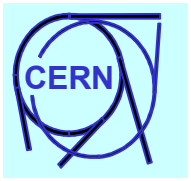


HPC at CERN and the Grid

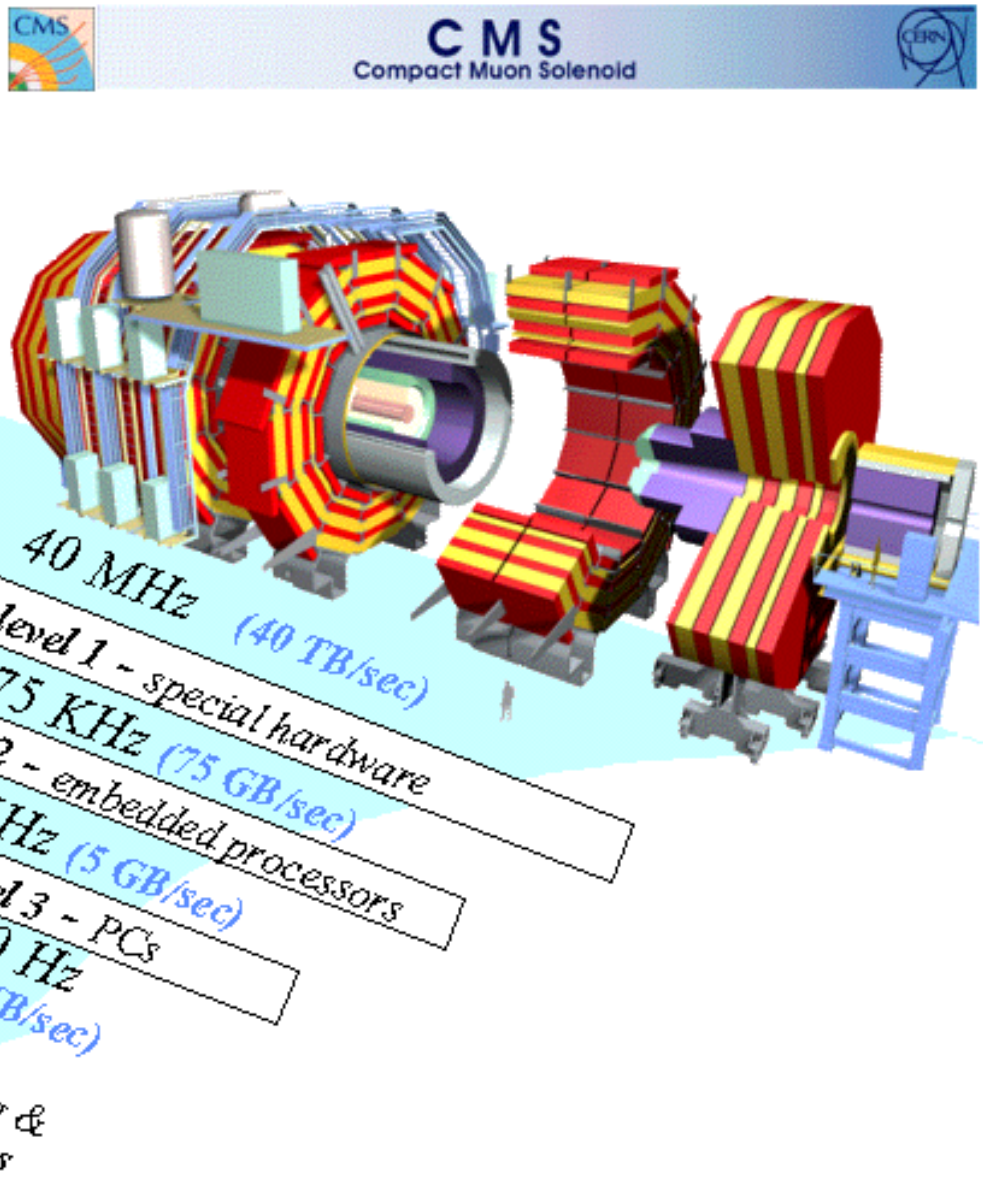
Fabrizio Gagliardi

CERN

Information Technology Division

October, 2000

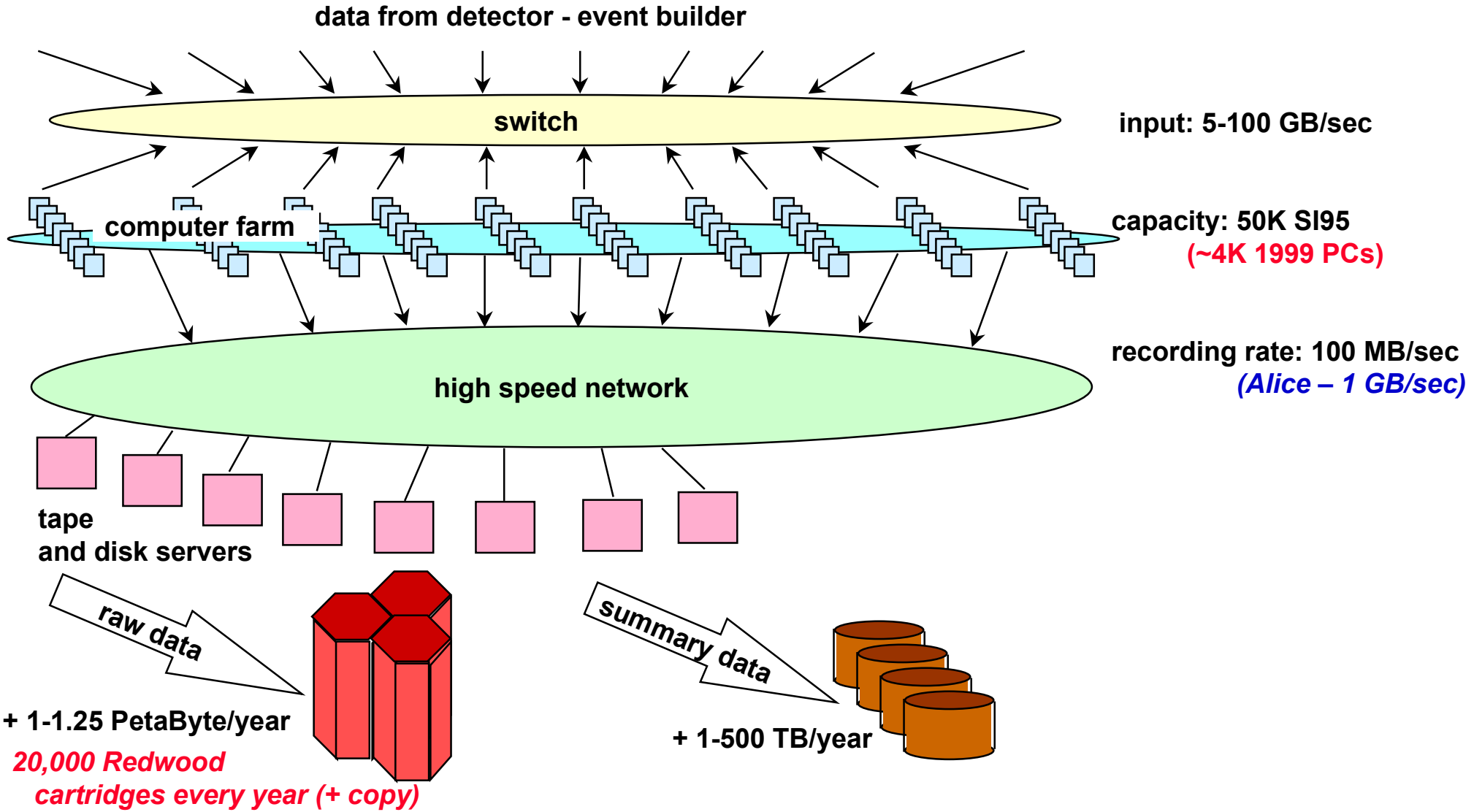
