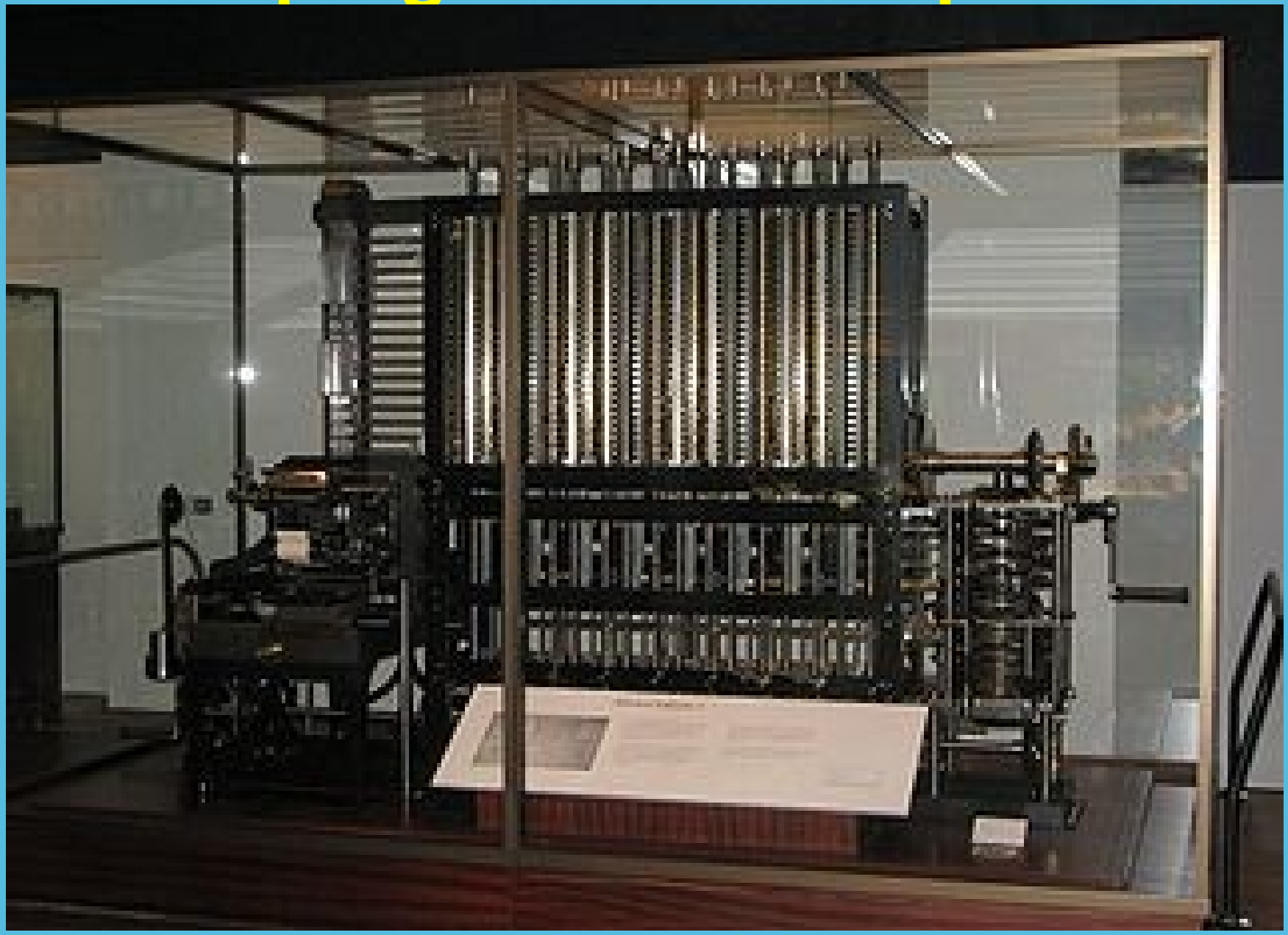
Course: Cluster, Grid, Cloud computing systems
Introduction lecture

- History of computing
- Demands for computing
  - Science, techniques, education, business, government
- **Large scale computing**
  - Equipment, programs
  - Grid and Cloud computing
  - Computing in Physics experiments
- **Scientific large scale computing**

