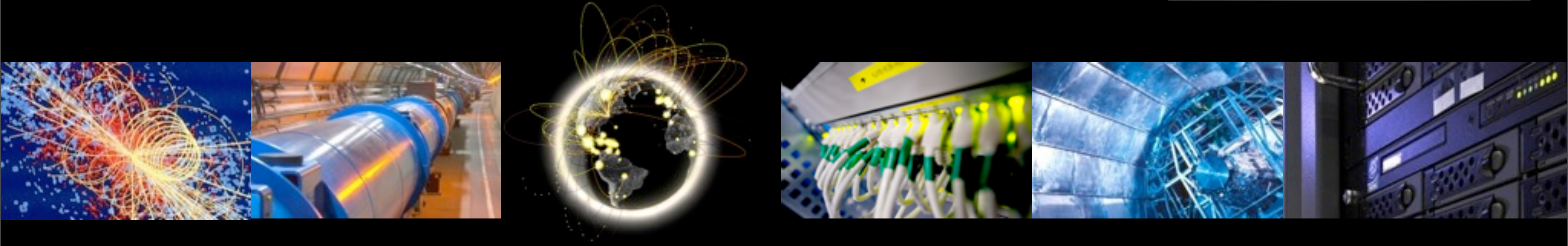
Grid Computing in HEP and other Fields

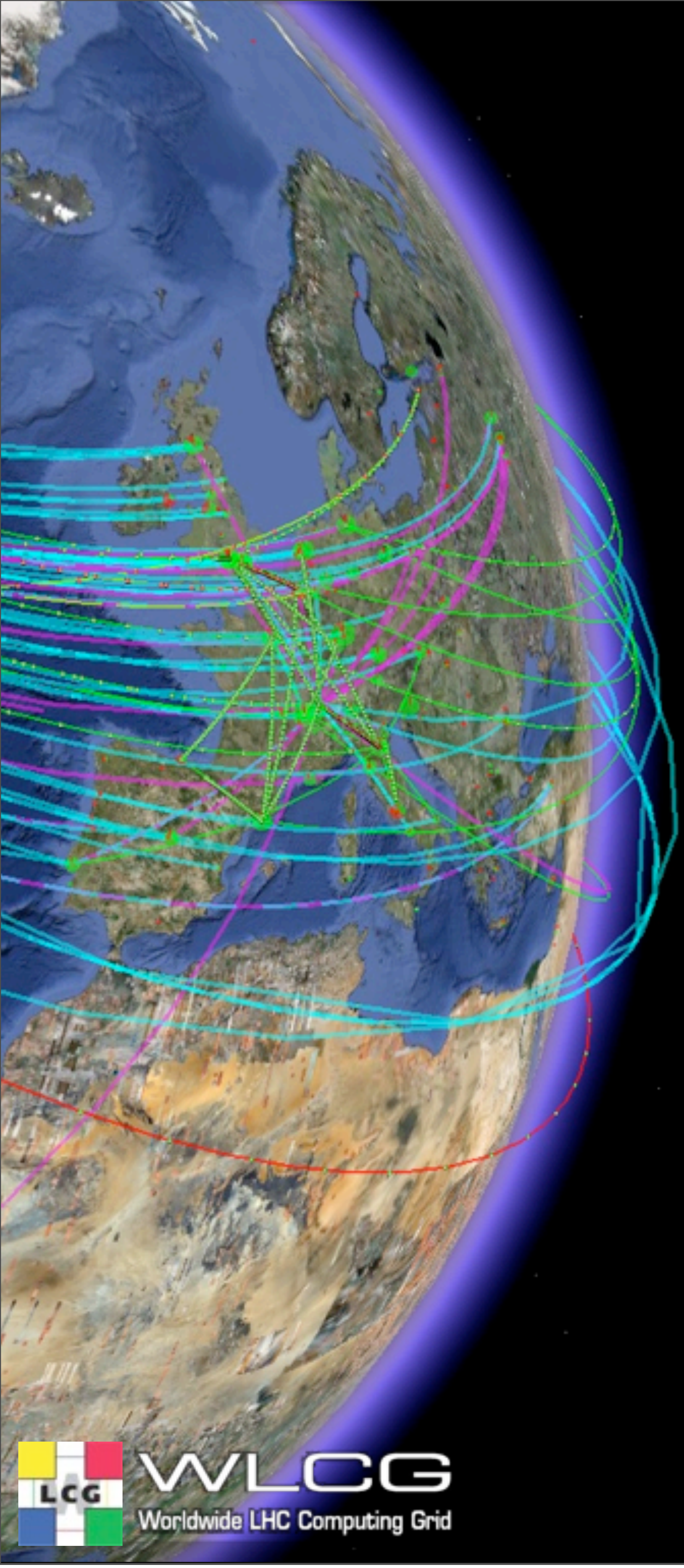
Dirk Duellmann, CERN

CERN School Thailand 2010

12 October 2010

Chulalongkorn University, Bangkok

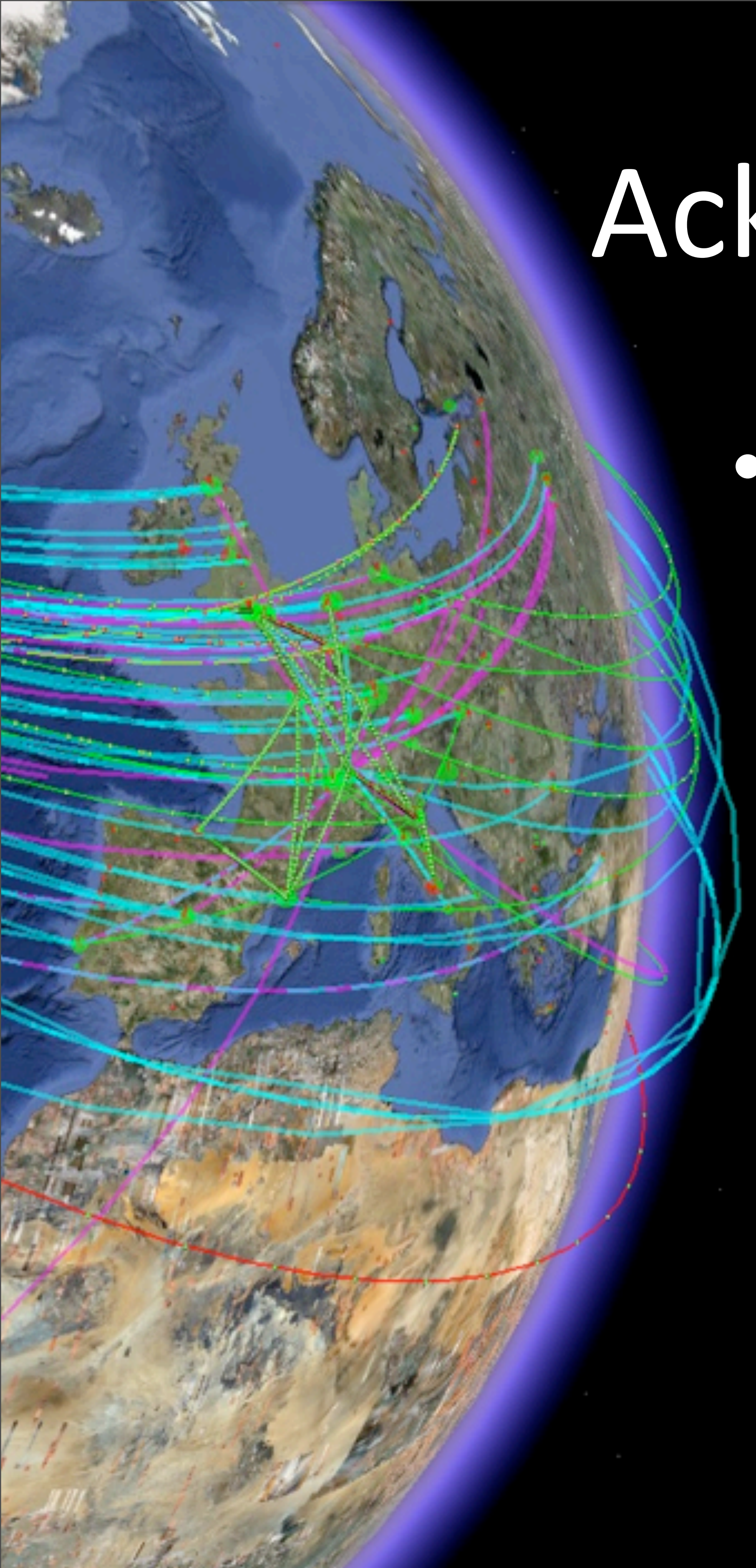




Outline

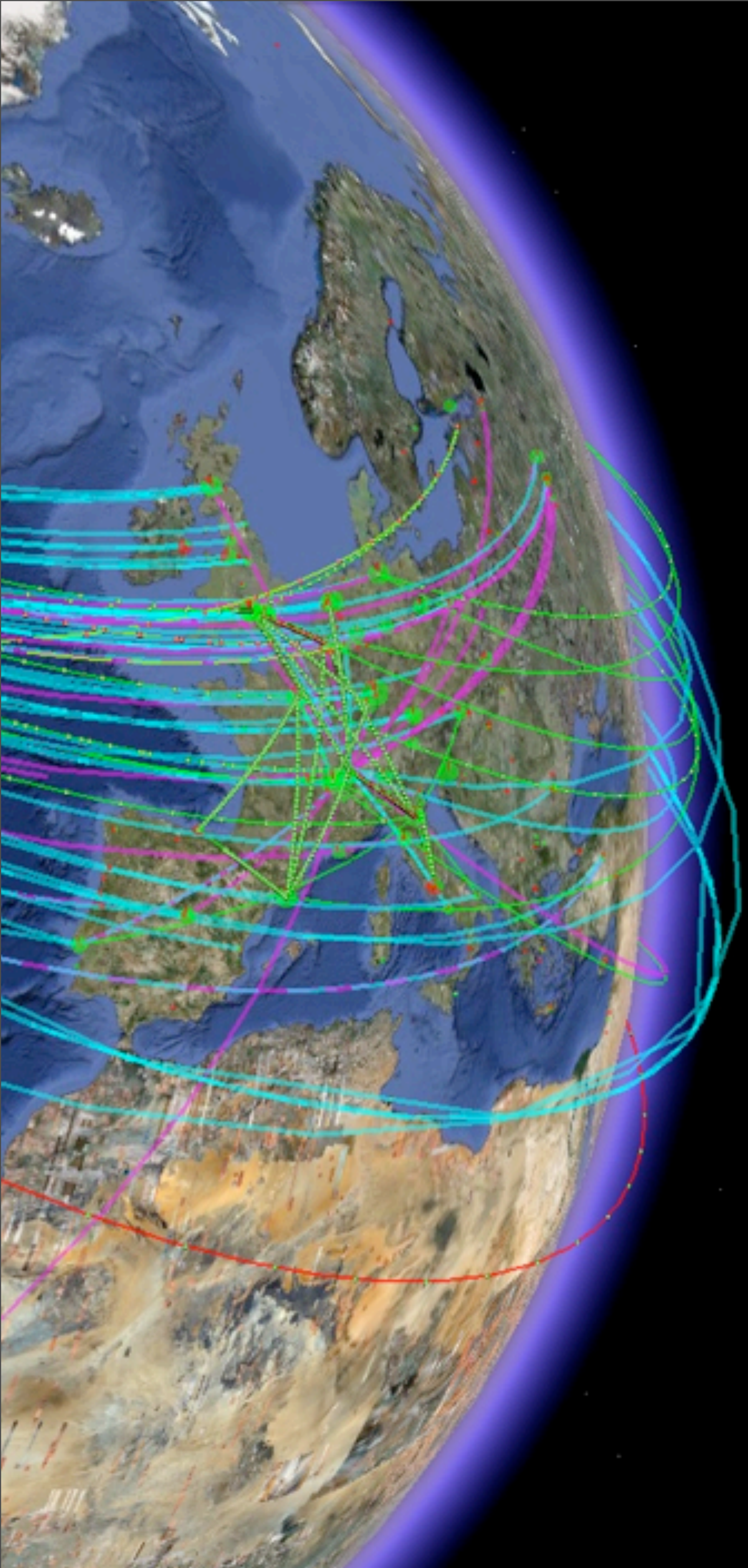
- Computing requirements of large scale projects
  - Eg LHC data processing
  - Hardware and Software base in HEP
- What is grid computing?
  - Why is it needed?
- Which problems are solved and how?
  - Concepts introduced by the grid
  - Components of today's production grids
  - How are large scale grids organised/coordinated?
  - How are they operated?
- What has be achieved up until today?
  - Which problems are still being addressed?
- Which future changes are expected?
  - Technology, market and user changes

Acknowledgements



- Many of my slides are collected from other presentations
  - –credits go to M. Schulz, B. Panzer-Steindel, I. Bird, J.F. Grosse-Oetringhaus, A. Pace and many others

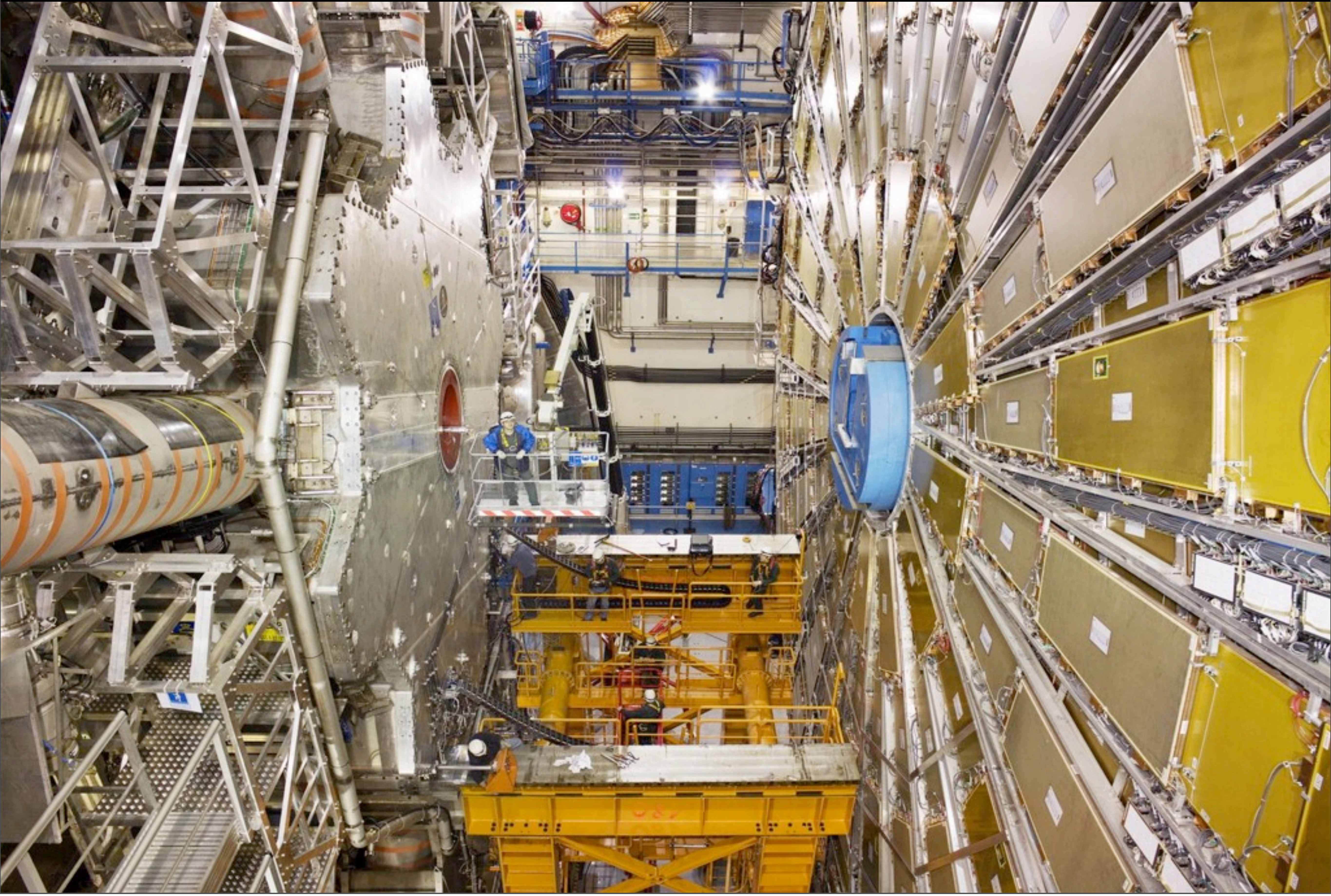
WLCG
Worldwide LHC Computing Grid



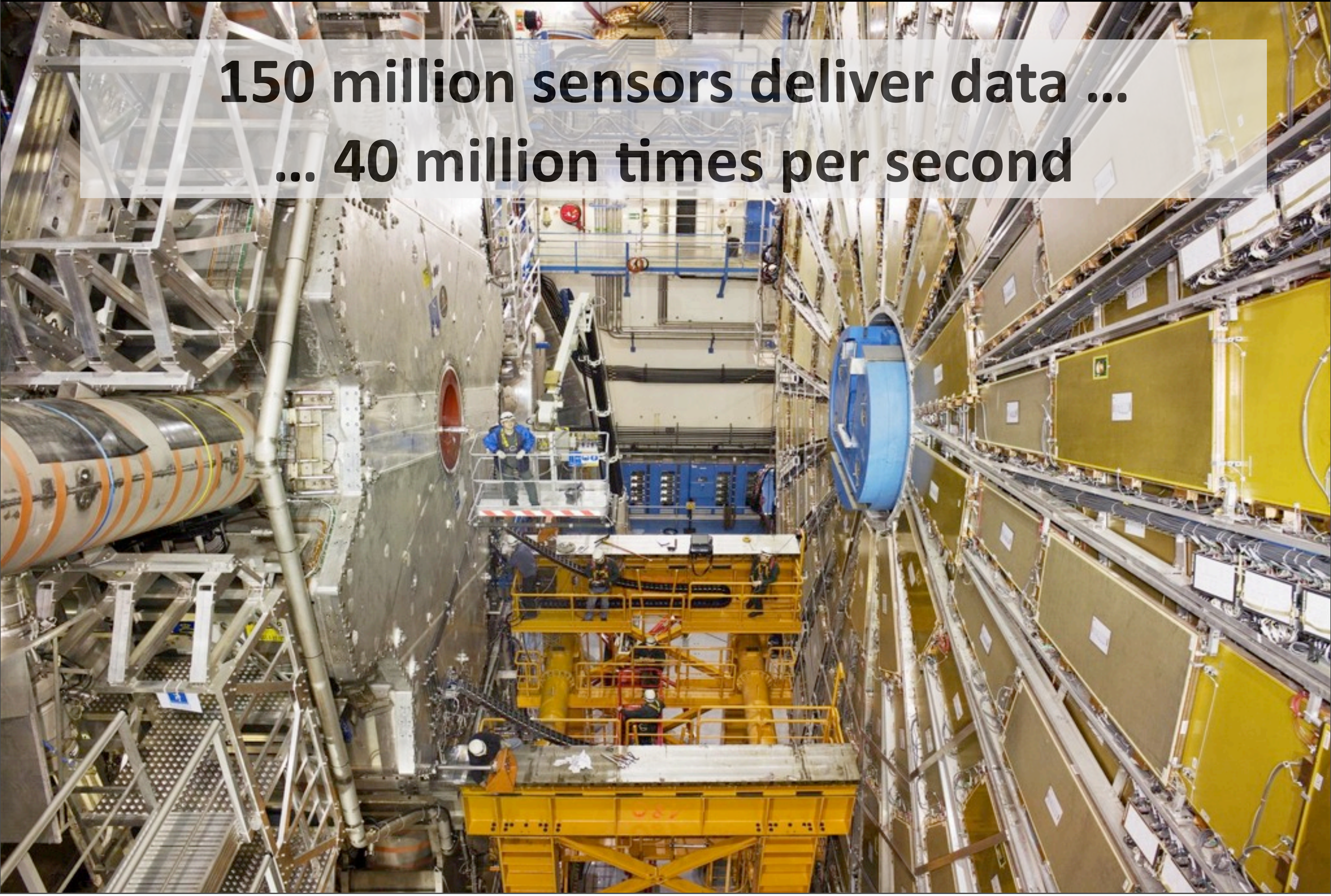
Computing for HEP - What do we need?

WLCG
Worldwide LHC Computing Grid

One of our data generators: ATLAS



One of our data generators: ATLAS

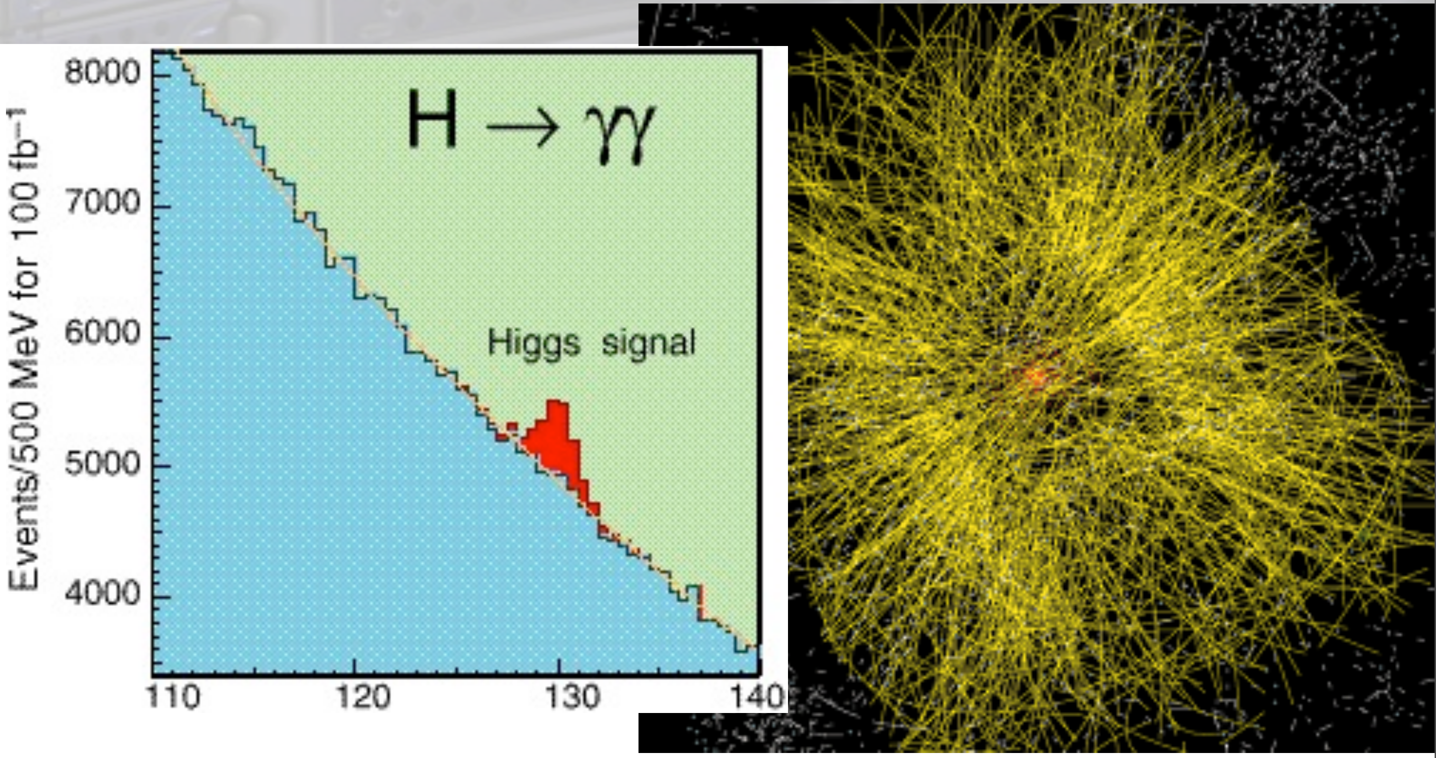


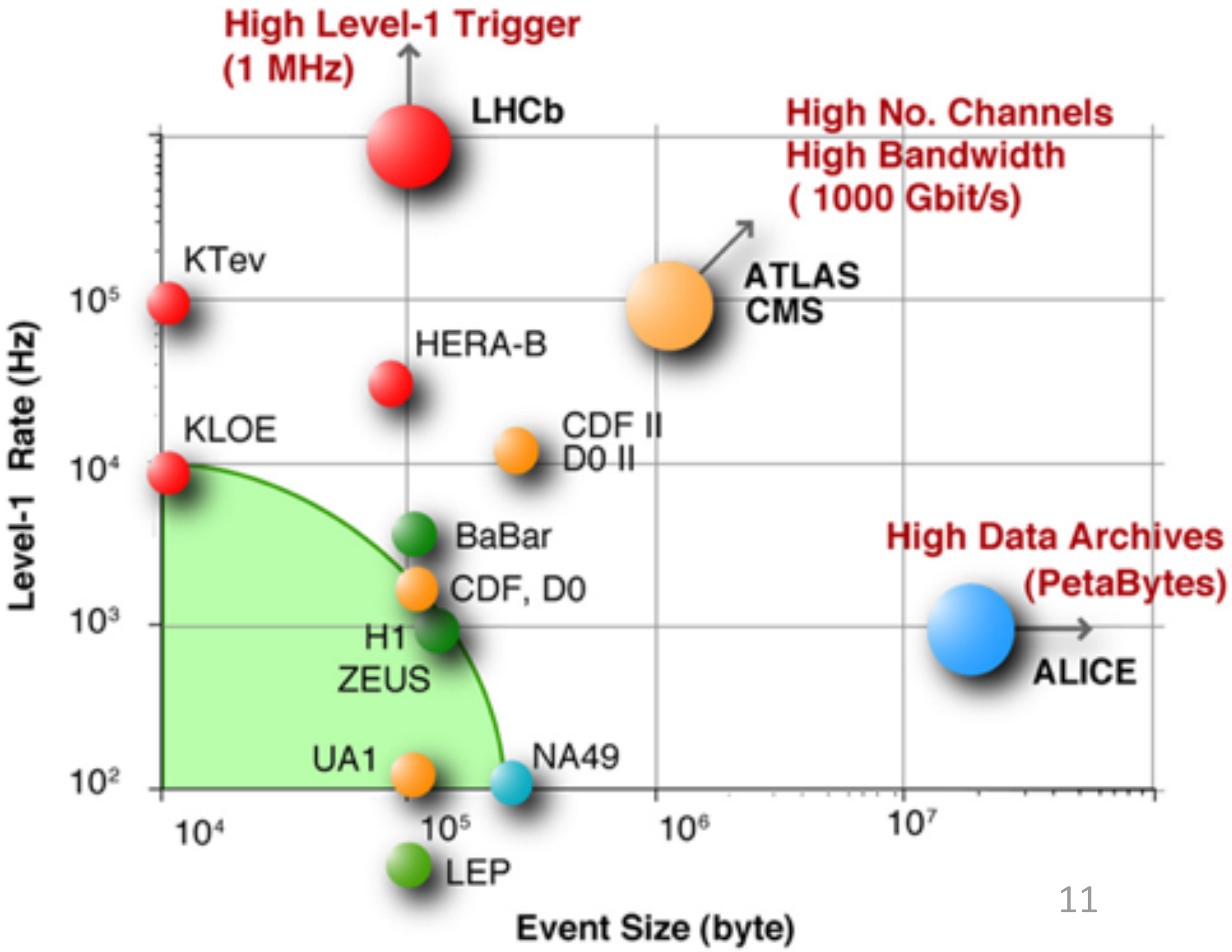
150 million sensors deliver data …

… 40 million times per second

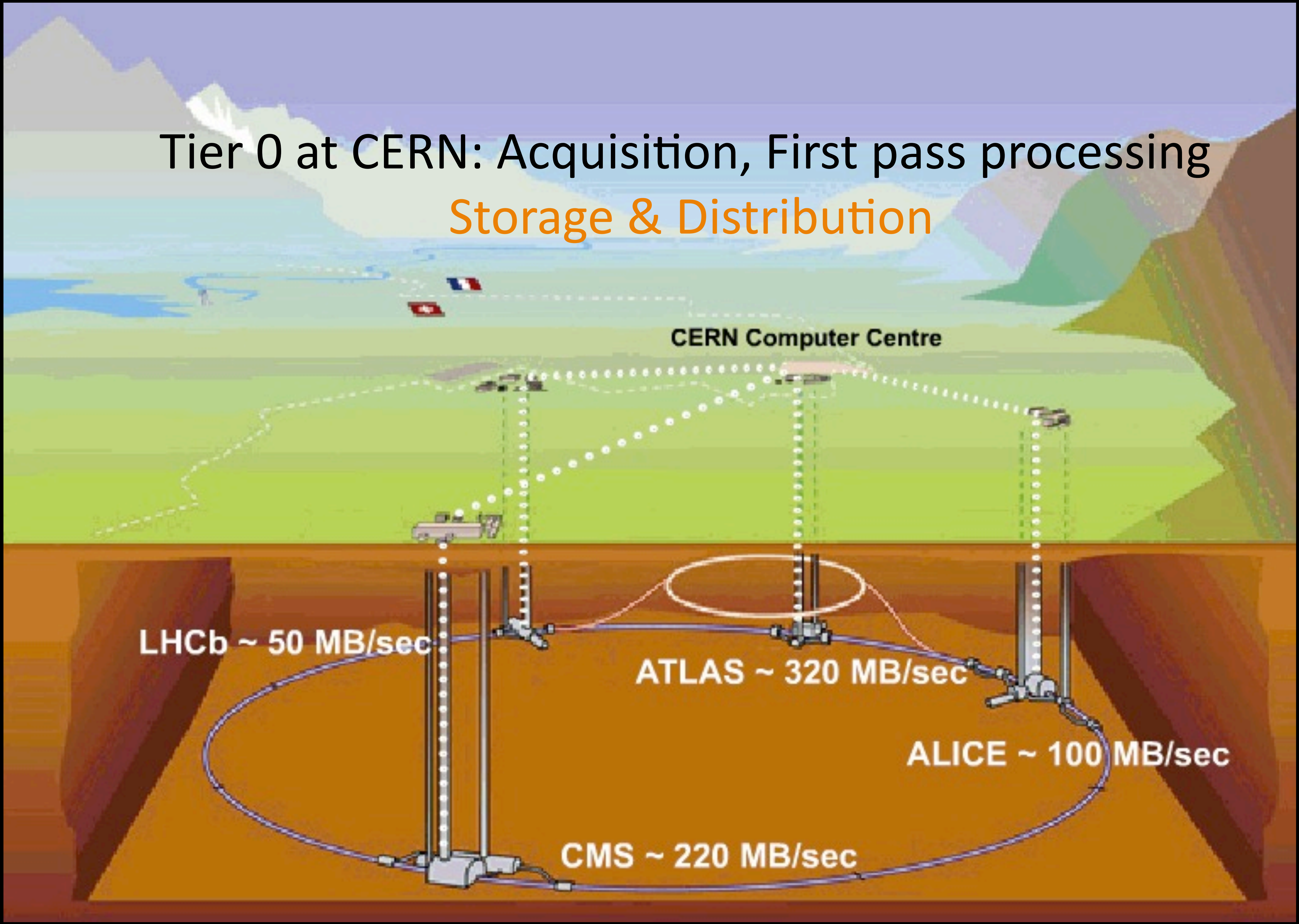
The LHC Computing Challenge

- Signal/Noise: 10^{-13} (10^{-9} offline)
- Data volume
  - High rate * large number of channels * 4 experiments
  - ➔ **15 Peta Bytes of new data each year**
- Compute power
  - Event complexity * Nb. events * thousands users
  - ➔**200 k of (today's) fastest CPUs**
  - ➔**45 PB of disk storage**
- Worldwide analysis & funding
  - Computing funding locally in major regions & countries
  - Efficient analysis everywhere
  - ➔ **GRID technology**



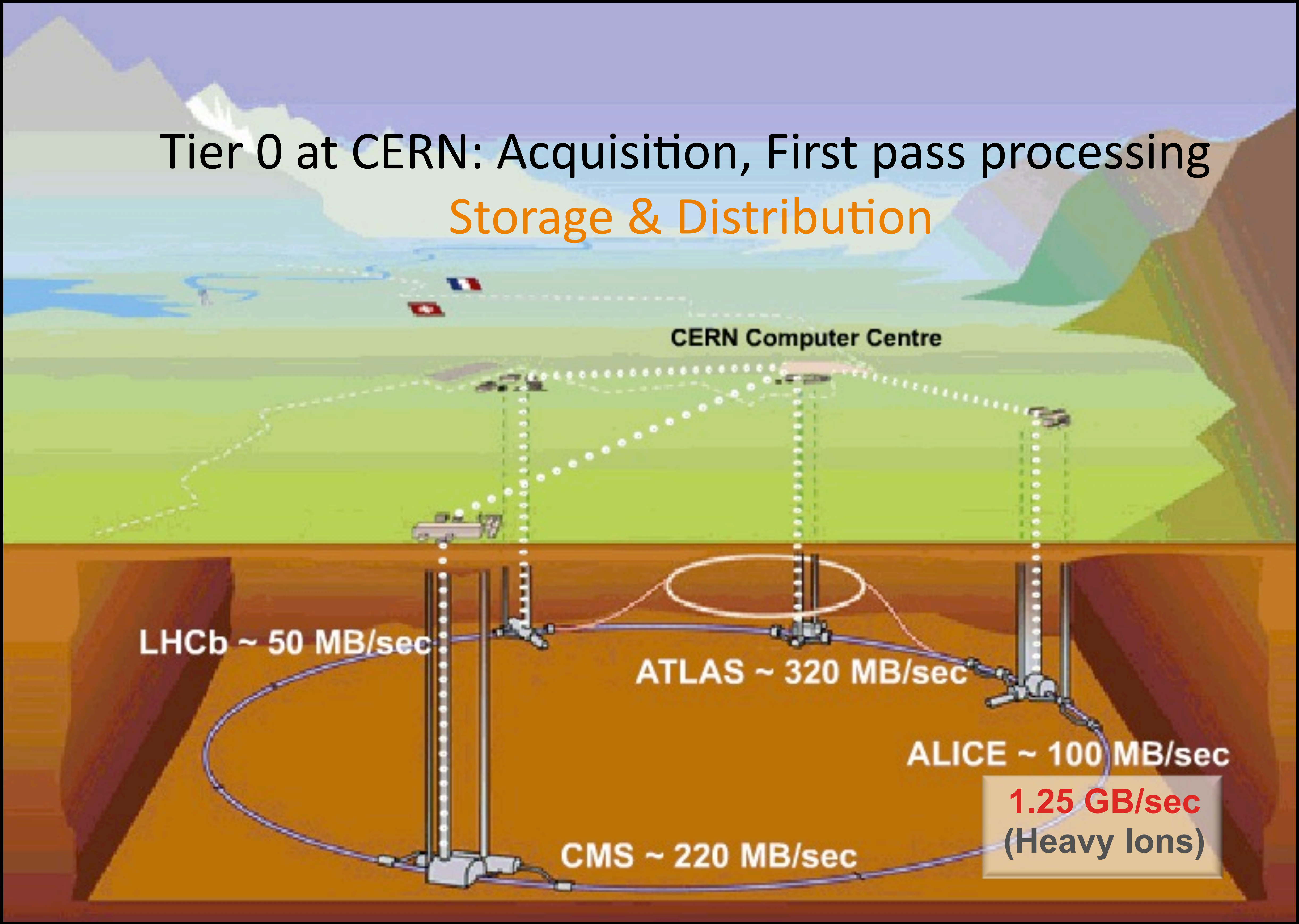


Tier 0 at CERN: Acquisition, First pass processing
Storage & Distribution



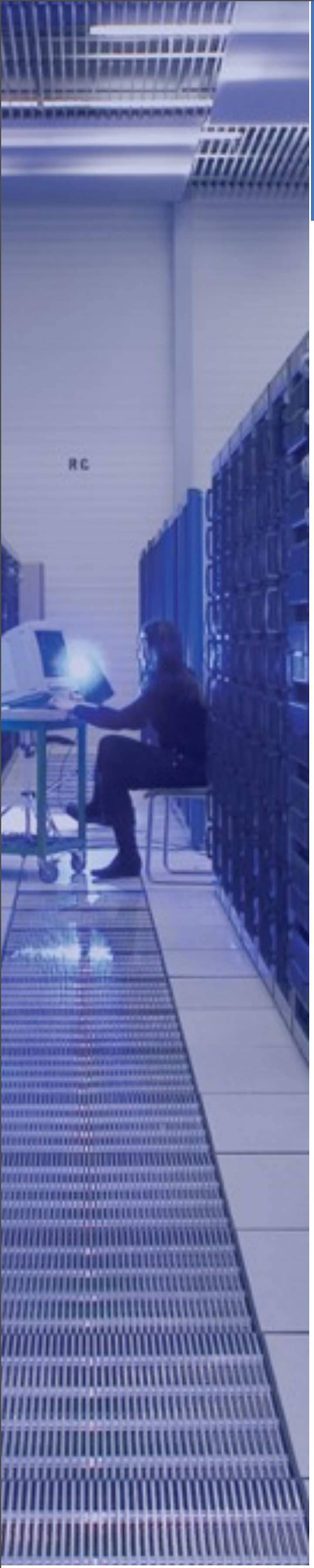
Tier 0 at CERN: Acquisition, First pass processing

