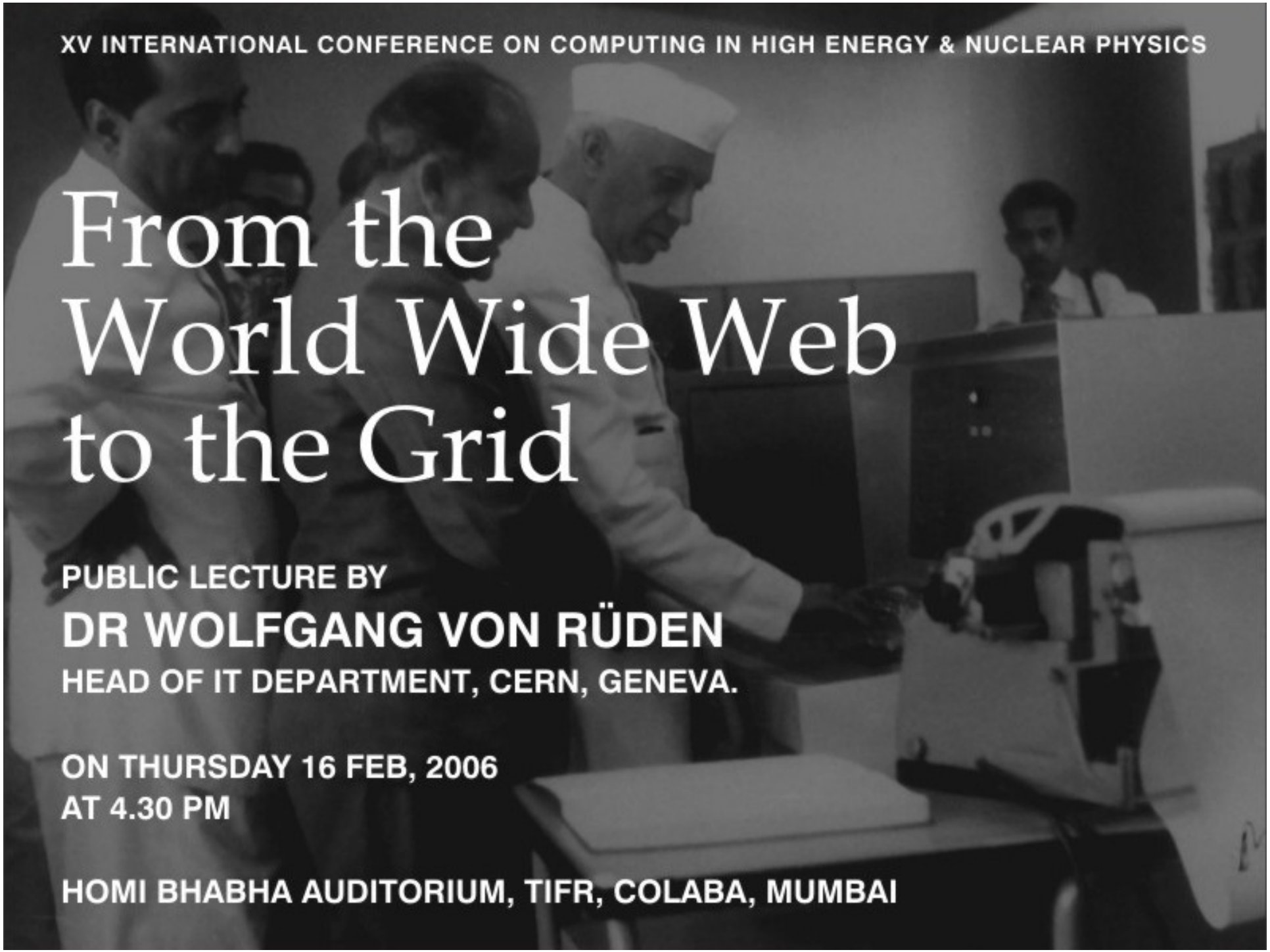

XV INTERNATIONAL CONFERENCE ON COMPUTING IN HIGH ENERGY & NUCLEAR PHYSICS
From the World Wide Web to the Grid
PUBLIC LECTURE BY
DR WOLFGANG VON RÜDEN
HEAD OF IT DEPARTMENT, CERN, GENEVA.
ON THURSDAY 16 FEB, 2006
AT 4.30 PM
HOMI BHABHA AUDITORIUM, TIFR, COLABA, MUMBAI