First programmed computer



Charles Babbage (1791 - 1871) –
Inventor of first programmed computer

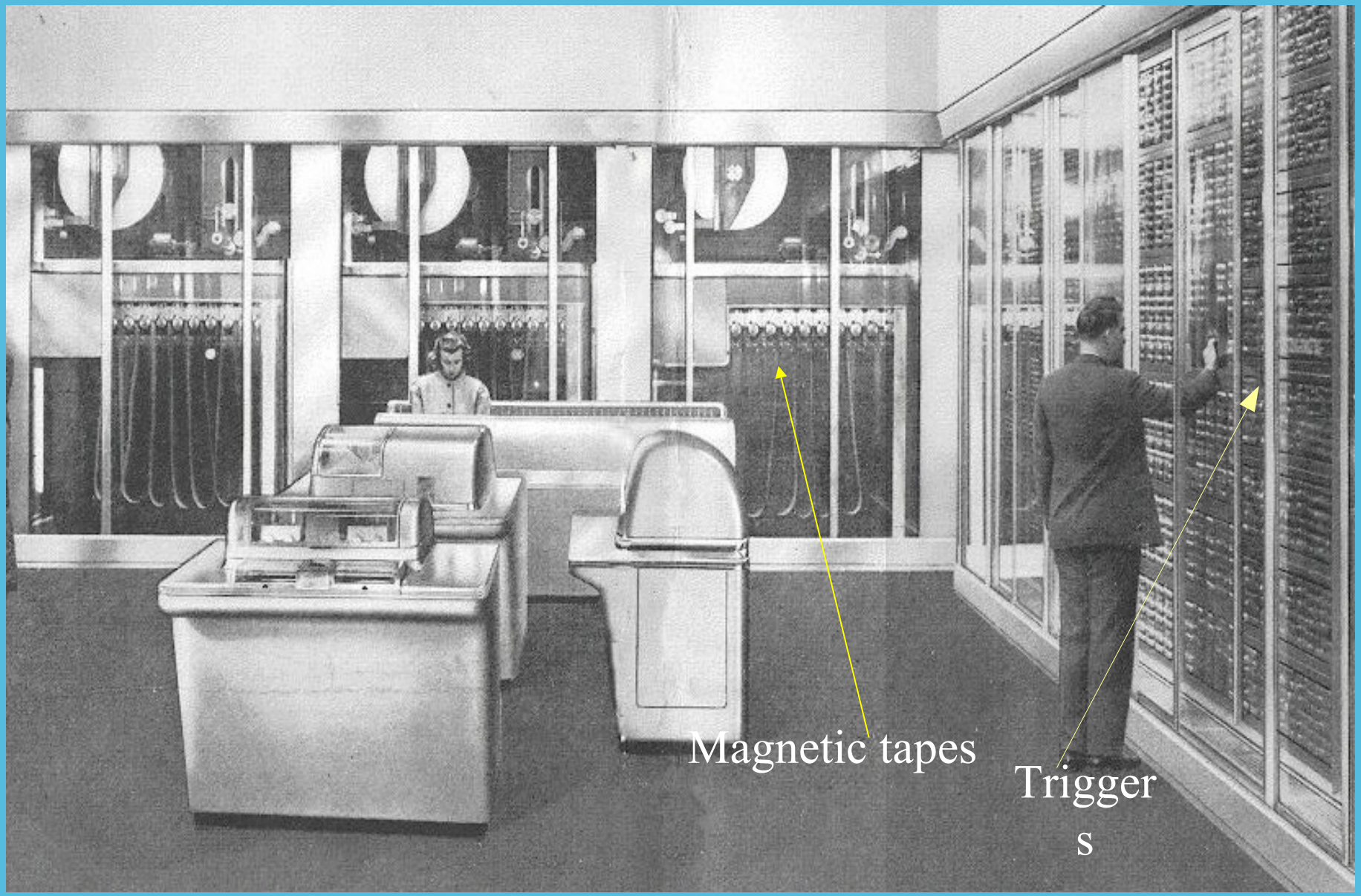
Ada Lovelace (Byron)

1815 - 1852

first programmer in the world

She (Ada) speculated that the Engine 'might act upon other things besides numbers... the Engine might compose elaborate and scientific pieces of music of any degree of complexity or extent'.

