen Knobloch

Middleware explained

Security

- Virtual Organization Management (VOMS)
- MyProxy

Data Management

- File catalogue (LFC)
- File transfer service (FTS)
- Storage Element (SE)
- Storage Resource Management (SRM)

Extracted slide Jurgen Knobloch

Middleware explained (cont'd)

Job Management

- Work Load Management System(WMS)
- Logging and Bookeeping (LB)
- Computing Element (CE)
- Worker Nodes (WN)

Information System

- Monitoring: BDII (Berkeley Database Information Index), RGMA (Relational Grid Monitoring Architecture) → aggregate service information from multiple Grid sites, now moved to SAM (Site Availability Monitoring)
- Monitoring & visualization (Gridview, Dashboard, Gridmap etc.)

Middleware explained (cont'd)

Job Management

- Work Load Management System(WMS)
- Logging and Bookeeping (LB)
- Computing Element (CE)
- Worker Nodes (WN)

Information System

- Monitoring: BDII (Berkeley Database Information Index), RGMA (Relational Grid Monitoring Architecture) → aggregate service information from multiple Grid sites, now moved to SAM (Site Availability Monitoring)
- Monitoring & visualization (Gridview, Dashboard, Gridmap etc.)

gLite, middleware for grid computing

Born from the collaborative efforts of more than 80 people in 12 different academic and industrial research centers as part of the EGEE Project, gLite provides a framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet. The gLite services are currently adopted by more than **250 Computing Centres** and used by more than 15000 researchers in Europe and around the world (Taiwan, Latin America etc.)

Services

- Access
- Security
- Data
- Job Management
- Information& Monitoring

gLite, middleware for grid computing

Born from the collaborative efforts of more than 80 people in 12 different academic and industrial research centers as part of the EGEE Project, gLite provides a framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet. The gLite services are currently adopted by more than **250 Computing Centres** and used by more than 15000 researchers in Europe and around the world (Taiwan, Latin America etc.)

Services

- Access
- Security
- Data
- Job Management
- Information& Monitoring

Extremely Large Fabric management system

Quattor

A system administration toolkit providing a powerful, portable, and modular set of tools for the automated installation, configuration, and management of clusters, grids and farms.

Lemon

A server/client based monitoring system. On every monitored node, a monitoring agent launches and communicates using a push/pull protocol with sensors which are responsible for retrieving monitoring information.

Extremely Large Fabric management system (cont'd)

Service Level Status

A web-based tool that dynamically shows availability, basic information and/or statistics about IT services, as well as dependencies between them.

State and Hardware management (LEAF)

LHC Era Automated Fabric is a collection of workflows for automated node hardware and state management: HMS=Hardware Management System, SMS=State Management System.

The experiment

Summary

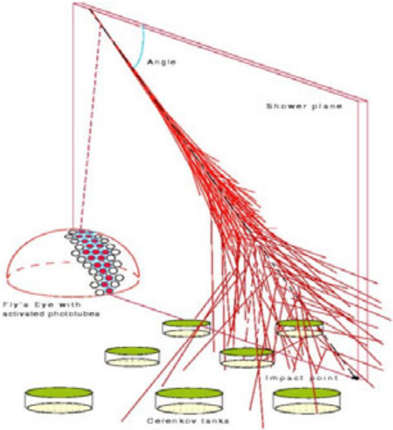
The Pierre Auger Cosmic Ray Observatory is studying ultra-high energy cosmic rays, the most energetic and rarest of particles in the universe. When these particles strike the earth's atmosphere, they produce extensive air showers made of billions of secondary particles. While much progress has been made in nearly a century of research in understanding cosmic rays with low to moderate energies, those with extremely high energies remain mysterious.

Cosmic ray

A high-energy particle that strikes the Earth's atmosphere from space, producing many secondary particles, also called cosmic rays.