F.Gagliardi@cern.ch



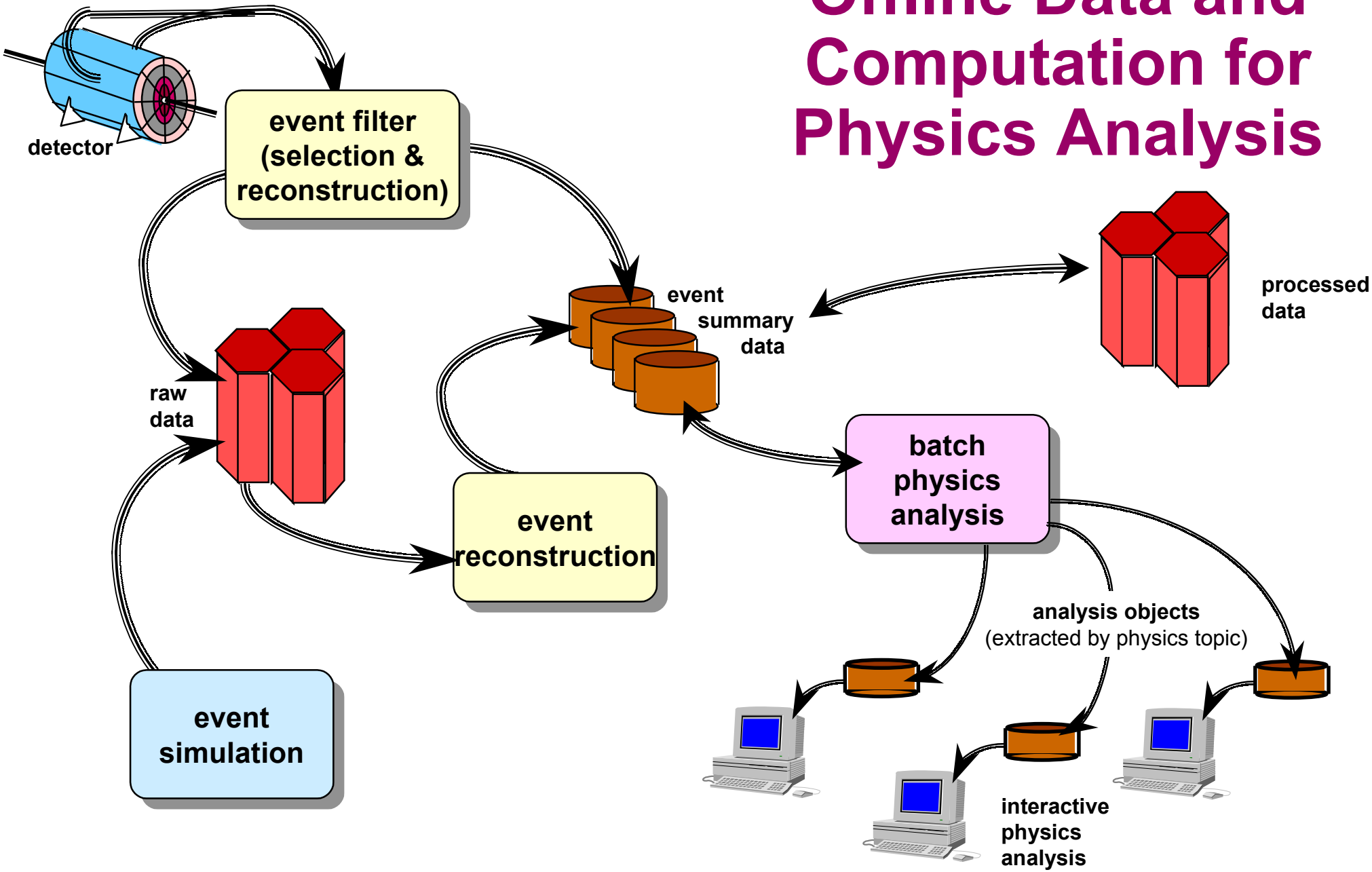
online system

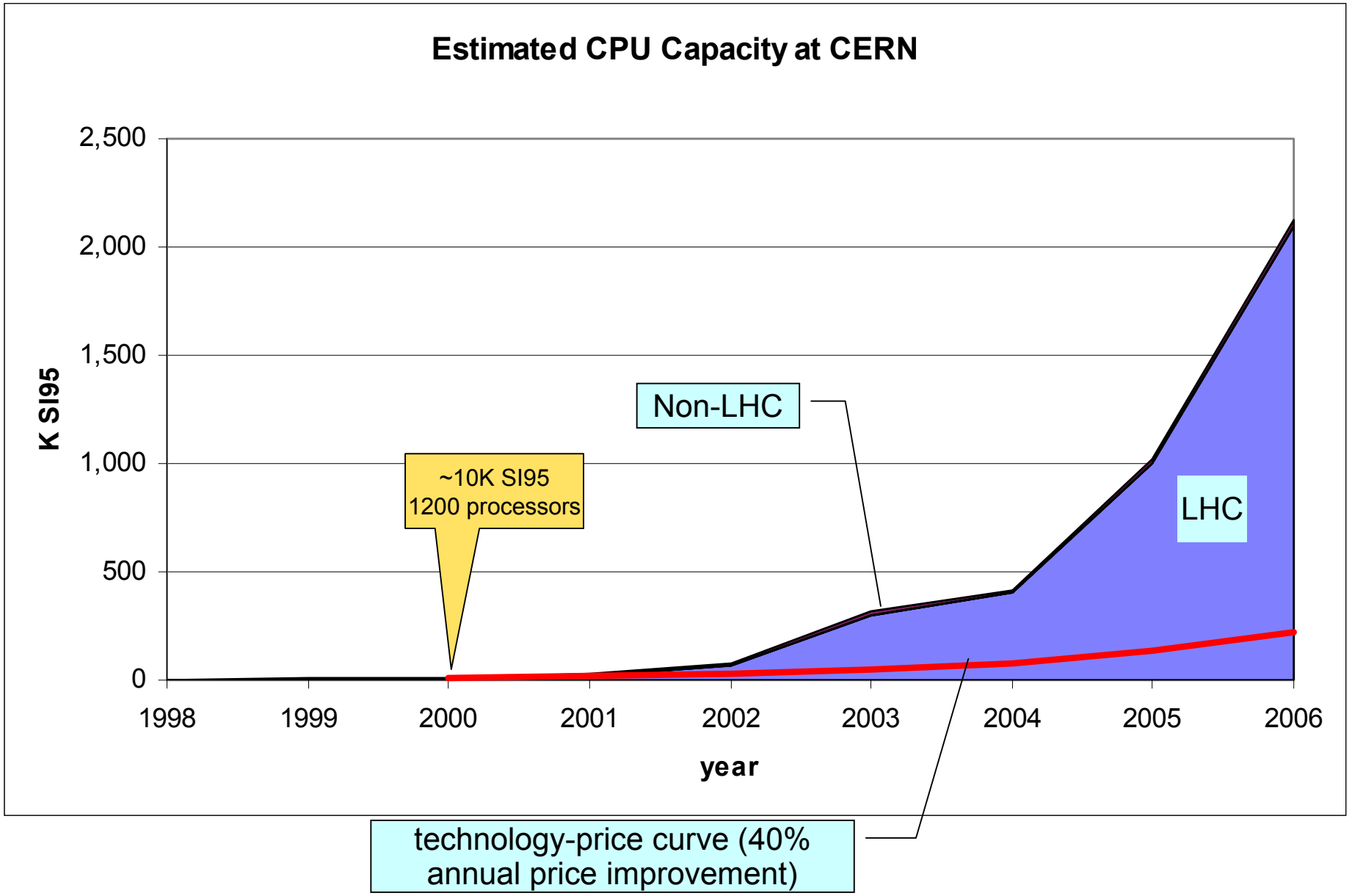
multi-level trigger
filter out background
reduce data volume

Event Filter & Reconstruction
(figures are for one experiment)