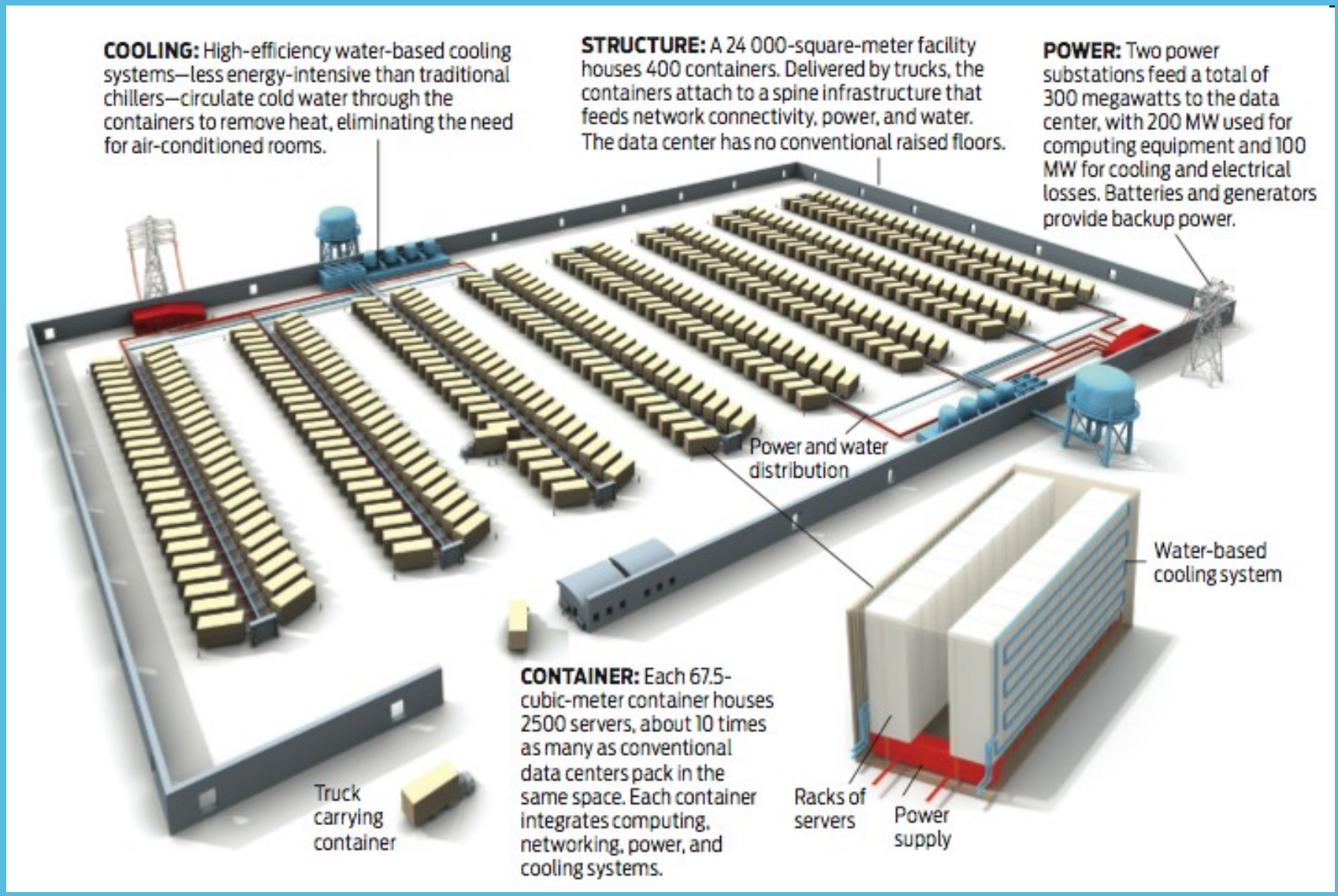
First elecronic: **Electronic Numerical Integrator And Computer (~1943)**



Modern computing systems

The main feature is using the heterogeneous architectures (servers, accelerators: GPGPU - General-Purpose Computing on Graphics Processing Units, FPGA - field-programmable gate array) and Open Source Software.