TIFRAC – The Tata Institute of Fundamental Research Automatic Calculator

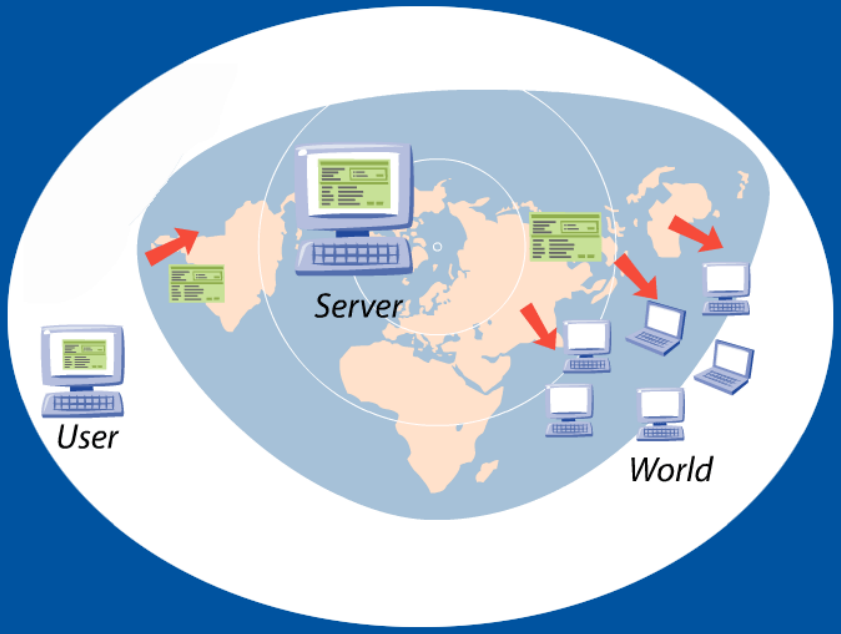
The first full-scale, general purpose, electronic digital computer designed and built in India, the TIFRAC was completed in 1959 and commissioned in February 1960. The computer was designed by R. Narasimhan and built by a team of six people. TIFRAC was named in 1962, when India's first Prime Minister, Jawaharlal Nehru, inaugurated the new buildings of the TIFR.

TIFRAC Characteristics

- Pilot model, completed in November 1956:
  - Ferrite core memory of 256 words, word length of 12 bits
  - CPU was parallel, asynchronous, fixed point, single address
  - I/O via paper tape and Teletype
  - Total power consumption of 10kW
- Production Model:
  - 2700 vacuum tubes, 1700 Germanium Diodes, 12,500 Resistors
  - Ferrite core memory of 2048 words, 15μs cycle time, 40 bit word length (better than the IBM 701)
  - I/O via paper tape and teletype, later completed by CRT display and tape storage
  - Total power consumption of 20kW
- Operation:
  - Used from 1960 to 1964 by many research scientists all over India

Note: none of the team, apart from Narasimhan, had had any previous experience of using or operating a computer, let alone building one.

From the Web to the Grid



Public lecture by Dr. Wolfgang von Rüden
CERN - Geneva (Switzerland)

Special thanks to R. Mondardini and F. Grey

What I will talk about:

- What is the Internet?
- What is the Web?
- Why was it invented at CERN?
- What is the Grid?
- Why is the High Energy Physics community developing the Grid?
- Will the Grid be available to all of you?


What is the Internet?

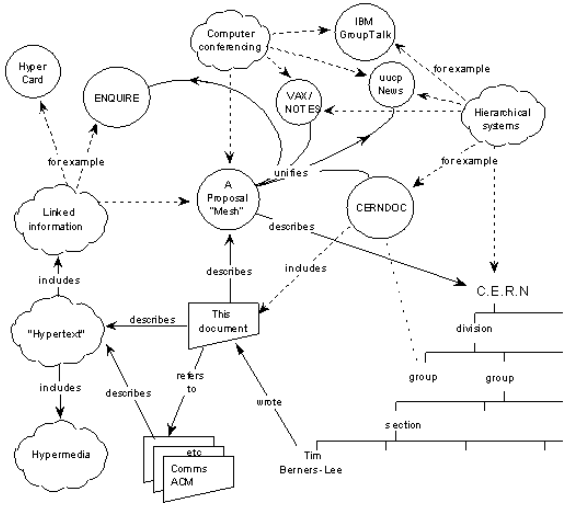

What is the Web?
Server
User
World

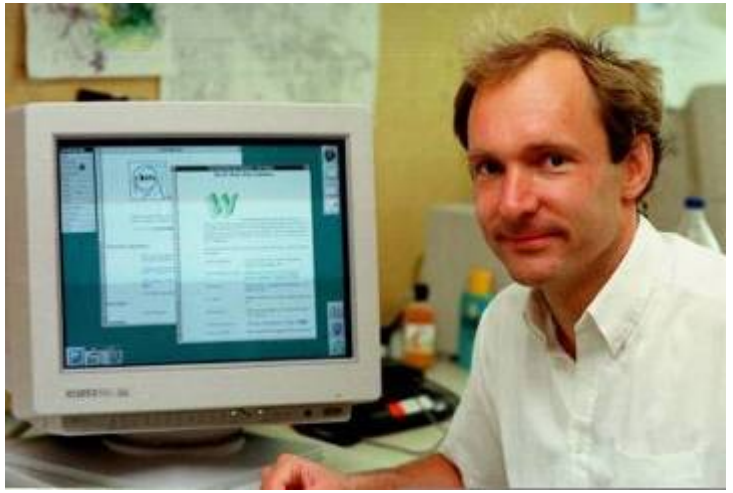
Why was the Web invented at CERN?