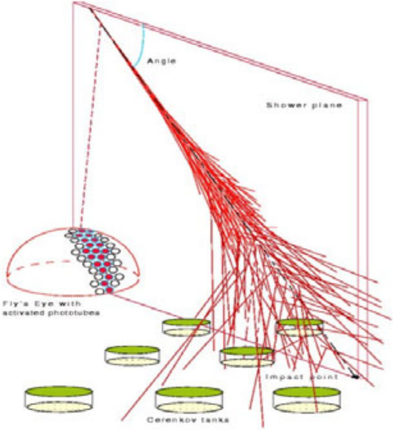
A hybrid detector

- two independent methods to detect and study high-energy cosmic rays
- one detects high energy particles through their interaction with water placed in surface detector tanks (1600)
- the other tracks the development of air showers by observing ultraviolet light emitted high in the Earth's atmosphere with 4 optical fluorescence detectors.



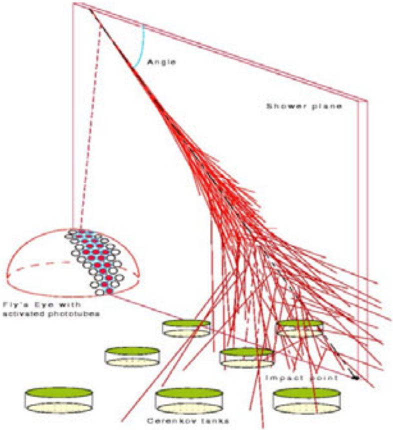
A hybrid detector

- two independent methods to detect and study high-energy cosmic rays
- one detects high energy particles through their interaction with water placed in surface detector tanks (1600)
- the other tracks the development of air showers by observing ultraviolet light emitted high in the Earth's atmosphere with 4 optical fluorescence detectors.



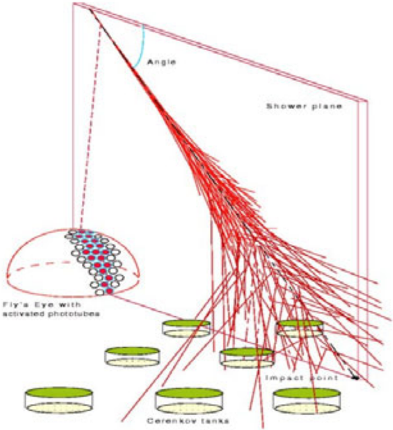
A hybrid detector

- two independent methods to detect and study high-energy cosmic rays
- one detects high energy particles through their interaction with water placed in surface detector tanks (1600)
- the other tracks the development of air showers by observing ultraviolet light emitted high in the Earth's atmosphere with 4 optical fluorescence detectors.