Modern big cluster plan



Large Data Centers

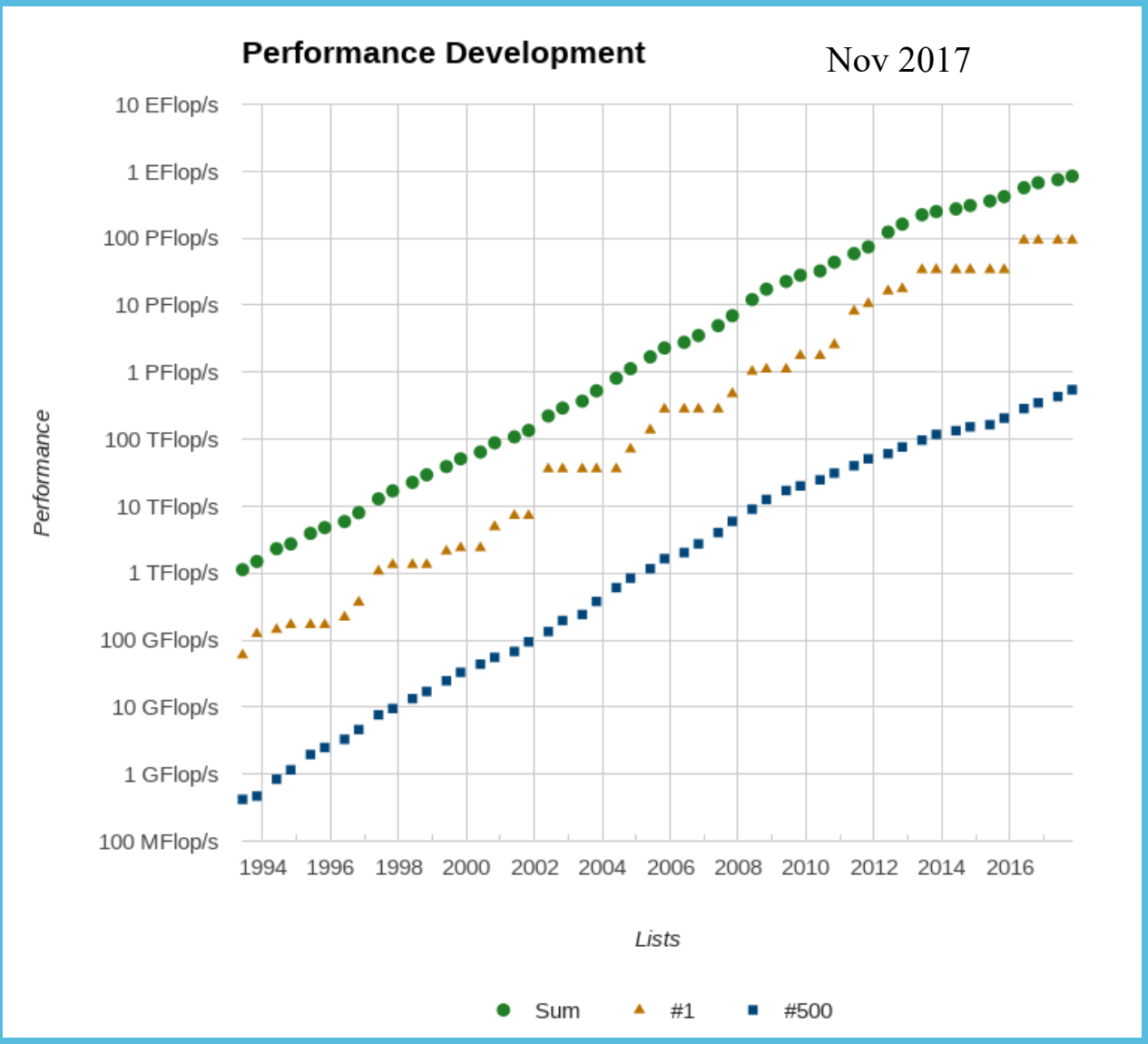
WAREHOUSE-SIZE COMPUTERS: Google has built a sprawling data center on the banks of the Columbia River, in The Dalles, Ore. The site, with two server-packed buildings and space for a third, houses tens of thousands of computers—the exact number is a closely guarded secret. Microsoft, Yahoo, and Amazon are also building data centers in the region, enticed by its readily available fiber-optic connectivity and cheap electricity. *PHOTO: MELANIE CONNER*