- Science depends on free **access to information** and exchange of ideas. CERN is the hub of a worldwide community of **6500 scientists** in **80 countries.**
- CERN has a long history of being at the forefront of **scientific computing** and **networking** (first lab on Internet outside the US).
- During the preparation of the previous large project LEP, the need to share documents in a global way became vital.

How did it start?

- 1989: **Tim Berners-Lee** circulates “Information Management: A proposal” to help with future Large Hadron Collider project.
- 1991: Early **www system released** to high energy physics via the CERN program library. First web servers located in European physics laboratories.
- 1993: First Mosaic browser; web reaches 500 servers and 1% of Internet traffic; CERN places **WWW in the public domain**.

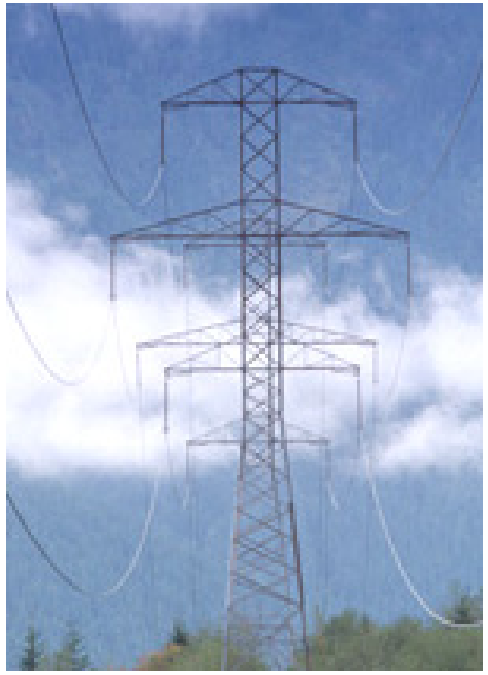





What is the Grid?

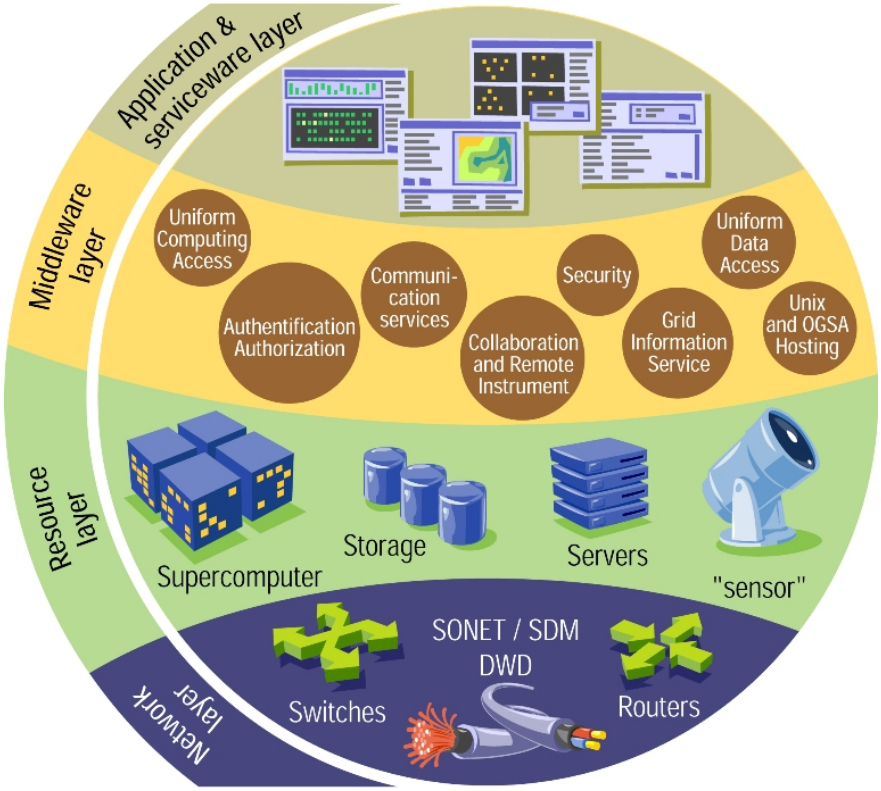
Grid history

- Name “Grid” chosen by analogy with electric power grid (Foster and Kesselman 1997)
- Vision: plug-in computer for processing power just like plugging in toaster for electricity.
- Concept has been around for decades (distributed computing, metacomputing)
- Key difference with the Grid is to realise the vision on a global scale.