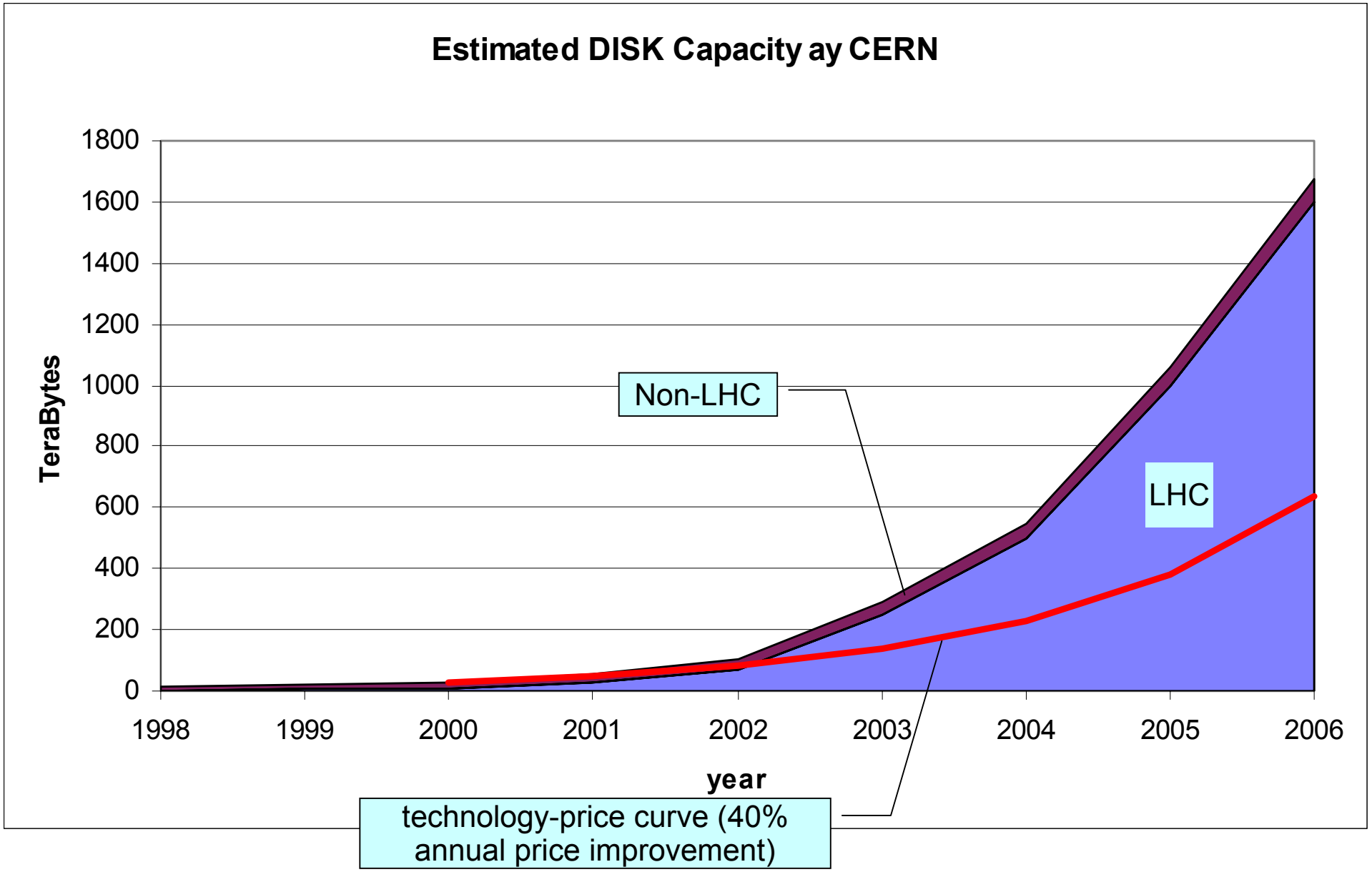


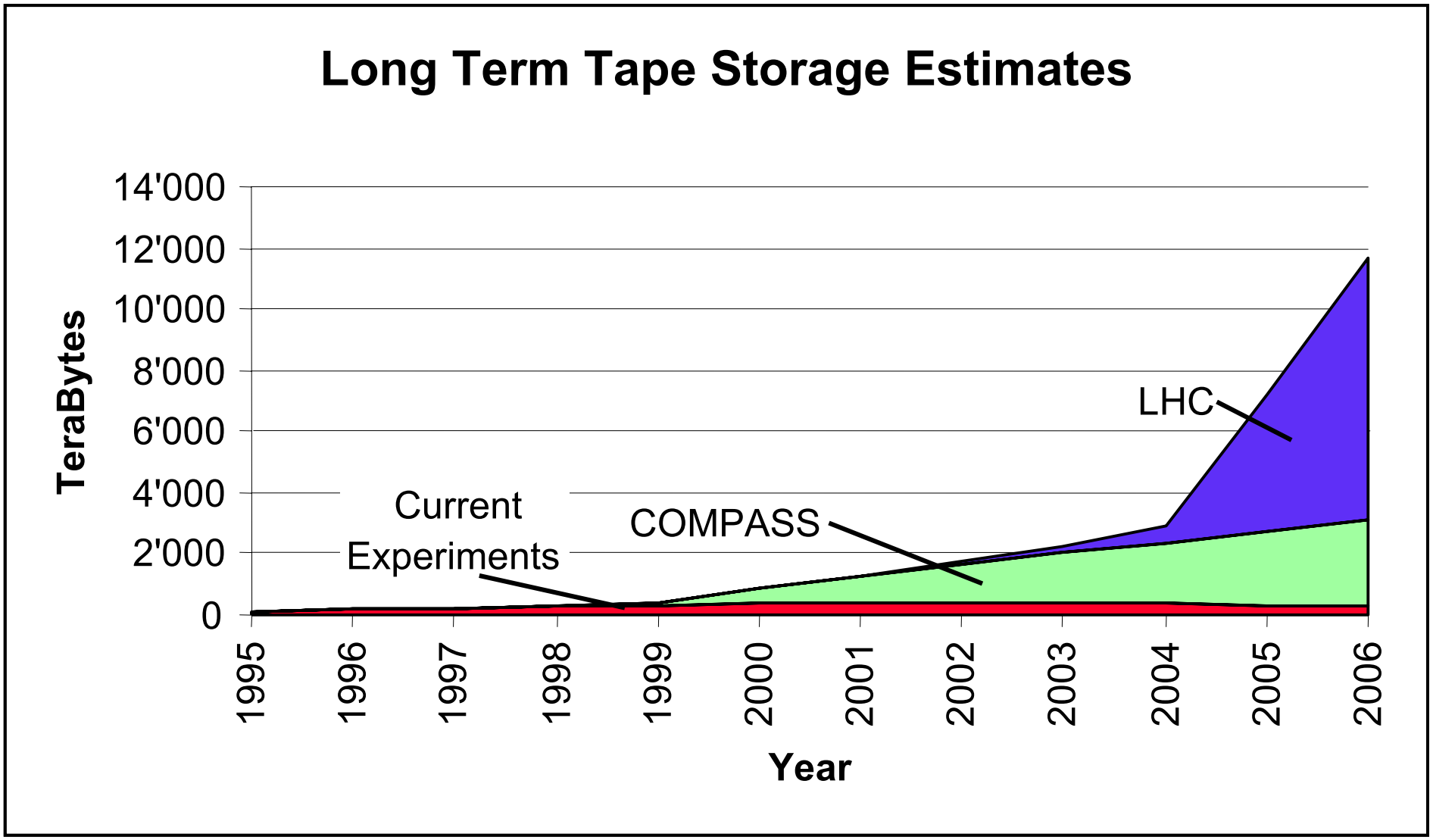
Offline Data and Computation for Physics Analysis

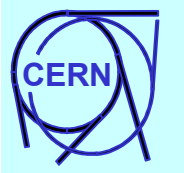




Capacity that can be purchased for the value of the equipment present in 2000


Estimated DISK Capacity ay CERN
TeraBytes
1800
1600
1400
1200
1000
800
600
400
200
0
1998
1999
2000
2001
2002
2003
2004
2005
2006
year
Non-LHC
LHC
technology-price curve (40% annual price improvement)


Long Term Tape Storage Estimates
TeraBytes
14'000
12'000
10'000
8'000
6'000
4'000
2'000
0
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
Year
LHC
COMPASS
Current
Experiments



HPC or HTC

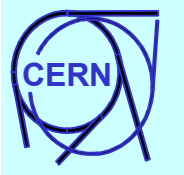
High Throughput Computing

- mass of modest problems
- throughput rather than performance
- resilience rather than ultimate reliability

Can exploit inexpensive mass market components

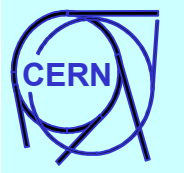
But we need to marry these with inexpensive highly scalable management tools

Much in common with data mining, Internet computing facilities, ……



History-1

- 1960s through 1980s
  - The largest scientific mainframes (Control Data, Cray, IBM, Siemens/Fujitsu)
  - Time-sharing interactive services on IBM & DEC-VMS
  - Scientific workstations from 1982 (Apollo) for development, final analysis
- 1988 -- On-line computing farms (Falcon) - joint project with Digital (microVax and Vaxstations)
- 1989 -- First batch services on RISC - joint project with HP (Apollo DN10.000)
- 1990 -- Central Simulation Facility (CSF) - 4 X mainframe capacity
- 1991 -- SHIFT - data intensive applications, distributed model
- 1993 -- First central interactive service on RISC