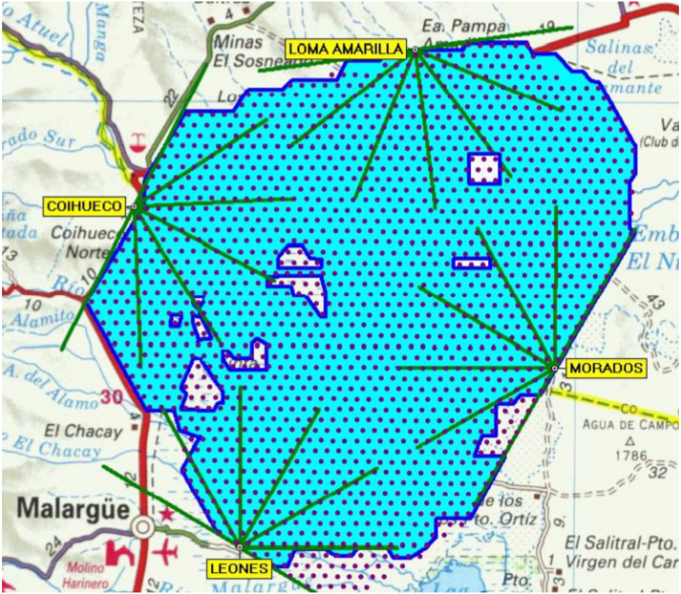


A hybrid detector

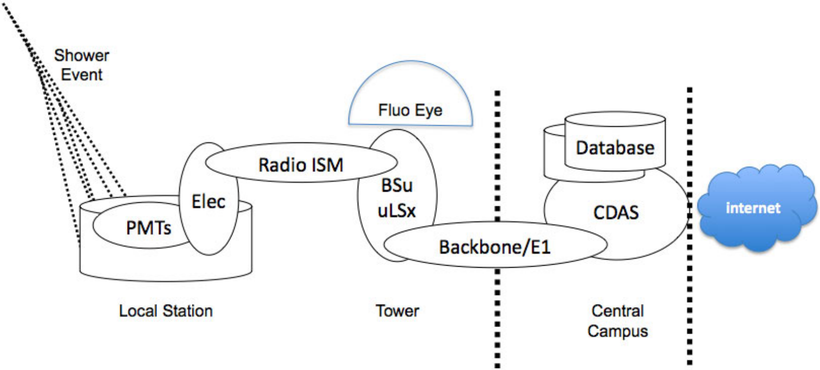
- two independent methods to detect and study high-energy cosmic rays
- one detects high energy particles through their interaction with water placed in surface detector tanks (1600)
- the other tracks the development of air showers by observing ultraviolet light emitted high in the Earth's atmosphere with 4 optical fluorescence detectors.



Pierre Auger Surface & Flouresence Array

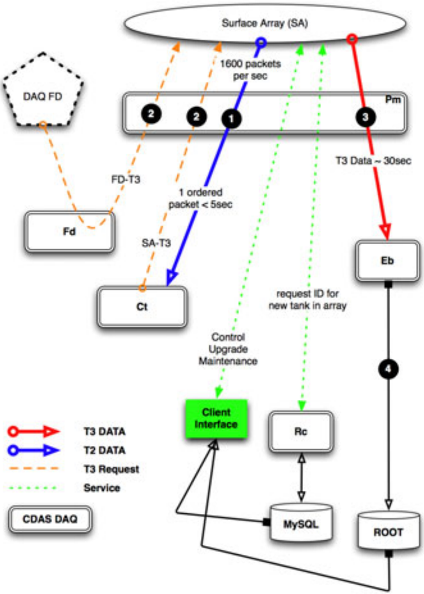


Particle life cycle



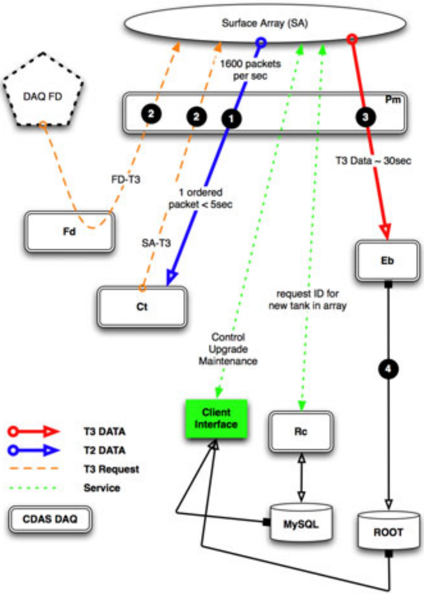
Central Data Acquisition System (CDAS)

- an hybrid client/server architecture
- a postmaster process **to dispatch** raw/event data (**2Mb/sec** via TCP) to/from clients/array.
- a central trigger process **to decide** which tanks should be in a run array to send event raw data.
- a event builder process **to build** from event raw data an auger event data stored in a database.
- a event displayer **to read** auger event data for analyze.



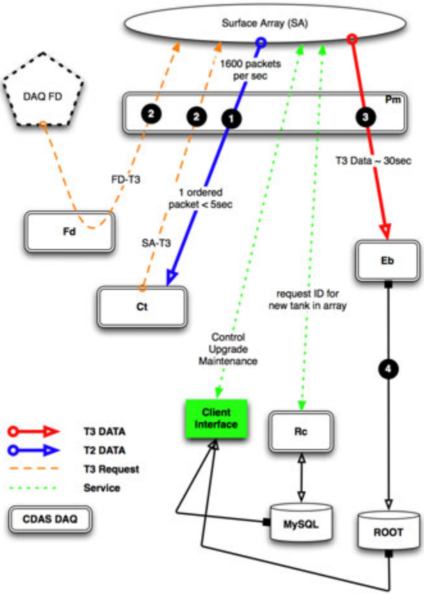
Central Data Acquisition System (CDAS)

- an hybrid client/server architecture
- a postmaster process **to dispatch** raw/event data (**2Mb/sec** via TCP) to/from clients/array.
- a central trigger process **to decide** which tanks should be in a run array to send event raw data.
- a event builder process **to build** from event raw data an auger event data stored in a database.
- a event displayer **to read** auger event data for analyze.