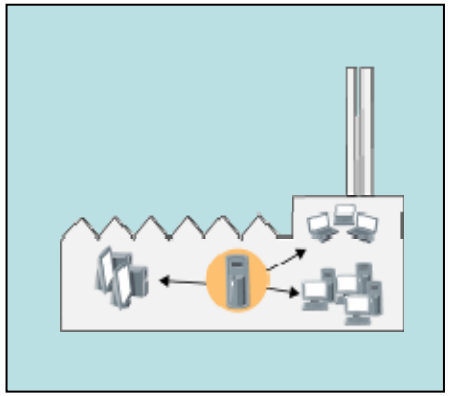
How does the Grid work?

- It relies on advanced 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.



Different Grids for different needs

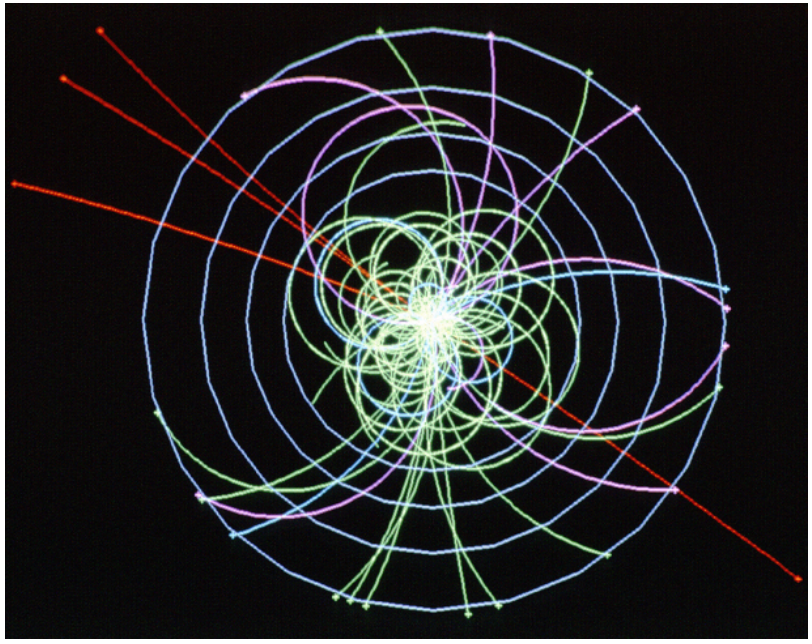
- There is as yet no unified Grid (like there is a single web) rather **there are many Grids** for many applications.
- The word Grid is used to signify different types of distributed computing for example **Enterprise Grids** (within one company) and **public resource Grids** (volunteer your own PC).
- In this talk, focus is on **scientific Grids** that link together major computing centres in research labs and universities.
- Latest trend is to **federate national Grids** to achieve a global Grid infrastructure. High Energy Physics is a driving force for this.