Storage & Distribution



1.25 GB/sec
(Heavy Ions)

Physics dataflow



One Experiment

Detector →
150 million electronic channels

1 PBytes/s

Level 1 Filter and Selection

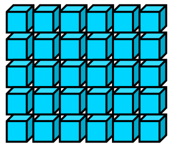
Fast response electronics,
FPGA, embedded processors,
very close to the detector

Limits :
Essentially the budget
and the downstream
data flow pressure

150 GBytes/s

High Level Filter and Selection

O(1000) PC's for processing,
Gbit Ethernet Network



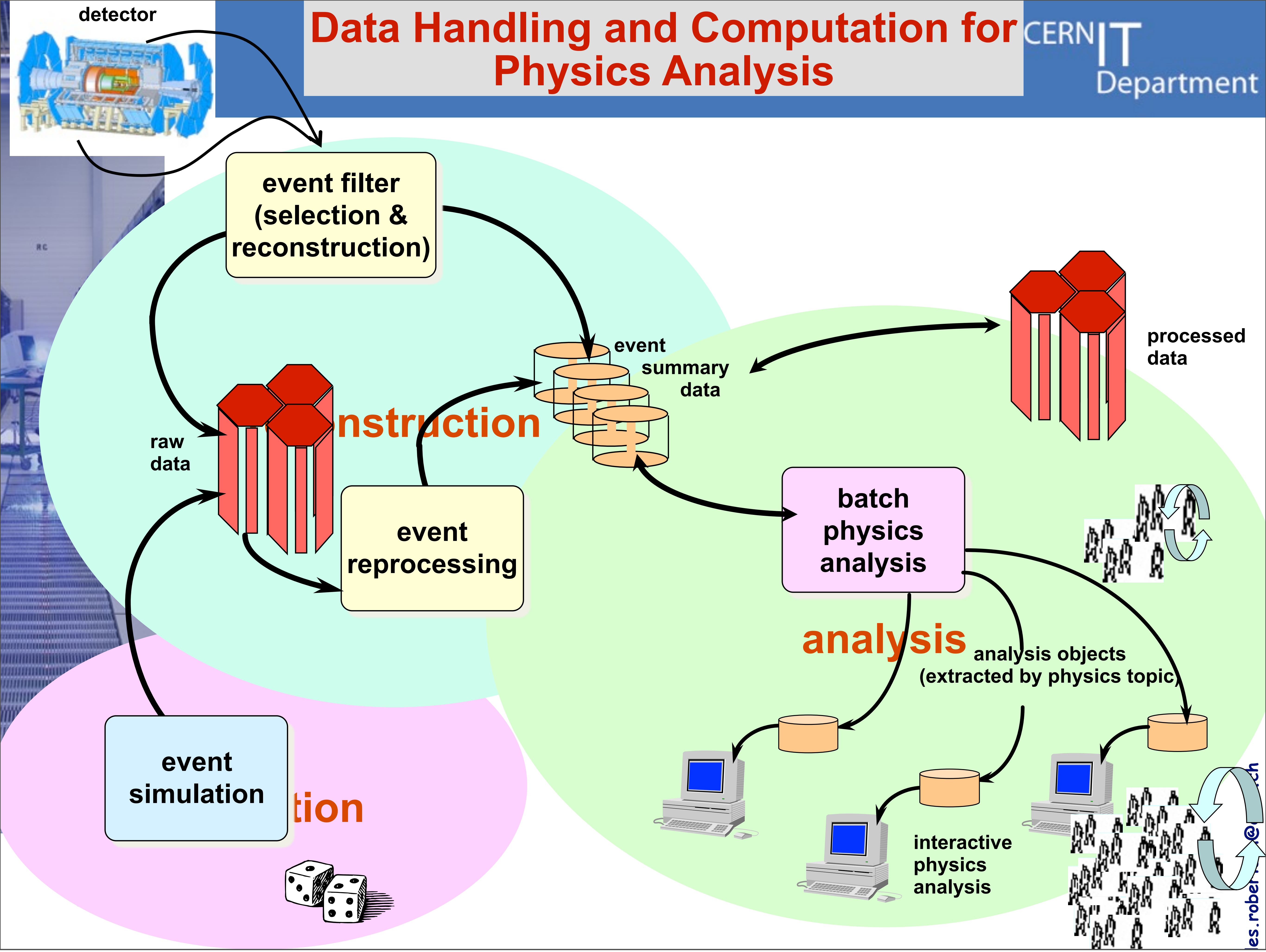
N x 10 Gbit links to
the Computer Center

0.6 GBytes/s

CERN Computer Center

Data Handling and Computation for Physics Analysis

CERN IT Department



detector

event filter (selection & reconstruction)

raw data

event summary data

processed data

event reprocessing

nstruction

batch physics analysis

analysis

analysis objects (extracted by physics topic)

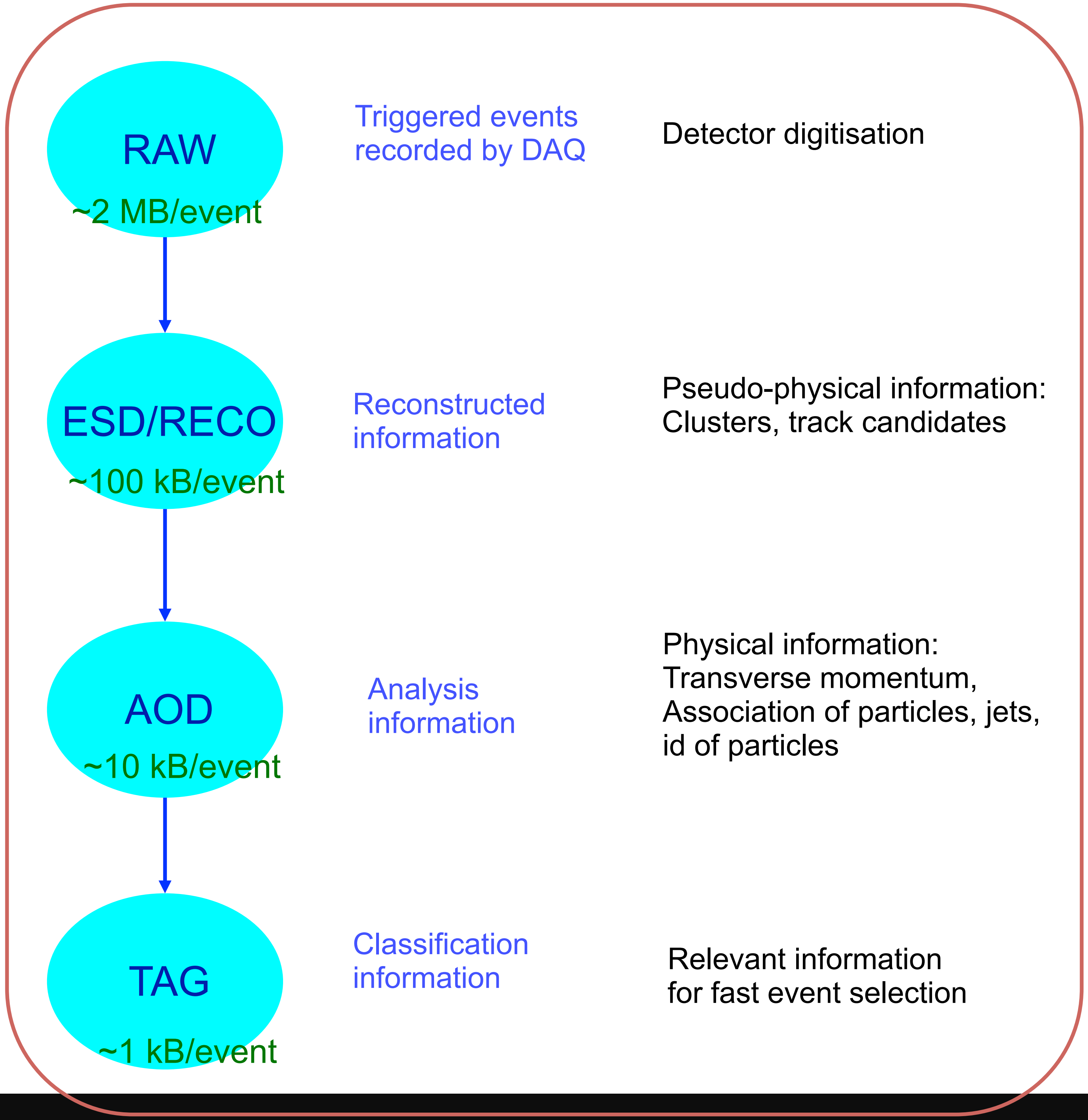
interactive physics analysis

event simulation

tion

Data and Algorithms

- HEP data are organised as *Events* (particle collisions)
- Simulation, Reconstruction and Analysis programs process “one event at a time”
  - events are independent → trivial parallel processing
- Event processing programs are composed of algorithms selecting and transforming “raw” events into “processed” (reconstructed) events



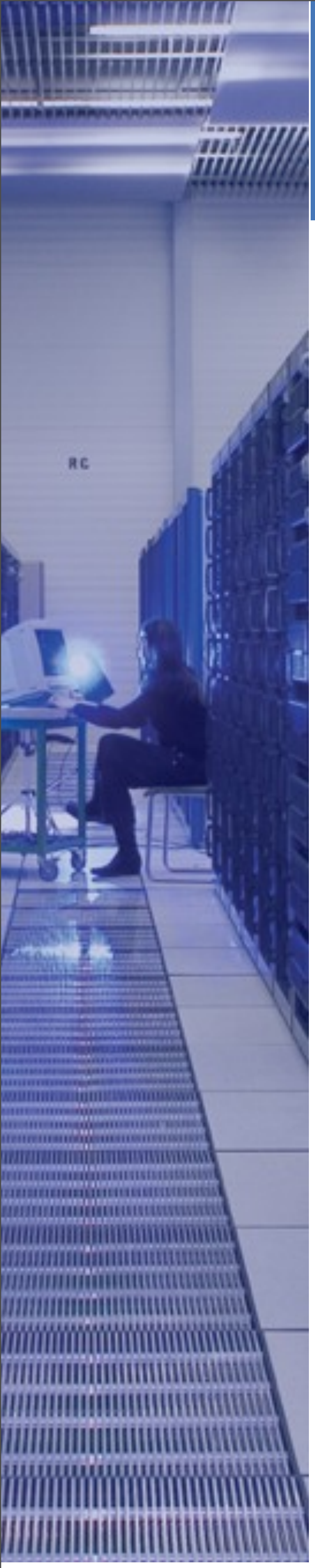
RAW ~2 MB/event — Triggered events recorded by DAQ — Detector digitisation

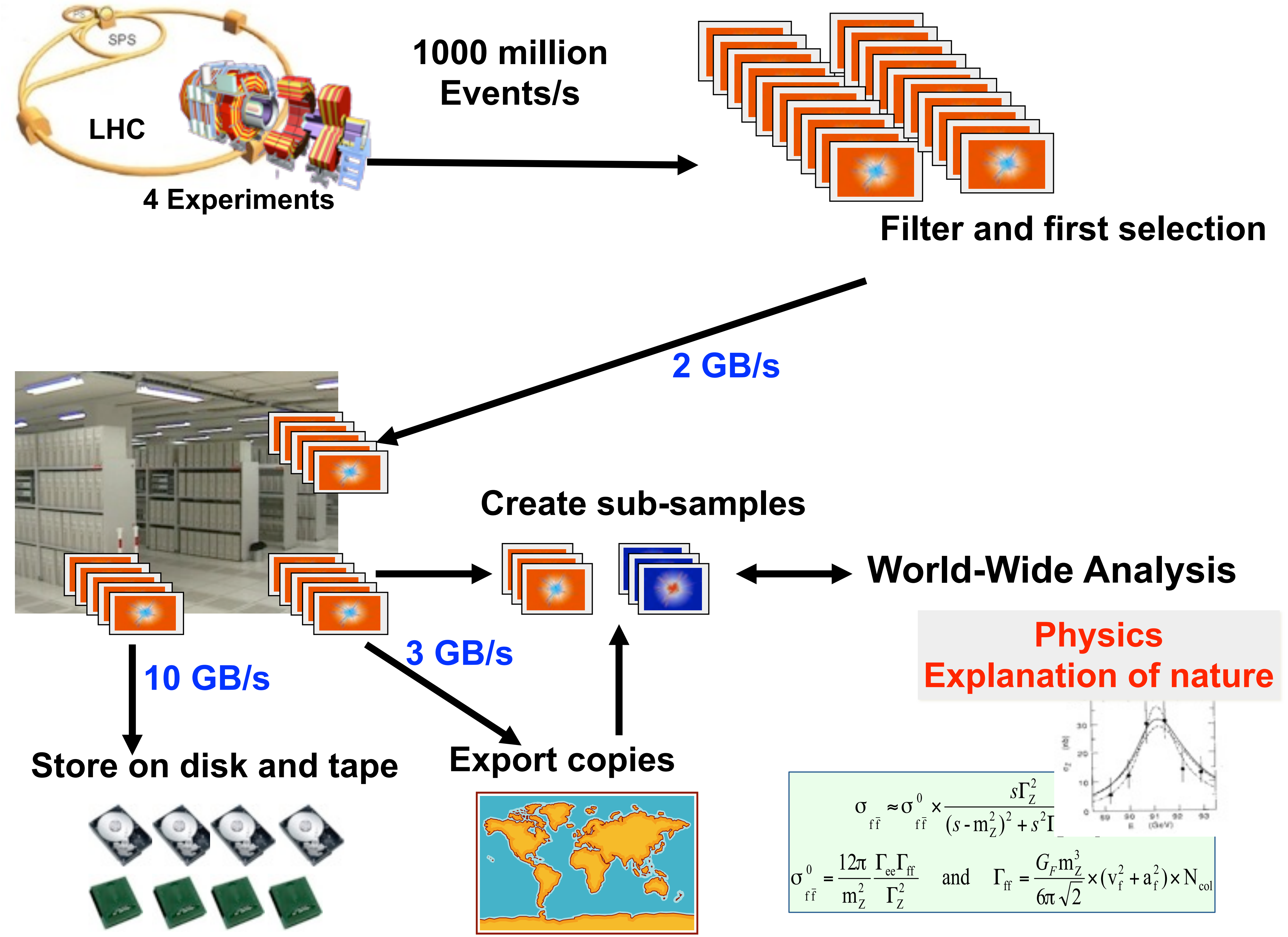
ESD/RECO ~100 kB/event — Reconstructed information — Pseudo-physical information: Clusters, track candidates

AOD ~10 kB/event — Analysis information — Physical information: Transverse momentum, Association of particles, jets, id of particles

TAG ~1 kB/event — Classification information — Relevant information for fast event selection

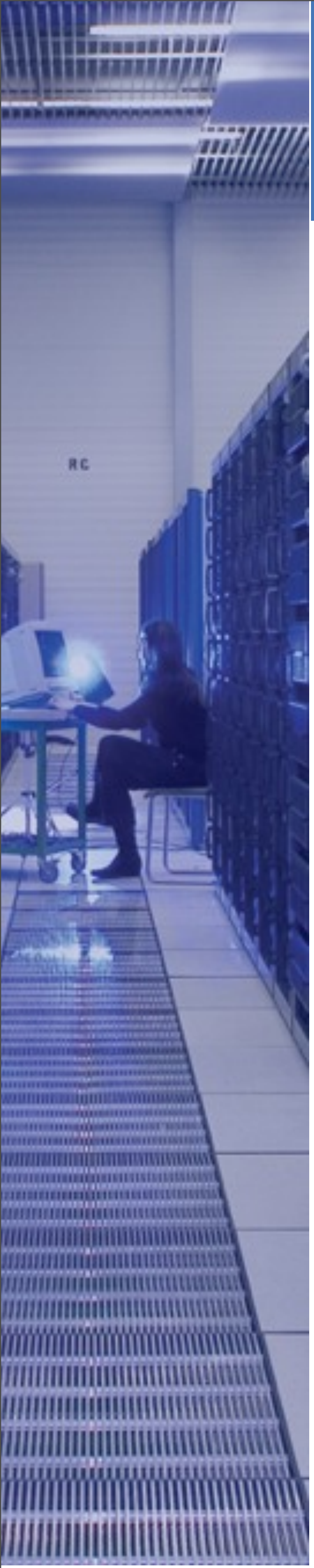
Physics dataflow II





$$\sigma_{f\bar f} \approx \sigma^0_{f\bar f} \times \frac{s\Gamma_Z^2}{(s-m_Z^2)^2+s^2\Gamma}$$

$$\sigma^0_{f\bar f} = \frac{12\pi}{m_Z^2}\,\frac{\Gamma_{ee}\Gamma_{ff}}{\Gamma_Z^2} \quad \text{and} \quad \Gamma_{ff} = \frac{G_F m_Z^3}{6\pi\sqrt{2}} \times (v_f^2 + a_f^2) \times N_{col}$$

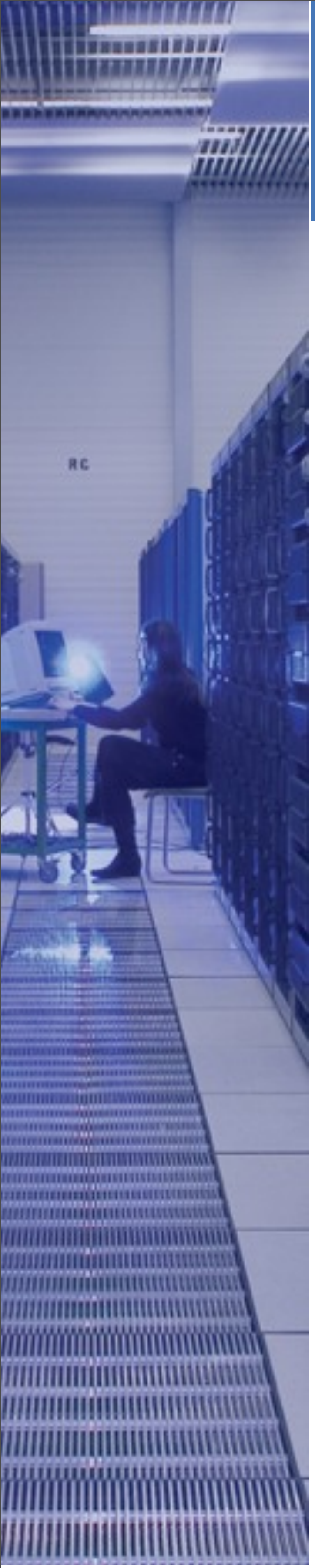


Information growth

Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections
- One photo ~ 2 Mbytes yearly → 500 billion photos
  one Exabyte
  25 billion photos on Facebook
- World Wide Web 25 billion pages (searchable)
  Deep web estimates > 1 trillion documents ~ one Exabyte
- Blue-ray disks ~ 25 Gbytes yearly → 100 million units
  2.5 Exabytes, copies

World-wide telephone calls → ~ 50 Exabyte (ECHELON)

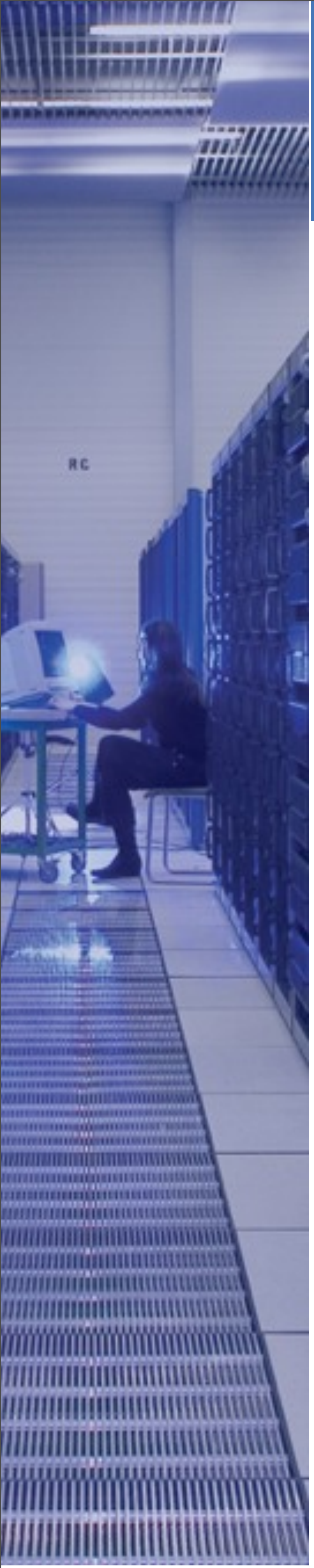


Information growth

Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

- Library of congress ~ 200 TB
- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections
- One photo ~ 2 Mbytes yearly → 500 billion photos
  one Exabyte
  25 billion photos on Facebook
- World Wide Web 25 billion pages (searchable)
  Deep web estimates > 1 trillion documents ~ one Exabyte
- Blue-ray disks ~ 25 Gbytes yearly → 100 million units
  2.5 Exabytes, copies

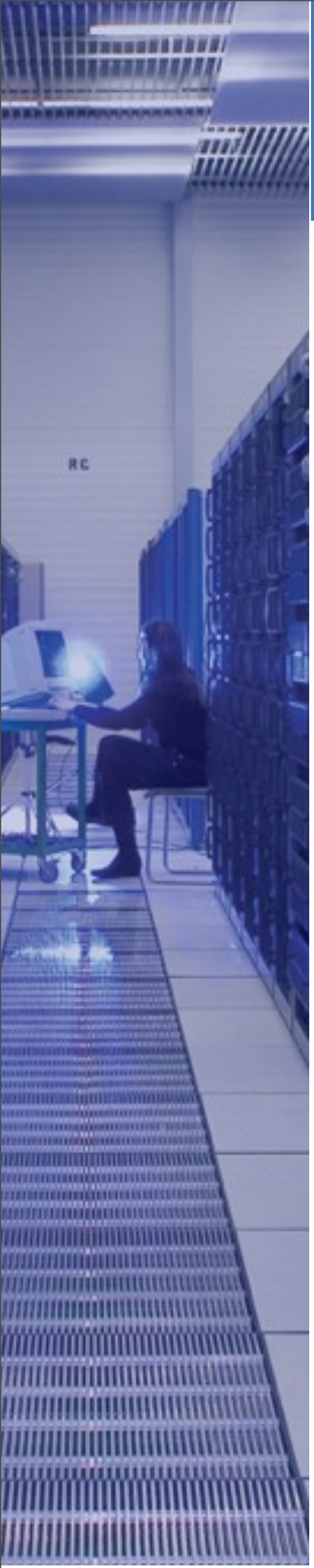
World-wide telephone calls → ~ 50 Exabyte (ECHELON)



Information growth

Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

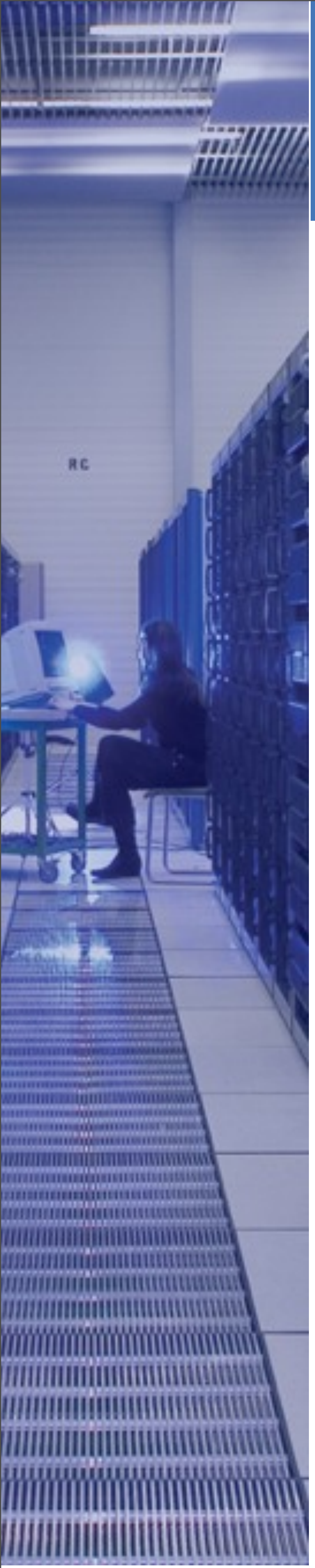
- Library of congress ~ 200 TB
- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections



Information growth

Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

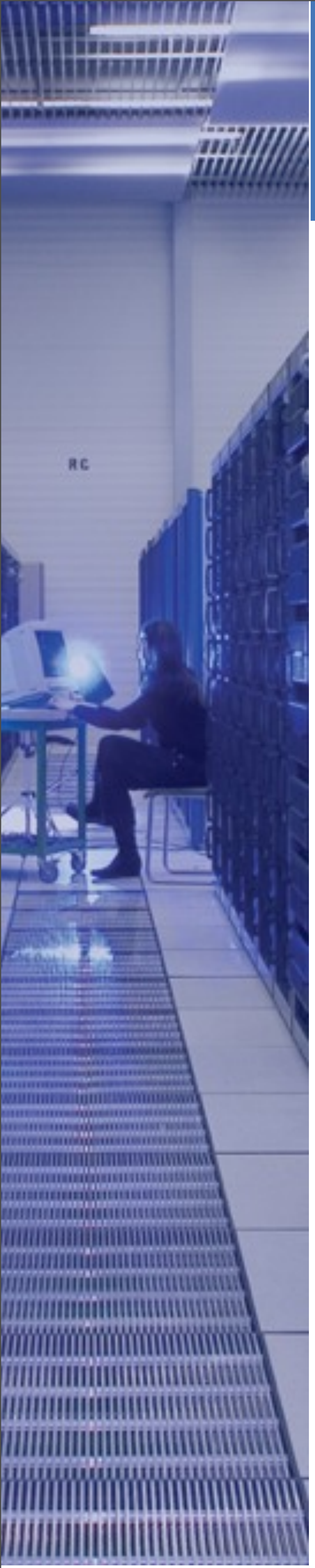
- Library of congress ~ 200 TB
- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections
- One photo ~ 2 Mbytes yearly → 500 billion photos
  one Exabyte
  25 billion photos on Facebook



Information growth

Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

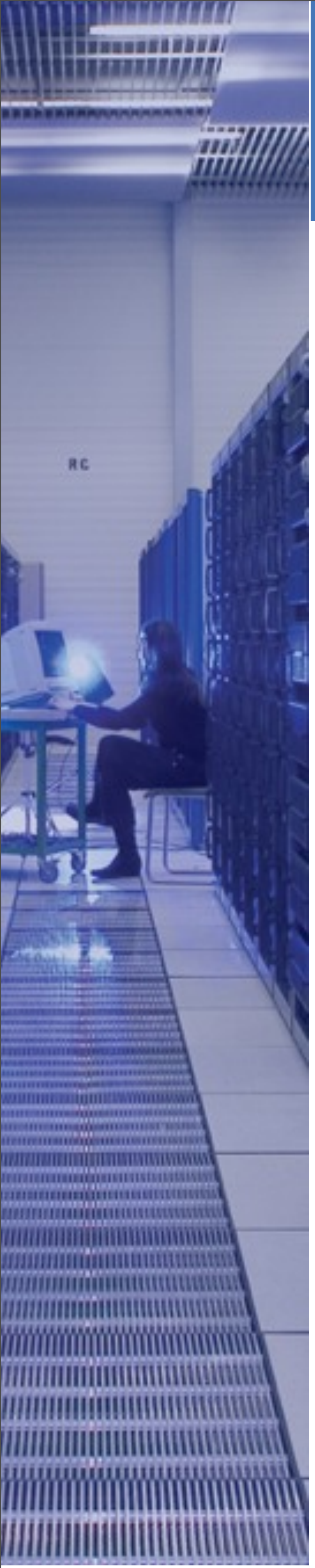
- Library of congress ~ 200 TB
- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections
- One photo ~ 2 Mbytes yearly → 500 billion photos
  one Exabyte
  25 billion photos on Facebook
- World Wide Web 25 billion pages (searchable)



Information growth

Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

- Library of congress ~ 200 TB
- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections
- One photo ~ 2 Mbytes yearly → 500 billion photos
  one Exabyte
  25 billion photos on Facebook
- World Wide Web 25 billion pages (searchable)
  Deep web estimates > 1 trillion documents ~ one Exabyte
- Blue-ray disks ~ 25 Gbytes yearly → 100 million units



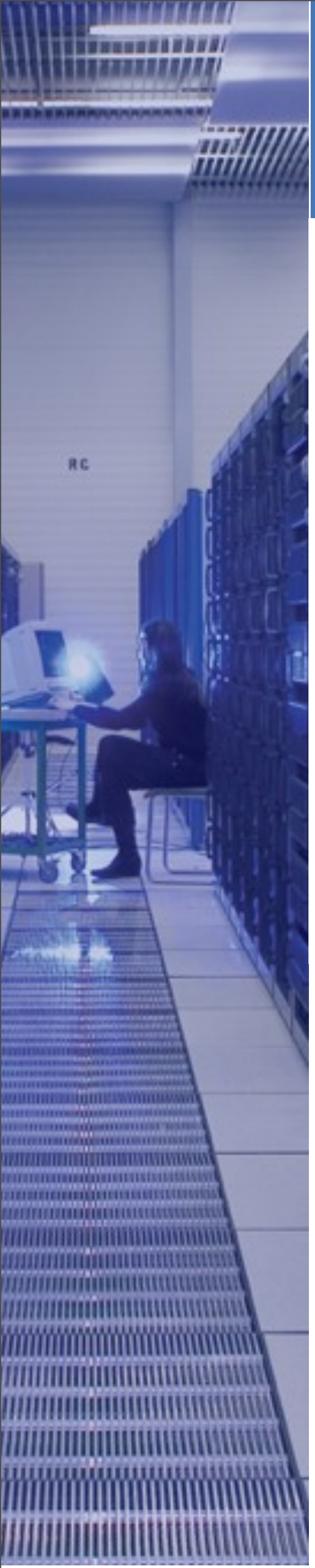
Information growth

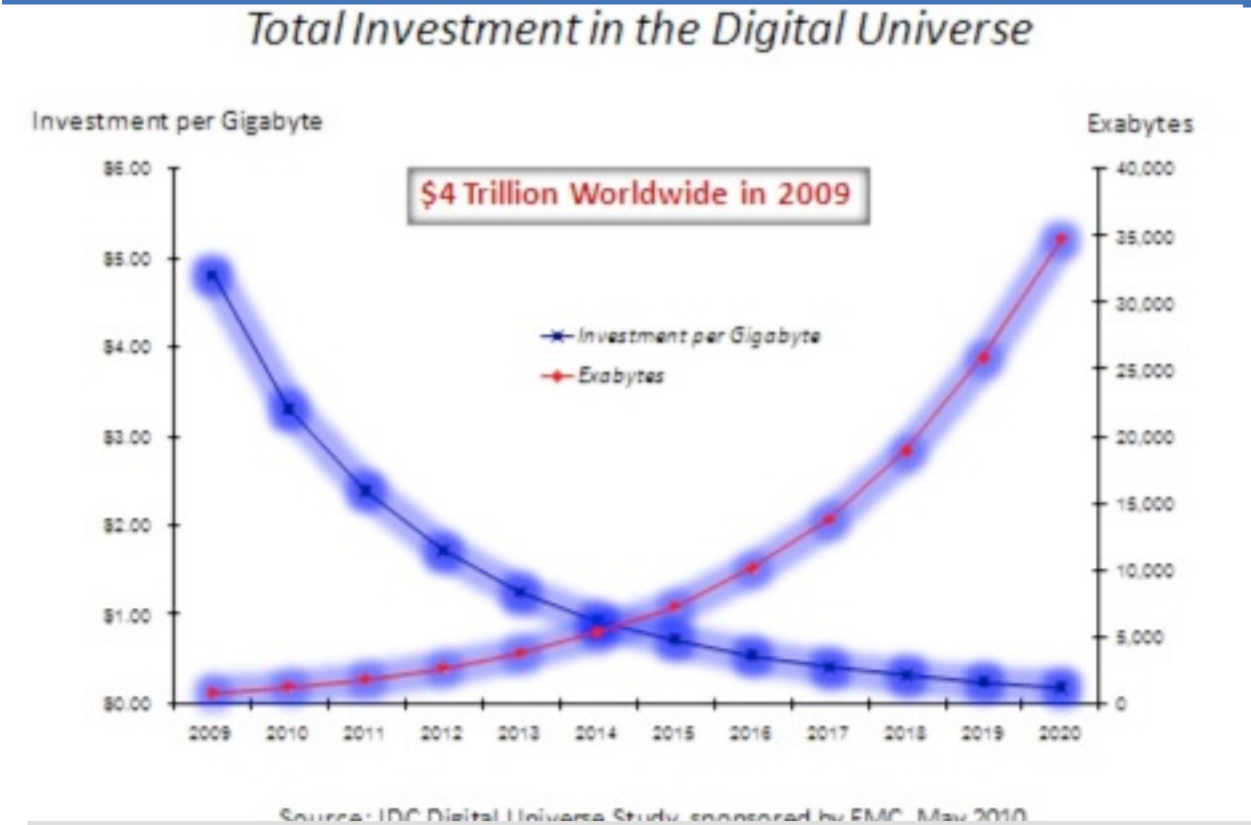
Zetta = 1000 Exa = 1000000 Peta = 1000000000 Tera

- Library of congress ~ 200 TB
- One email ~ 1 Kbyte yearly → 30 trillion emails (no spam)
  30 Petabyte * N for copies
- One LHC event ~ 1.0 Mbyte yearly → 10 billion RAW events
  ~ 70 Petabyte with copies
  and selections
- One photo ~ 2 Mbytes yearly → 500 billion photos
  one Exabyte
  25 billion photos on Facebook
- World Wide Web 25 billion pages (searchable)
  Deep web estimates > 1 trillion documents ~ one Exabyte
- Blue-ray disks ~ 25 Gbytes yearly → 100 million units
  2.5 Exabytes, copies