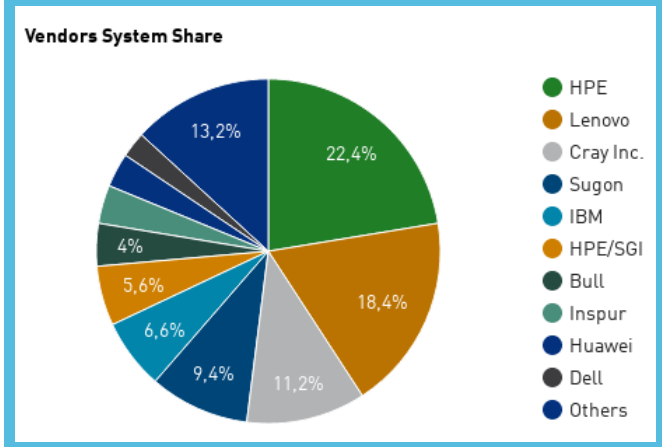
Google data center in Finland

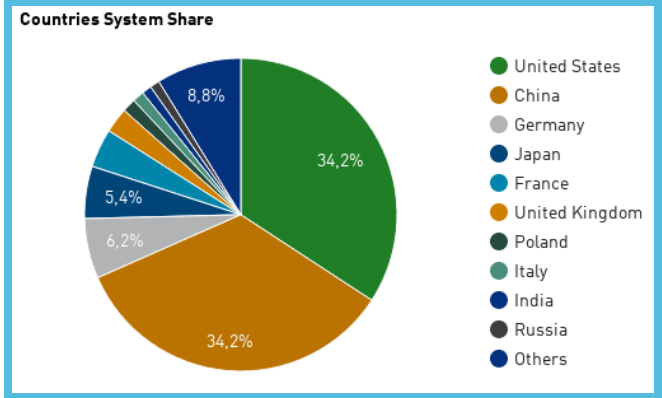
Green clusters, large clusters

Green500.org – “greenest» clusters
Top500.org – largest clusters

From site top500.org

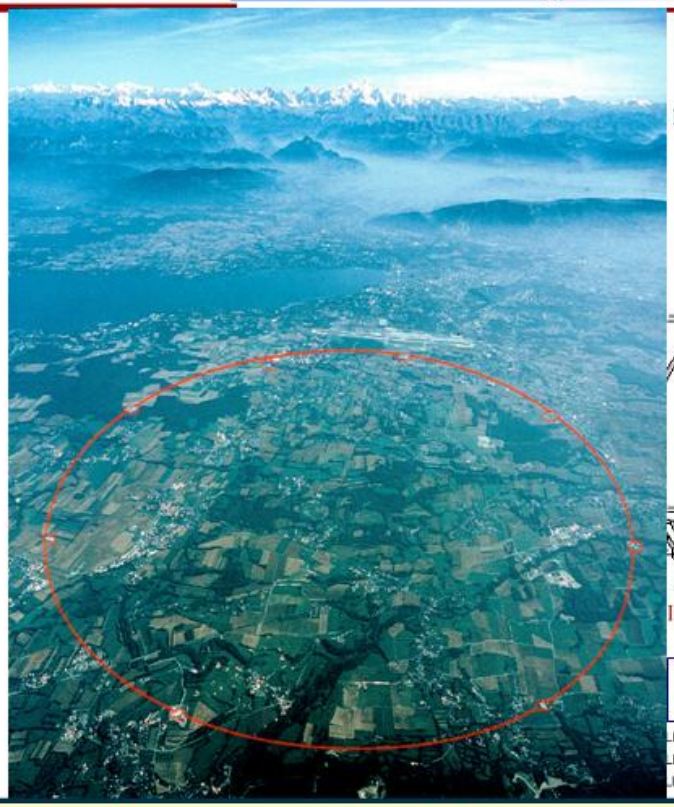




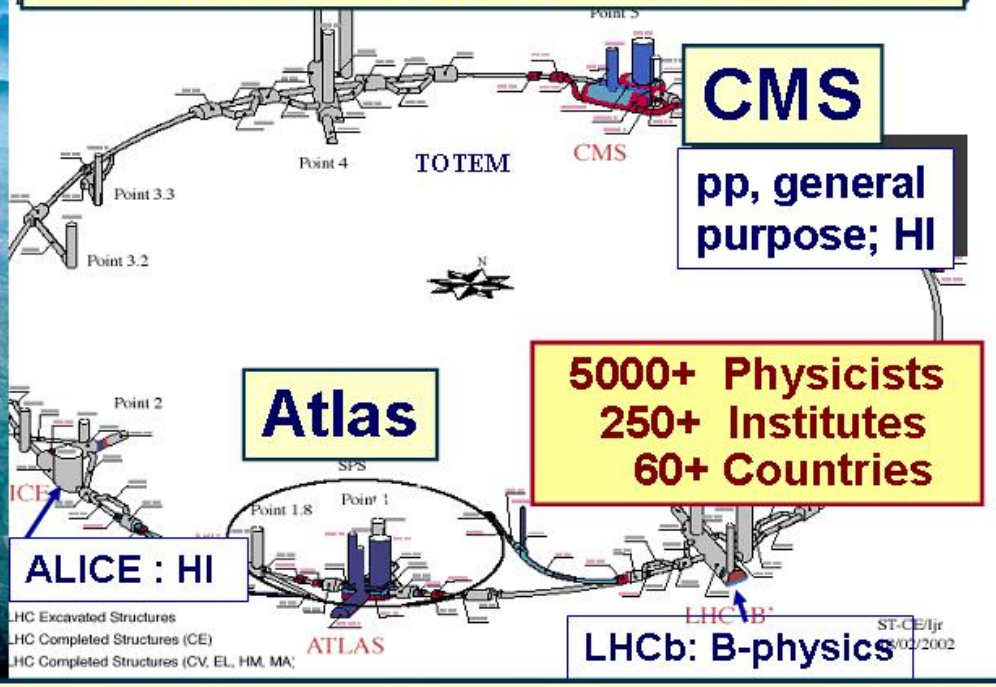


Harvey B. Newman

Large Hadron Collider
CERN, Geneva: 2007 Start



- pp $\sqrt{s}$ =14 TeV L=10^{34} cm^{-2} s^{-1}
- 27 km Tunnel in Switzerland & France