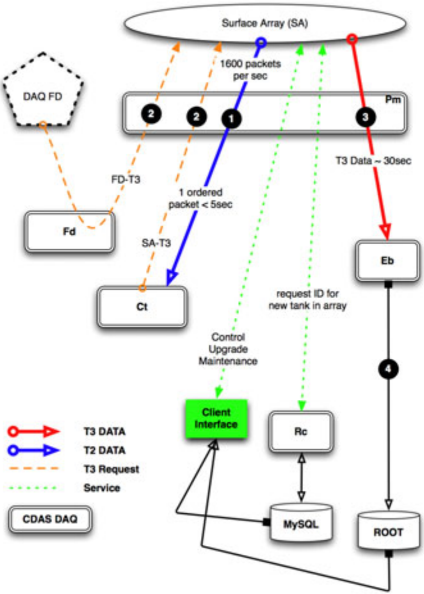
Central Data Acquisition System (CDAS)

- an hybrid client/server architecture
- a postmaster process **to dispatch** raw/event data (**2Mb/sec** via TCP) to/from clients/array.
- a central trigger process **to decide** which tanks should be in a run array to send event raw data.
- a event builder process **to build** from event raw data an auger event data stored in a database.
- a event displayer **to read** auger event data for analyze.



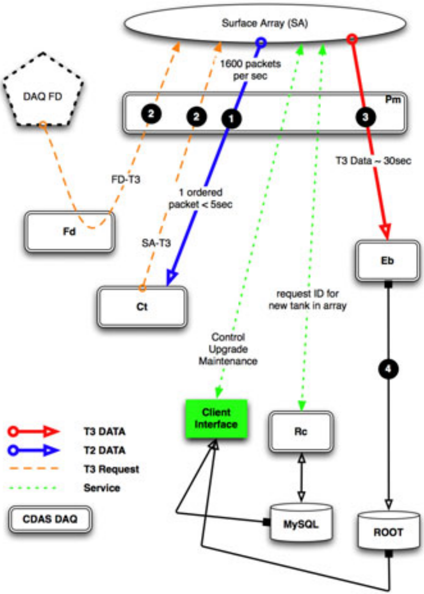
Central Data Acquisition System (CDAS)

- an hybrid client/server architecture
- a postmaster process **to dispatch** raw/event data (**2Mb/sec** via TCP) to/from clients/array.
- a central trigger process **to decide** which tanks should be in a run array to send event raw data.
- a event builder process **to build** from event raw data an auger event data stored in a database.
- a event displayer **to read** auger event data for analyze.



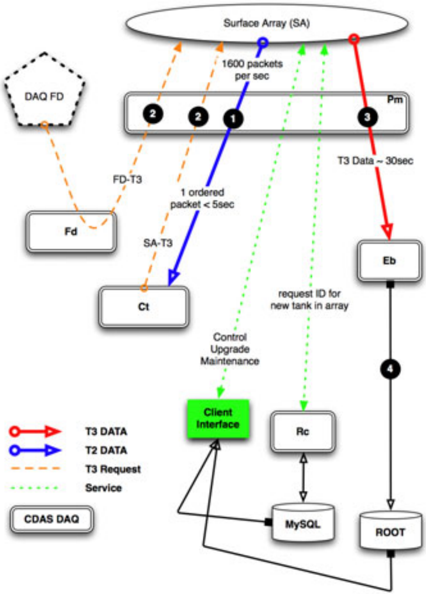
Central Data Acquisition System (CDAS)

- an hybrid client/server architecture
- a postmaster process **to dispatch** raw/event data (**2Mb/sec** via TCP) to/from clients/array.
- a central trigger process **to decide** which tanks should be in a run array to send event raw data.
- a event builder process **to build** from event raw data an auger event data stored in a database.
- a event displayer **to read** auger event data for analyze.