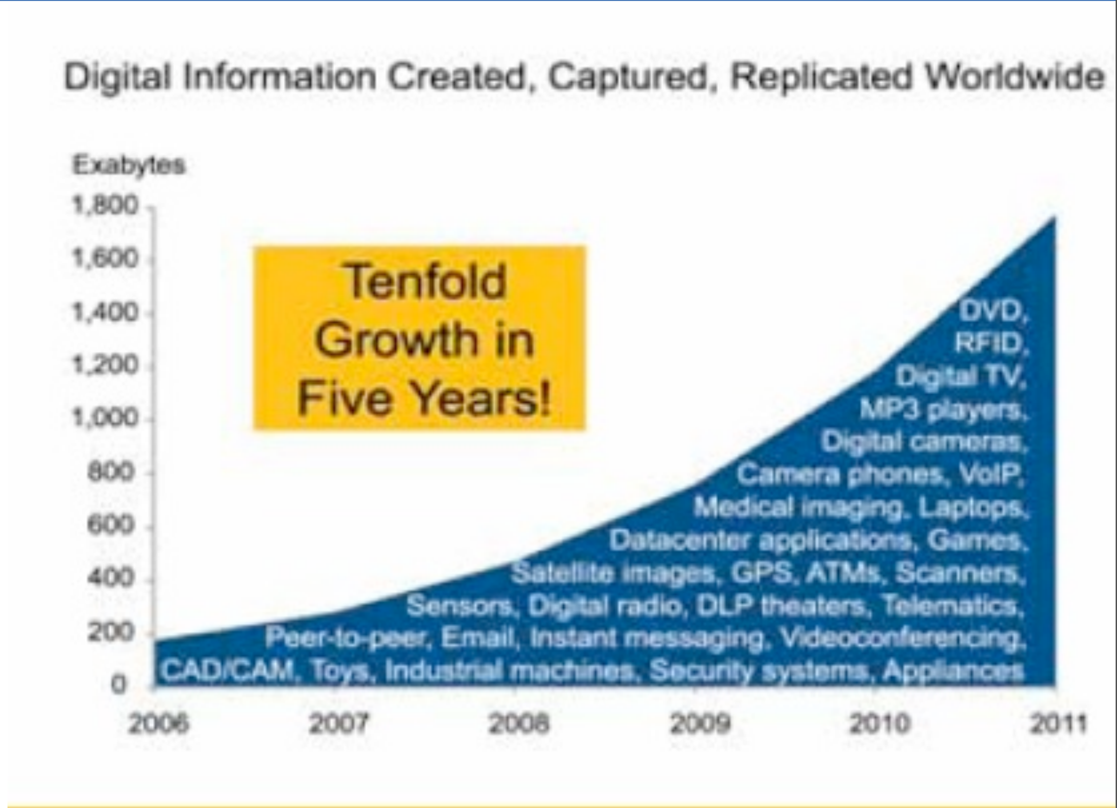
World-wide telephone calls → ~ 50 Exabyte (ECHELON)

The Digital Universe

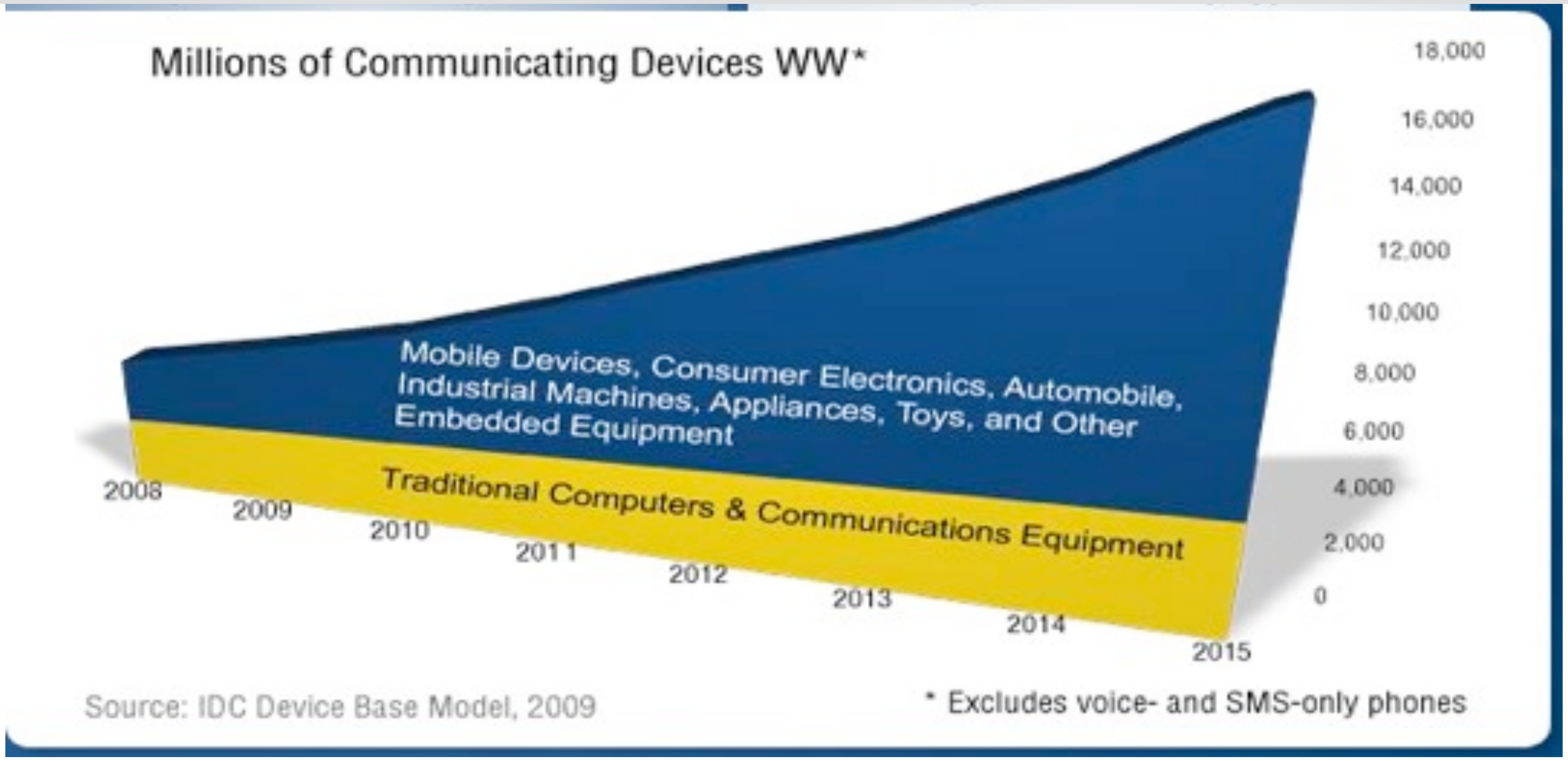




Source: IDC Digital Universe Study, sponsored by EMC, May 2010



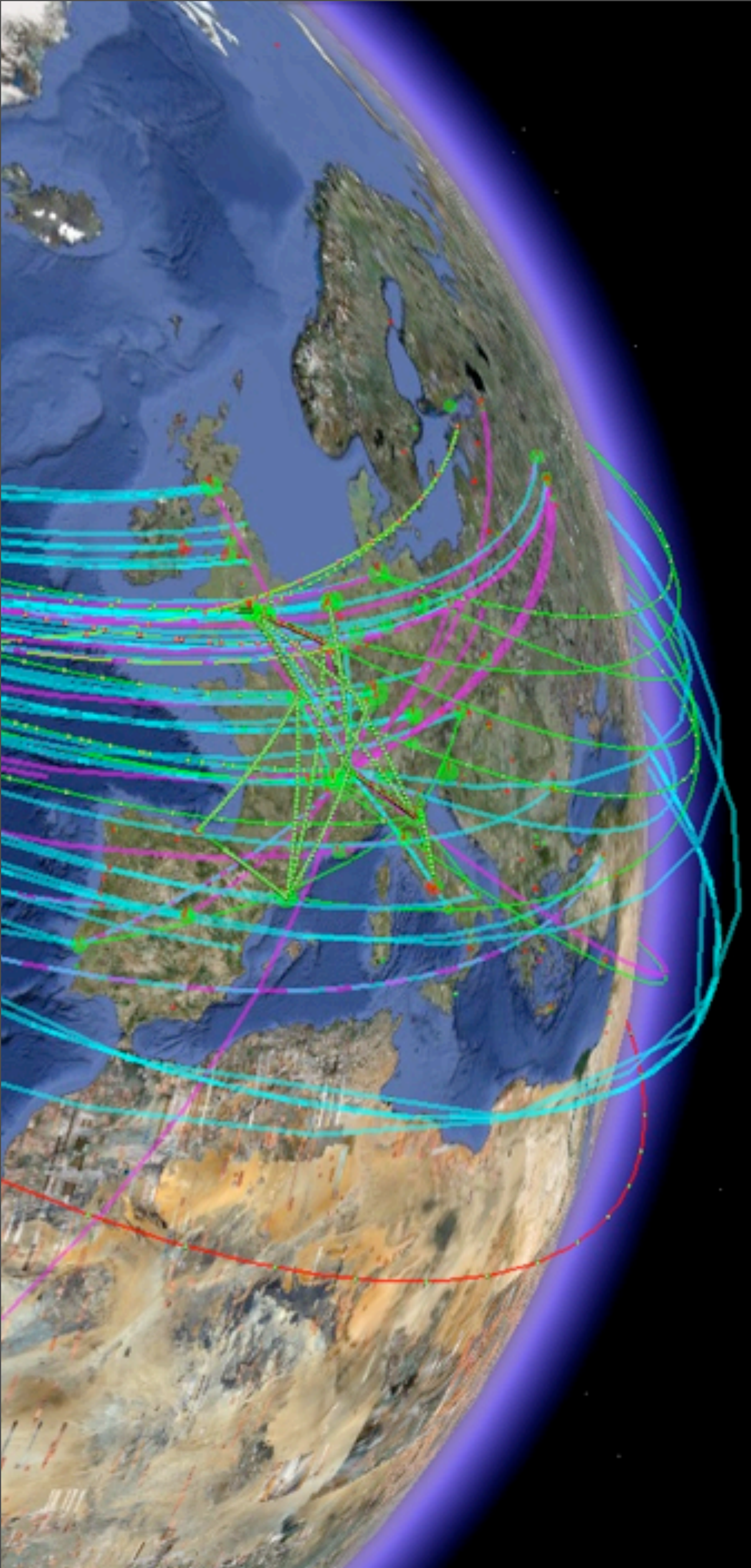
We will reach more than 1 Zettabyte of global information produced in 2010



Source: IDC Device Base Model, 2009

* Excludes voice- and SMS-only phones

http://indonesia.emc.com/leadership/digital-universe/expanding-digital-universe.htm

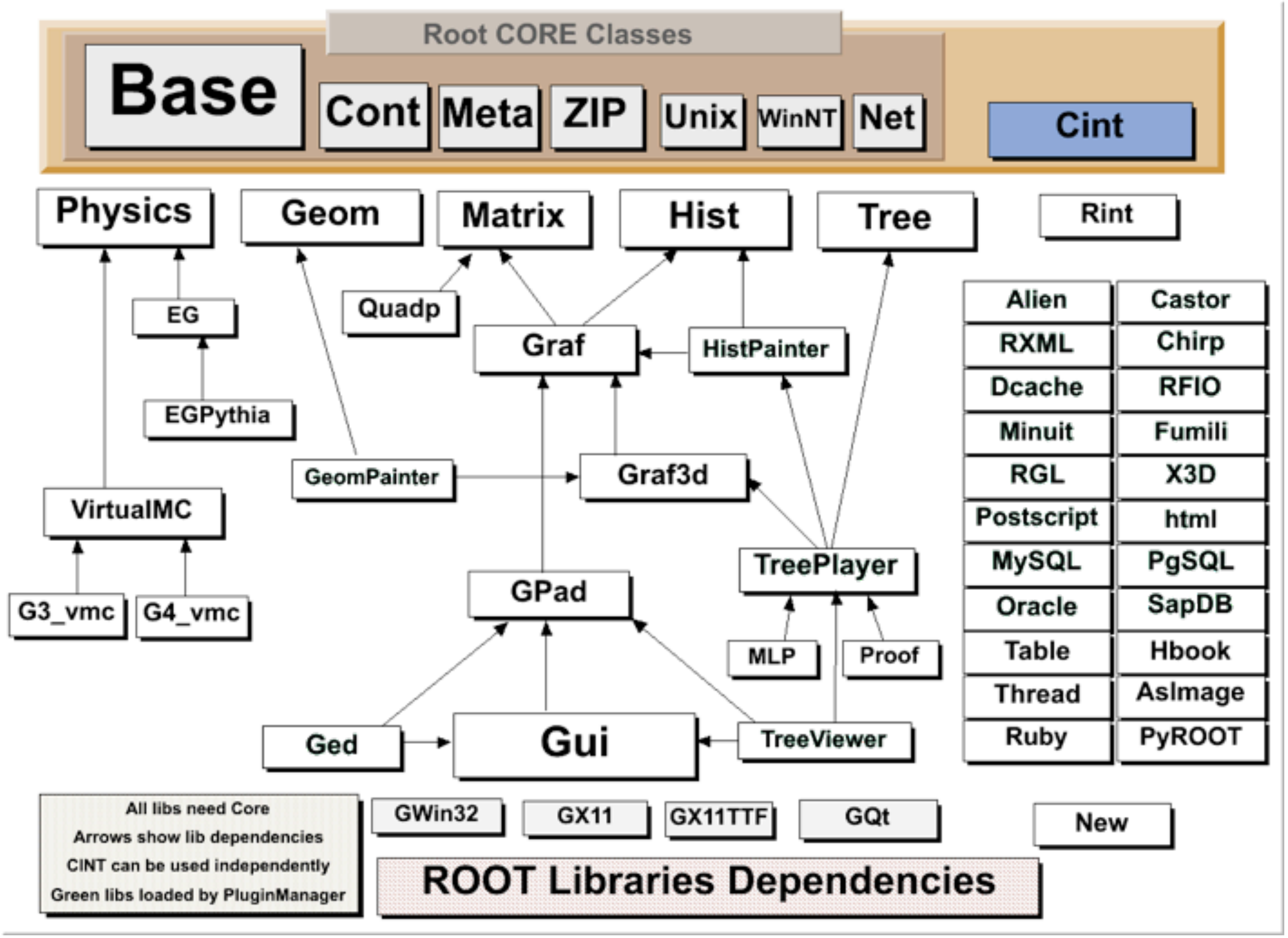


The HEP Software Environment

ROOT in a Nutshell

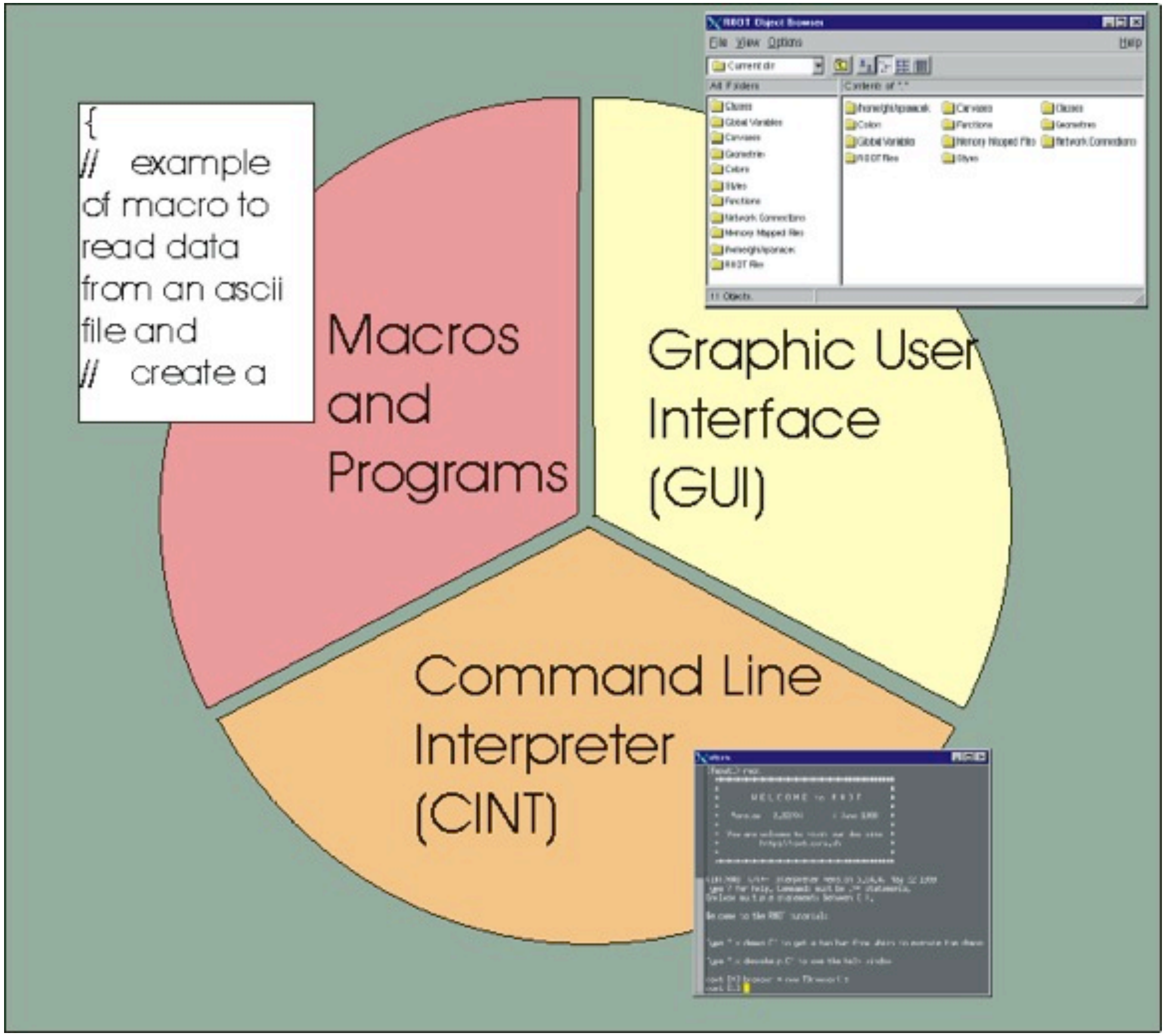
- ROOT is a large Object-Oriented data handling and analysis framework
- Efficient object store scaling from KB's to PB's
- C++ interpreter
- Extensive 2D+3D scientific data visualisation capabilities
- Extensive set of multi-dimensional histograming, data fitting, modelling and analysis methods
- Complete set of GUI widgets
- Classes for threading, shared memory, networking, etc.
- Parallel version of analysis engine runs on clusters and multi-core
- Fully cross platform: Unix/Linux, MacOS X and Windows

The ROOT Libraries



- Over 2500 classes
- 3,000,000 lines of code
- CORE (8 Mbytes)
- CINT (2 Mbytes)
- Most libraries linked on demand via plug-in manager (only a subset shown)
- 100 shared libs

Three User Interfaces



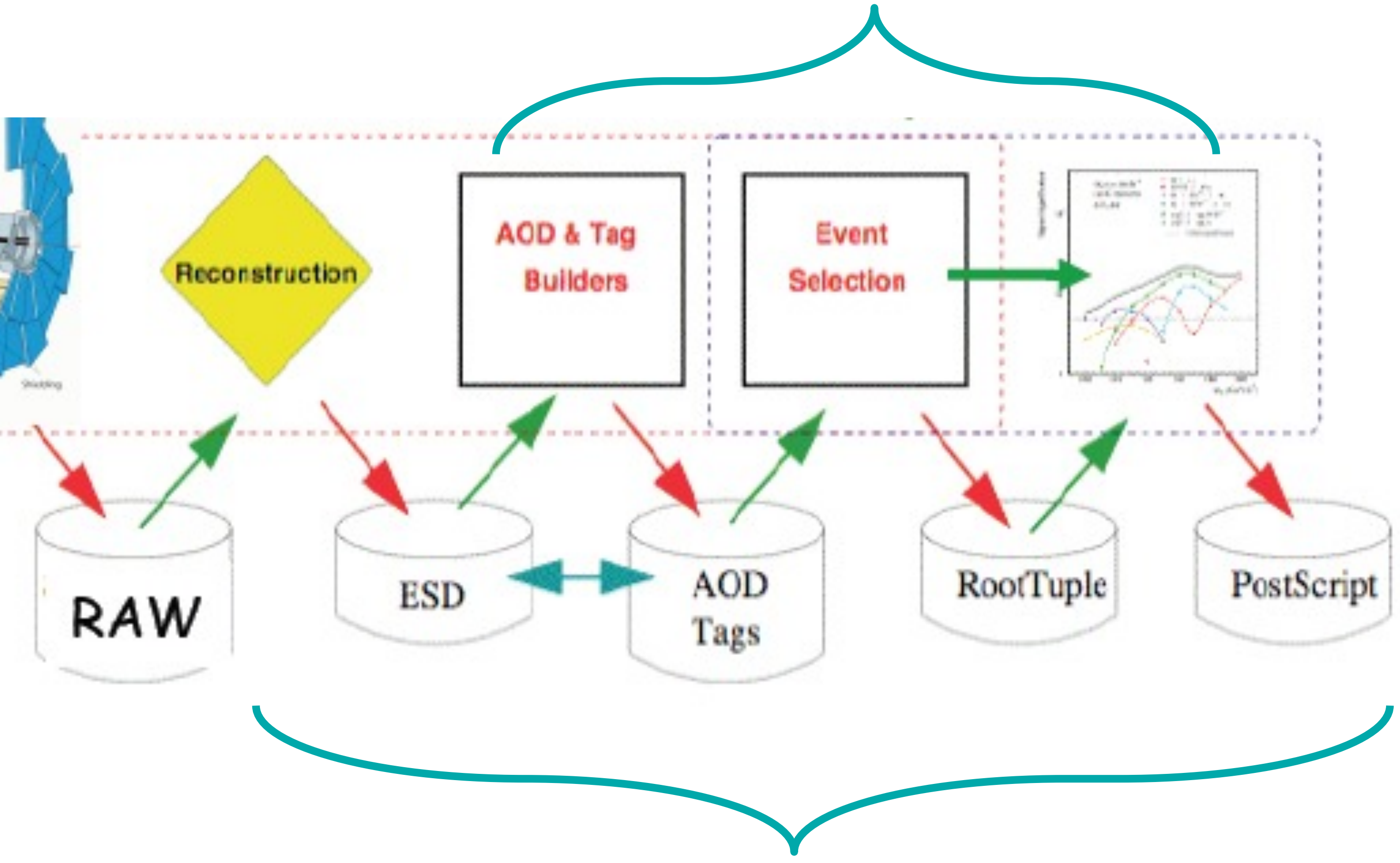
```
{
//   example
of macro to
read data
from an ascii
file and
//   create a
```

- GUI windows, buttons, menus
- Command line CINT (C++ interpreter)
- Macros, applications, libraries (C++ compiler and interpreter)

ROOT Application Domains

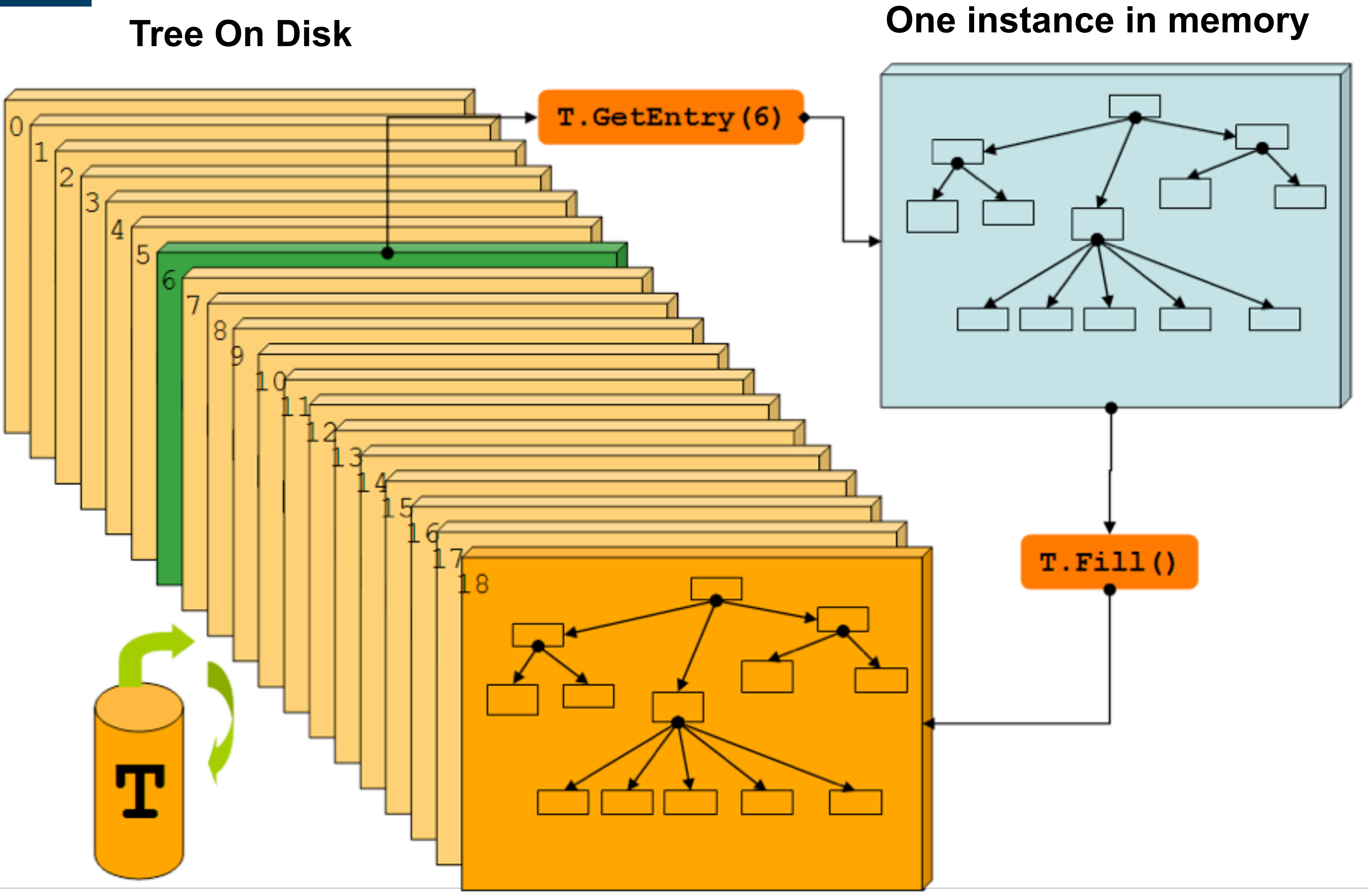
General Framework

Data Analysis & Visualization



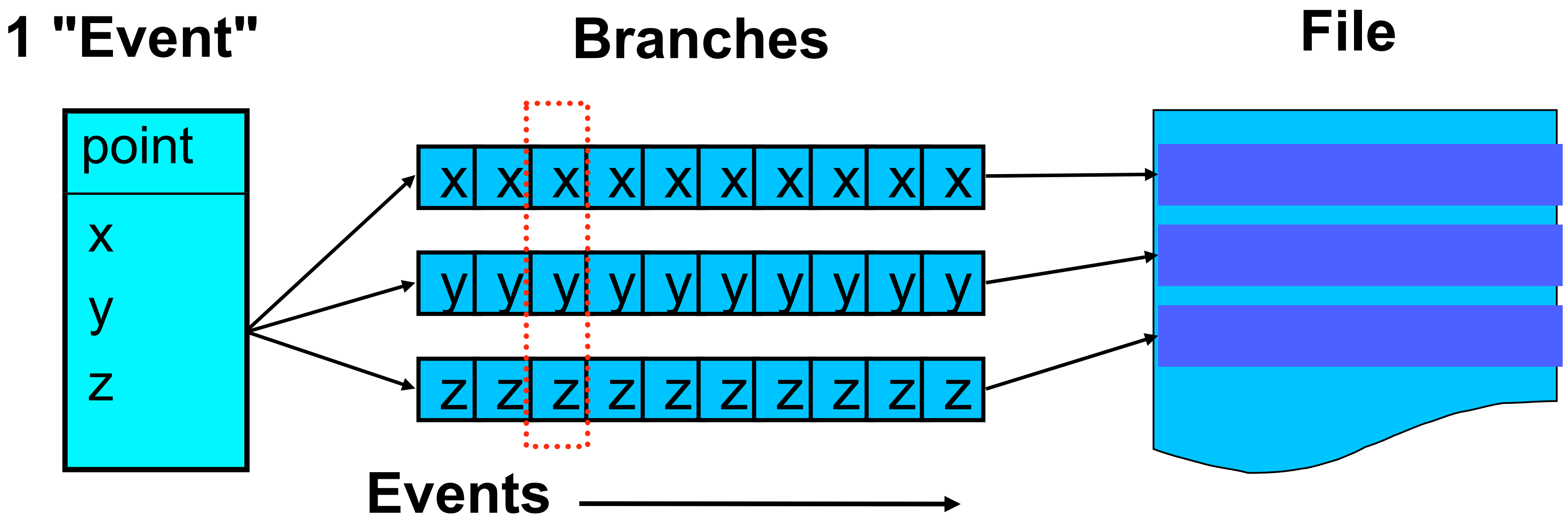
Data Storage: Local, Network

Stored Trees vs. Memory



Trees: Split Mode

- The tree is partioned in branches
  - Each class member is a branch (in split mode)
  - When reading a tree, certain branches can be switched off
    → speed up of analysis when not all data is needed



1 "Event"