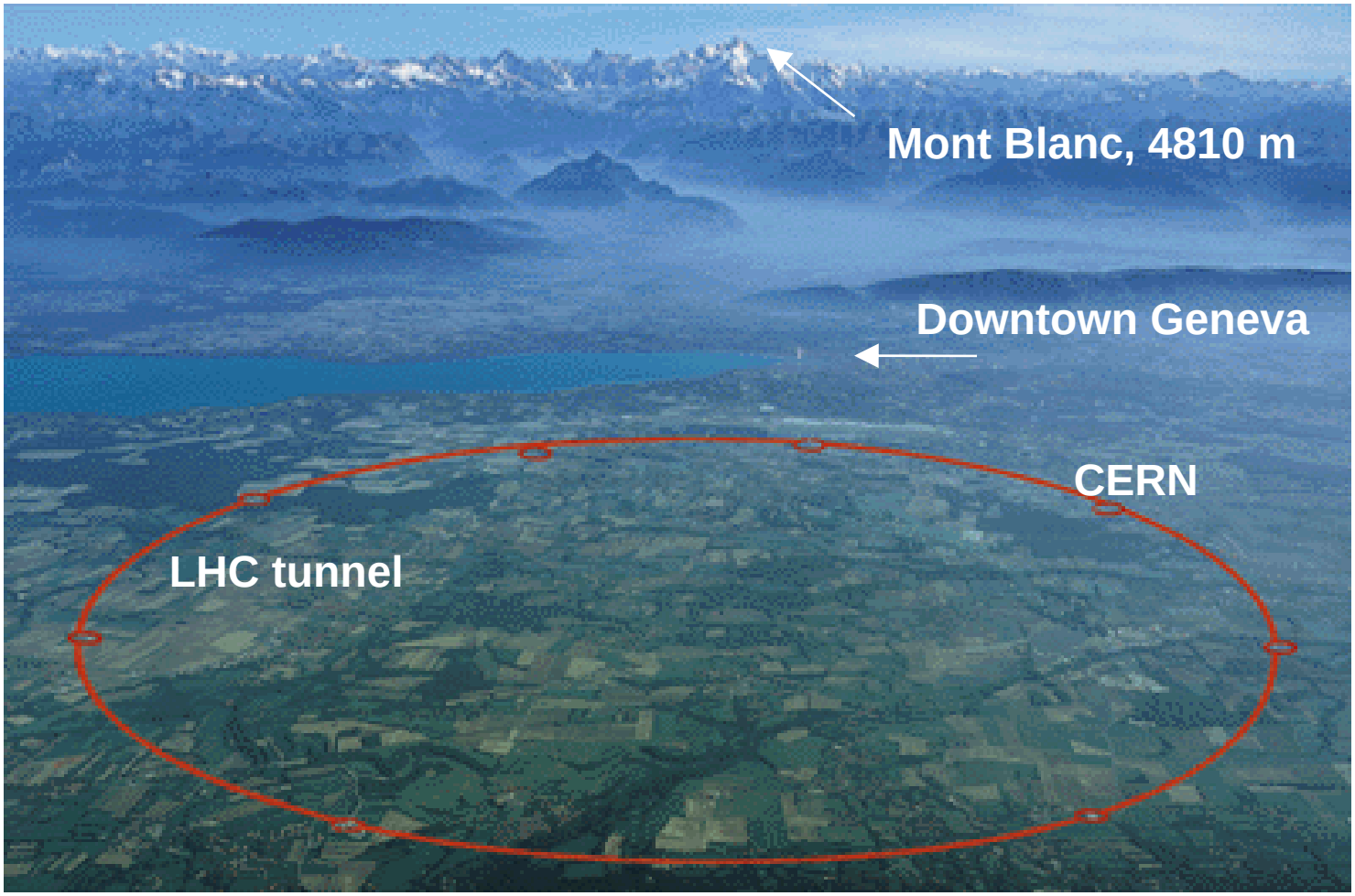

Why do
scientists
need the Grid?

Why the Grid at CERN?

- CERN is the world's largest particle physics centre, and is building the world's largest scientific instrument, the Large Hadron Collider (LHC) to answer fundamental questions about the Universe like:
  - *how did the Universe begin?*
  - *what is the origin of mass?*
  - *what is the nature of antimatter?*



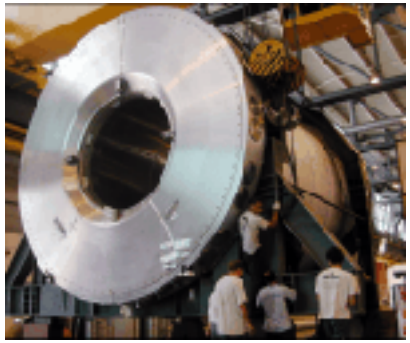
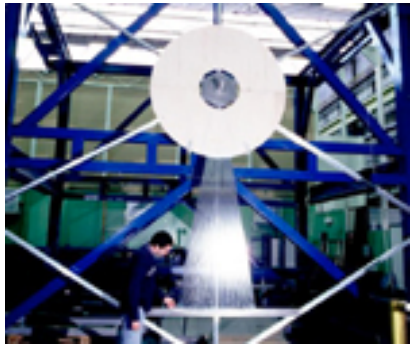
CERN Site