Central Data Acquisition System (CDAS)

- an hybrid client/server architecture
- a postmaster process **to dispatch** raw/event data (**2Mb/sec** via TCP) to/from clients/array.
- a central trigger process **to decide** which tanks should be in a run array to send event raw data.
- a event builder process **to build** from event raw data an auger event data stored in a database.
- a event displayer **to read** auger event data for analyze.



Free softwares & opensource tools everywhere

Data acquisition and management system

DAQ local softs coded only in C++ using ROOT and Gtk. C codes for embedded softs in tanks. MySQL for webservices.

Development framework & programming language

cvs/svn, bugzilla, c/c++, python, perl, shell, php, java.

System, Security & Network

Debian and SLC, nagios, cfengine, netfilter, openssh, rsync/tar, lvm/mdadm and Xen.

Sommaire

Thanks Tim!

- Tim Berners-Lee, a scientist at CERN, invented the **World Wide Web (WWW) in 1989**. The Web was originally conceived and developed to meet the demand for automatic information sharing between scientists all over the world.
- The basic idea of the WWW was to merge the technologies of personal computers, computer networking and hypertext into a powerful and easy to use global information system.

How it began

In 1991, an early WWW system was released to the high energy physics community via the CERN program library.

info.cern.ch

Was the address of the world's first-ever web site and web server, running on a NeXT computer at CERN (a hypermedia browser and a web editor).

Thanks Tim!

- Tim Berners-Lee, a scientist at CERN, invented the **World Wide Web (WWW) in 1989**. The Web was originally conceived and developed to meet the demand for automatic information sharing between scientists all over the world.
- The basic idea of the WWW was to merge the technologies of personal computers, computer networking and hypertext into a powerful and easy to use global information system.

How it began

In 1991, an early WWW system was released to the high energy physics community via the CERN program library.

info.cern.ch

Was the address of the world's first-ever web site and web server, running on a NeXT computer at CERN (a hypermedia browser and a web editor).

Thanks Tim!

- Tim Berners-Lee, a scientist at CERN, invented the **World Wide Web (WWW) in 1989**. The Web was originally conceived and developed to meet the demand for automatic information sharing between scientists all over the world.
- The basic idea of the WWW was to merge the technologies of personal computers, computer networking and hypertext into a powerful and easy to use global information system.

How it began

In 1991, an early WWW system was released to the high energy physics community via the CERN program library.

info.cern.ch

Was the address of the world's first-ever web site and web server, running on a NeXT computer at CERN (a hypermedia browser and a web editor).

Thanks Tim!

- Tim Berners-Lee, a scientist at CERN, invented the **World Wide Web (WWW) in 1989**. The Web was originally conceived and developed to meet the demand for automatic information sharing between scientists all over the world.
- The basic idea of the WWW was to merge the technologies of personal computers, computer networking and hypertext into a powerful and easy to use global information system.

How it began

In 1991, an early WWW system was released to the high energy physics community via the CERN program library.

info.cern.ch

Was the address of the world's first-ever web site and web server, running on a NeXT computer at CERN (a hypermedia browser and a web editor).

A linux tailored for HEP needs

Extracted from www.scientificlinux.org and linuxsoft.cern.ch

- A Linux release put together by Fermilab, CERN, and various other labs and universities around the world.
- An redhat entreprise linux (RHEL) distribution recompiled from source. SLC, a cern SL's flavour.
- A fully compatible RHEL with minor add-ons and changes, eg. pine/alpine, cern libs, openafs, etc.
- Allow easy high-level customization for a site/lab using anaconda without disturbing SL core.
- A growing team of developers and mainteners team inside HEP community.

A linux tailored for HEP needs

Extracted from www.scientificlinux.org and linuxsoft.cern.ch

- A Linux release put together by Fermilab, CERN, and various other labs and universities around the world.
- An redhat entreprise linux (RHEL) distribution recompiled from source. SLC, a cern SL's flavour.
- A fully compatible RHEL with minor add-ons and changes, eg. pine/alpine, cern libs, openafs, etc.
- Allow easy high-level customization for a site/lab using anaconda without disturbing SL core.
- A growing team of developers and mainteners team inside HEP community.