Branches

File

point

x
y
z

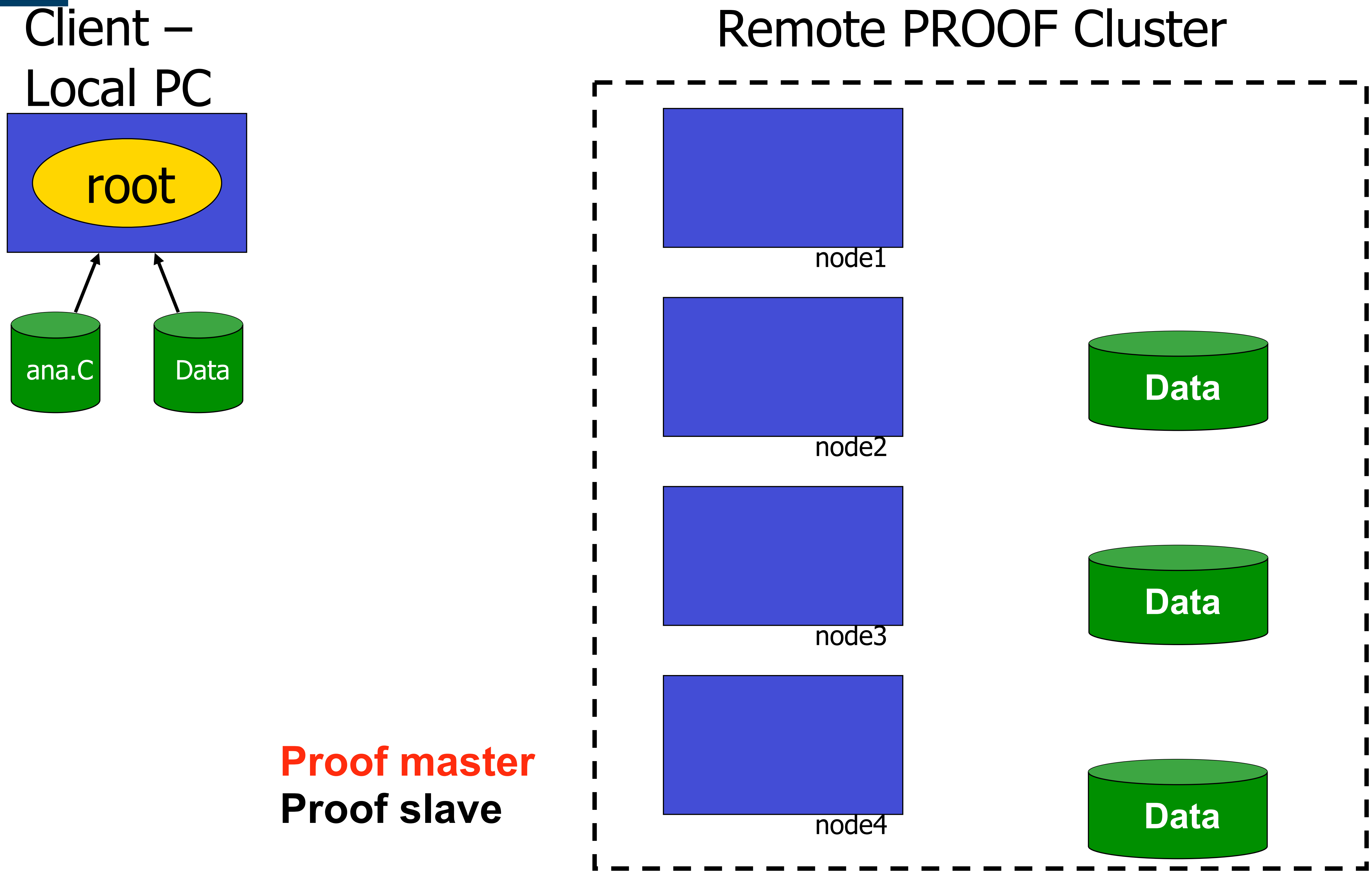
x x x x x x x x x x

y y y y y y y y y y

z z z z z z z z z z

Events

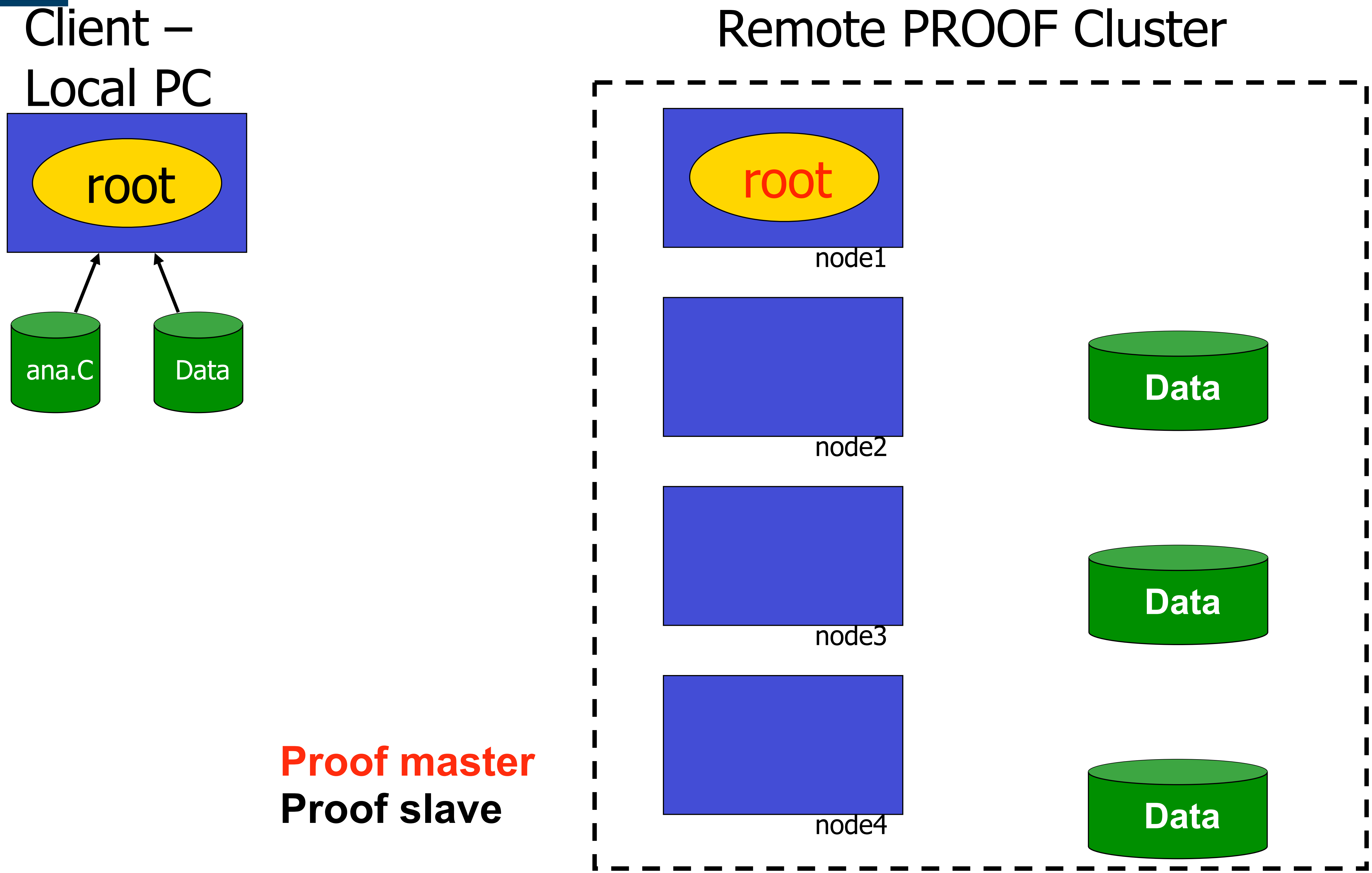
PROOF Schema



Proof master

Proof slave

PROOF Schema



Client –
Local PC

root

ana.C

Data

Remote PROOF Cluster

root

node1

node2

node3

node4

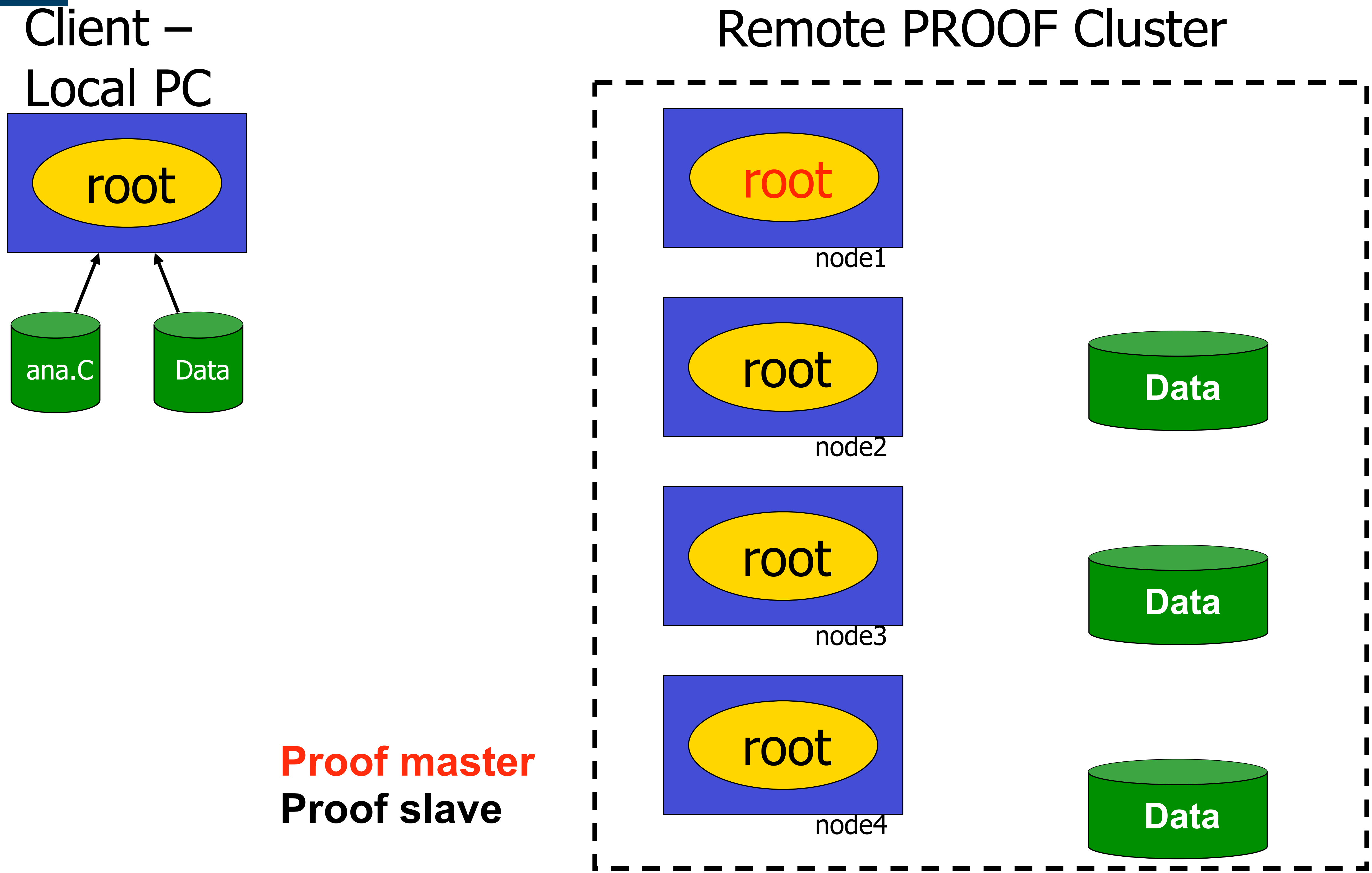
Data

Data

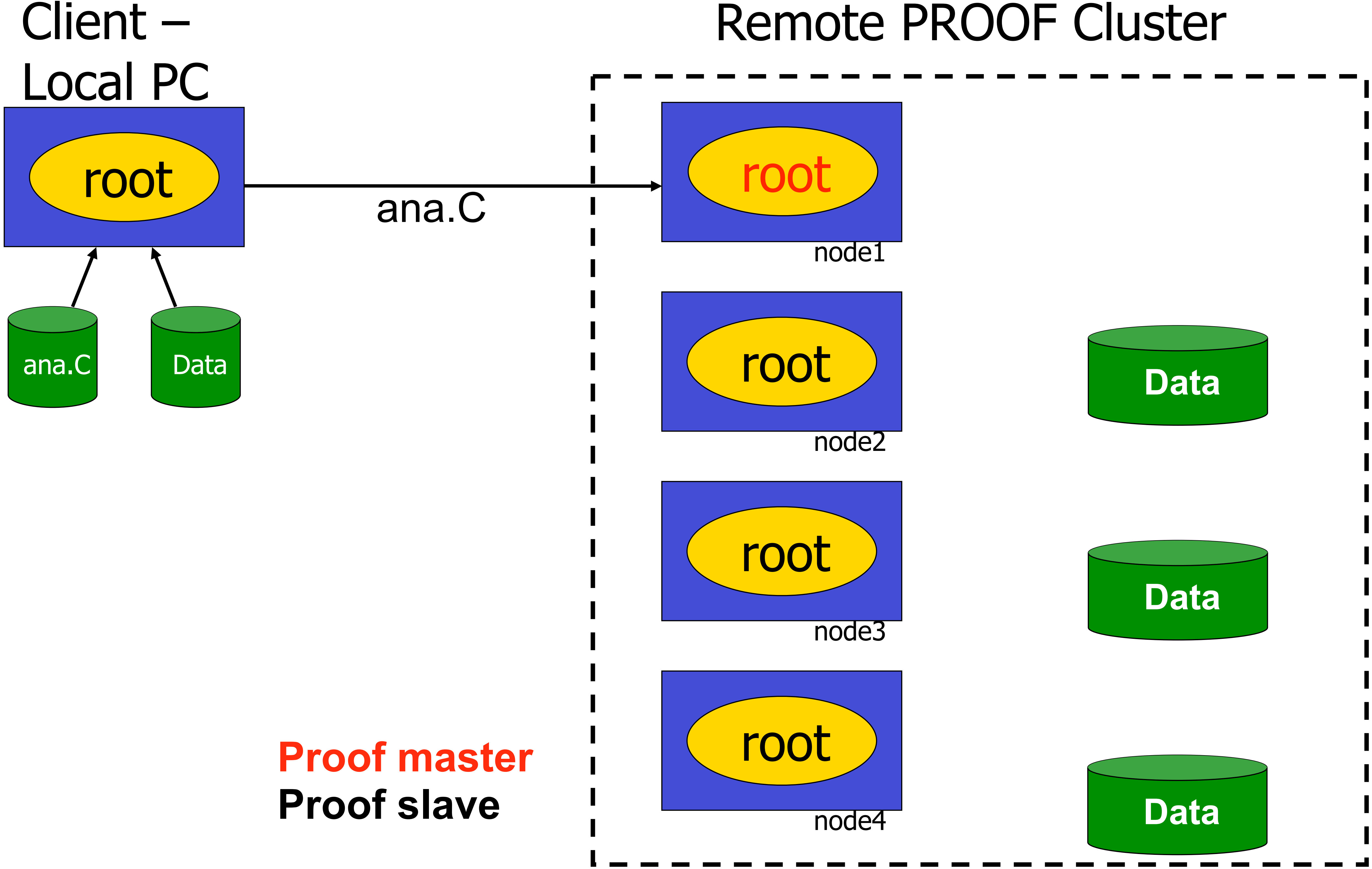
Data

Proof master
Proof slave

PROOF Schema



PROOF Schema

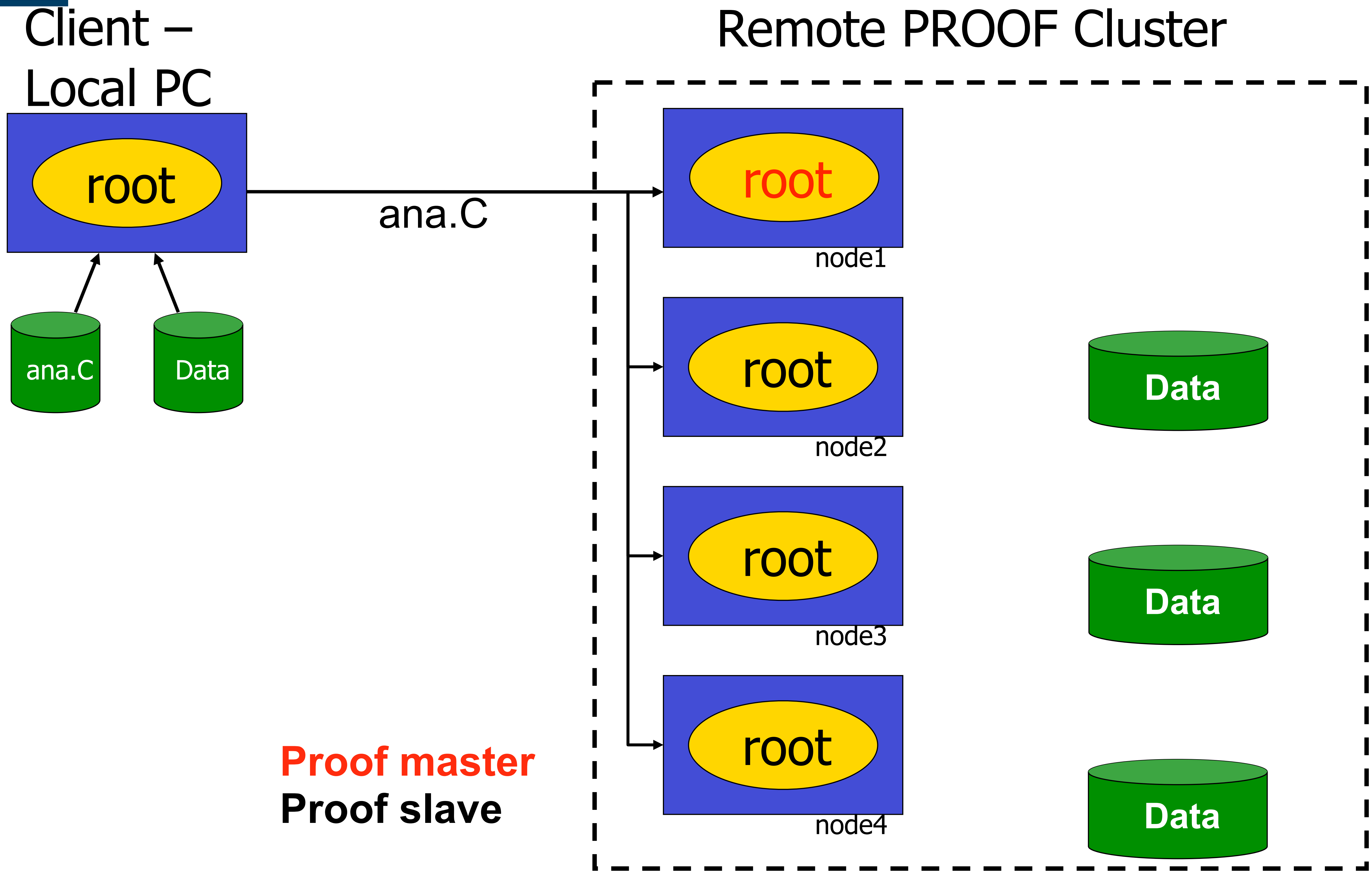


Client – Local PC

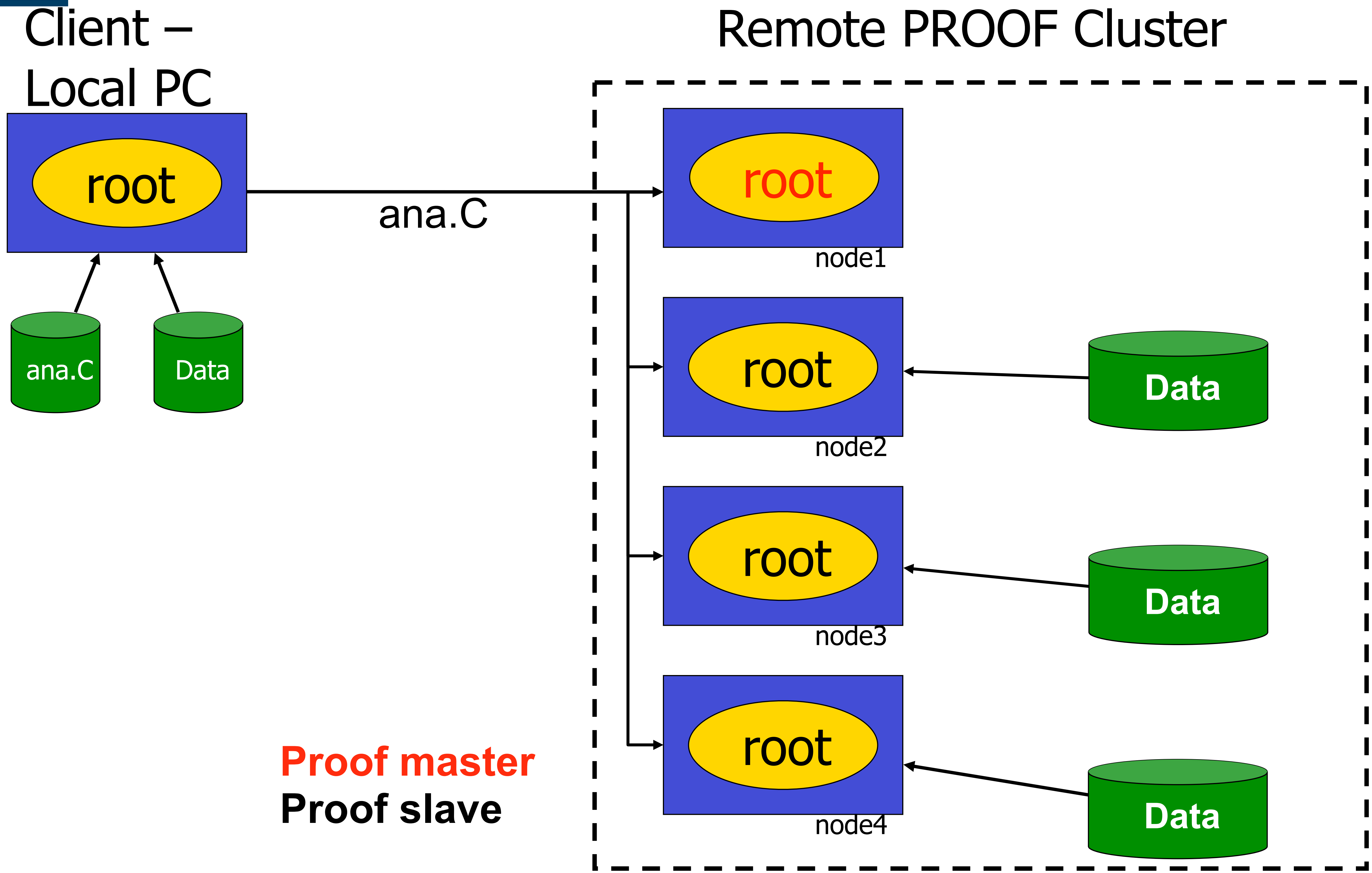
Remote PROOF Cluster

Proof master
Proof slave

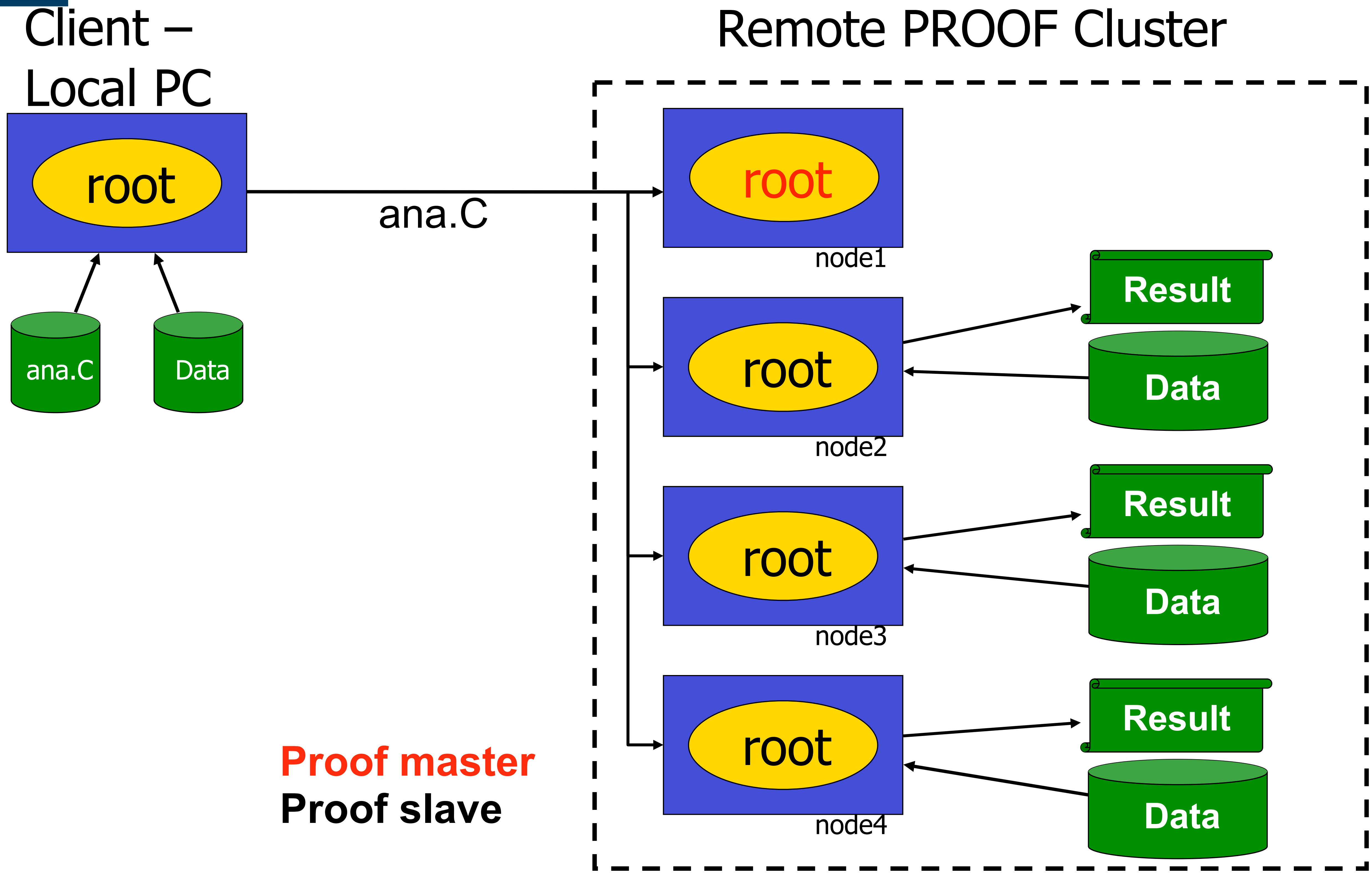
PROOF Schema



PROOF Schema

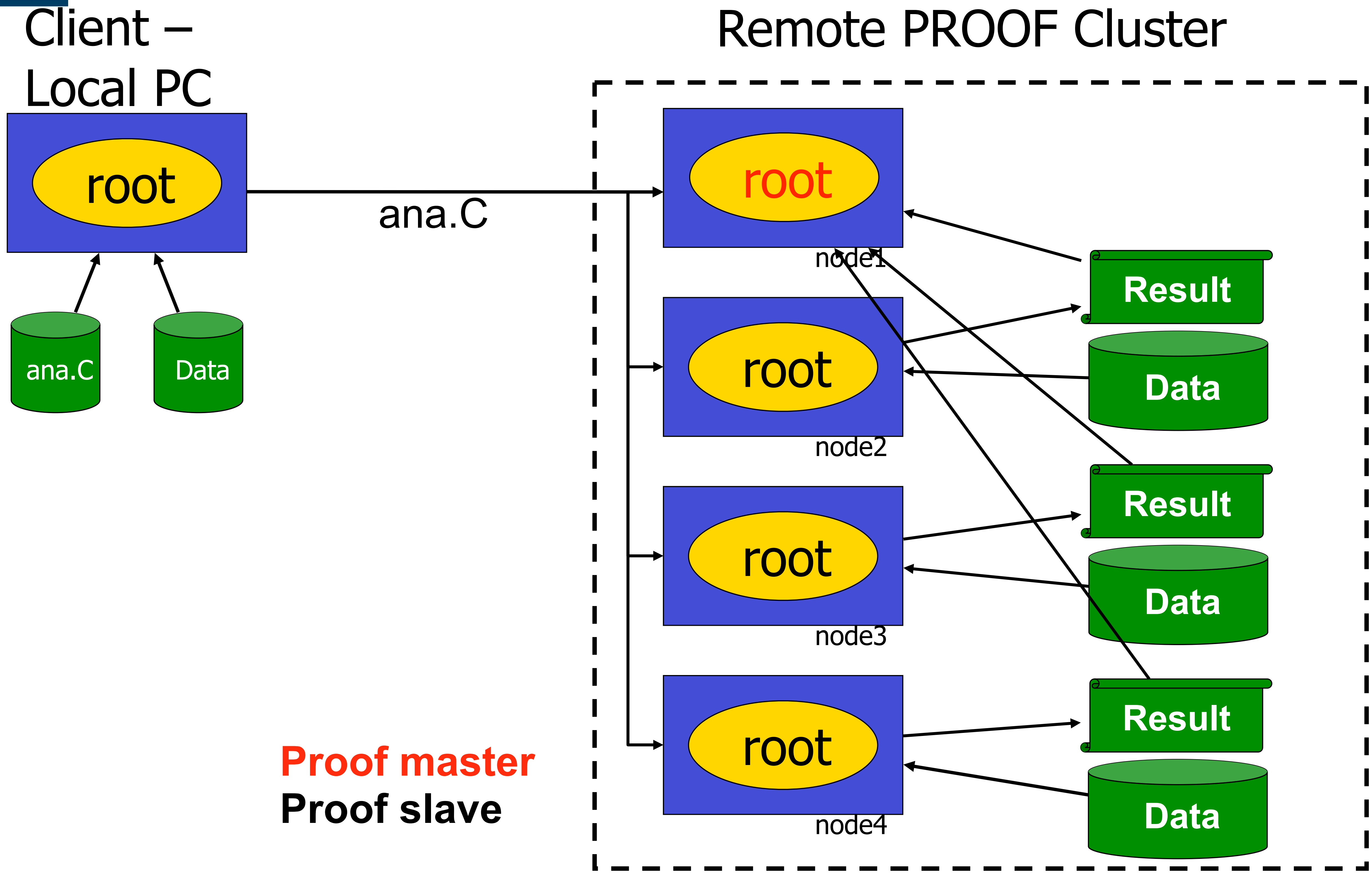


PROOF Schema



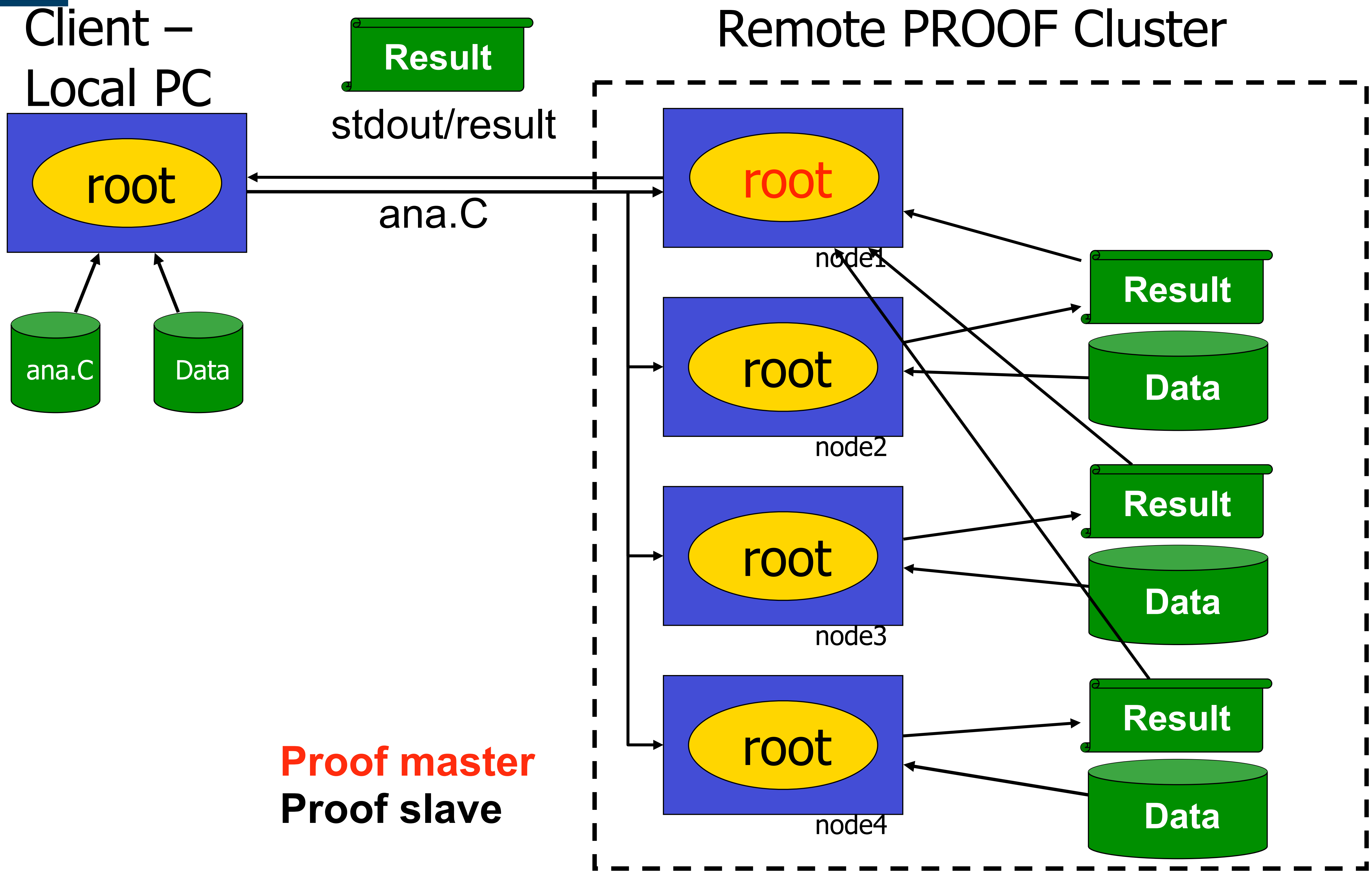
Proof master
Proof slave

PROOF Schema

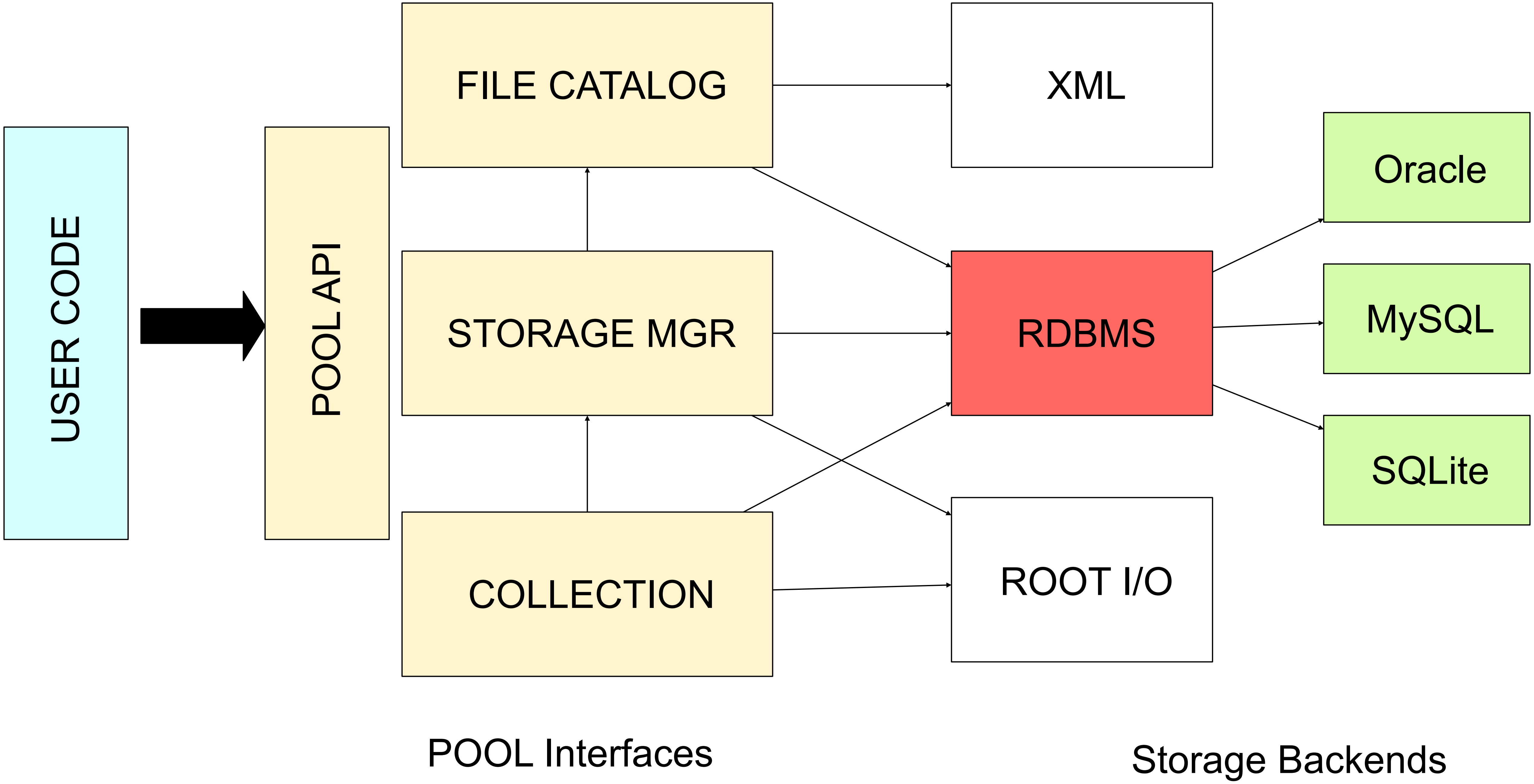


Proof master
Proof slave

PROOF Schema

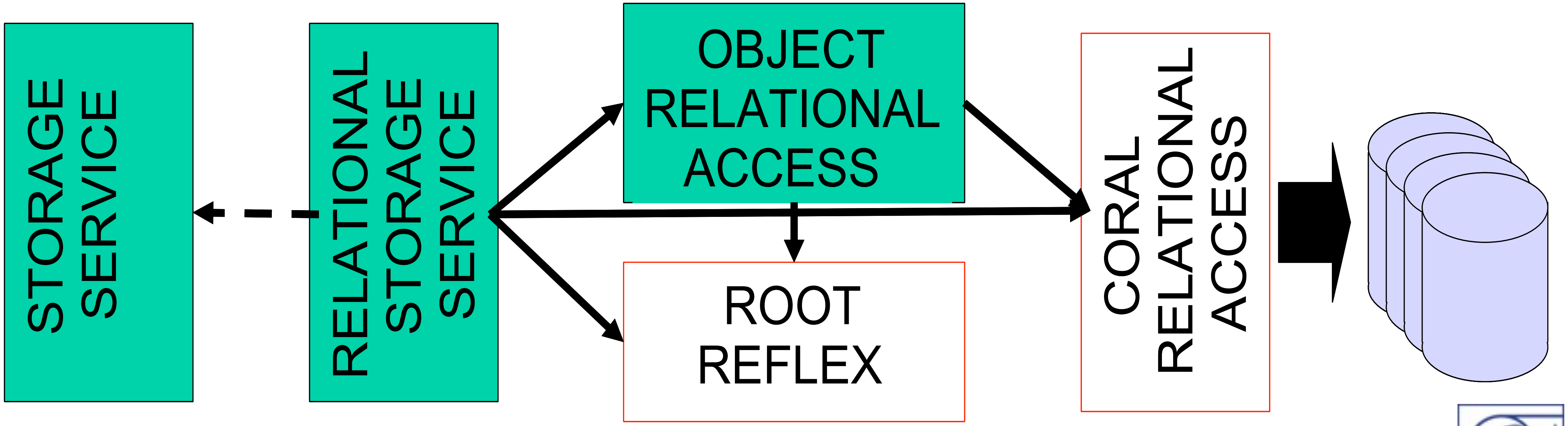


POOL Architecture



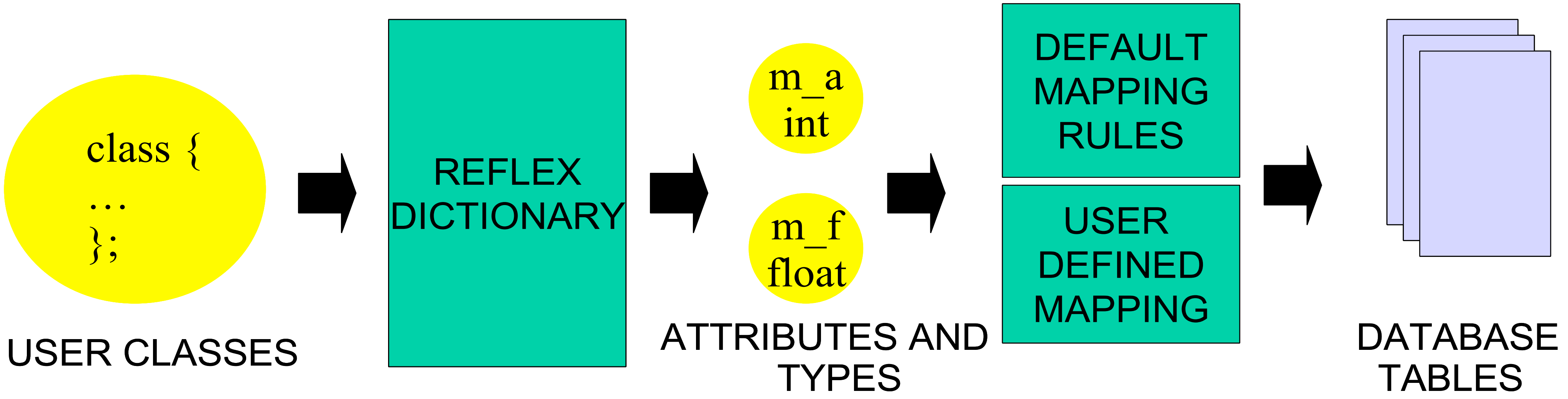
POOL Relational back-end (ORA)

- A POOL Storage Manager backend for **Relational Databases.**
- Users access data through the **same POOL interface** as for the ROOT backend.
- Implementation and database access based on **CORAL** package - **independent from the RDBMS technology**
  - Oracle, MySQL, SQLite and FronTier (Web cache) supported as plugins.
- Supports **writing** and **reading** of arbitrary C++ data structures into/from relational tables.