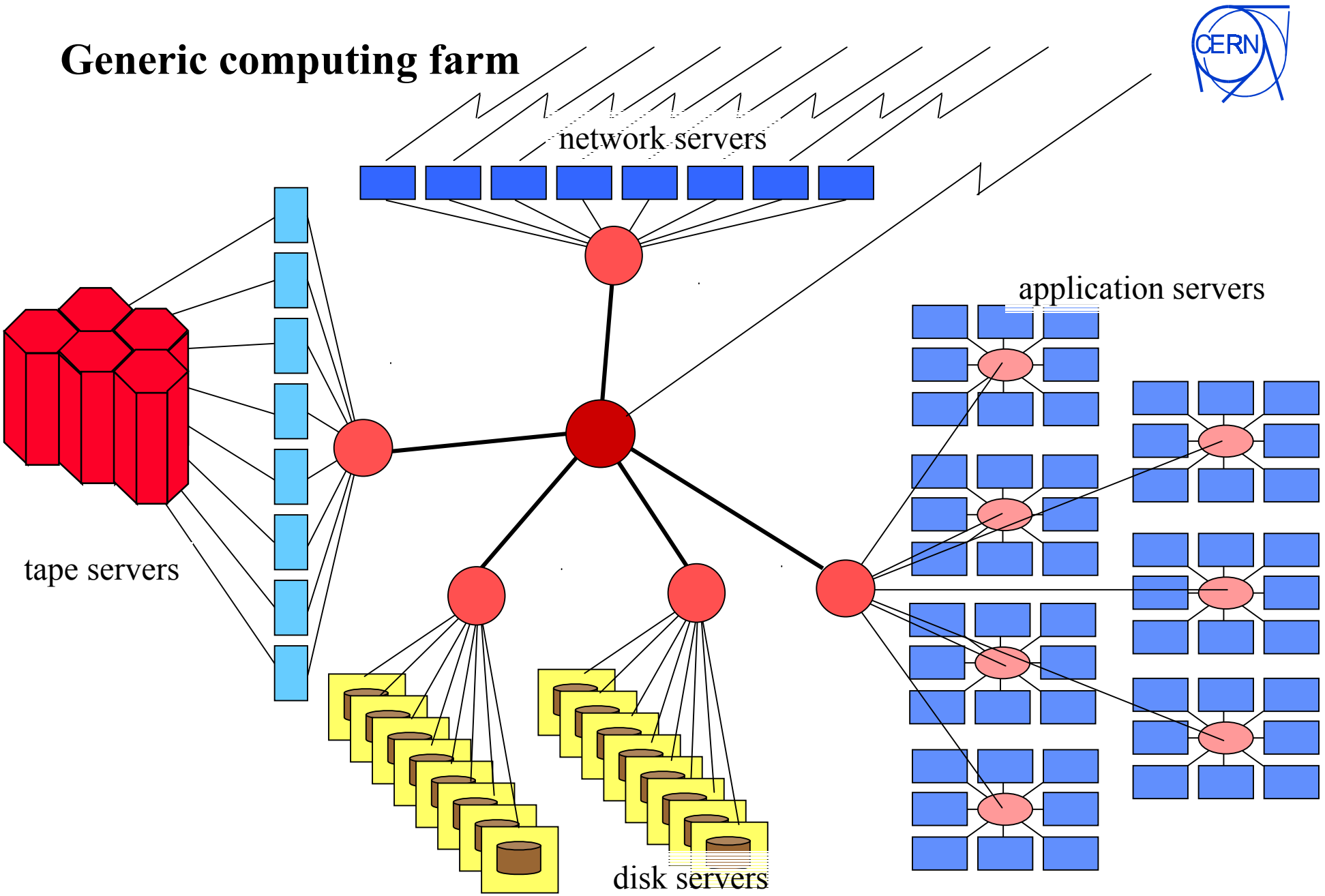
History-2

- 1994 -- 128 processor QSW (Meiko/QSW) CS2 and 72 processor IBM SP-2
- 1996 -- Last mainframe de-commissioned
- 1997 -- First batch services on PCs
- 1998 -- NA48 record 70 TeraBytes of data in one year

LHC Computing Fabric —

Can we scale up the current commodity-component based approach?

Generic computing farm



Standard components

Computing & Storage Fabric

built up from commodity components

- Simple PCs
- Inexpensive network-attached disk
- Standard network interface
  (whatever Ethernet happens to be in 2006)

with a minimum of high(er)-end components

- LAN backbone
- WAN connection

HEP's not special, just more cost conscious

Computing & Storage Fabric

built up from commodity components

- Simple PCs
  - Inexpensive network-attached disk
  - Standard network interface

with a minimum of high(er)-end components

- LAN backbone
- WAN connection

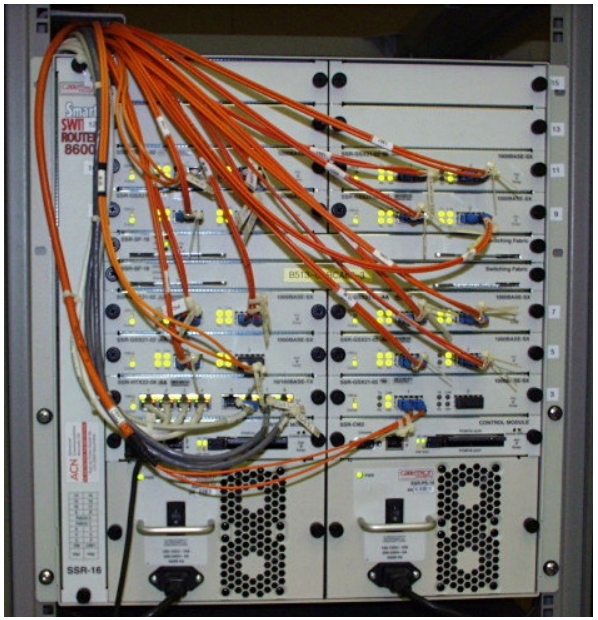
Limited role of high end equipment

Computing & Storage Fabric

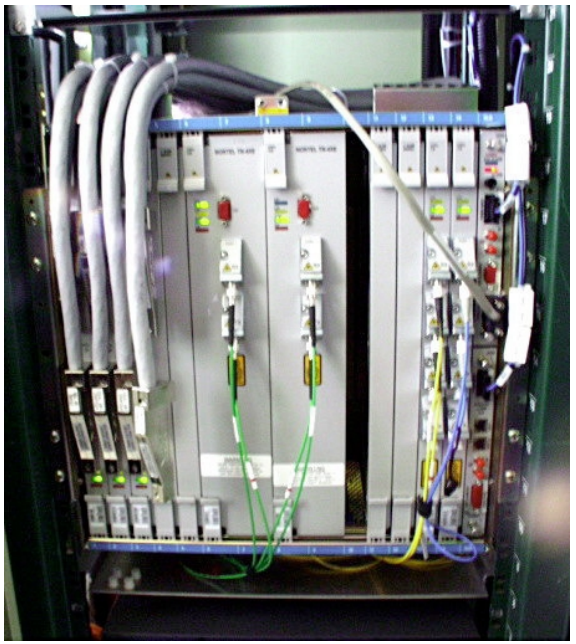
built up from commodity components

- Simple PCs
- Inexpensive network-attached disk
- Standard network interface
  (whatever Ethernet happens to be in 2006)