A linux tailored for HEP needs

Extracted from www.scientificlinux.org and linuxsoft.cern.ch

- A Linux release put together by Fermilab, CERN, and various other labs and universities around the world.
- An redhat entreprise linux (RHEL) distribution recompiled from source. SLC, a cern SL's flavour.
- A fully compatible RHEL with minor add-ons and changes, eg. pine/alpine, cern libs, openafs, etc.
- Allow easy high-level customization for a site/lab using anaconda without disturbing SL core.
- A growing team of developers and mainteners team inside HEP community.

A linux tailored for HEP needs

Extracted from www.scientificlinux.org and linuxsoft.cern.ch

- A Linux release put together by Fermilab, CERN, and various other labs and universities around the world.
- An redhat entreprise linux (RHEL) distribution recompiled from source. SLC, a cern SL's flavour.
- A fully compatible RHEL with minor add-ons and changes, eg. pine/alpine, cern libs, openafs, etc.
- Allow easy high-level customization for a site/lab using anaconda without disturbing SL core.
- A growing team of developers and mainteners team inside HEP community.

A linux tailored for HEP needs

Extracted from www.scientificlinux.org and linuxsoft.cern.ch

- A Linux release put together by Fermilab, CERN, and various other labs and universities around the world.
- An redhat entreprise linux (RHEL) distribution recompiled from source. SLC, a cern SL's flavour.
- A fully compatible RHEL with minor add-ons and changes, eg. pine/alpine, cern libs, openafs, etc.
- Allow easy high-level customization for a site/lab using anaconda without disturbing SL core.
- A growing team of developers and mainteners team inside HEP community.

Why ROOT ?

Extracted from www.cern.ch

- Growth of maintenance of old products developed in FORTRAN, with librairies old over 20years.
- Large scale in amount of data to be simulated and analyzed needed by LHC and other HEP experiments.
- Provide a basic and oriented-object framework for High Energy Physics computing with extensions to other domains, like simulation, reconstruction, event displays and DAQ.

Why ROOT ?

Extracted from www.cern.ch

- Growth of maintenance of old products developed in FORTRAN, with librairies old over 20years.
- Large scale in amount of data to be simulated and analyzed needed by LHC and other HEP experiments.
- Provide a basic and oriented-object framework for High Energy Physics computing with extensions to other domains, like simulation, reconstruction, event displays and DAQ.

Why ROOT ?

Extracted from www.cern.ch

- Growth of maintenance of old products developed in FORTRAN, with librairies old over 20years.
- Large scale in amount of data to be simulated and analyzed needed by LHC and other HEP experiments.
- Provide a basic and oriented-object framework for High Energy Physics computing with extensions to other domains, like simulation, reconstruction, event displays and DAQ.

How it works ?

- An OO data analysis framework that can handle and analyze large amounts of data in a very efficient way.
- Use sets of objects, specialized storage methods to get direct access to the separate attributes of the selected objects, without having to touch the bulk of the data.
- a powerful histograming methods in an arbitrary number of dimensions, curve fitting, function evaluation, minimization, graphics and visualization classes to allow the easy setup of an analysis system that can query and process the data interactively or in batch mode.
- designed in such a way that it can query its databases in parallel on clusters of workstations or many-core machines
- the premier platform on which to build data acquisition, simulation and data analysis systems.

How it works ?

- An OO data analysis framework that can handle and analyze large amounts of data in a very efficient way.
- Use sets of objects, specialized storage methods to get direct access to the separate attributes of the selected objects, without having to touch the bulk of the data.
- a powerful histograming methods in an arbitrary number of dimensions, curve fitting, function evaluation, minimization, graphics and visualization classes to allow the easy setup of an analysis system that can query and process the data interactively or in batch mode.
- designed in such a way that it can query its databases in parallel on clusters of workstations or many-core machines
- the premier platform on which to build data acquisition, simulation and data analysis systems.

How it works ?