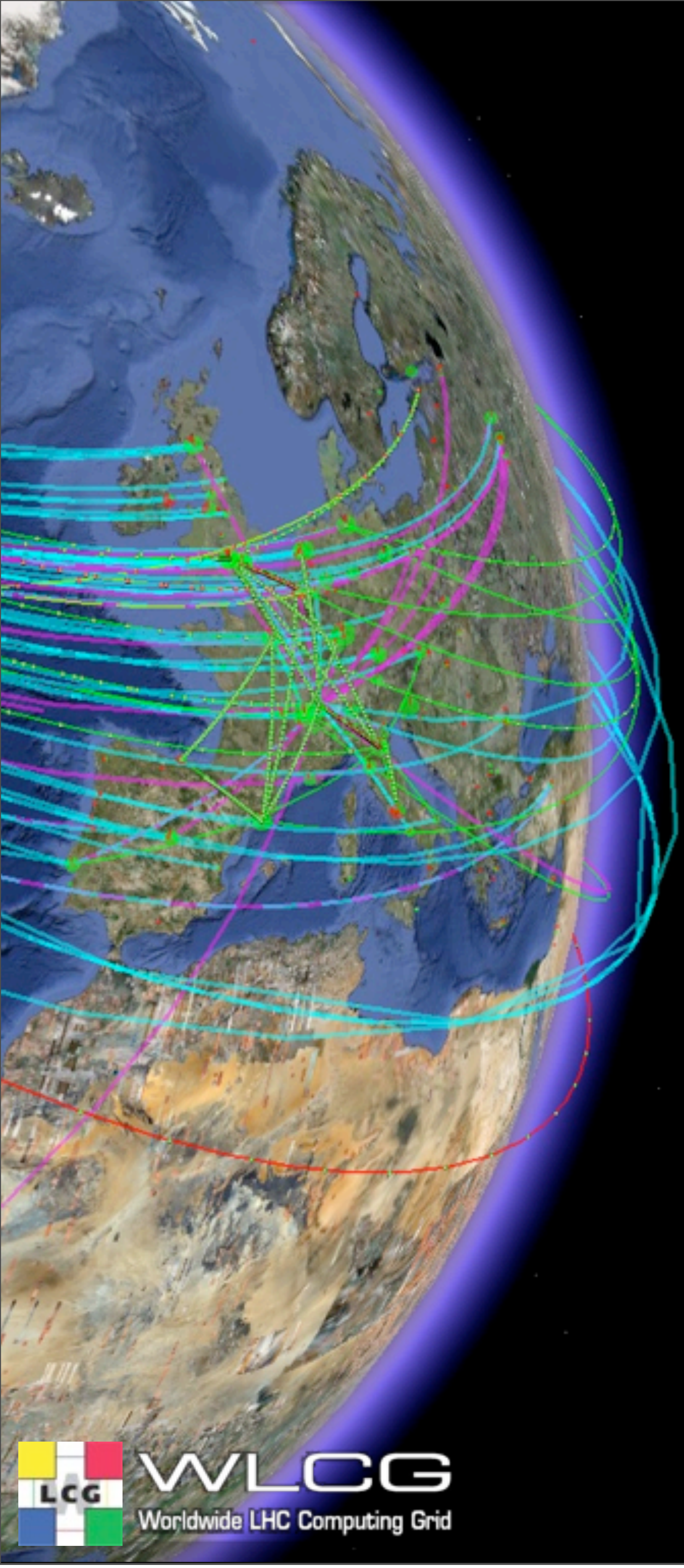


Object storing into RDBMS

- How to map C++ classes ↔ RDBMS tables?
  - Objects need an **unique identifier** (persistent address).
  - **Reflex dictionary** necessary to decompose the class hierarchy in simple elements (primitives/strings).
  - **Tables** and **columns** are associated to the fields with simple 'default' rules.
  - Mapping information is stored in the database.



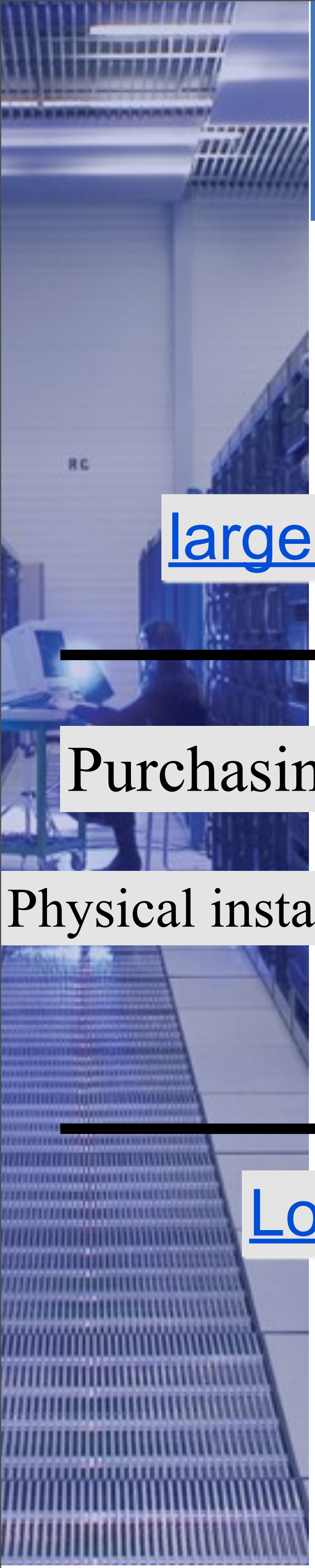
- User-defined mapping:
  - A **command line tool** based on a XML driver generates the mapping altering the default rules according to the user specifications.



What fabric is required for HEP computing tasks?

Worldwide LHC Computing Grid

CERN Computing Fabric

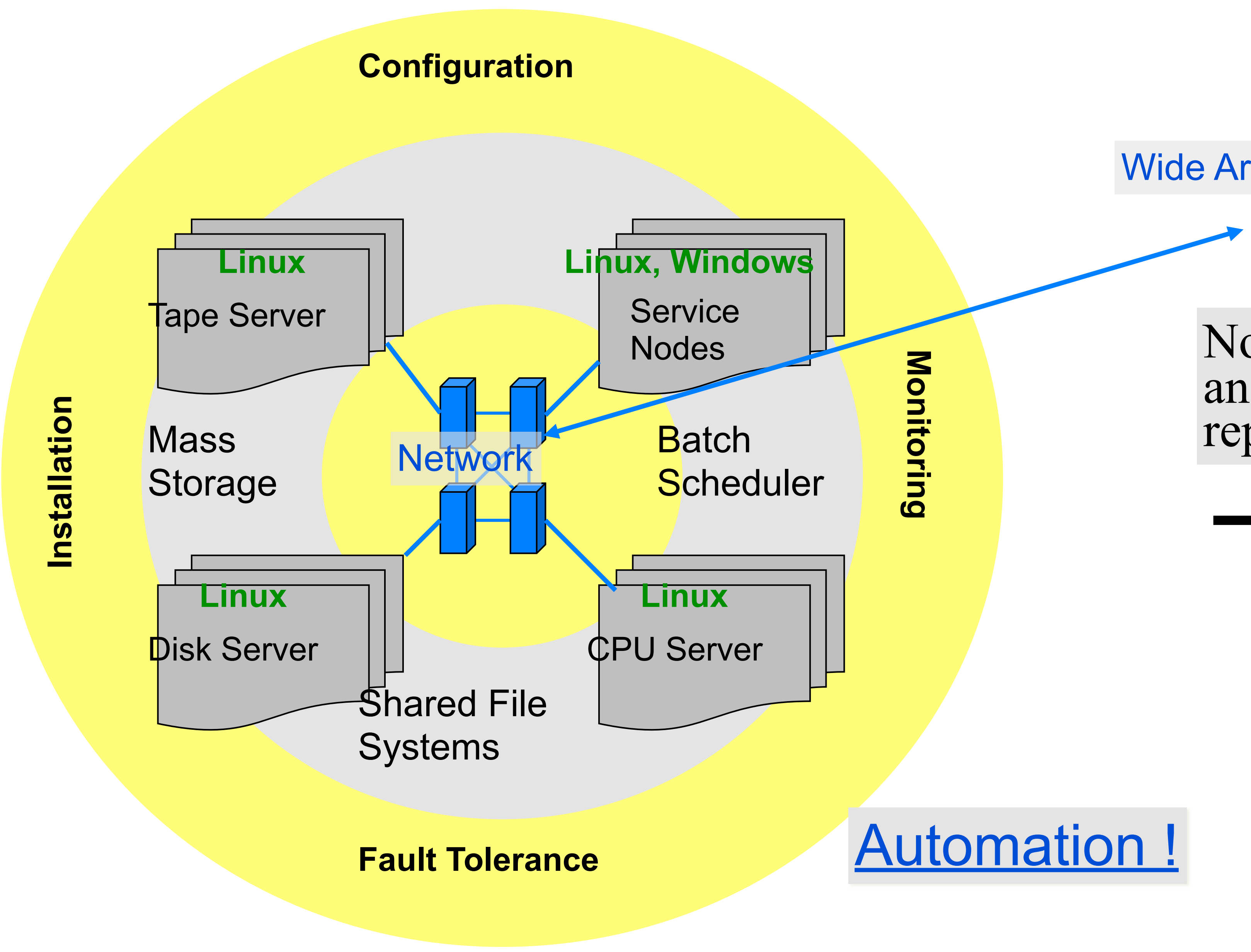


large scale !

Purchasing

Physical installation

Logistic !



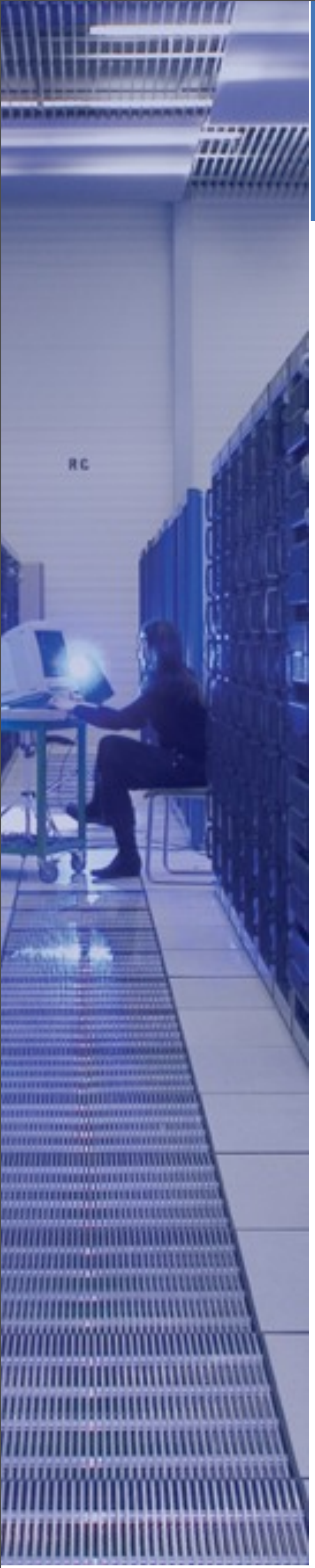
Node repair and replacement

Automation !

Electricity

Space

Cooling



Logistics at CERN

- Today we have about 7000 PC's installed in the center

- Assume 3 years lifetime for the equipment
  → Key factors = power consumption, performance, reliability

- Experiment requests require investments of ~ 15 MCHF/year for new PC hardware and infrastructure

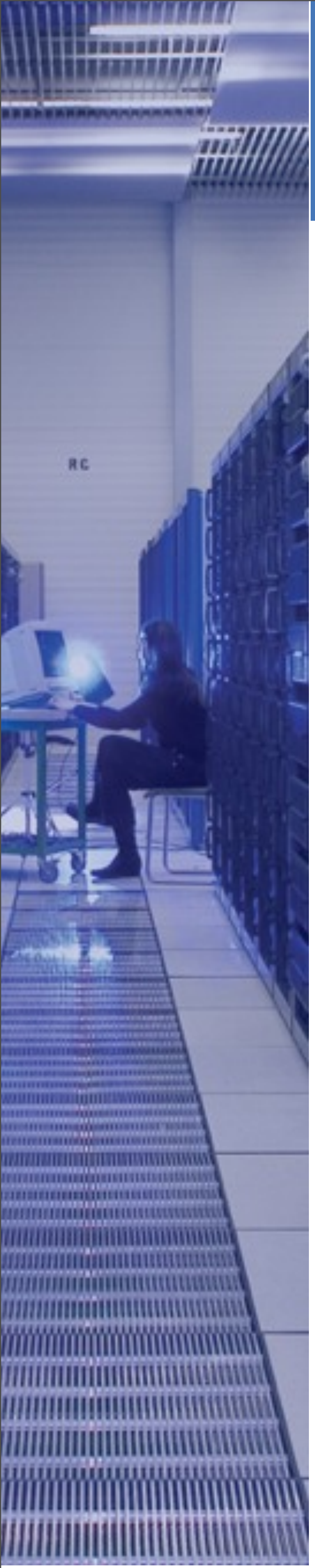
Infrastructur and operation setup needed for :

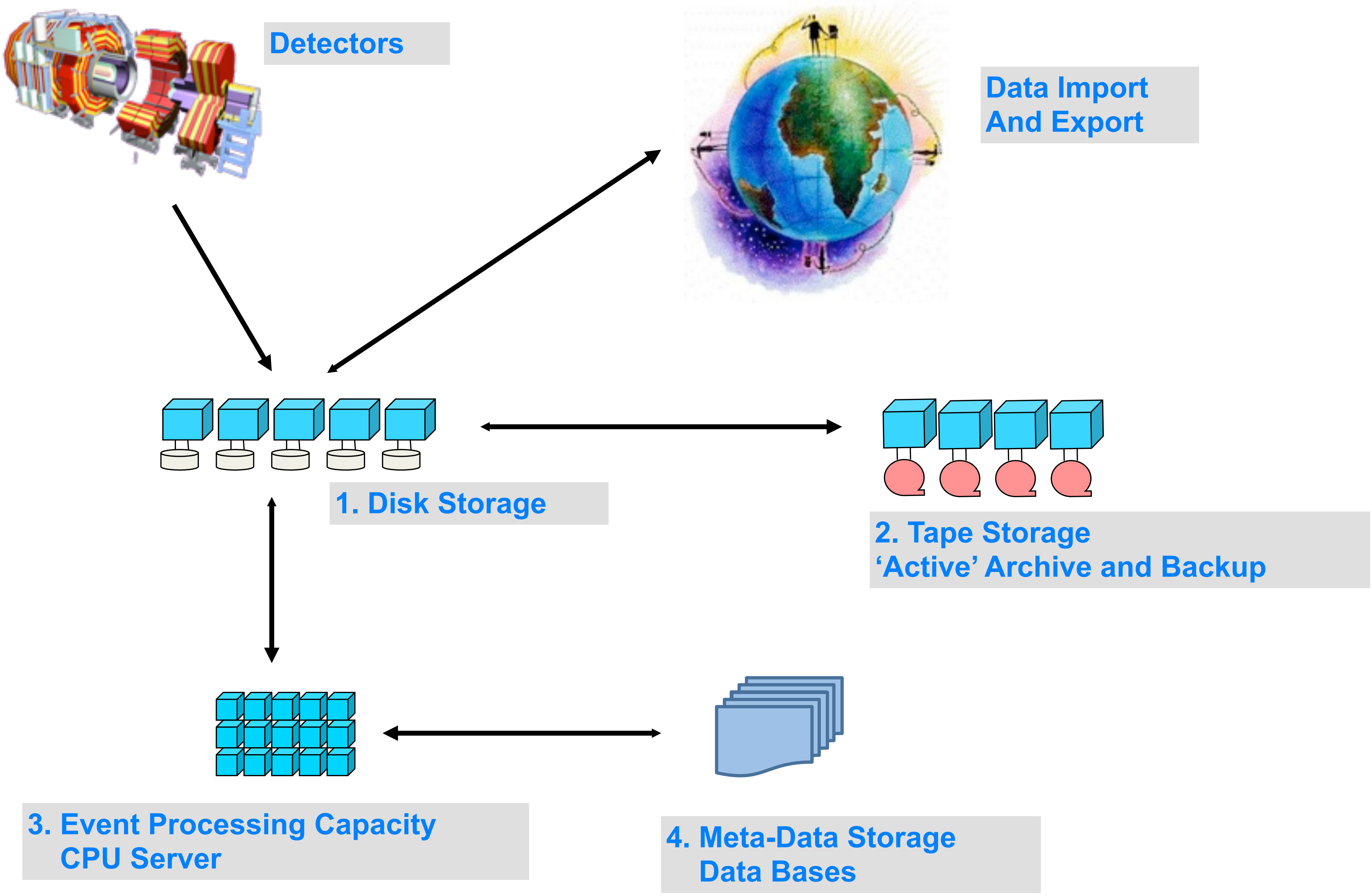
- ~2000 nodes installed per year and ~2000 nodes removed per year

- Installation in racks, cabling, automatic installation, Linux software environment

- Equipment replacement rate, e.g. 2 disk errors per day
  several nodes in repair per week 50 node crashes per day

Functional Computing Units





Detectors

Data Import
And Export

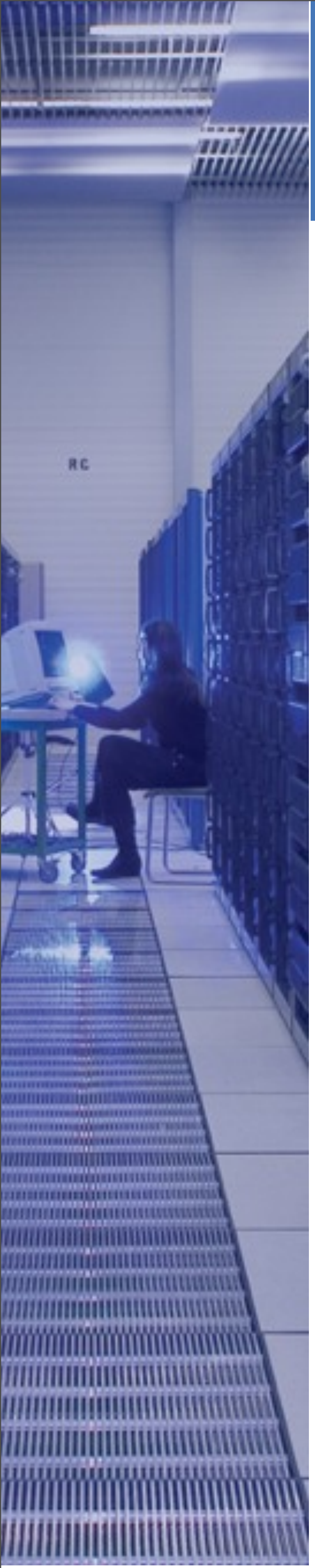
1. Disk Storage

2. Tape Storage
'Active' Archive and Backup

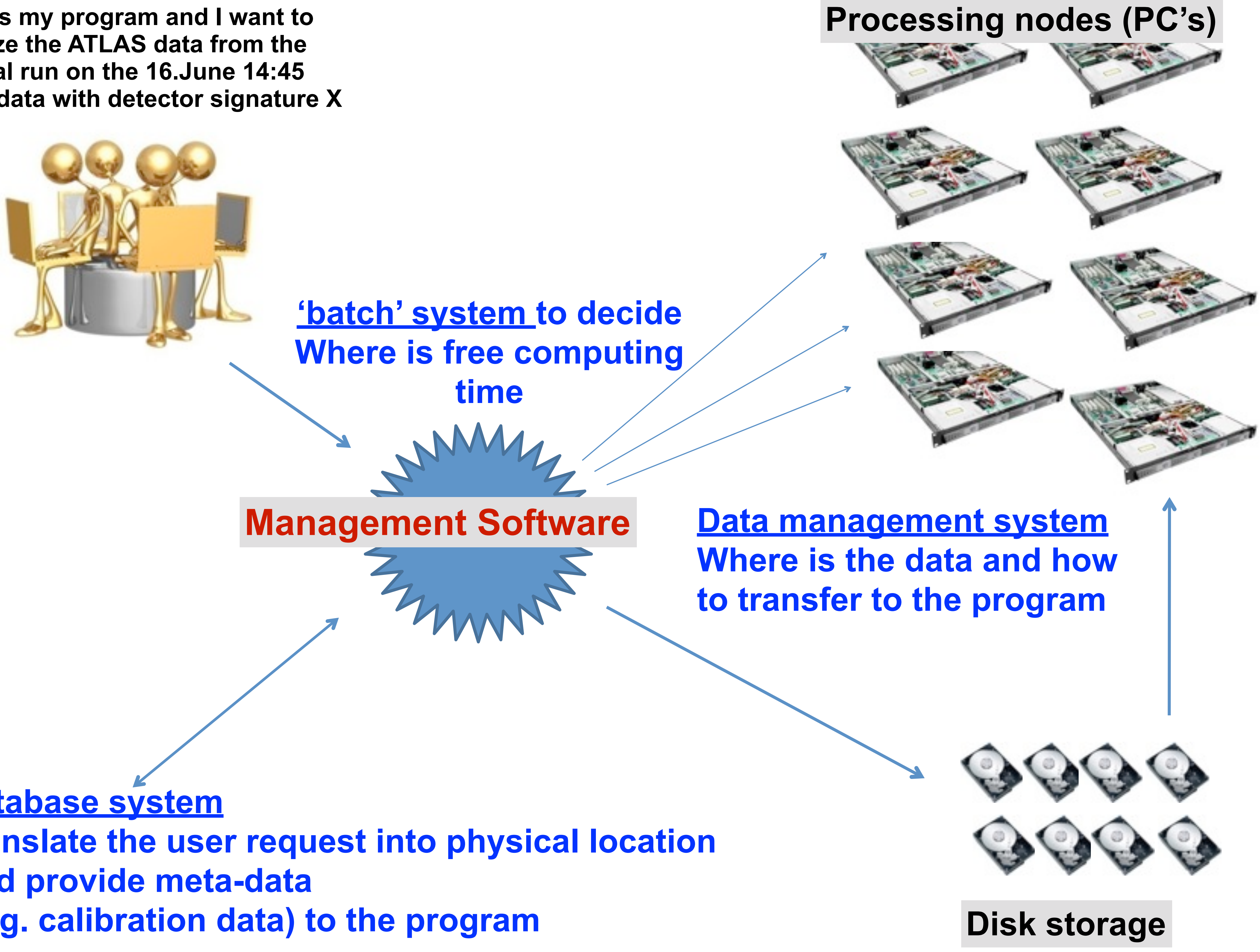
3. Event Processing Capacity
CPU Server

4. Meta-Data Storage
Data Bases

Data and control flow I



Here is my program and I want to analyze the ATLAS data from the special run on the 16.June 14:45
or all data with detector signature X



Processing nodes (PC's)

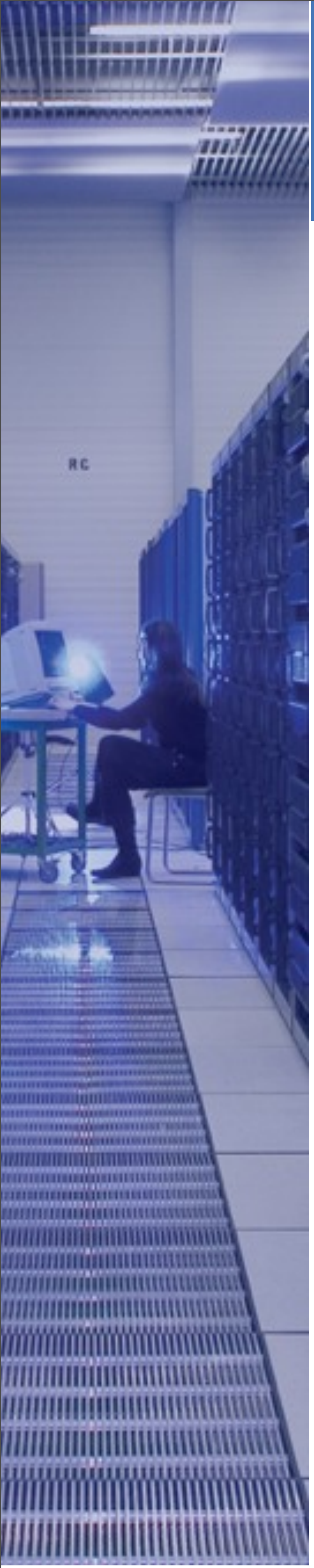
'batch' system to decide
Where is free computing time

Management Software

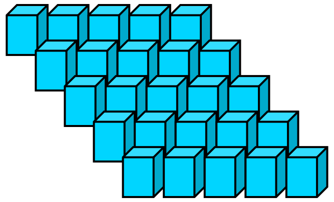
Data management system
Where is the data and how to transfer to the program

Database system
Translate the user request into physical location
And provide meta-data
(e.g. calibration data) to the program

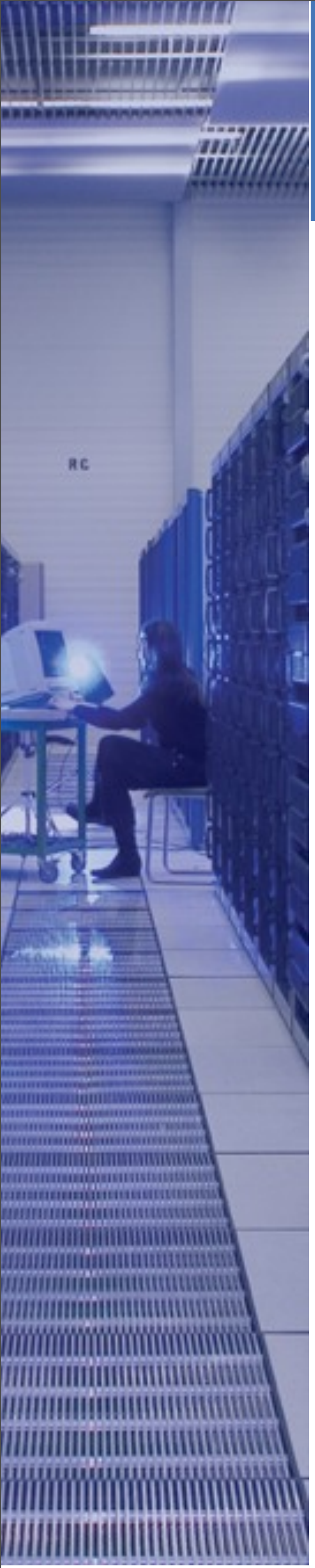
Disk storage



Processing Facility at CERN: Lxbatch

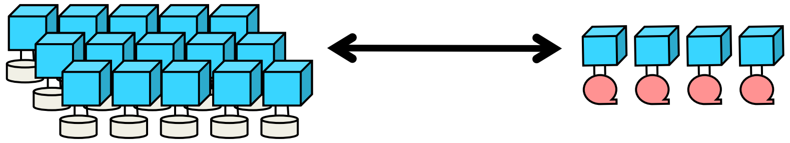


- Jobs are submitted from Lxplus or channeled through GRID interfaces world-wide
- Today about 4000 nodes with 31000 processors (cores)
- About 150000 user jobs are run per day
- Reading and writing > 1 PB per day
- Uses LSF as a management tool to schedule the various jobs from a large number of users.
- Expect a resource growth rate of ~30% per year

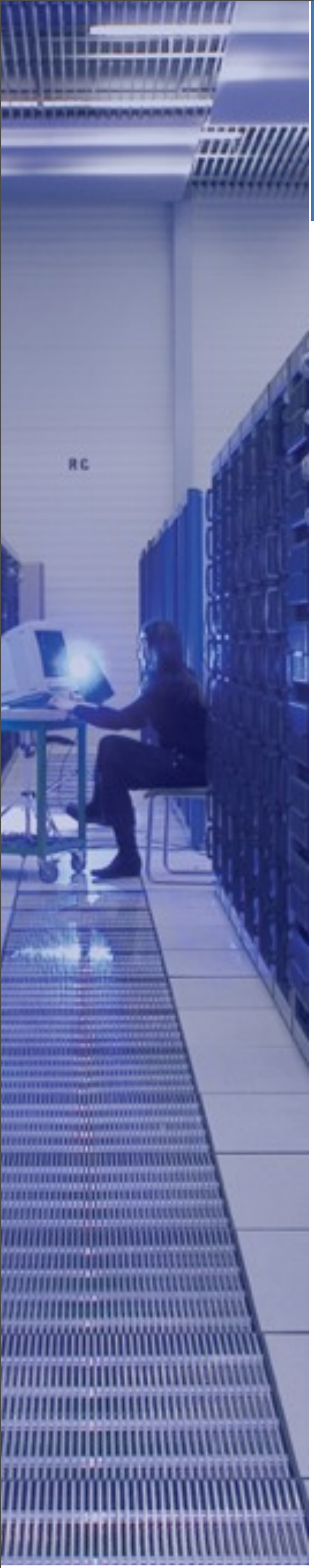


Mass Storage at CERN

CERN IT Department

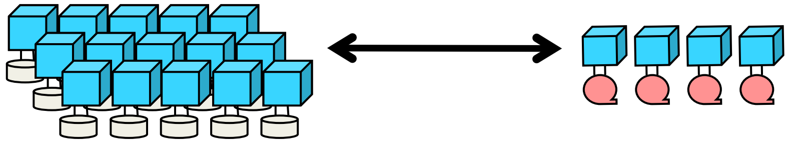


- Large disk cache in front of a long term storage tape system
- 1800 disk servers with 21 PB usable capacity
- Redundant disk configuration, 2-3 disk failures per day needs to be part of the operational procedures
- Logistics again : need to store all data forever on tape 20 PB storage added per year, plus a complete copy every 4 years (change of technology)
- CASTOR data management system, developed at CERN, manages the user IO requests
- Expect a resource growth rate of ~30% per year