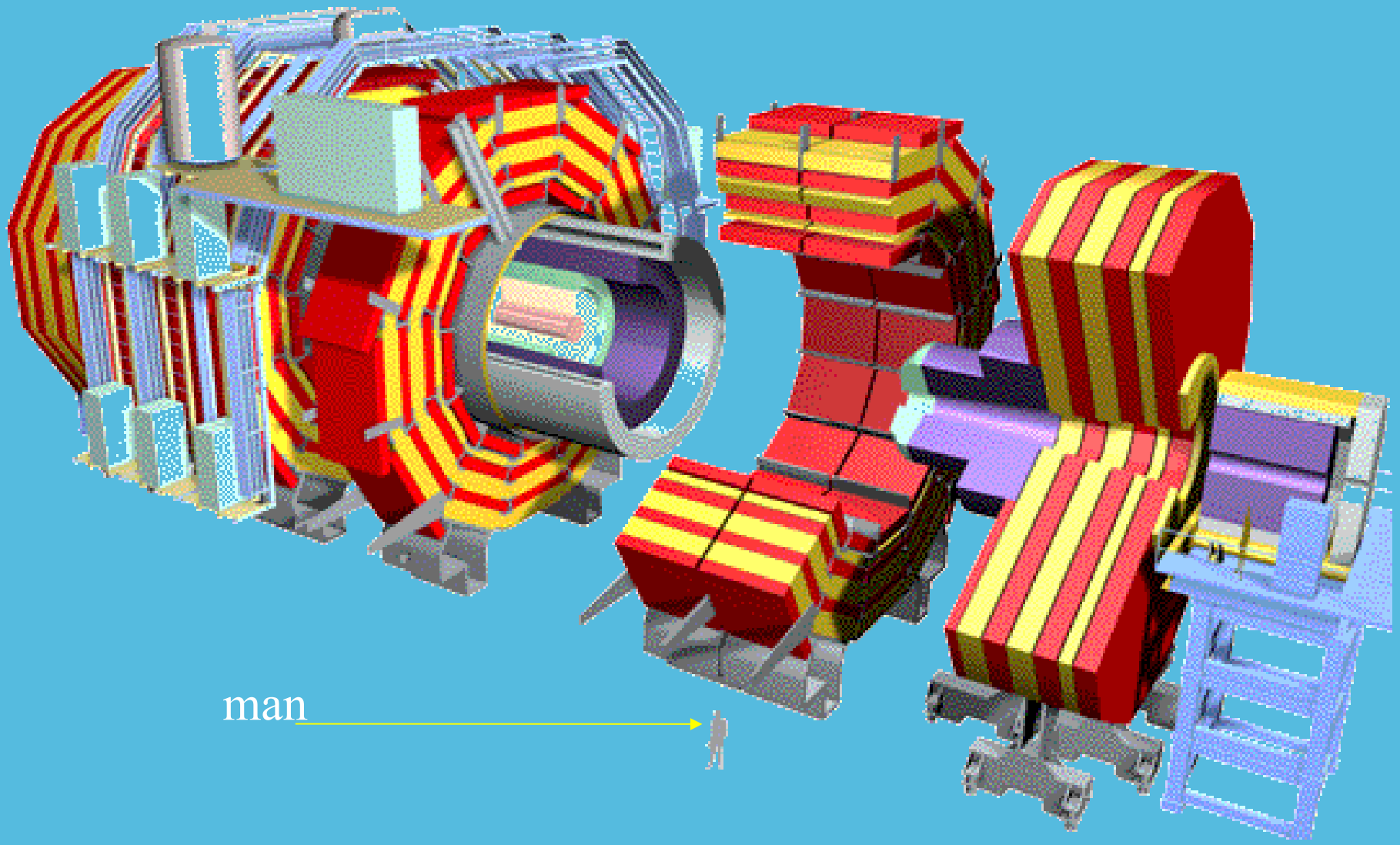


Challenges: Analyze petabytes of complex data cooperatively
Harness global computing, data & network resources