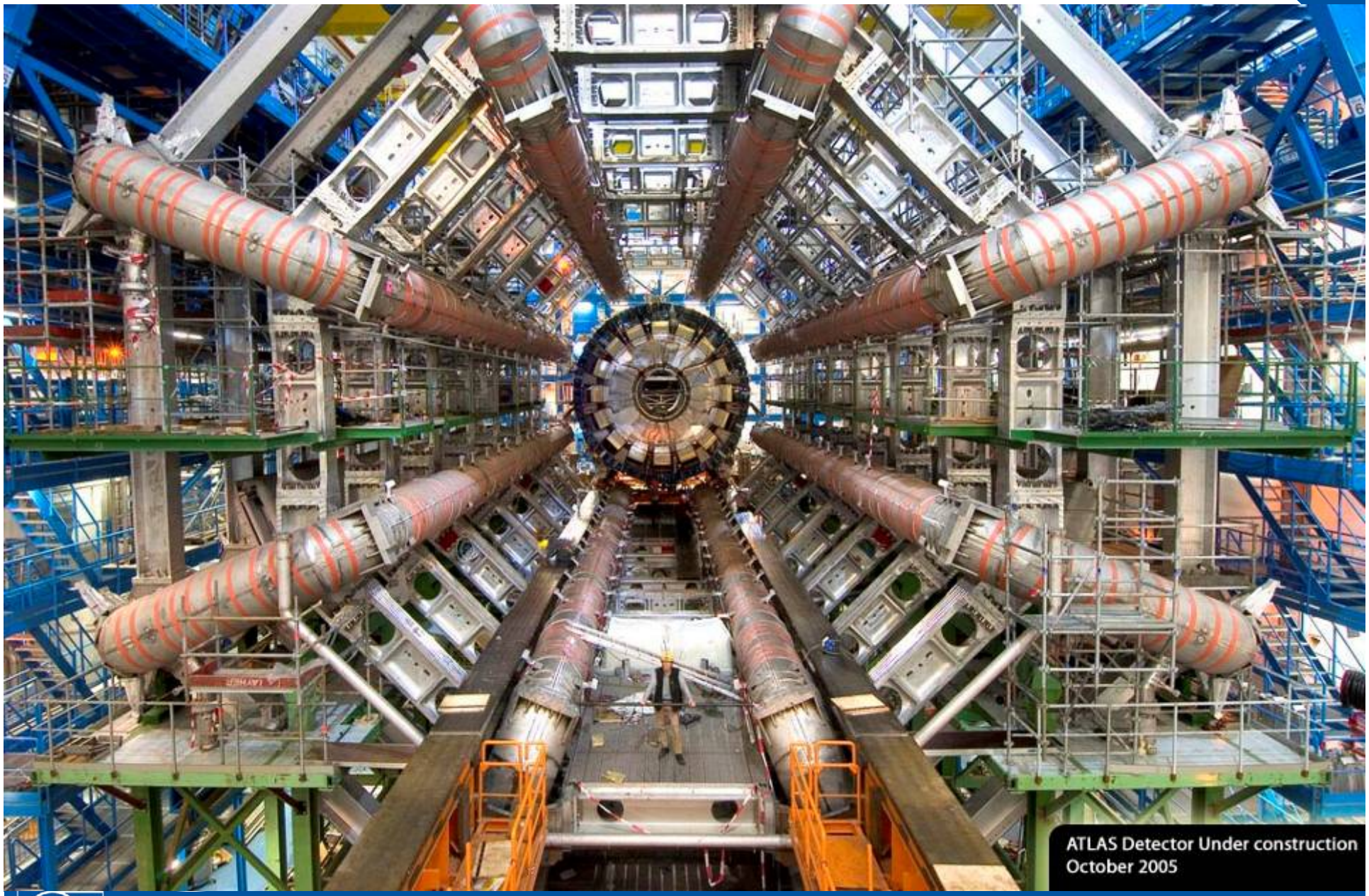
The Large Hadron Collider (LHC)

- LHC is a new particle accelerator due to be switched on in 2007, colliding beams of protons at the highest energy ever.
- Using the latest superconducting technologies, it will operate at about -270°C, just above absolute zero.
- With its 27 km circumference, the accelerator will be the largest superconducting installation in the world.
- Four experiments, with detectors as ‘big as cathedrals’: ALICE, ATLAS, CMS, LHCb.



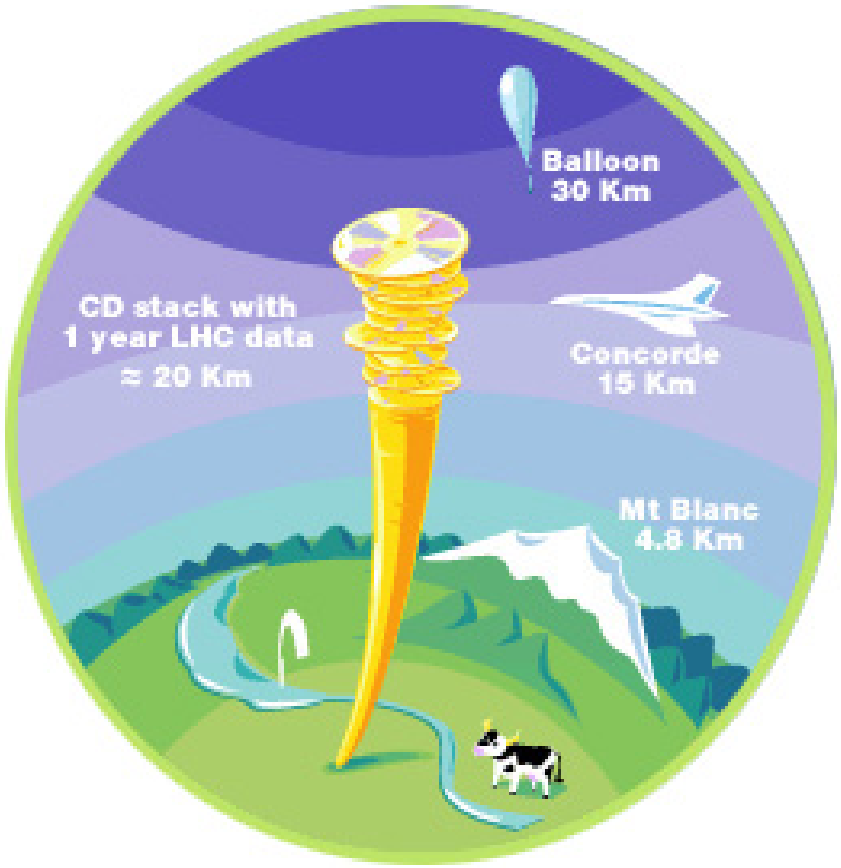



View of the ATLAS detector (under construction) along the beam axis



The LHC Data Challenge

- LHC experiments will produce **10-15 million Gigabytes** of data each year (about 20 million CDs!)
- LHC data analysis requires a computing power equivalent to **~ 100,000 of today's fastest PC processors**.
- Requires many cooperating computer centres, CERN proving only ~20% of the CPU power