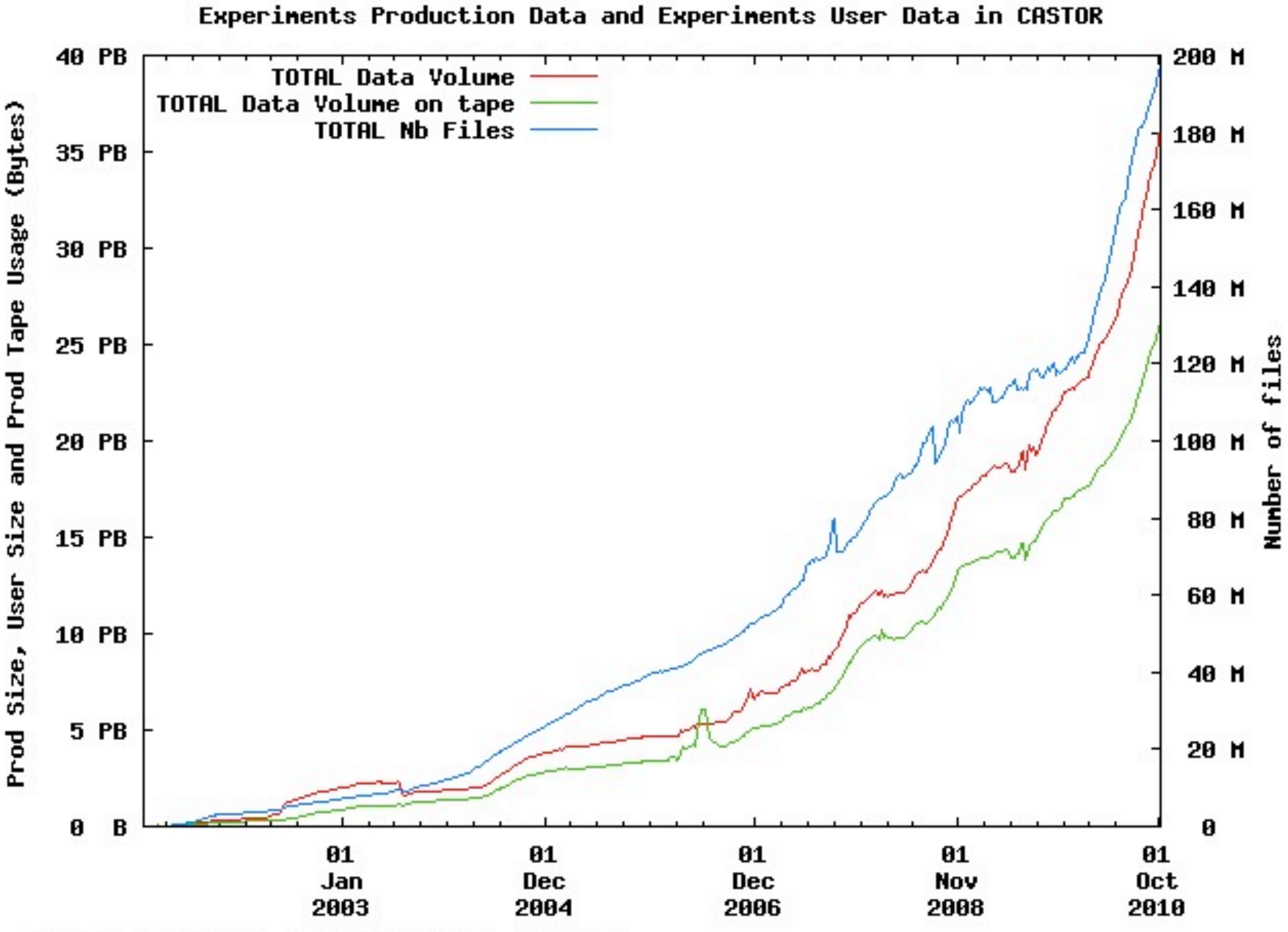


Mass Storage at CERN

CERN IT Department



200 million files

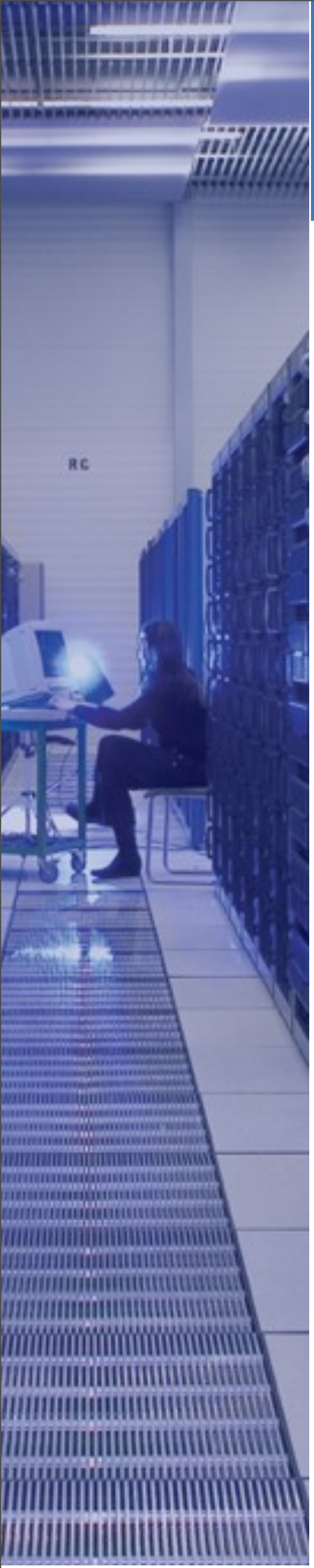


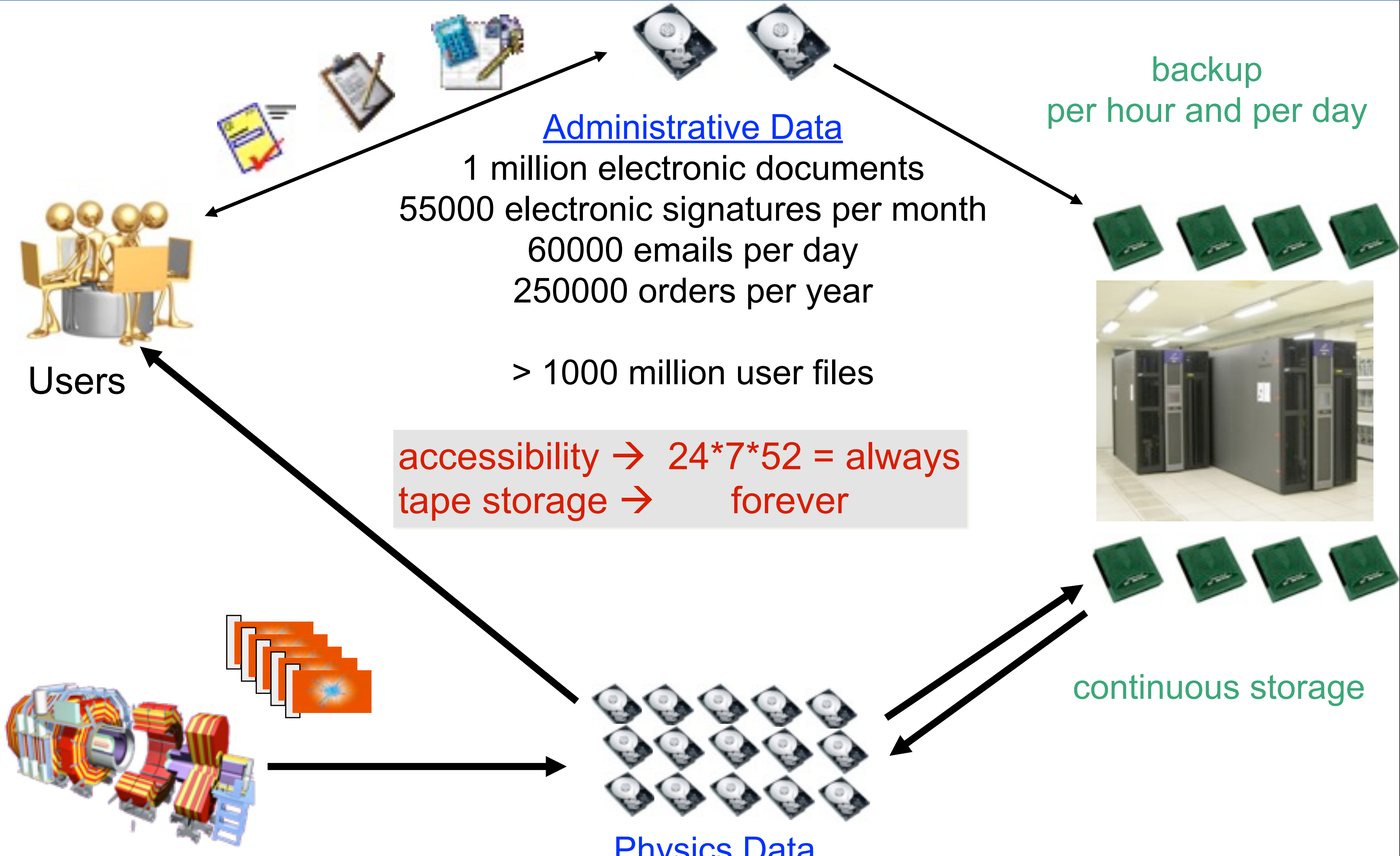
Generated Oct 05, 2010 CASTOR (c) CERN/IT

33 PB of data already today!

Steep growth - expecting some 15 PB/y

Storage





Users

Administrative Data

1 million electronic documents
55000 electronic signatures per month
60000 emails per day
250000 orders per year

> 1000 million user files

accessibility → 24*7*52 = always
tape storage → forever

backup
per hour and per day

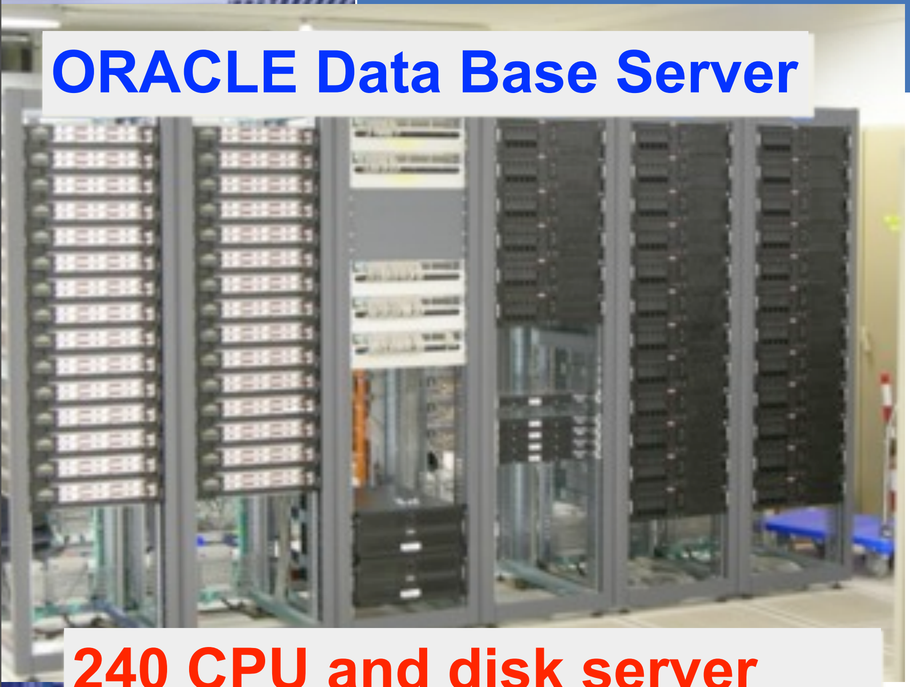
continuous storage

Physics Data

21000 TB and
50 million files per year

CERN Computer Center

ORACLE Data Base Server



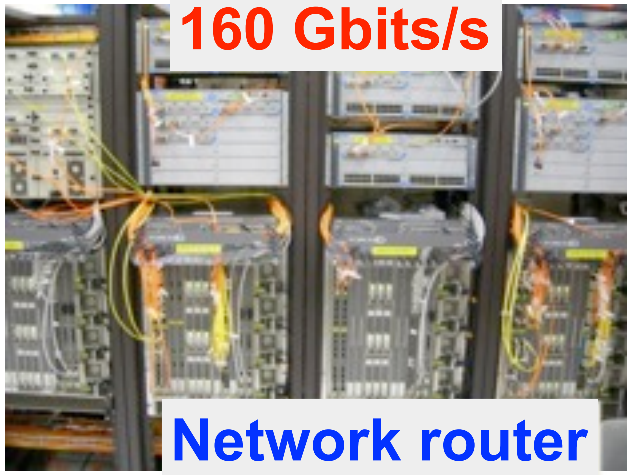
240 CPU and disk server
200 TB

2.9 MW Electricity
and Cooling
2700 m2

CPU Server

31000 processors

160 Gbits/s



Network router

Tape Server
and
Tape Library

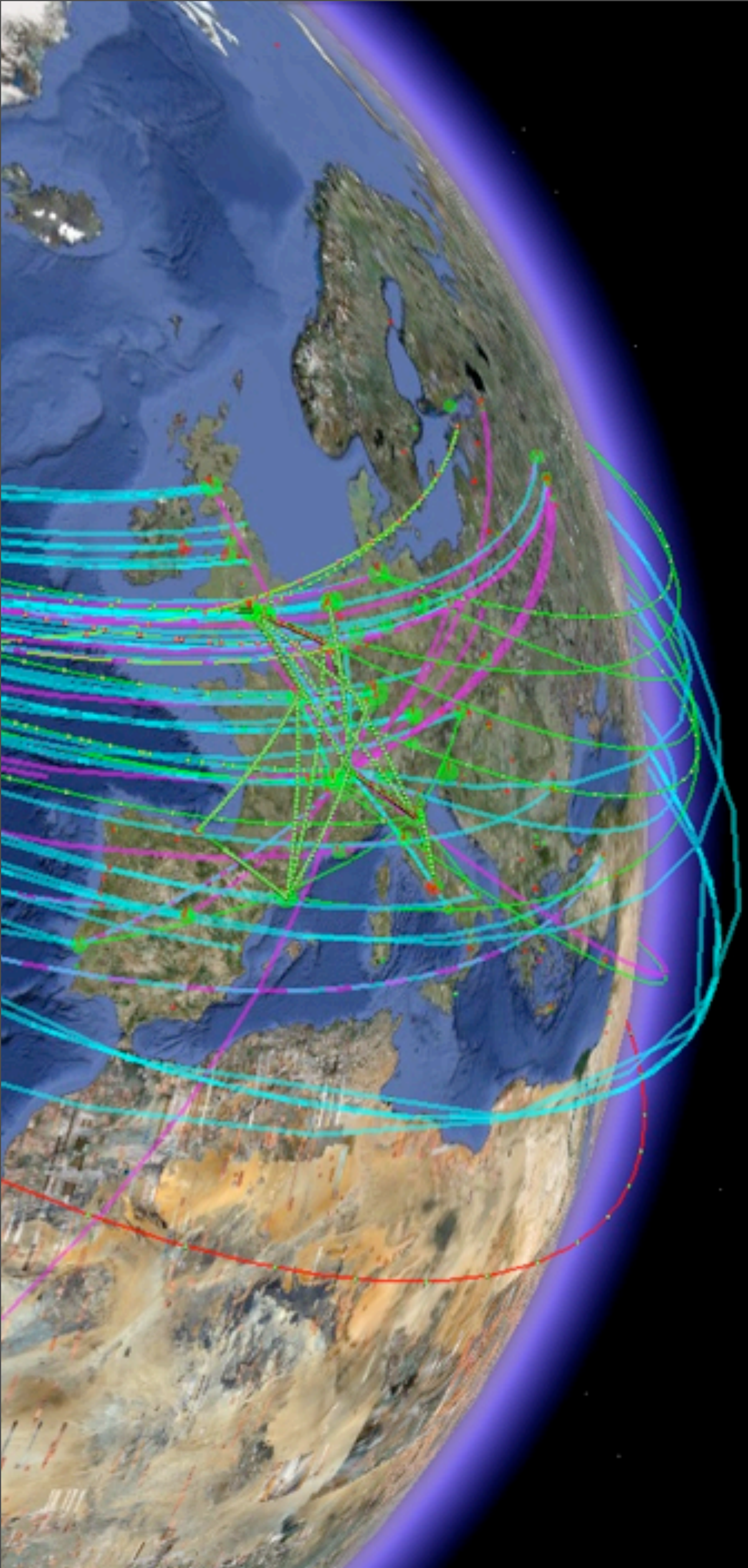
160 tape drives , 50000 tapes
40000 TB capacity

Ch

www.cern.ch/it

Disk Server

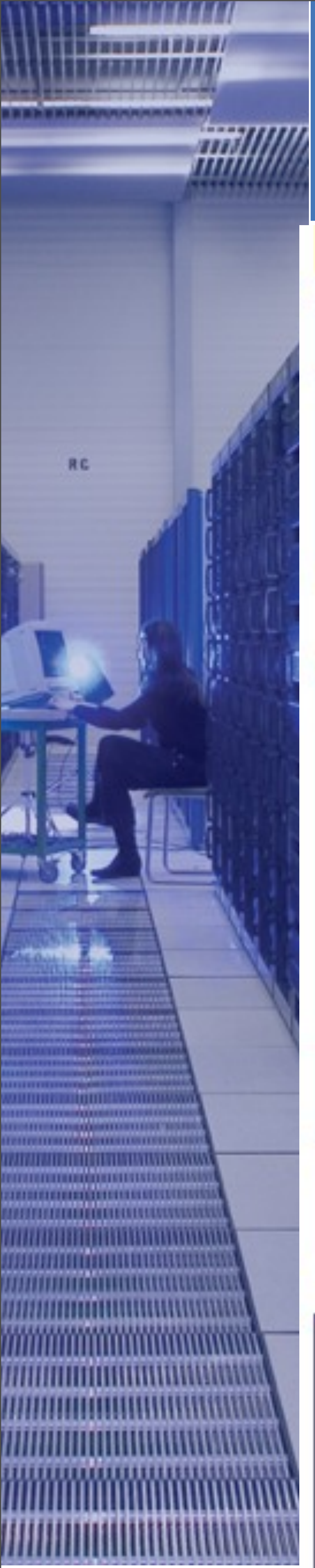
1800 NAS server, 21000 TB
30000 disks



CERN alone is by far not sufficient!

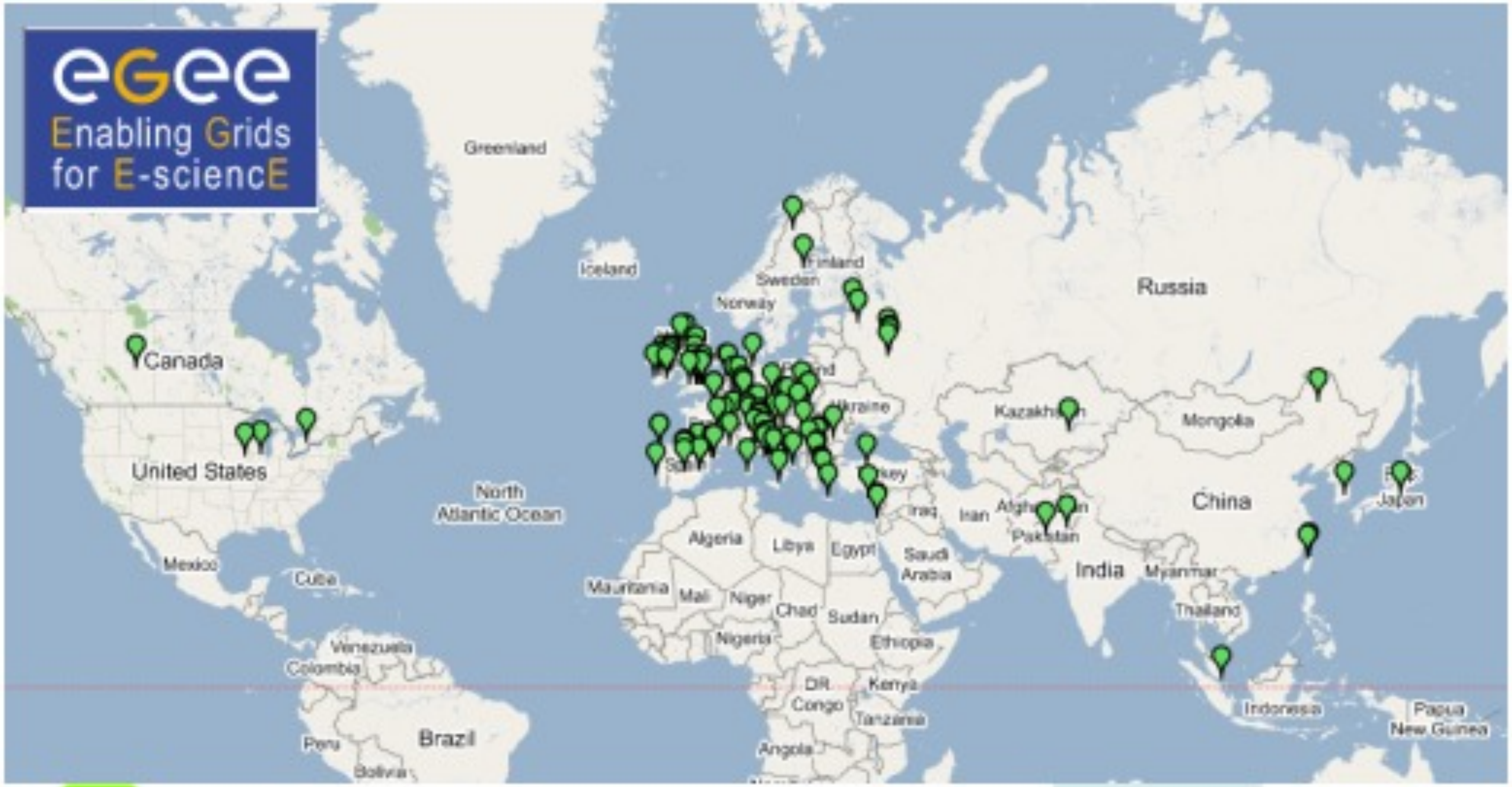
WLCG Worldwide LHC Computing Grid

LHC Computing Grid : LCG



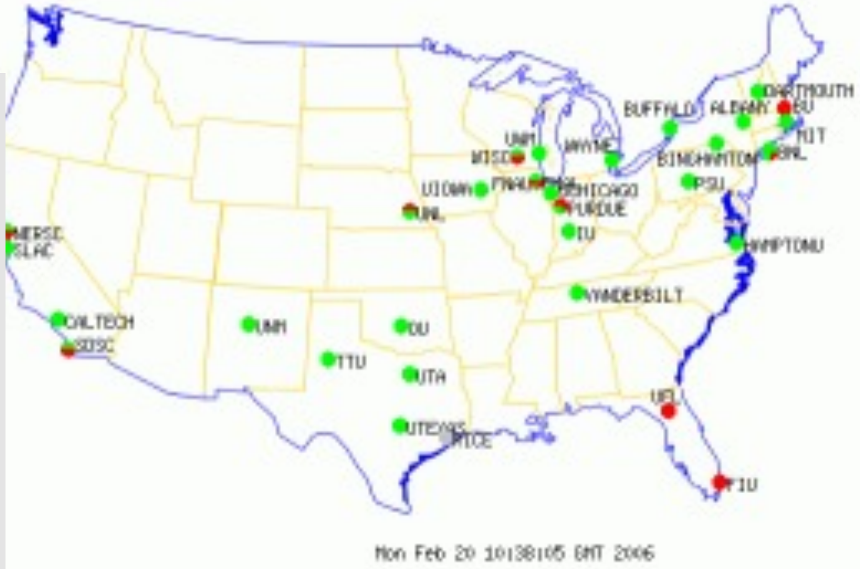
Grid Projects Collaborating in LHC Computing Grid

LCG



NORDUGRID
Grid Solution for Wide Area Computing and Data Handling

Open Science Grid



ALICE

ATLAS

CMS

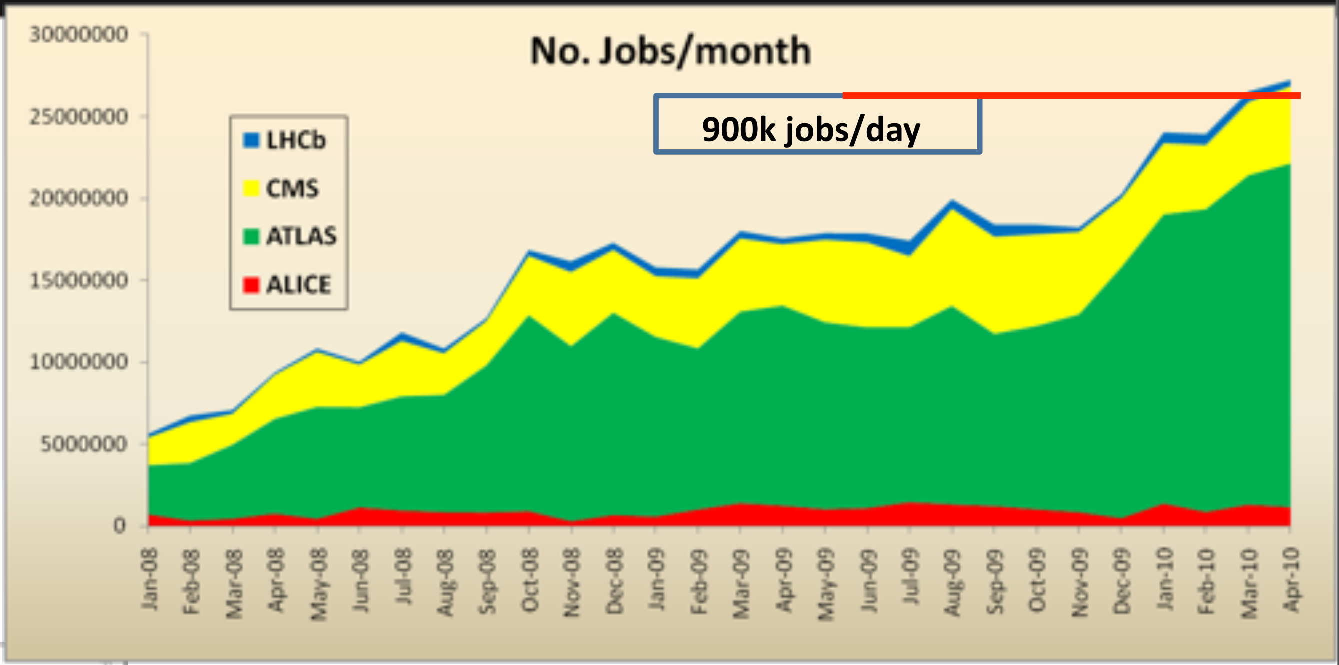
LHCb

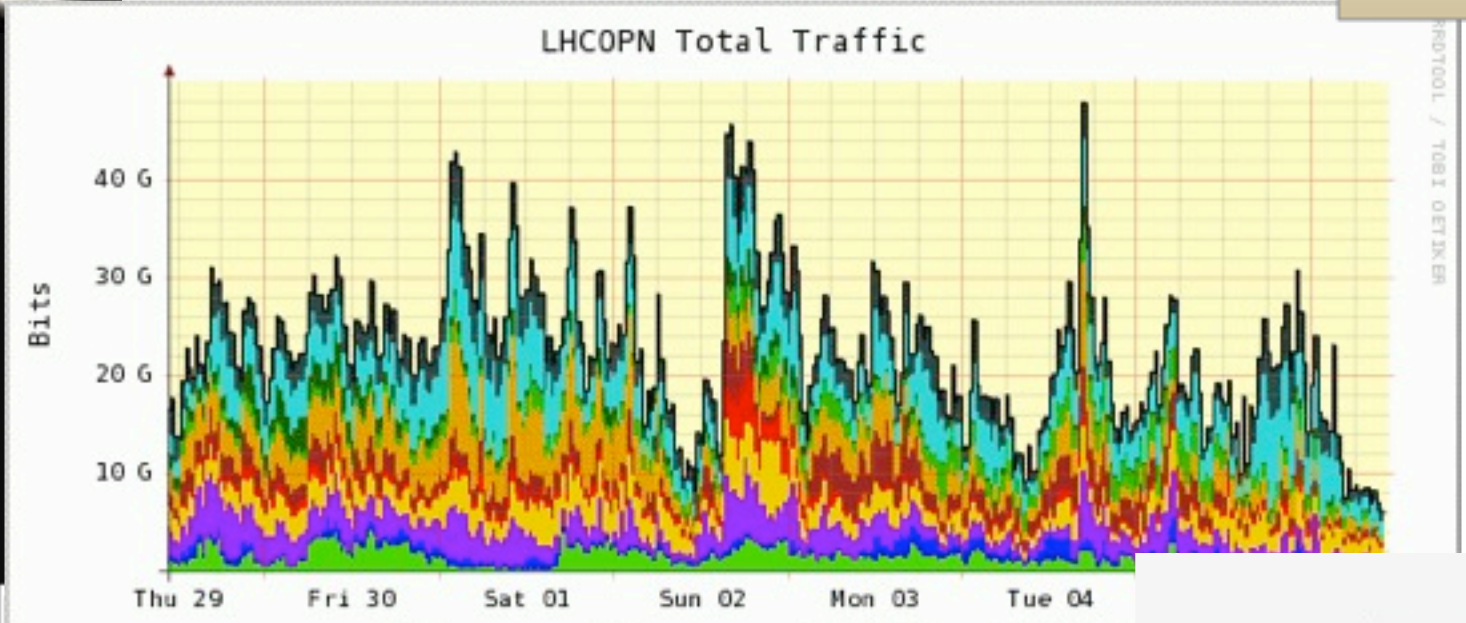
Today

Active sites	:	> 300
Countries involved	:	~ 70
Available processors	:	~ 100000
Available space	:	~ 100 PB

LHC Computing Grid : LCG

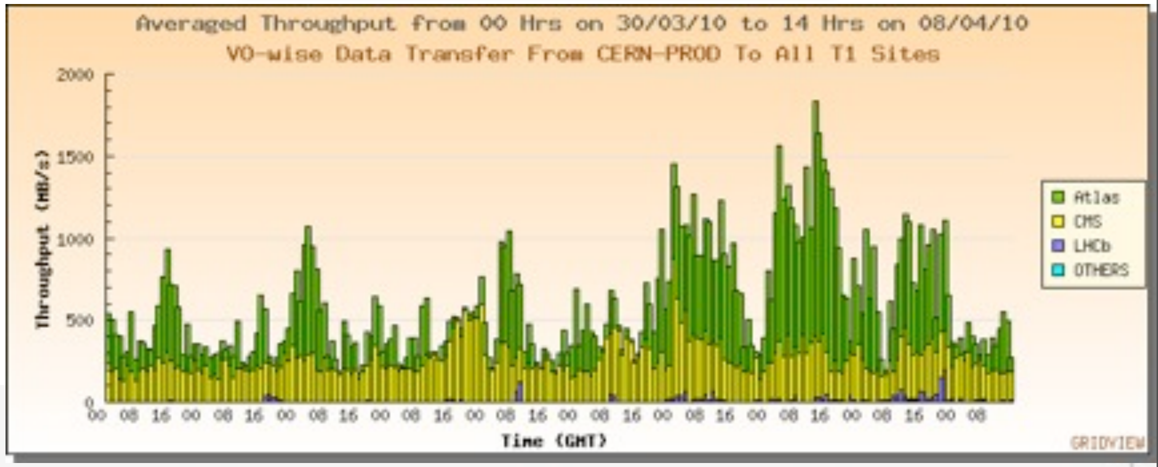
- Running increasingly high
- workloads:
  - Jobs in excess of 900k / day;
  - Anticipate millions / day soon
  - CPU equiv. ~100k cores



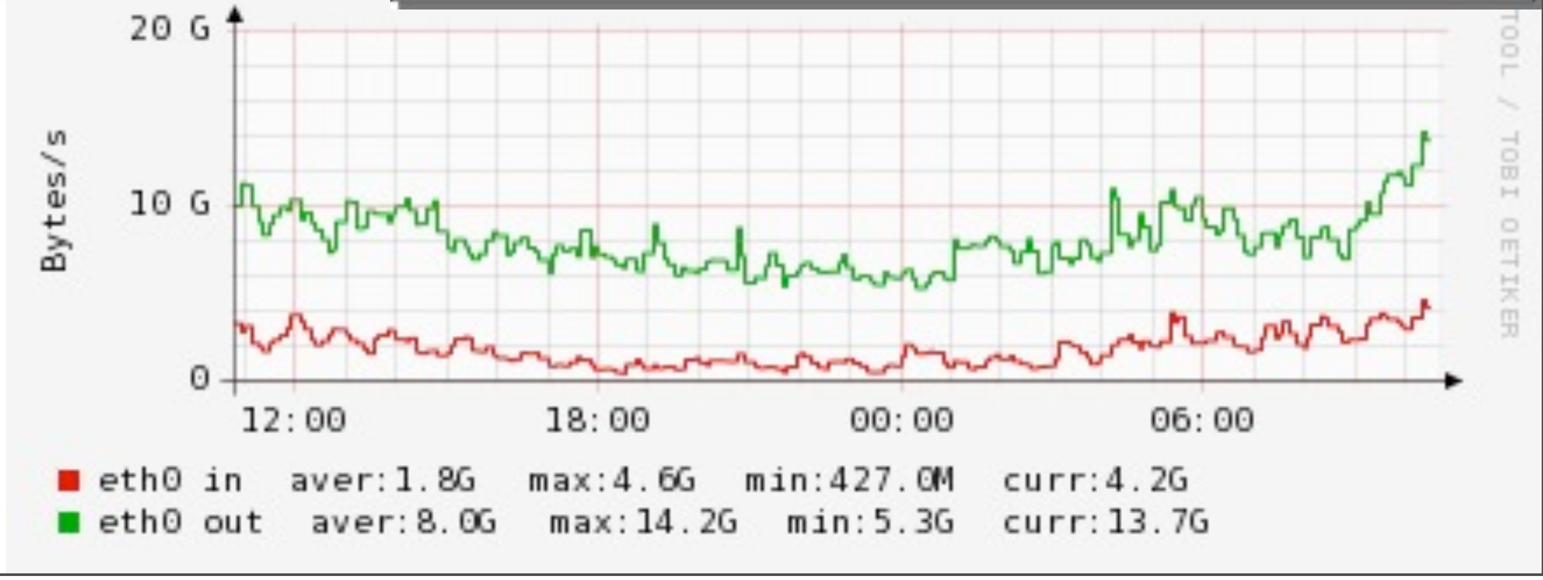


OPN traffic – last 7 days

Real data – from 30/3



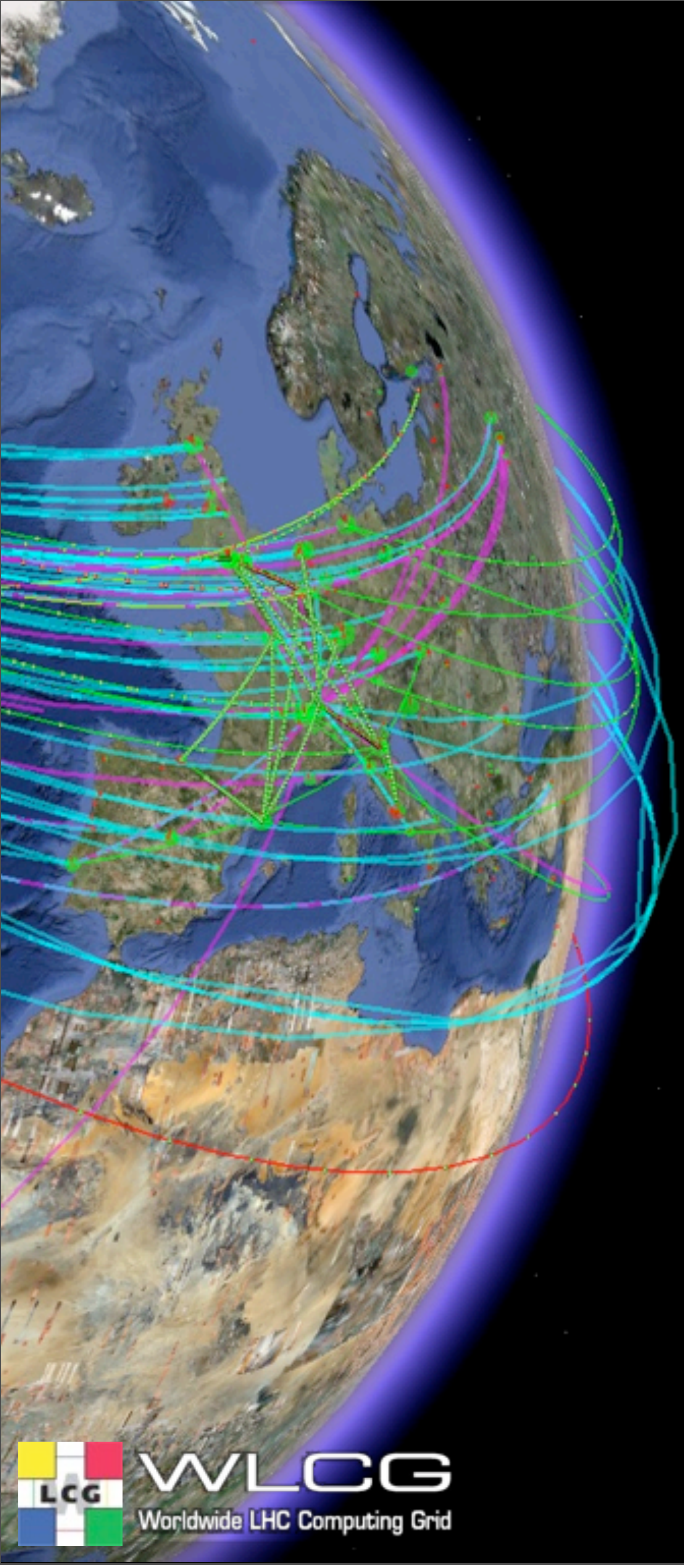
Castor traffic last month:
> 4 GB/s input
> 13 GB/s served



```
■ eth0 in    aver:1.8G   max:4.6G   min:427.0M  curr:4.2G
■ eth0 out   aver:8.0G   max:14.2G  min:5.3G    curr:13.7G
```

WLCG Worldwide LHC Computing Grid

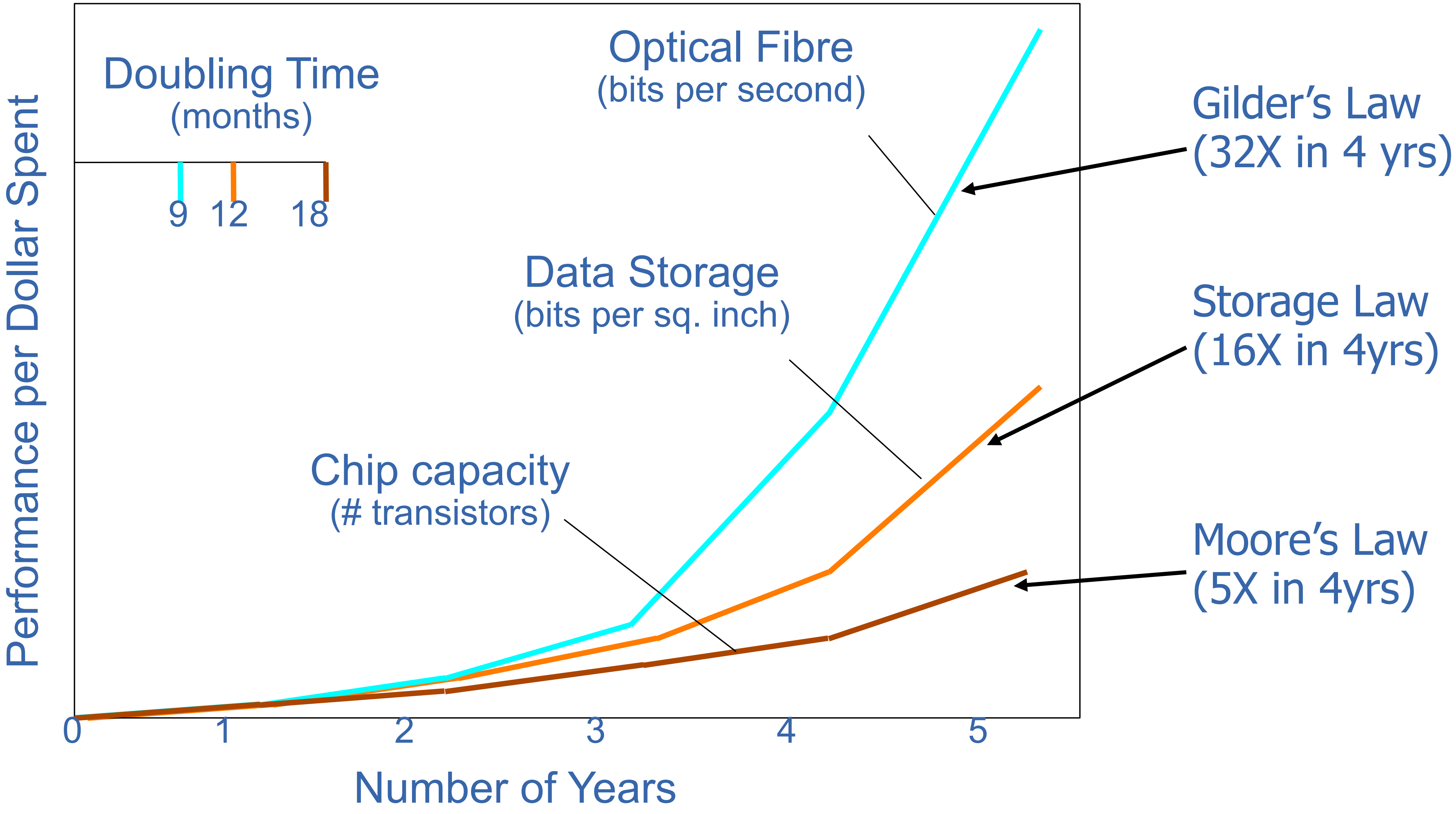
LCG



So what is the Grid?