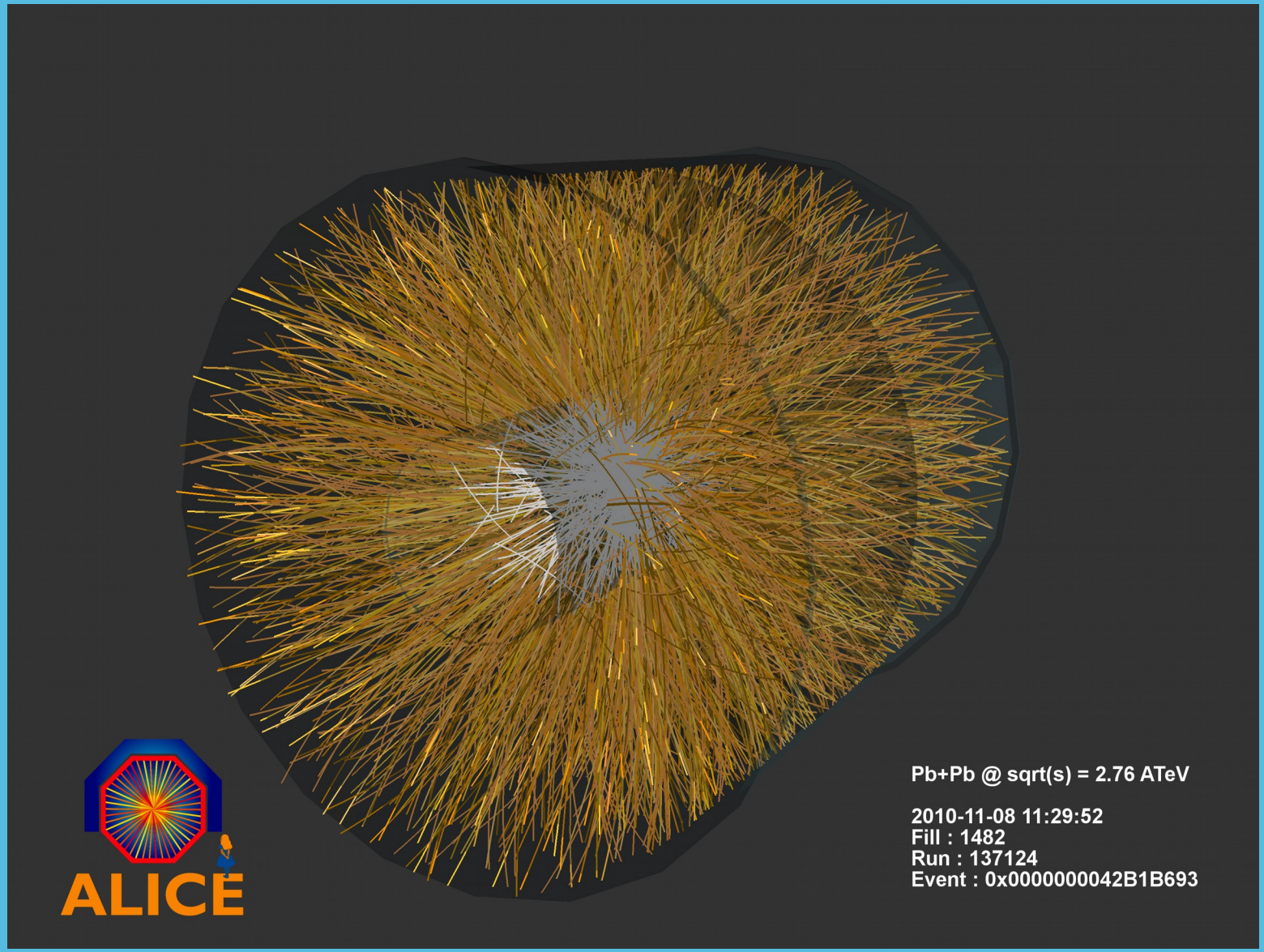
View of the tunnel at LHC

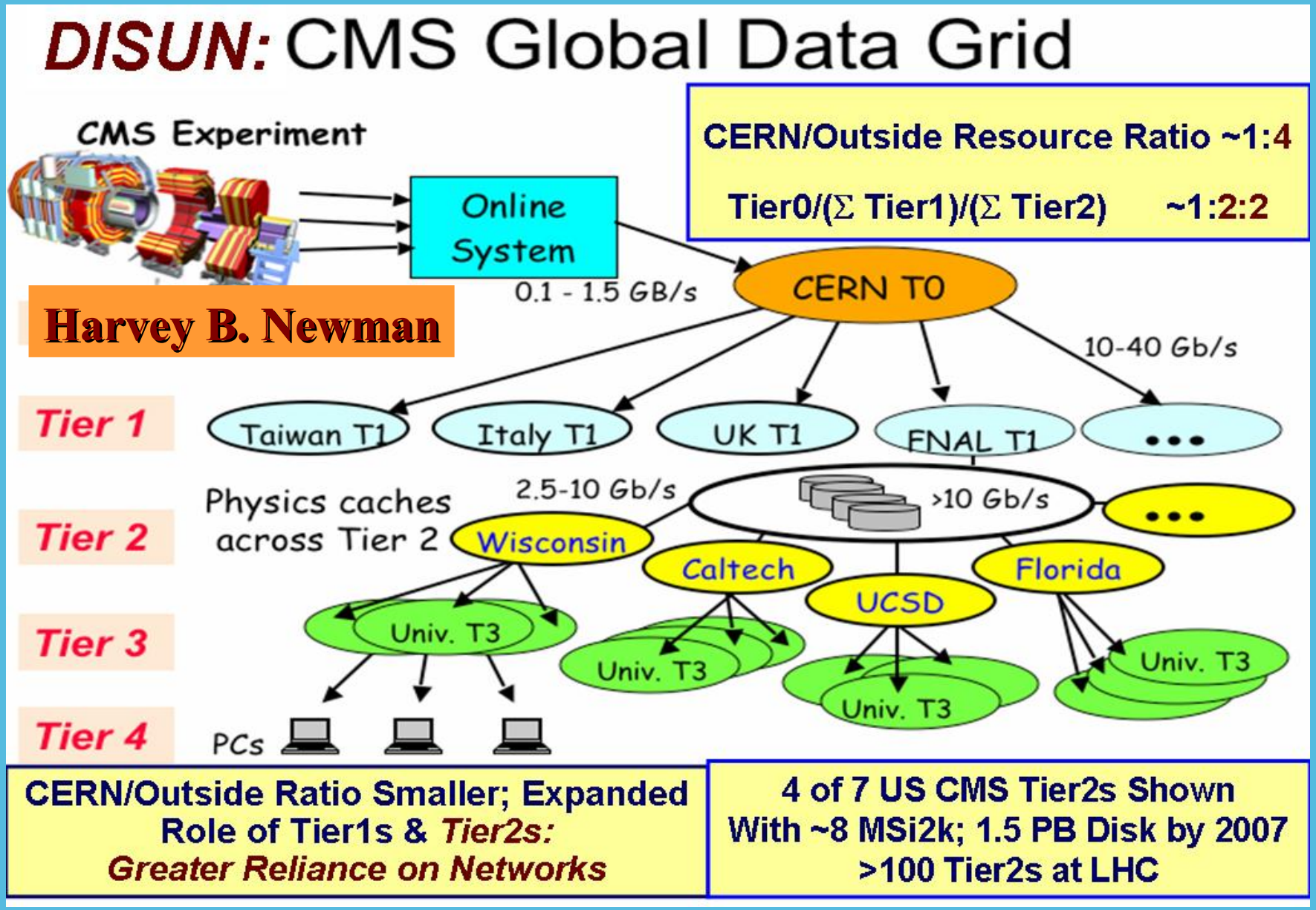

ERN

Example of the detector size at LHC



Colliding of the particles




DISUN: CMS Global Data Grid
CMS Experiment
Online System
0.1 - 1.5 GB/s
CERN/Outside Resource Ratio ~1:4
Tier0/(Σ Tier1)/(Σ Tier2) ~1:2:2
Harvey B. Newman
CERN T0
10-40 Gb/s
Tier 1
Taiwan T1
Italy T1
UK T1
FNAL T1
...
Physics caches across Tier 2
2.5-10 Gb/s
>10 Gb/s
Tier 2
Wisconsin
Caltech
UCSD
Florida
Tier 3
Univ. T3
Univ. T3
Univ. T3
Univ. T3
Univ. T3
Tier 4
PCs
CERN/Outside Ratio Smaller; Expanded Role of Tier1s & Tier2s: Greater Reliance on Networks
4 of 7 US CMS Tier2s Shown With ~8 MSi2k; 1.5 PB Disk by 2007 >100 Tier2s at LHC

Grid and Clouds

Grid — to gather fraction of computing power from different organizations and different clusters in order to accumulate large computing resource.

Cloud — to share large computing resources between many users without system administrator intervention. From user point of view — to get required resources by clicking on web panel.

Conclusion

- Modern features of the computing systems
  - Heterogeneous computing clusters;
  - Free and Open Source Software as important factor;
  - Virtualization;
  - Grid and Cloud architectures.

End of Lecture

Is it possible to take participation in looking the features of the matter in LHC ?

- YES – pls see http://lhcathome.web.cern.ch/LHCathome/