with a minimum of high(er)-end components



LAN backbone

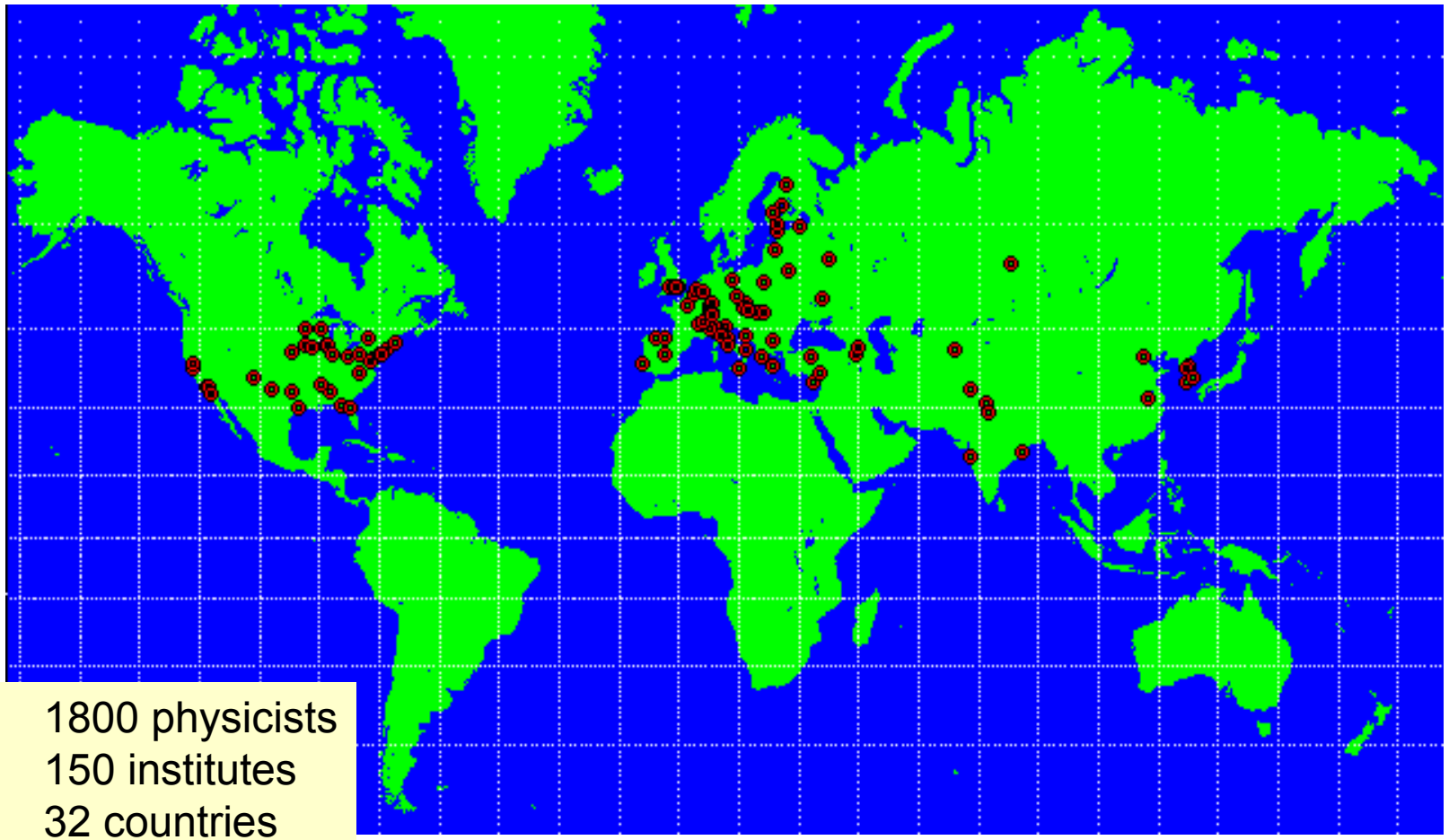


WAN connection

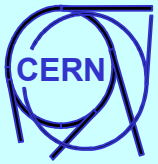
Not everything has been commoditised yet

World Wide Collaboration

⇒ distributed computing & storage capacity



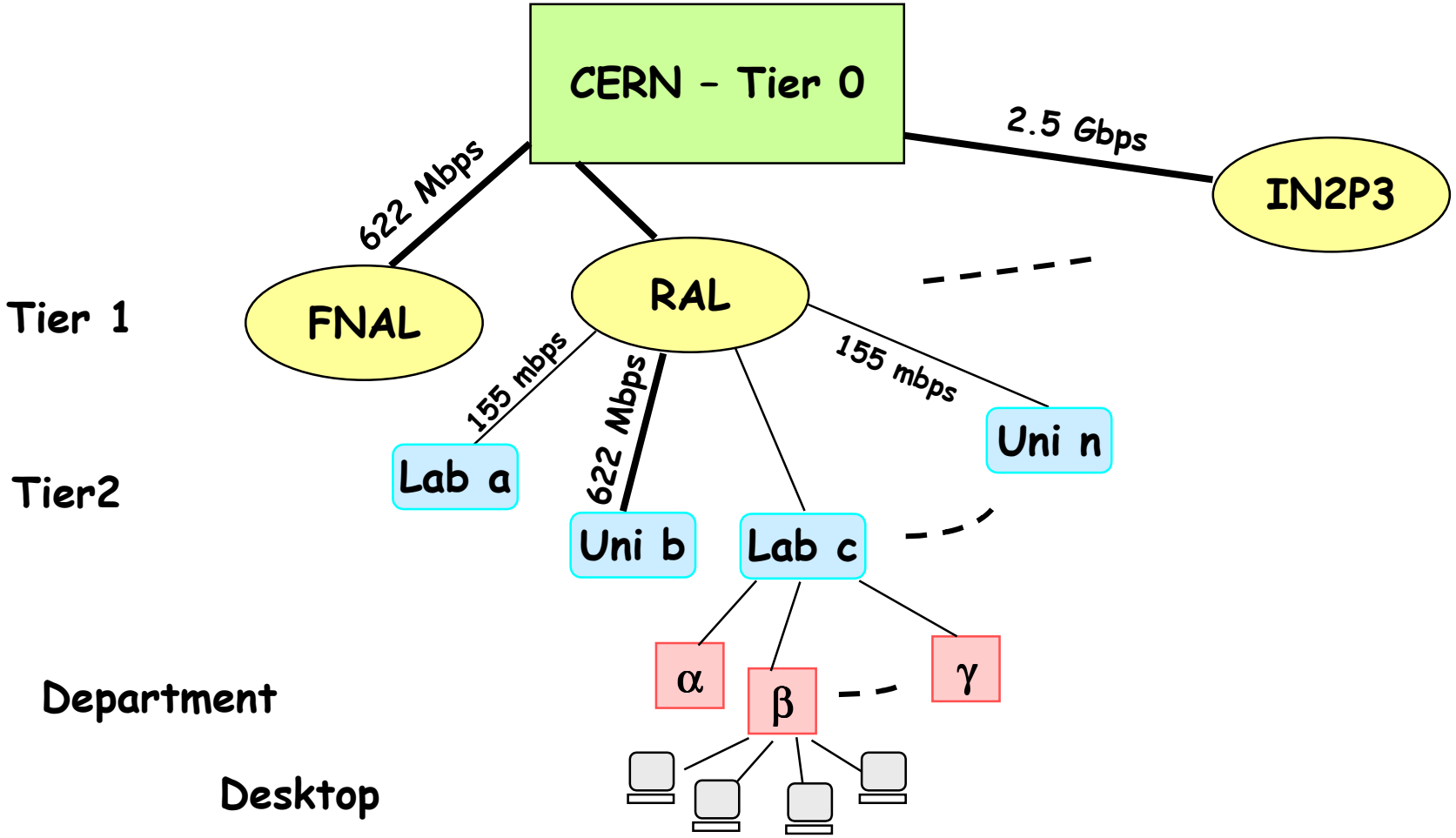
CMS: 1800 physicists
150 institutes
32 countries