WLCG
Worldwide LHC Computing Grid

Exponential Growth ...

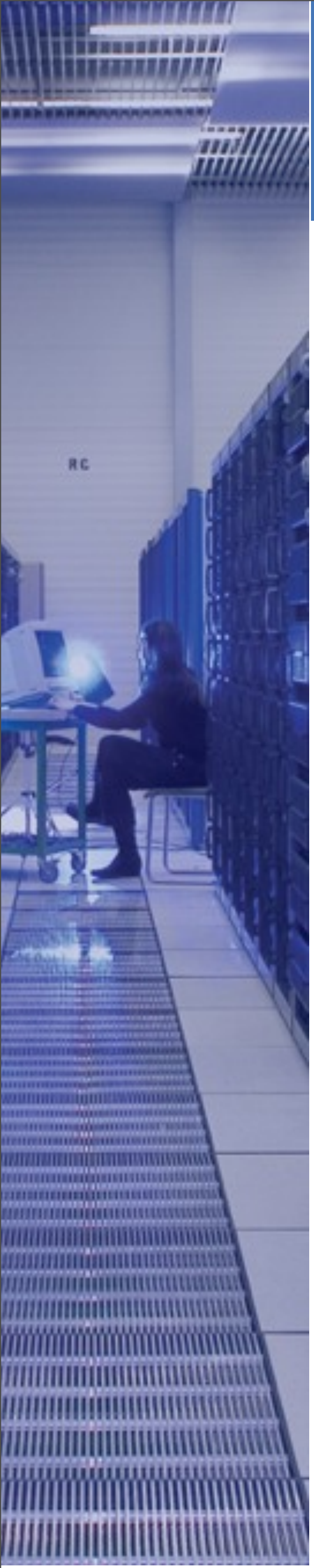


Triumph of Light – *Scientific American*. George Stix, January 2001

Evolution

„When the network is as fast as the computer's internal links, the machine dis-integrates across the net into a set of special purpose appliances“

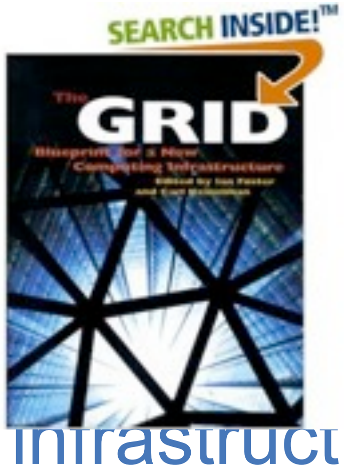
(Gilder Technology Report, June 2001)



Web and Grid

The World Wide Web provides seamless access to information that is stored in many millions of different geographical locations

Tim Berners-Lee invented the World Wide Web at CERN in 1989


SEARCH INSIDE!™
The GRID
Blueprint for a New Computing Infrastructure
Edited by Ian Foster and Carl Kesselman

The Grid is an infrastructure that provides seamless access to computing power and data storage capacity distributed over the globe

Foster and Kesselman 1997

An Analogy to the Power Grid

- Power on demand
  - User is not aware of producers
- Simple Interface
  - Few types of sockets
- Standardized protocols
  - Voltage, Frequency
- Resilience
  - Re-routing
  - Redundancy
- Can't be stored, has to be consumed as produced
  - Use it or loose it
  - Pricing models for:
    - Large/small scale users

An Analogy to the Power Grid

- Power on demand
  - User is not aware of producers

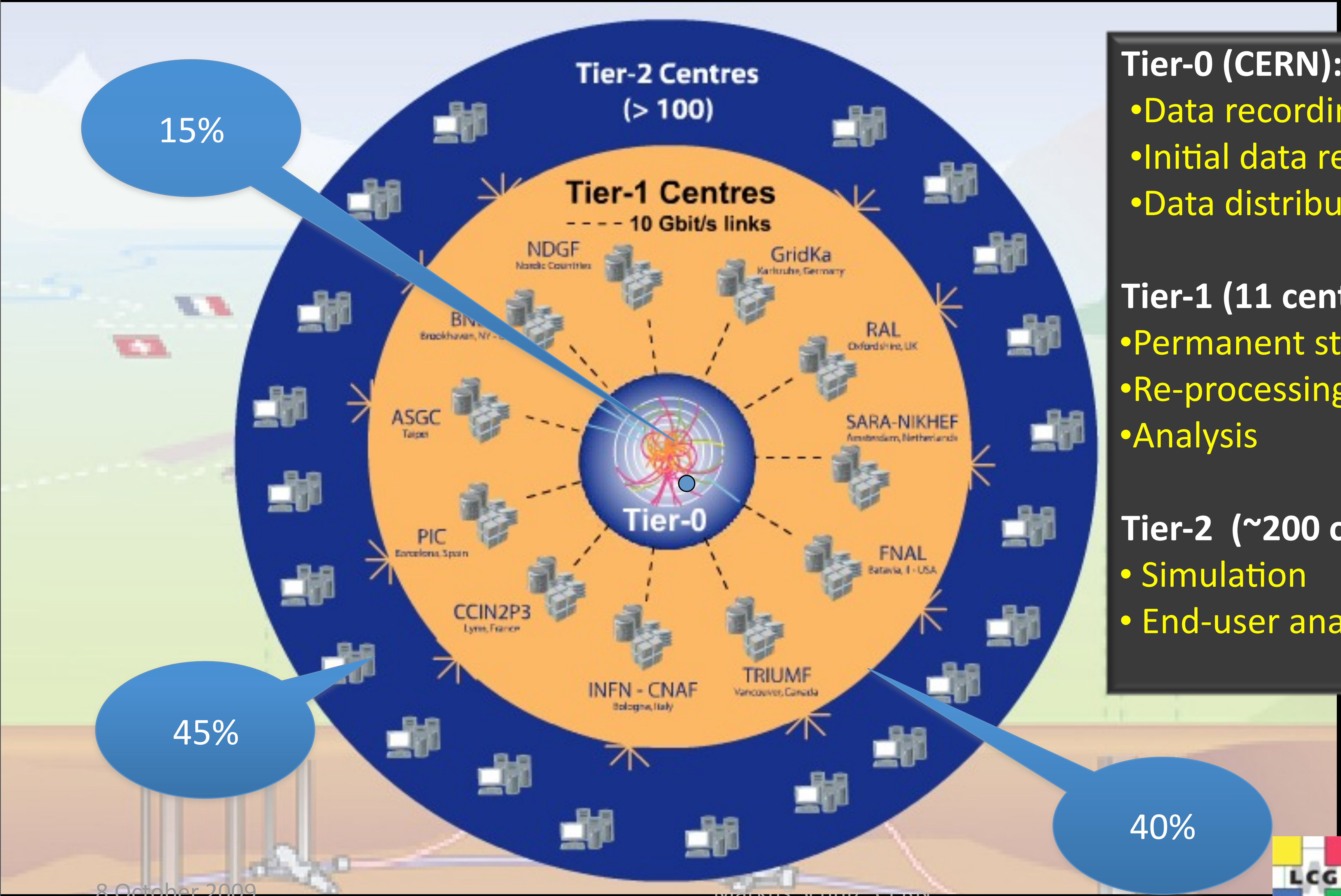
- Simple Interface
  - Few types of sockets

- Standardized protocols
  - Voltage, Frequency
- Resilience
  - Re-routing
  - Redundancy
- Can’t be stored, has to be consumed as produced
  - Use it or loose it
  - Pricing models for:
    - Large/small scale users

What is a Computing Grid?

- There are many conflicting definitions
  - Has been used for several years for marketing…
    - Now they use Cloud
- Ian Foster and Karl Kesselman
  - "coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations. "
  - These are the people who started globus, the first grid middleware project
- From the user's perspective:
  - I want to be able to use computing resources as I need
  - I don't care who owns resources, or where they are
  - Have to be secure
  - My programs have to run there
- The owners of computing resources (CPU cycles, storage, bandwidth)
  - My resources can be used by any authorized person (not for free)
  - Authorization is not tied to my administrative organization
- **– NO centralized control of resources or users**

Tier 0 – Tier 1 – Tier 2
the Service Hierarchy



Tier-0 (CERN): (15%)

- •Data recording
- •Initial data reconstruction
- •Data distribution

Tier-1 (11 centres): (40%)

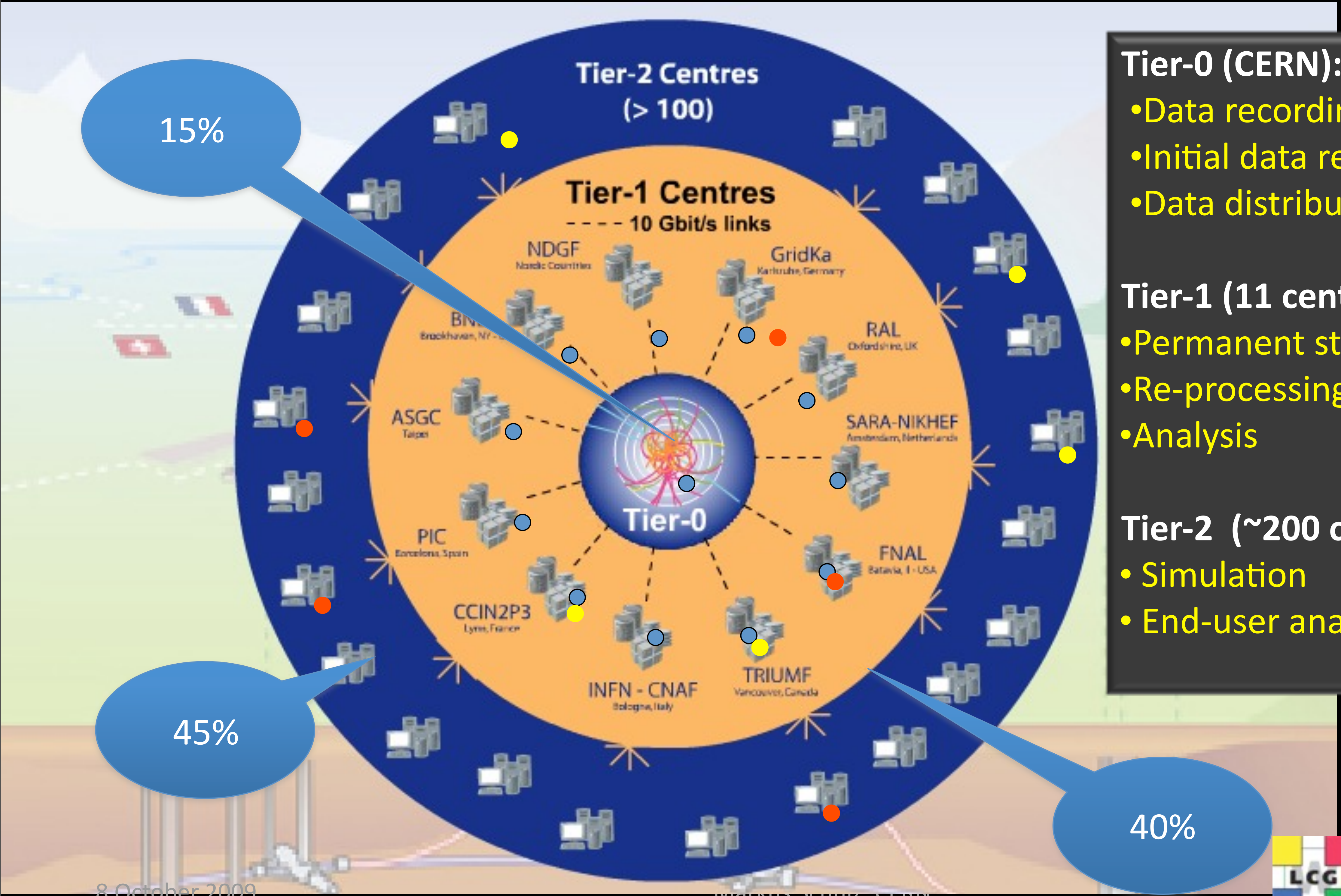
- •Permanent storage
- •Re-processing
- •Analysis

Tier-2 (~200 centres): (45%)

- • Simulation
- • End-user analysis

WLCG
Worldwide LHC Computing Grid

Tier 0 – Tier 1 – Tier 2 the Service Hierarchy



Tier-0 (CERN): (15%)

- Data recording
- Initial data reconstruction
- Data distribution

Tier-1 (11 centres): (40%)

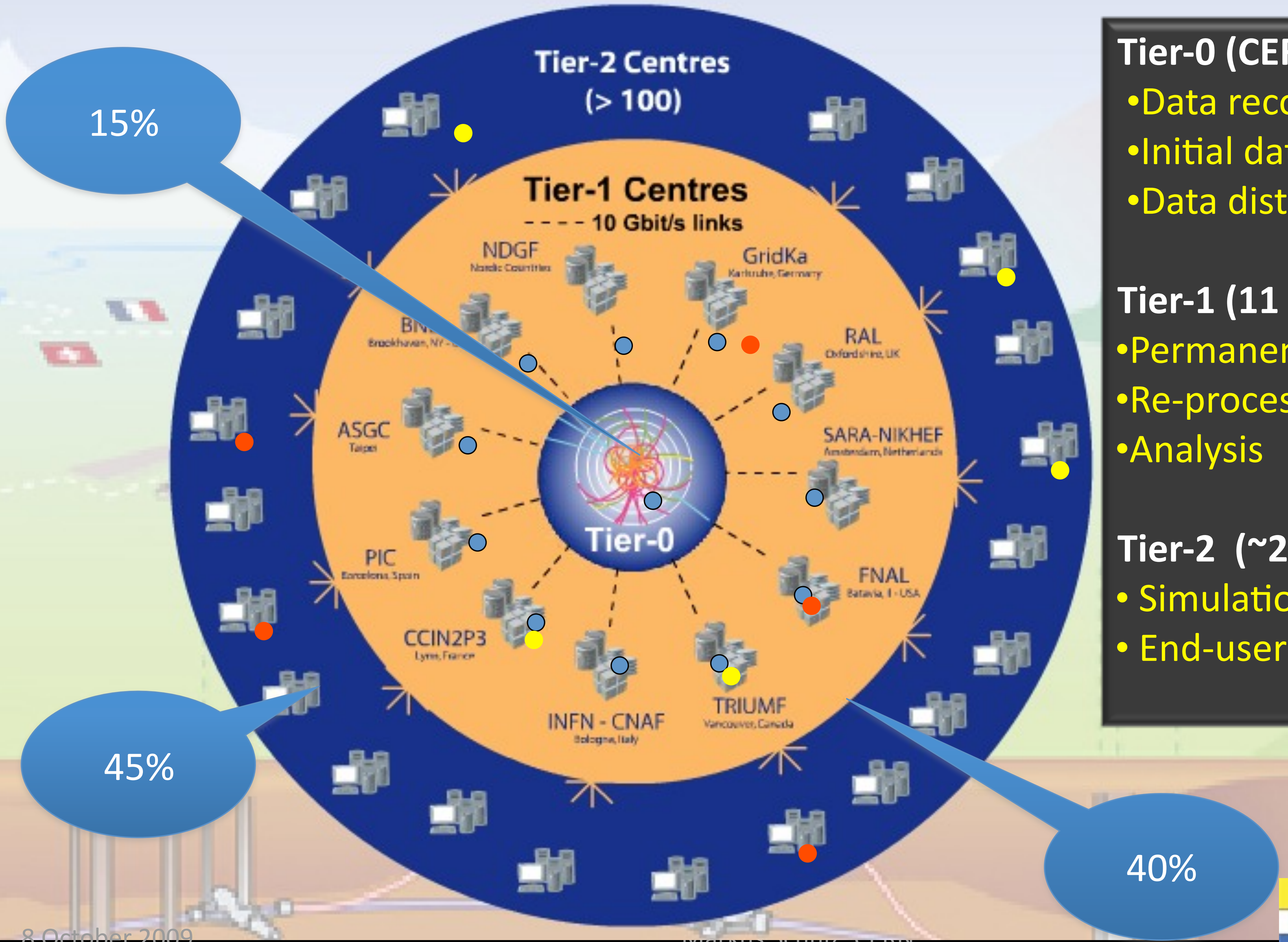
- Permanent storage
- Re-processing
- Analysis

Tier-2 (~200 centres): (45%)

- Simulation
- End-user analysis

WLCG
Worldwide LHC Computing Grid

Tier 0 – Tier 1 – Tier 2 the Service Hierarchy



Tier-0 (CERN): (15%)

- Data recording
- Initial data reconstruction
- Data distribution

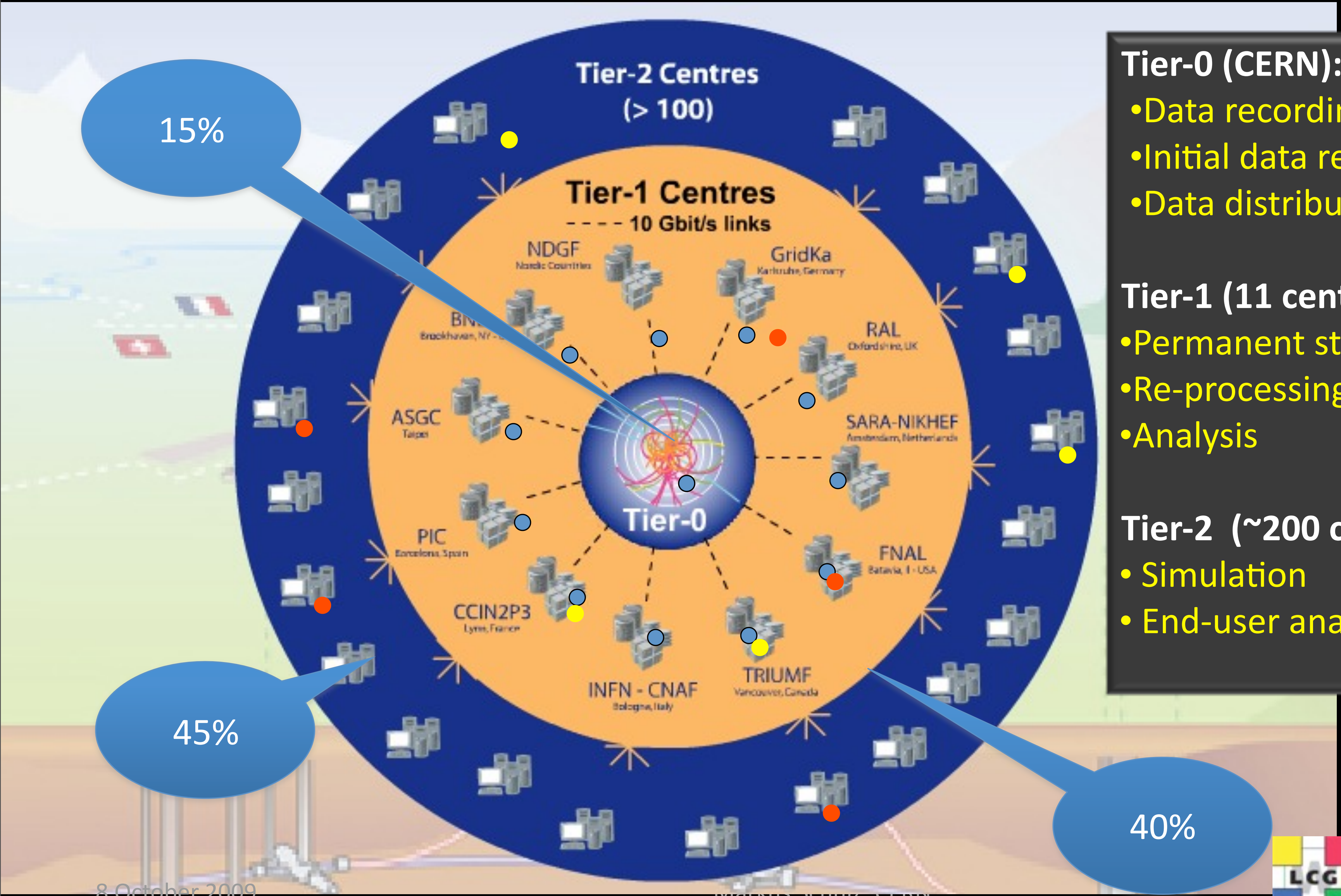
Tier-1 (11 centres): (40%)

- Permanent storage
- Re-processing
- Analysis

Tier-2 (~200 centres): (45%)

- Simulation
- End-user analysis

Tier 0 – Tier 1 – Tier 2 the Service Hierarchy



Tier-0 (CERN): (15%)

- •Data recording
- •Initial data reconstruction
- •Data distribution

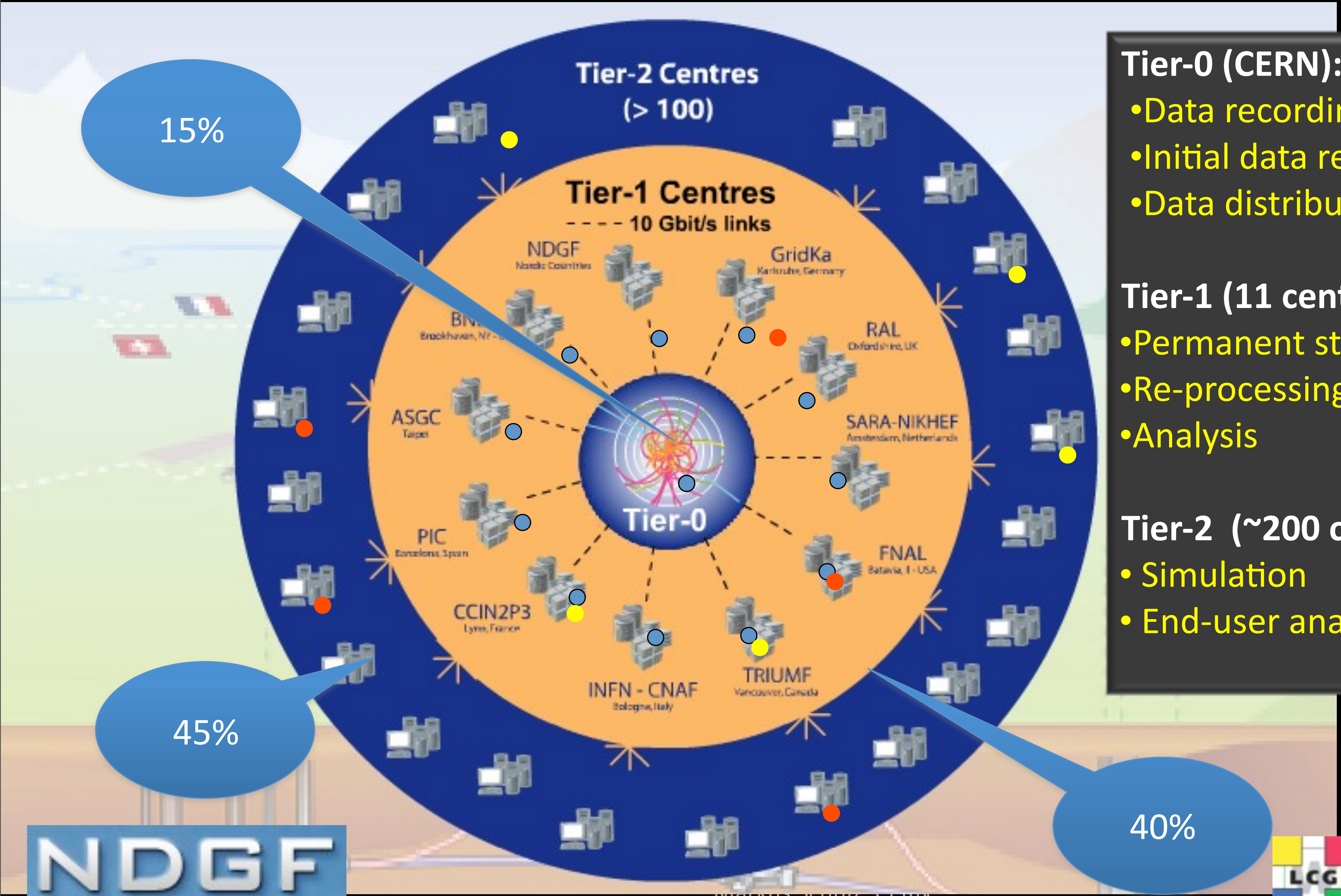
Tier-1 (11 centres): (40%)

- •Permanent storage
- •Re-processing
- •Analysis

Tier-2 (~200 centres): (45%)

- • Simulation
- • End-user analysis

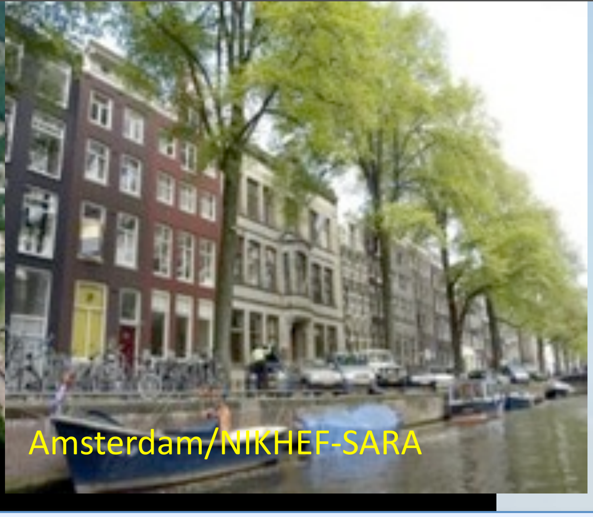
Tier 0 – Tier 1 – Tier 2 the Service Hierarchy

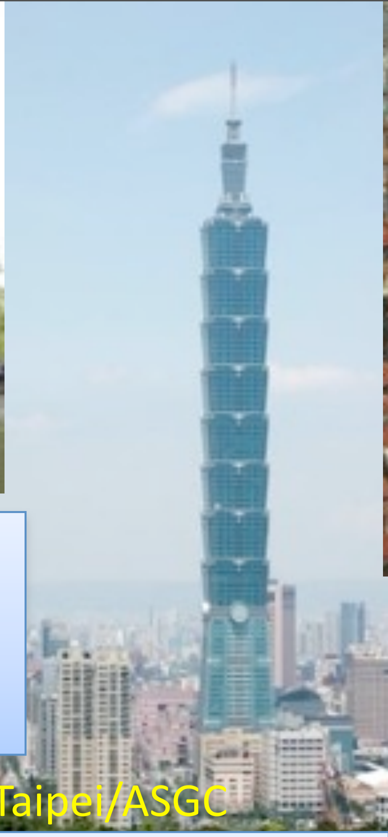


Tier-0 (CERN): (15%)
- Data recording
- Initial data reconstruction
- Data distribution

Tier-1 (11 centres): (40%)
- Permanent storage
- Re-processing
- Analysis

Tier-2 (~200 centres): (45%)
- Simulation
- End-user analysis



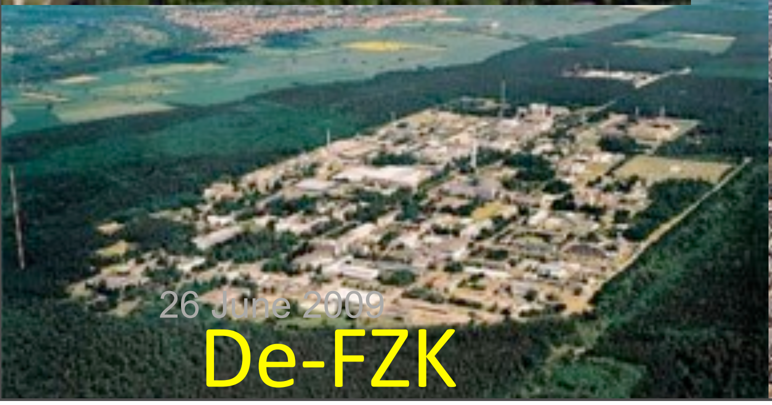


WLCG Collaboration Status

Tier 0; 11 Tier 1s; 64 Tier 2 federations (124 Tier 2 sites)

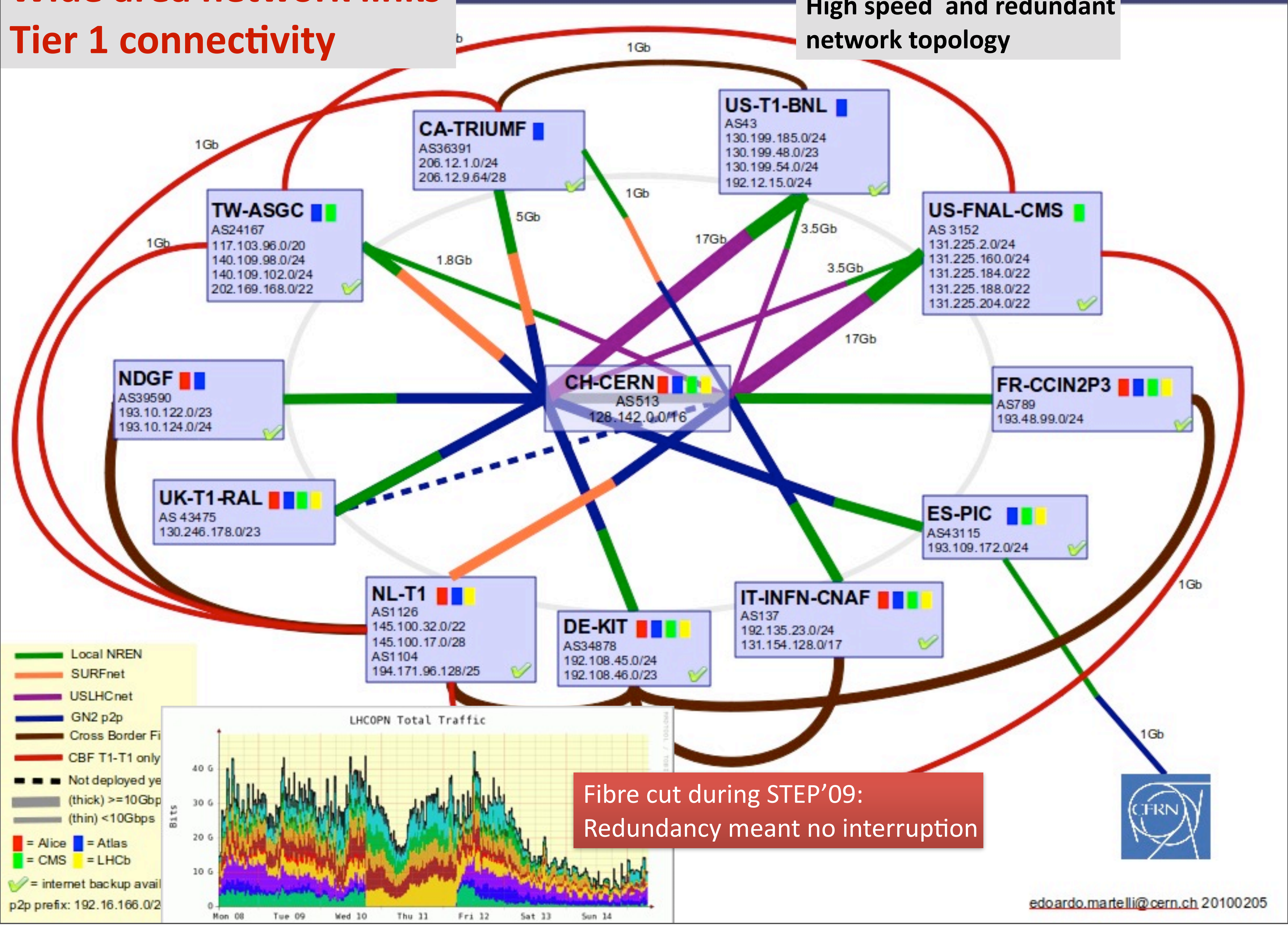
Today we have 49 MoU signatories, representing 34 countries:

Australia, Austria, Belgium, Brazil, Canada, China, Czech Rep, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, India, Israel, Japan, Rep. Korea, Netherlands, Norway, Pakistan, Poland, Portugal, Romania, Russia, Slovenia, Spain, Sweden, Switzerland, Taipei, Turkey, UK, Ukraine, USA.