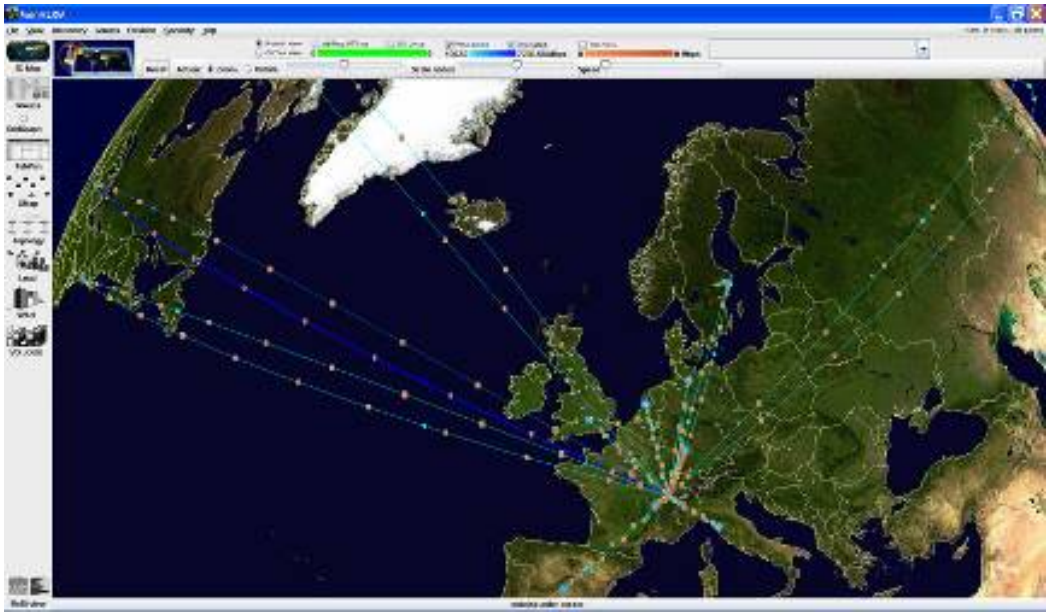


Computing @ CERN

- High-throughput computing based on reliable "commodity" technology
- More than 2500 dual processor PCs
- About 3 million Gigabytes of data on disk and tapes
- PROBLEM: **nowhere near enough!**

Grid projects

- SOLUTION: use the Grid to unite computing resources of particle physics institutes around the world.
- CERN leads two major global Grid projects:
  - **WLCG:** World-wide LHC Computing Grid Collaboration
  - **EGEE:** Enabling Grid for E-sciencE project for all sciences



- The LHC Computing Grid project launched a service with 1[illegible] in 2003. Today 200 sites in 30 countries with 16,000 PCs
- WLCG uses infrastructure of multi-science Grids [illegible] and **OSG**
- India is establishing two regional centres to [illegible] data analysis for two experiments
- EGEE launched in 2004, already s[illegible]20 applications in six scientific domains (biomedicin[illegible]ysics, quantum chemistry…)

CERN's LHC project cannot be done without the Grid

India and the LHC

- India's collaboration with CERN currently involves some 130 people.
- Indian scientists are participating in the CMS and ALICE detectors.
- Indian engineers are playing a key role in LHC magnet testing.
- Indian industry is delivering state-of-the-art equipment.
- India is a partner in developing a global Grid for the LHC.

May 25th, 2005 Visit of Dr Avul Pakir Jainulabdeen Abdul Kalam - President of India

Will the Grid be available to all of you?

- Hard to predict…

Jules Piccard, a professor at the University of Basel, installed the first telephone in the city, around 1880, between his home and his institute. He showed it proudly to other scientists and got the comment: **“Looks very good, but I doubt it will ever have any practical use”.**

"The world will only need five computers"
attributed to Thomas J. Watson, IBM

"There is absolutely no need for a computer in the home"
attributed to Ken Olsen, DEC (once a leading minicomputer manufacturer)

"640 kilobytes is all the memory you will ever need"
attributed to Bill Gates, Microsoft

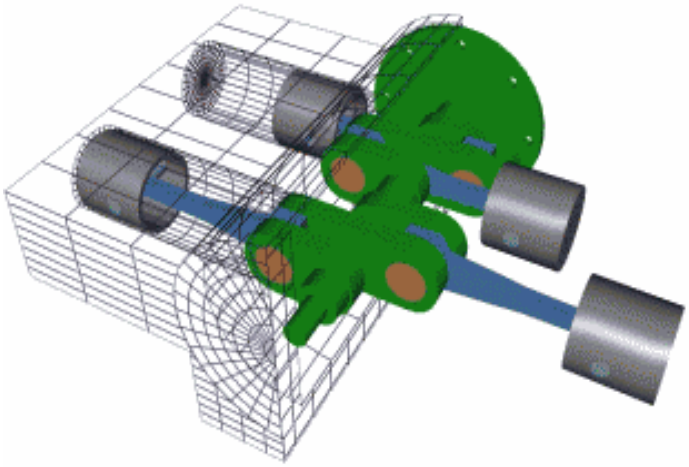
Look at the Web

- In December 2005, the one-billionth user went online!