Regional Computing Centres

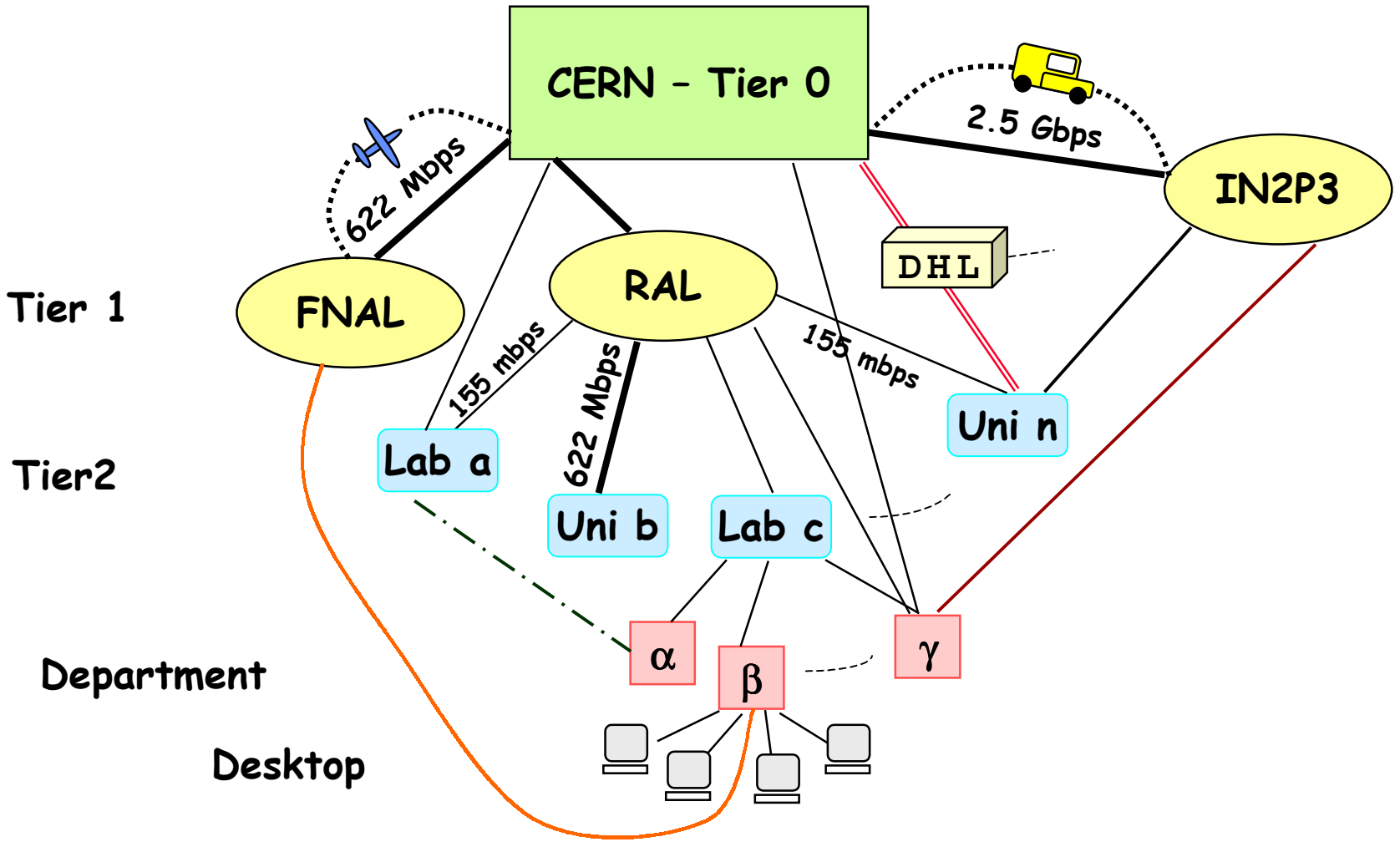
- Exploit established computing expertise & infrastructure
  - In national labs, universities
- Reduce dependence on links to CERN
  - full ESD available nearby - through a fat, fast, reliable network link
- Tap funding sources not otherwise available to HEP

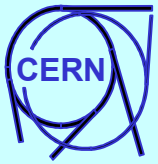
Regional Centres - a Multi-Tier Model



MONARC report: http://home.cern.ch/~barone/monarc/RCArchitecture.html

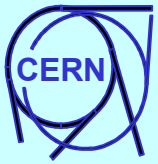
More realistically - a Grid Topology





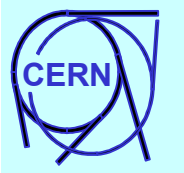
Summary - the basic problem

- **Scalability**
  - Thousands of processors, thousands of disks, PetaBytes of data, Terabits/second of I/O bandwidth, ….
- **Wide-area distribution**
  - WANs are and will be 1% of LANs
  - Distribute, replicate, cache, synchronise the data
  - Multiple ownership, policies, ….
  - integration of this amorphous collection of Regional Centres
  - With some attempt at optimisation
- **Adaptability**
  - We shall only know how analysis is done once the data arrives



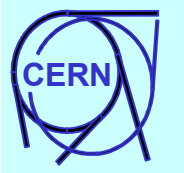
Are Grids a solution?