- An OO data analysis framework that can handle and analyze large amounts of data in a very efficient way.
- Use sets of objects, specialized storage methods to get direct access to the separate attributes of the selected objects, without having to touch the bulk of the data.
- a powerful histograming methods in an arbitrary number of dimensions, curve fitting, function evaluation, minimization, graphics and visualization classes to allow the easy setup of an analysis system that can query and process the data interactively or in batch mode.
- designed in such a way that it can query its databases in parallel on clusters of workstations or many-core machines
- the premier platform on which to build data acquisition, simulation and data analysis systems.

How it works ?

- An OO data analysis framework that can handle and analyze large amounts of data in a very efficient way.
- Use sets of objects, specialized storage methods to get direct access to the separate attributes of the selected objects, without having to touch the bulk of the data.
- a powerful histograming methods in an arbitrary number of dimensions, curve fitting, function evaluation, minimization, graphics and visualization classes to allow the easy setup of an analysis system that can query and process the data interactively or in batch mode.
- designed in such a way that it can query its databases in parallel on clusters of workstations or many-core machines
- the premier platform on which to build data acquisition, simulation and data analysis systems.

How it works ?

- An OO data analysis framework that can handle and analyze large amounts of data in a very efficient way.
- Use sets of objects, specialized storage methods to get direct access to the separate attributes of the selected objects, without having to touch the bulk of the data.
- a powerful histograming methods in an arbitrary number of dimensions, curve fitting, function evaluation, minimization, graphics and visualization classes to allow the easy setup of an analysis system that can query and process the data interactively or in batch mode.
- designed in such a way that it can query its databases in parallel on clusters of workstations or many-core machines
- the premier platform on which to build data acquisition, simulation and data analysis systems.

Overview

- based on a distributed file system originally developed at the Information Technology Center at Carnegie-Mellon University that was called the "Andrew File System".
- marketed, maintained, and extended by Transarc Corporation.(now IBM Pittsburgh Labs)
- this release is a branch source of the IBM AFS product, a copy of the source was made available for community development and maintenance.

How it works ?

Defintion

AFS is a distributed filesystem that enables co-operating hosts (clients and servers) to efficiently share filesystem resources across both local area and wide area networks. It implements layers such as a replicated read-only content distribution, providing location independence, scalability, security, and transparent migration capabilities.

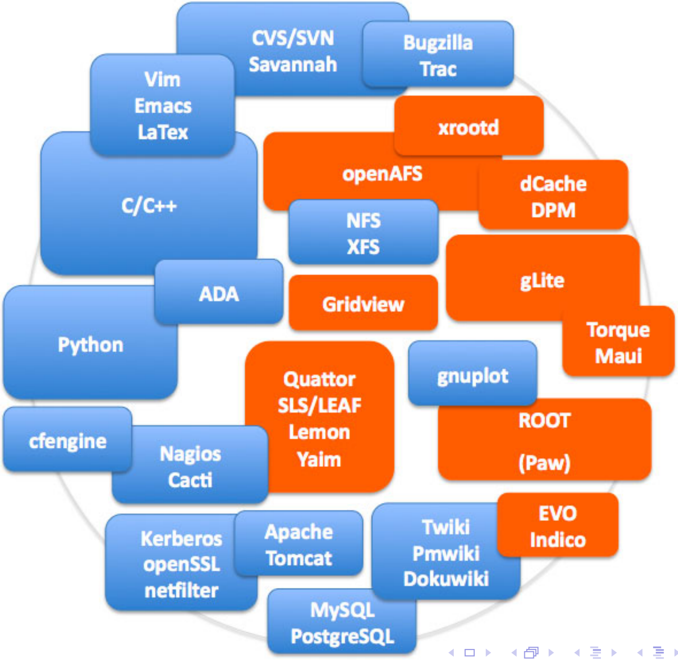
What is an AFS cell ?

A collection of servers grouped together administratively and presenting a single, cohesive filesystem. Typically, an AFS cell is a set of hosts that use the same Internet domain name.

Its strengths

Caching facility, security features with Kerberos, simplicity of addressing, scalability and communications protocol

HEP community thinks opensource, since always...



Thank you

Richard.Randria (at) LPNHE.IN2P3.FR
Computer Science Research Engineer - LPNHE/CNRS Paris
http://lpnhe.in2p3.fr - http://www.in2p3.fr