How does Google work?
Google

Possible “first customers” for the Grid

- Disaster relief: Help government organisations share data and forecasting resources (earthquakes, floods)
- Education: Help students access large computing resources for educational purposes (shrink digital divide)
- Business: Help small companies access large computing resources for drug design, materials simulation etc.

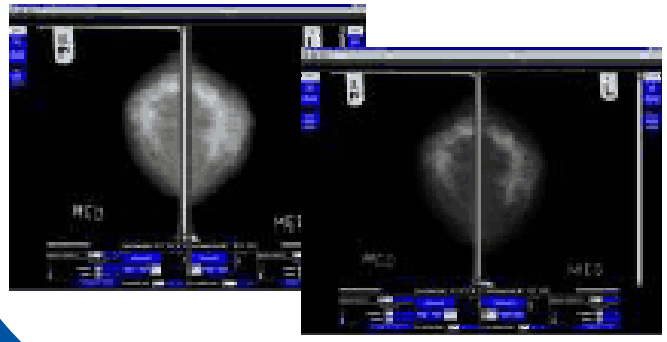


Medical/Healthcare Applications

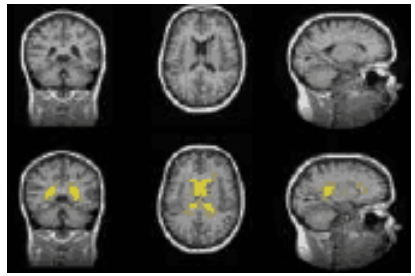
- Digital image archives
- Collaborative virtual environments
- On-line clinical conferences

"The Grid will enable a standardized, distributed digital mammography resource for improving diagnostic confidence"

"The Grid makes it possible to use large collections of images in new, dynamic ways, including medical diagnosis."



"The ability to visualise 3D medical images is key to the diagnosis of pathologies and pre-surgical planning"



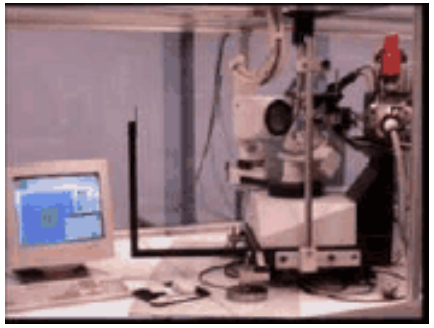
Bioinformatics

- Analysing genetic and proteomic information
- Determining the development of an embryo
- Modelling drugs for diseases and genetic disorders

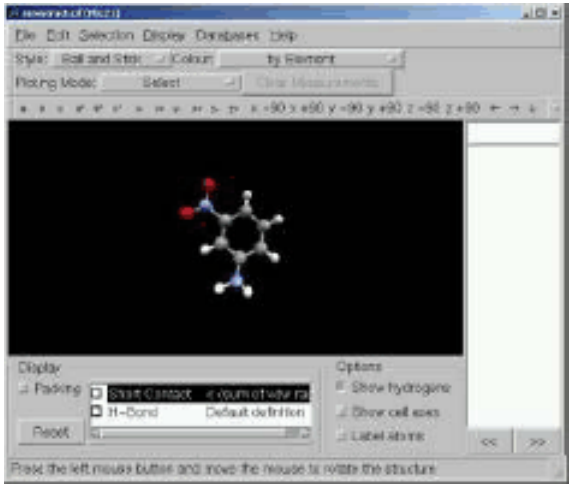
"Every time a new genome is sequenced the result is compared in a variety of ways with other genomes. Each code is made of 3.5 billion pairs of chemicals…"

Nanotechnology

- Computer-aided design of new functional materials
- Enable faster, cheaper discovery of new catalysts, metals, polymers, organic and inorganic materials



"The Grid has the potential to store and analyze data on a scale that will support faster, cheaper synthesis of a whole range of new materials."



Conclusions

- Fundamental science is a driving force for new information technologies and international cooperation.
- From the Web to the Grid: we have come a long way in just over a decade, but this may just be the beginning!
- The Web has changed our world in science, commerce and society at large.
- Will the Grid have a similar impact? What do you think?

- Important issues for Grids, that I did not address:
  - Security
  - Business Model
  - Reliability and simplicity of use

For more information about the Grid: www.gridcafe.org
If you want to contribute to the LHC: www.lhcathome.org

Thank you for your kind attention!
Thanks to TIFR for your kind hospitality!