- Change of orientation of US Meta-computing activity
  - From inter-connected super-computers
    … .. towards a more general concept of a computational Grid (The Grid – Ian Foster, Carl Kesselman)
- Has initiated a flurry of activity in HEP
  - US – Particle Physics Data Grid (PPDG)
  - Grid technology evaluation project in INFN
  - UK proposal for funding for a *prototype grid*
  - GriPhyN – data grid proposal just approved by NSF
  - NASA Information Processing Grid
- DataGrid initiative launched



The GRID metaphor

- Unlimited ubiquitous distributed computing
- Transparent access to multipetabyte distributed data bases
- Easy to plug in
- Hidden complexity of the infrastructure
- Analogy with the electrical power GRID



The Grid from a Services View

Applications

Chemistry

Cosmology

Environment

Biology

High Energy Physics

Application Toolkits

Distributed Computing Toolkit

Data-Intensive Applications Toolkit

Collaborative Applications Toolkit

Remote Visualization Applications Toolkit

Problem Solving Applications Toolkit

Remote Instrumentation Applications Toolkit

Grid Services (Middleware)

Resource-independent and application-independent services

authentication, authorization, resource location, resource allocation, events, accounting, remote data access, information, policy, fault detection

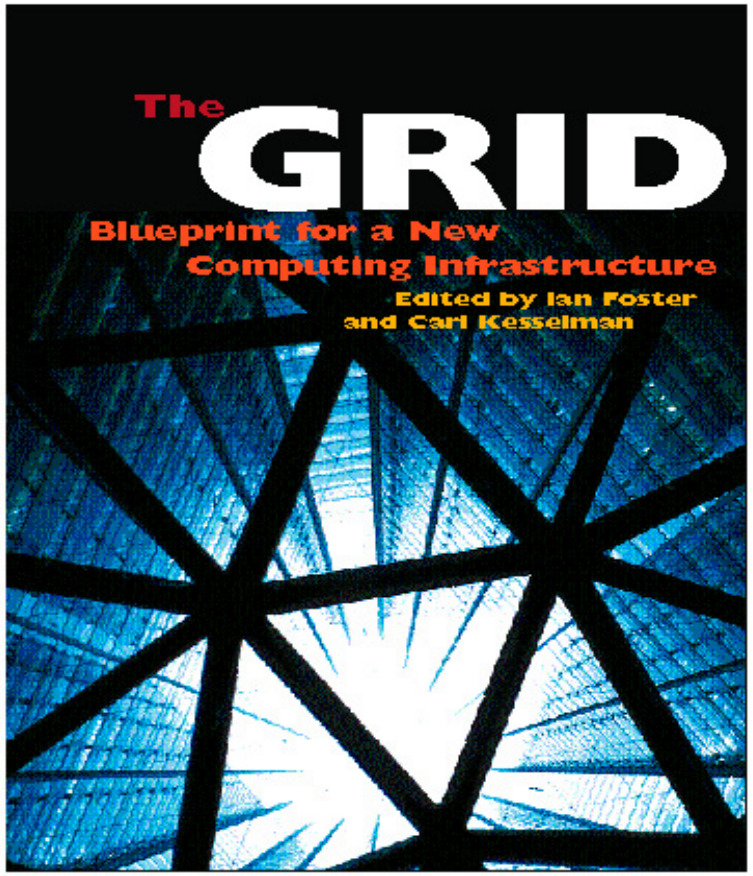
Grid Fabric (Resources)

Resource-specific implementations of basic services

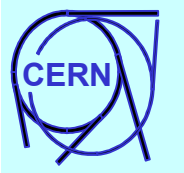
E.g., Transport protocols, name servers, differentiated services, CPU schedulers, public key infrastructure, site accounting, directory service, OS bypass

Five Emerging Models of Networked Computing From The Grid

- Distributed Computing
  - // synchronous processing
- High-Throughput Computing
  - // asynchronous processing
- On-Demand Computing
  - // dynamic resources
- Data-Intensive Computing
  - // databases
- Collaborative Computing
  - // scientists



Ian Foster and Carl Kesselman, editors, "The Grid: Blueprint for a New Computing Infrastructure," Morgan Kaufmann, 1999, http://www.mkp.com/grids



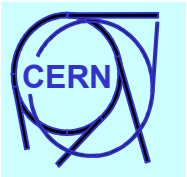
R&D required

Local fabric

- Management of giant computing fabrics
  - auto-installation, configuration management, resilience, self-healing
- Mass storage management
  - multi-PetaByte data storage, "real-time" data recording requirement, active tape layer – 1,000s of users

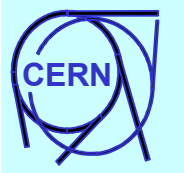
Wide-area - *building on an existing framework & RN (e.g.Globus, Geant and high performance network R&D)*

- workload management
  - no central status
  - local access policies
- data management
  - caching, replication, synchronisation
  - object database model
- application monitoring



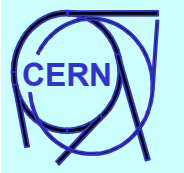
HEP Data Grid Initiative

- European level coordination of national initiatives & projects
- Principal goals:
  - Middleware for fabric & Grid management
  - Large scale testbed - major fraction of one LHC experiment
  - Production quality HEP demonstrations
    - "mock data", simulation analysis, current experiments
  - Other science demonstrations
- Three year phased developments & demos
- Complementary to other GRID projects
  - *EuroGrid:* Uniform access to parallel supercomputing resources
- Synergy being developed (GRID Forum, Industry and Research Forum)



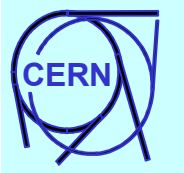
Participants

- Main partners: CERN, INFN(I), CNRS(F), PPARC(UK), NIKHEF(NL), ESA-Earth Observation
- Other sciences: KNMI(NL), Biology, Medicine
- Industrial participation: CS SI/F, DataMat/I, IBM/UK
- Associated partners: Czech Republic, Finland, Germany, Hungary, Spain, Sweden (mostly computer scientists)
- Formal collaboration with USA established
- Industry and Research Project Forum with representatives from:
  - Denmark, Greece, Israel, Japan, Norway, Poland, Portugal, Russia, Switzerland



Status

- Prototype work already started at CERN and in most of collaborating institutes (Globus initial installation and tests)
- Proposal to the EU positively reviewed at the end of July, 9.8 M Euros (covering 1/3 of total investment), 3 years contract being negotiated now
- Expect start of the project, January next year



Conclusions

- The Grid is a useful metaphor to describe an appropriate computing model for LHC and future HEP computing
- Middleware, APIs and interface general enough to accommodate many different models for science, industry and commerce
- Still important R&D to be done
- Perfect field for multidisciplinary collaboration (computer science, physics and other sciences)
- If successful could develop next generation Internet computing
- Major funding agencies prepared to fund large testbeds in USA, EU and Japan
- Excellent opportunity for HEP computing to deploy a sustainable HPC model