Wide area network links
Tier 1 connectivity

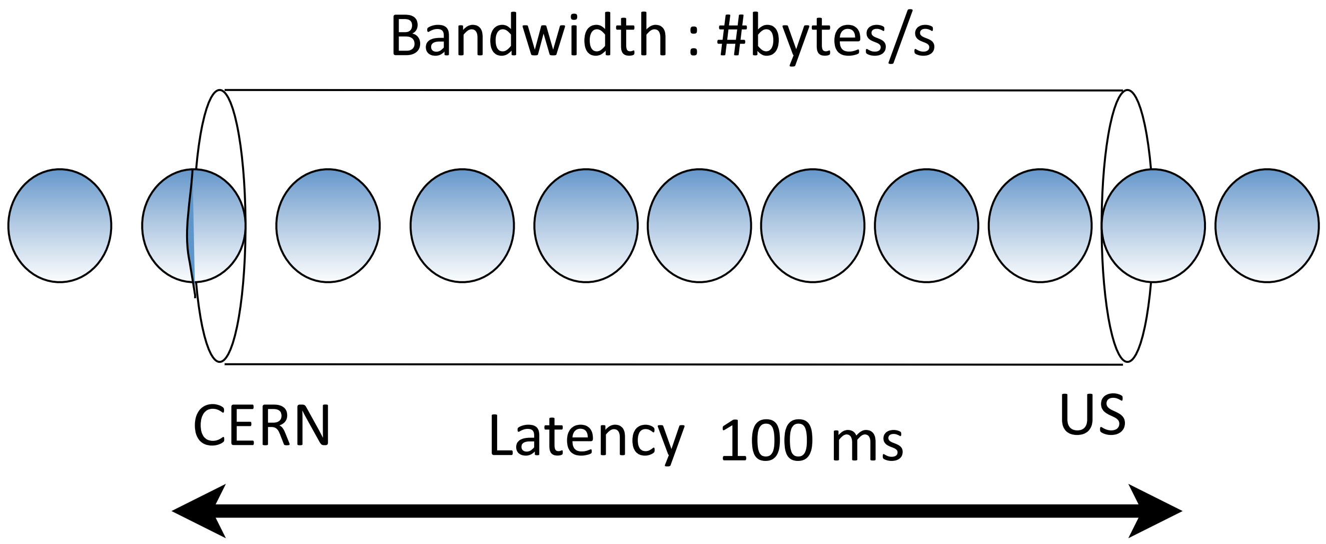
High speed and redundant network topology



Fibre cut during STEP'09:
Redundancy meant no interruption

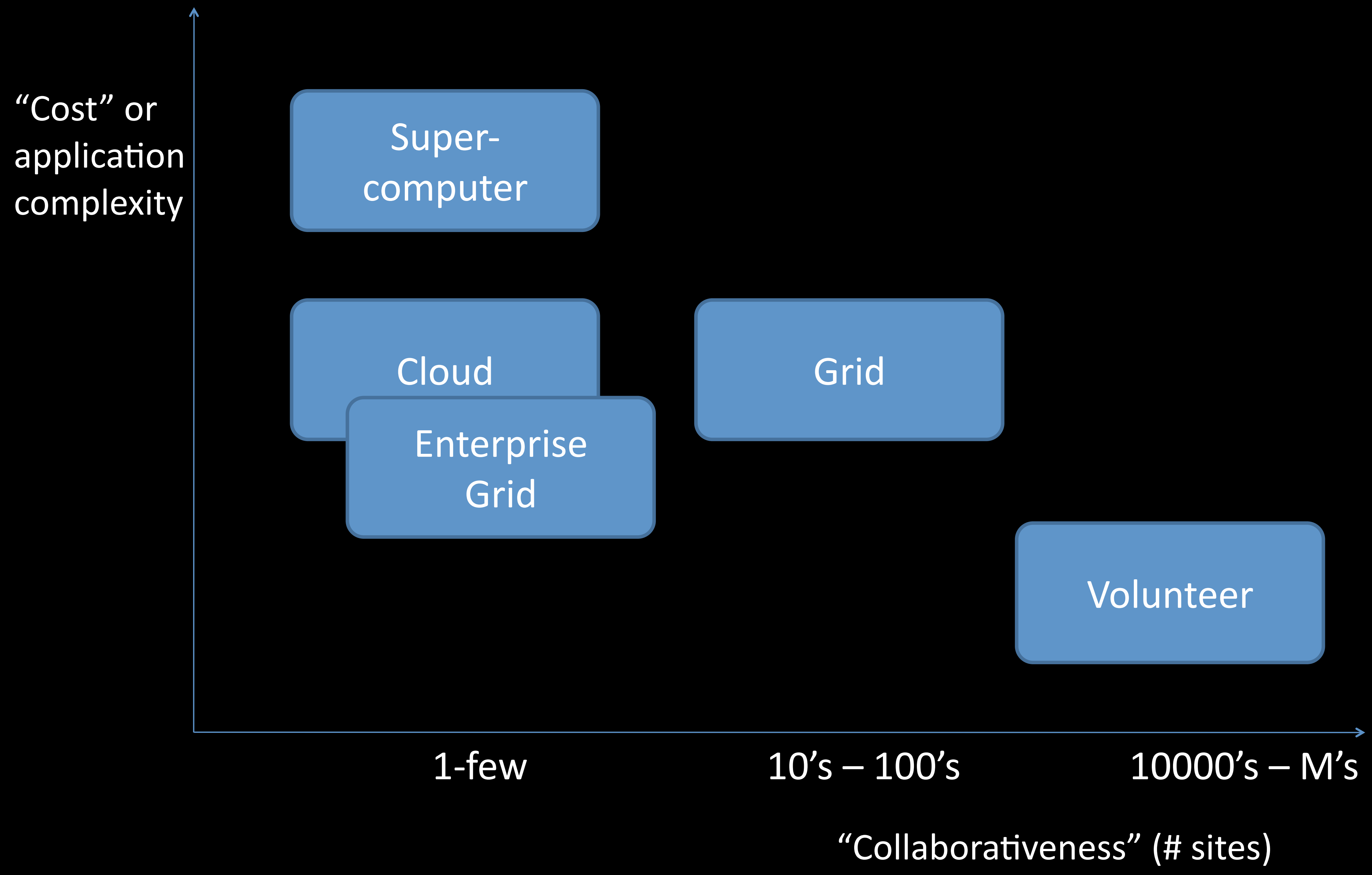
edoardo.martelli@cern.ch 20100205

Limitations of Distributed Systems



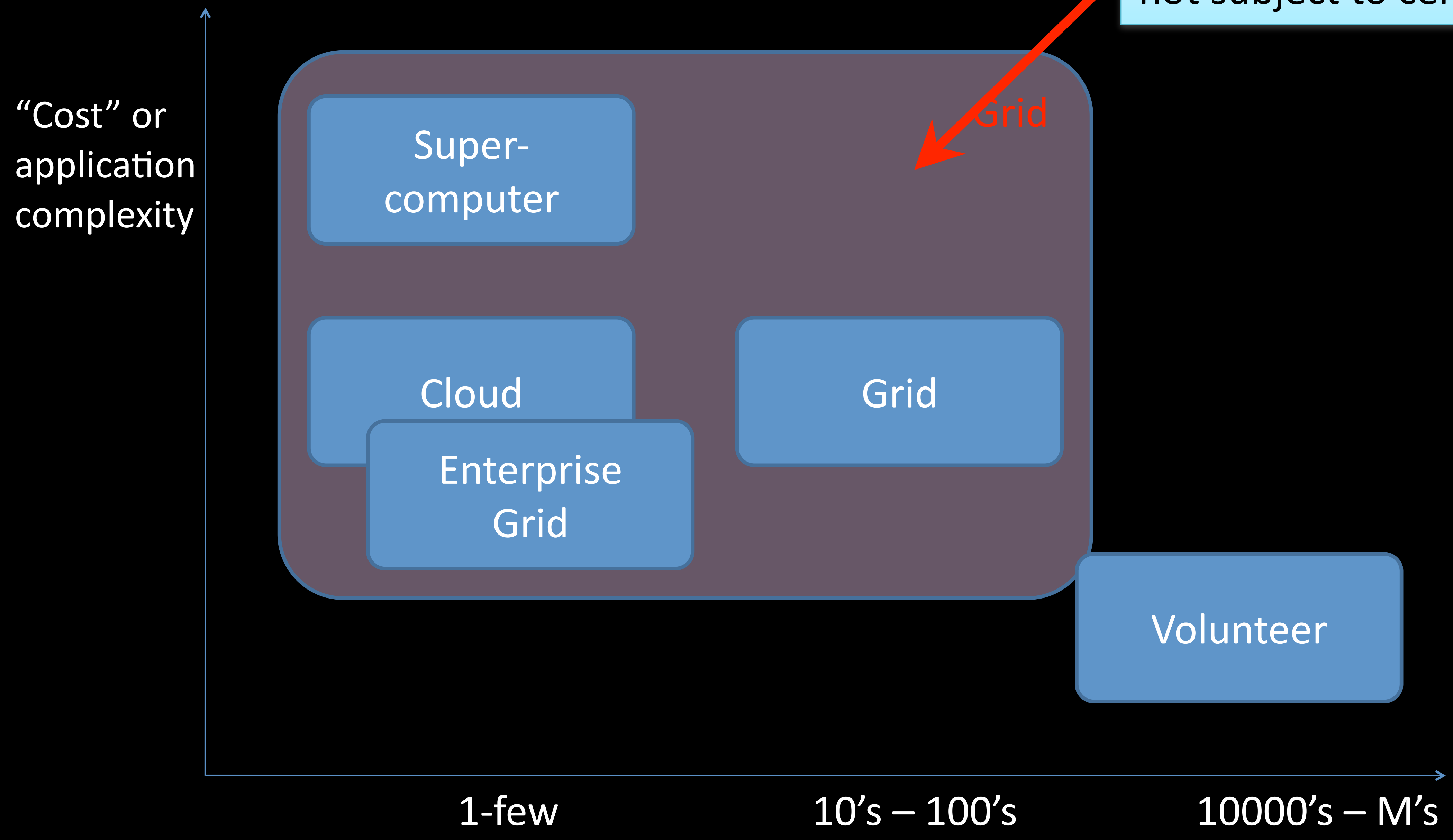
- Application performance depends on two main qualities
  - Bandwidth
    - technology improves quickly
  - Latency
    - but the speed of light is a constant (!)
- Not all applications immediately perform well in a wide area environment
  - several 100 ms latency per round trip accumulate quickly
    - needs to be taken into account when designing a protocol
  - large memory buffers are needed to drive a high bandwidth, high latency network pipe
    - Can limits sever scalability

So what is a Grid?



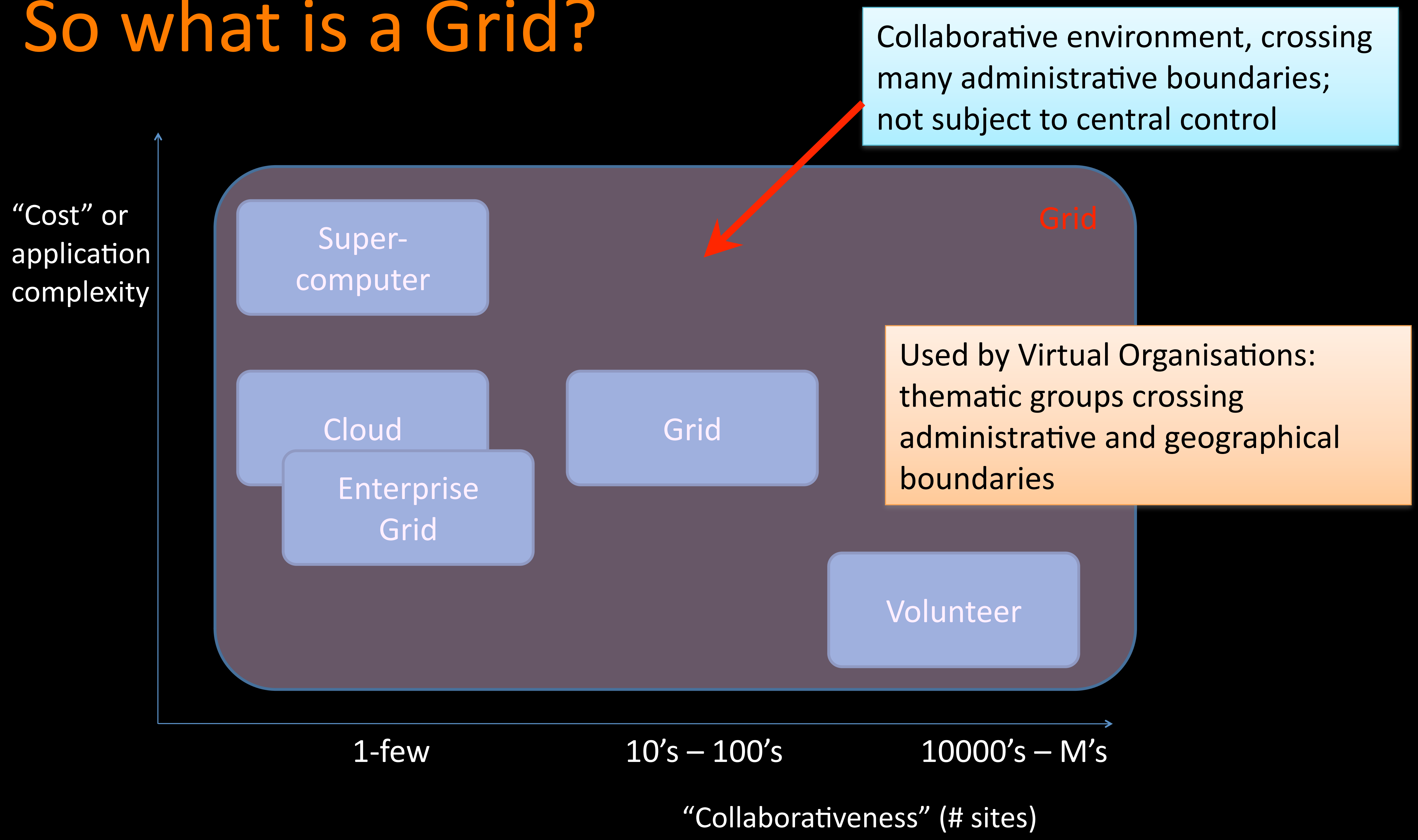
Markus Schulz, CERN

So what is a Grid?


Collaborative environment, crossing many administrative boundaries; not subject to central control
“Cost” or application complexity
Grid
Super-computer
Cloud
Enterprise Grid
Grid
Volunteer
1-few
10’s – 100’s
10000’s – M’s
“Collaborativeness” (# sites)

Markus Schulz, CERN

So what is a Grid?



Collaborative environment, crossing many administrative boundaries; not subject to central control

“Cost” or application complexity

Grid

Super-computer

Cloud

Enterprise Grid

Grid

Used by Virtual Organisations: thematic groups crossing administrative and geographical boundaries

Volunteer

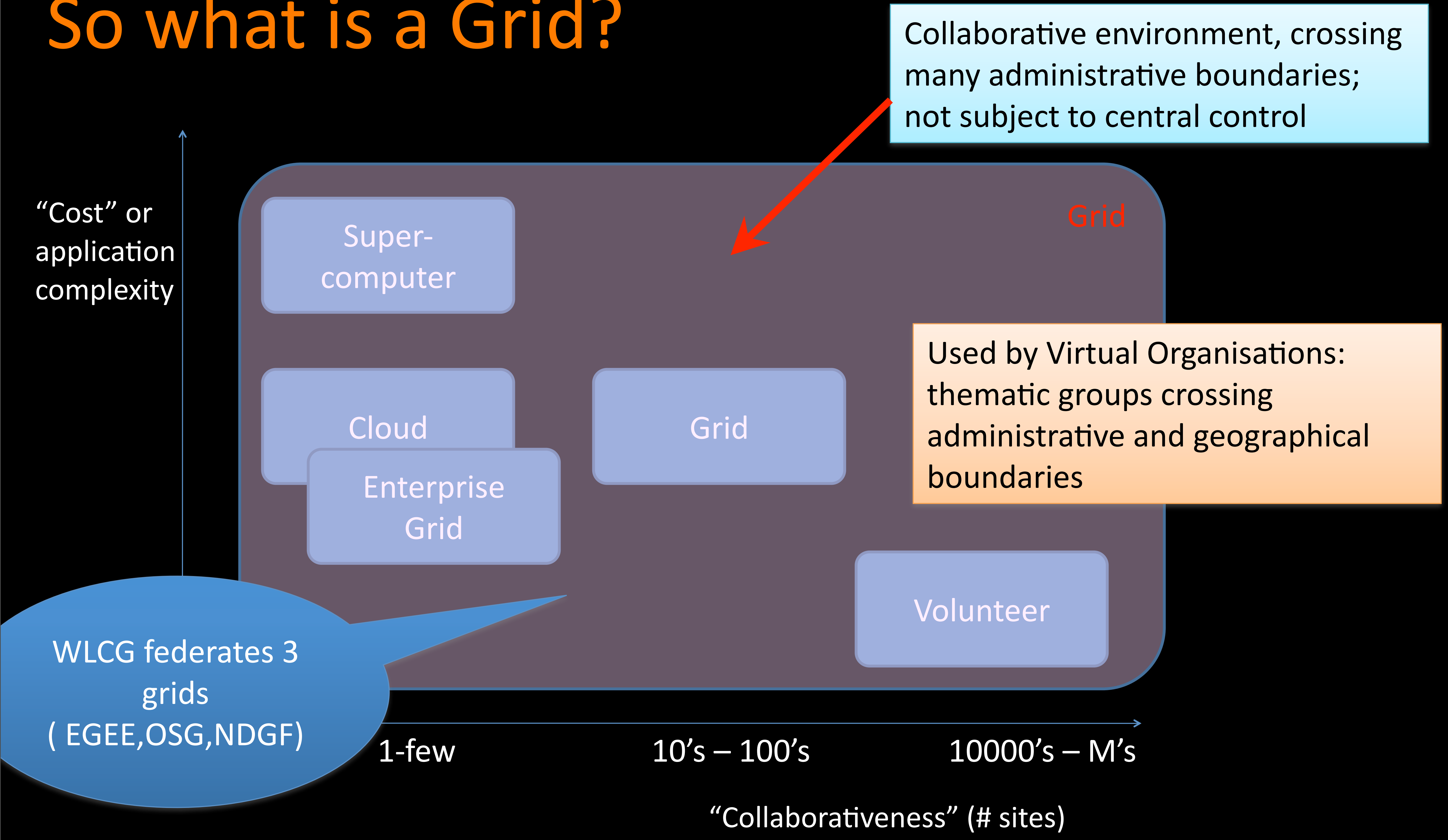
1-few

10’s – 100’s

10000’s – M’s

“Collaborativeness” (# sites)

So what is a Grid?



Collaborative environment, crossing many administrative boundaries; not subject to central control

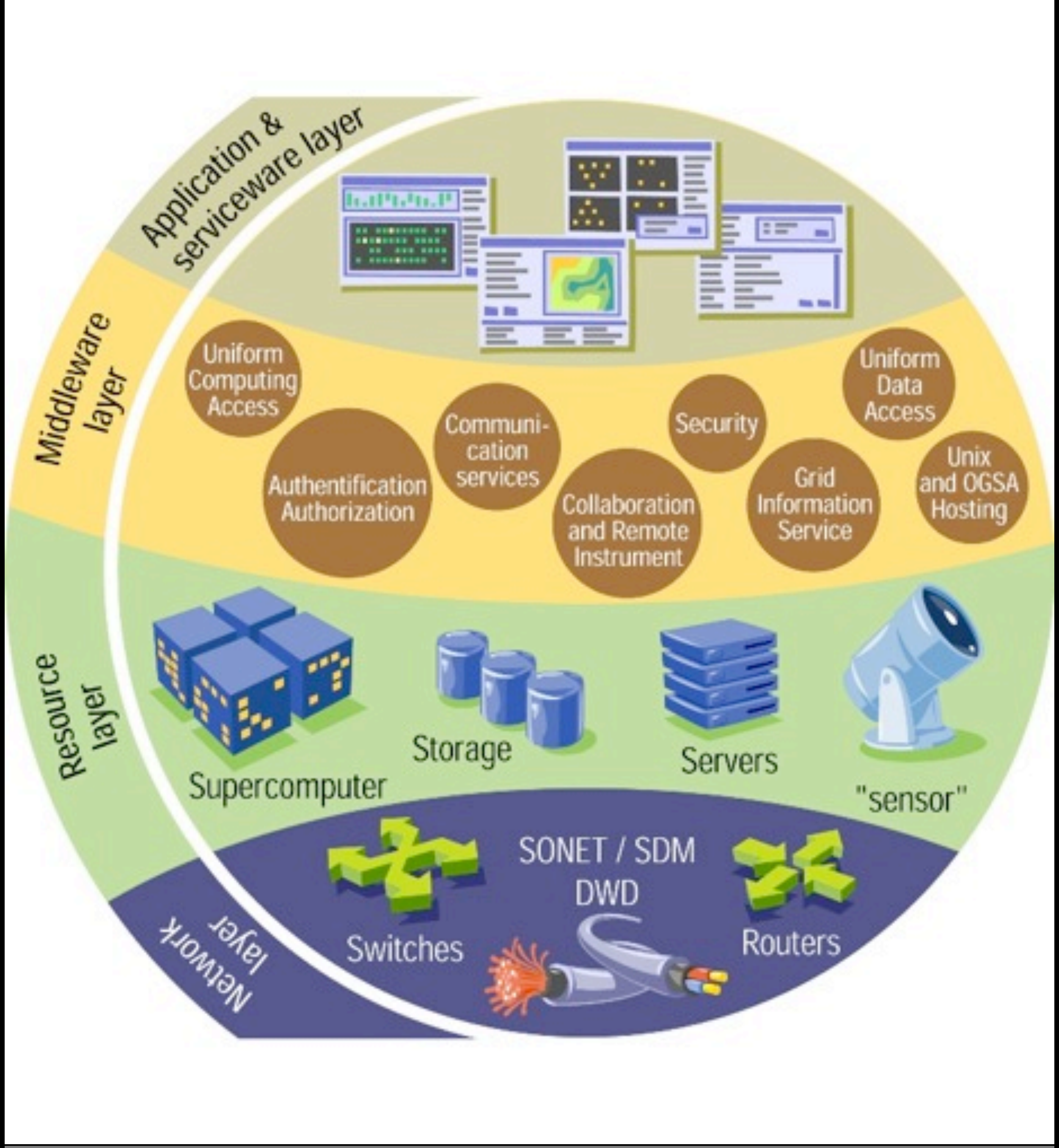
Used by Virtual Organisations: thematic groups crossing administrative and geographical boundaries

WLCG federates 3 grids (EGEE,OSG,NDGF)

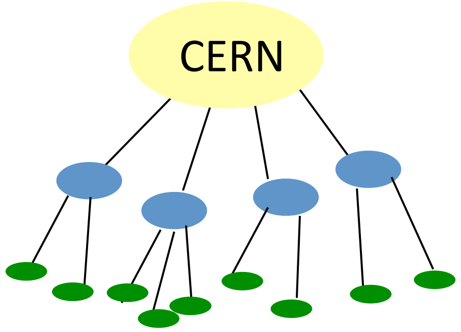
What is Grid Middleware?

- The glue that creates the illusion that a distributed infrastructure is a single resource

- If it works, no one will notice it

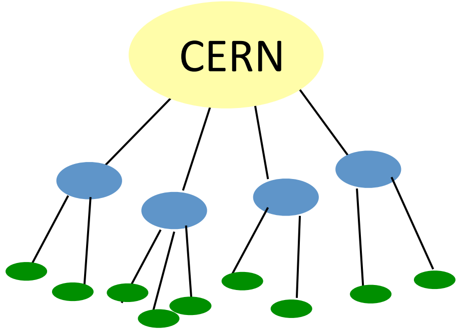


History of Grid Projects



History of Grid Projects

- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model

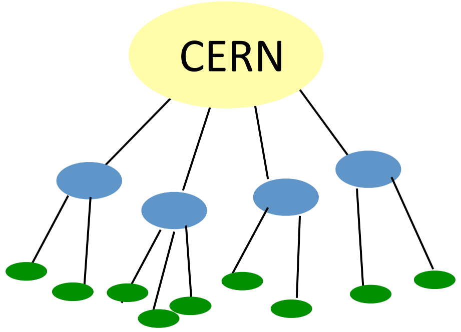


WLCG
Worldwide LHC Computing Grid

LCG

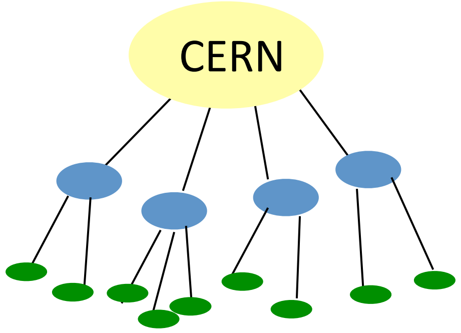
History of Grid Projects

- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project

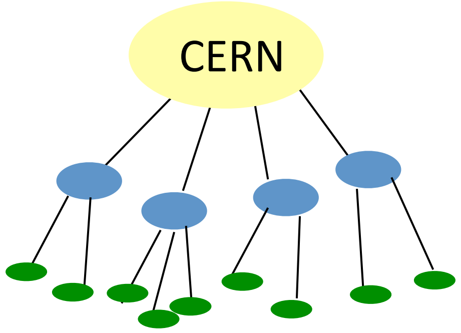


History of Grid Projects

- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project
- 2001-2004 - EU DataGrid project
  - middleware & testbed for an operational grid



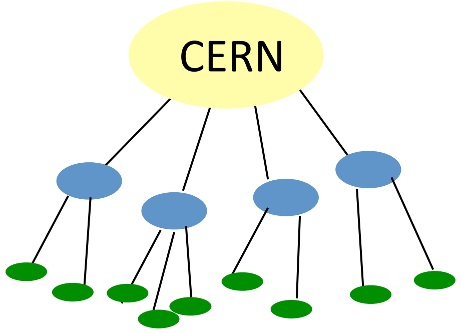
History of Grid Projects



- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project
- 2001-2004 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments

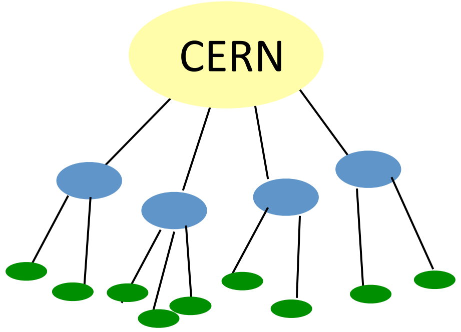
WLCG Worldwide LHC Computing Grid

History of Grid Projects



- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project
- 2001-2004 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments
- 2004-2006 – EU EGEE project phase 1
  - starts from the LCG grid
  - shared production infrastructure
  - expanding to other communities and sciences

History of Grid Projects

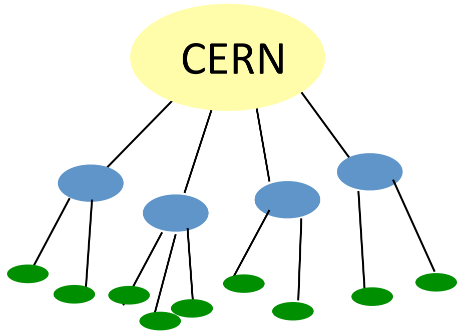


- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project
- 2001-2004 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments
- 2004-2006 – EU EGEE project phase 1
  - starts from the LCG grid
  - shared production infrastructure
  - expanding to other communities and sciences
- 2006-2008 – EU EGEE project phase 2
  - expanding to other communities and sciences
  - Scale and stability
  - Interoperations/Interoperability

WLCG Worldwide LHC Computing Grid

LCG

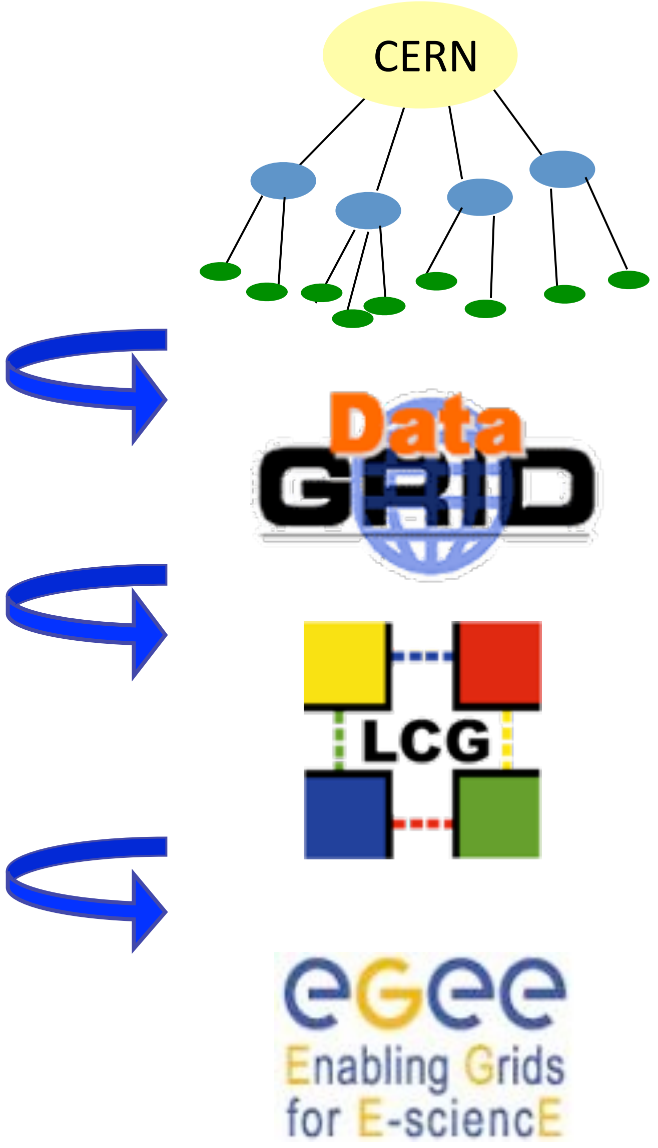
History of Grid Projects



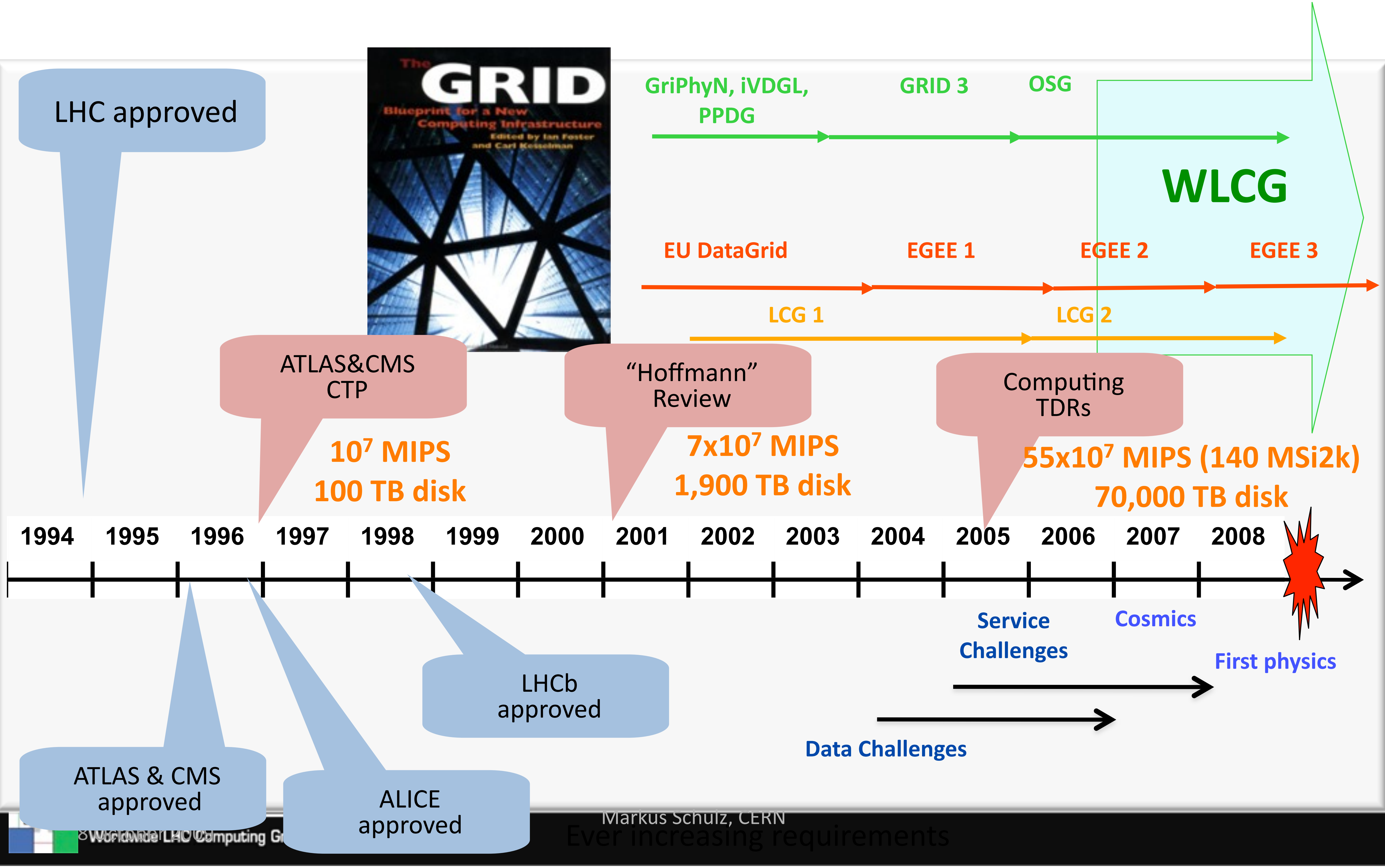
- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project
- 2001-2004 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments
- 2004-2006 – EU EGEE project phase 1
  - starts from the LCG grid
  - shared production infrastructure
  - expanding to other communities and sciences
- 2006-2008 – EU EGEE project phase 2
  - expanding to other communities and sciences
  - Scale and stability
  - Interoperations/Interoperability
- 2008-2010 – EU EGEE project phase 3
  - More communities
  - Efficient operations
  - Less central coordination

History of Grid Projects

- 1999 - MONARC project
  - First LHC computing architecture – hierarchical distributed model
- 2000 – growing interest in grid technology
  - HEP community main driver in launching the DataGrid project
- 2001-2004 - EU DataGrid project
  - middleware & testbed for an operational grid
- 2002-2005 – LHC Computing Grid – LCG
  - deploying the results of DataGrid to provide a production facility for LHC experiments
- 2004-2006 – EU EGEE project phase 1
  - starts from the LCG grid
  - shared production infrastructure
  - expanding to other communities and sciences
- 2006-2008 – EU EGEE project phase 2
  - expanding to other communities and sciences
  - Scale and stability
  - Interoperations/Interoperability
- 2008-2010 – EU EGEE project phase 3
  - More communities
  - Efficient operations
  - Less central coordination
- 2010 – 201x EGI and EMI
  - Sustainable infrastructures based on National Grid Infrastructures
  - Decoupling of middleware development and infrastructure



Evolution of (W)LCG



LHC approved

ATLAS&CMS CTP

10^7 MIPS
100 TB disk

"Hoffmann" Review

$7x10^7$ MIPS
1,900 TB disk

Computing TDRs

$55x10^7$ MIPS (140 MSi2k)
70,000 TB disk

GriPhyN, iVDGL, PPDG

GRID 3

OSG

WLCG

EU DataGrid

EGEE 1

EGEE 2

EGEE 3

LCG 1

LCG 2

1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008

ATLAS & CMS approved

ALICE approved

LHCb approved

Service Challenges

Cosmics

First physics

Data Challenges

Ever increasing requirements