Computing at Fermilab

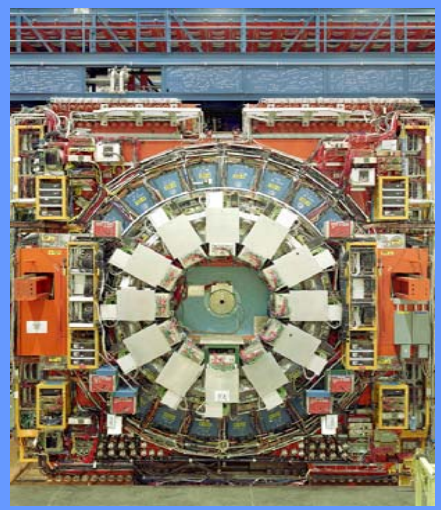
David J. Ritchie
Computing Division
May 26, 2009

Computing At Fermilab is…

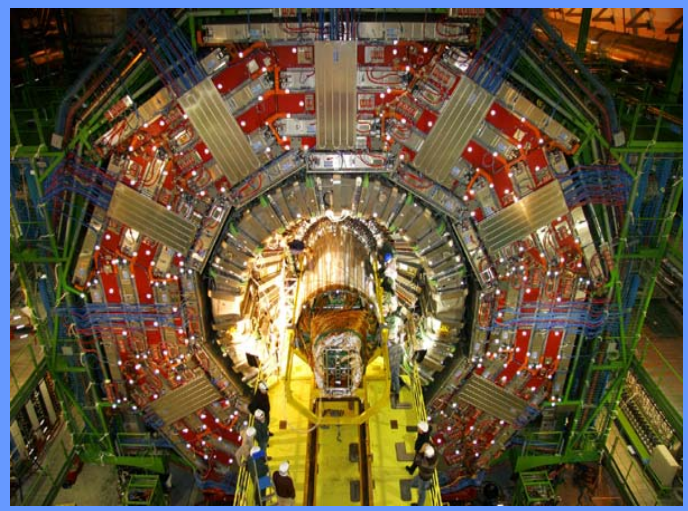
- In support of the mission…
  - To advance the understanding of the fundamental nature of
    - Matter and
    - Energy

The Fermilab Accelerators

The DZero Experiment



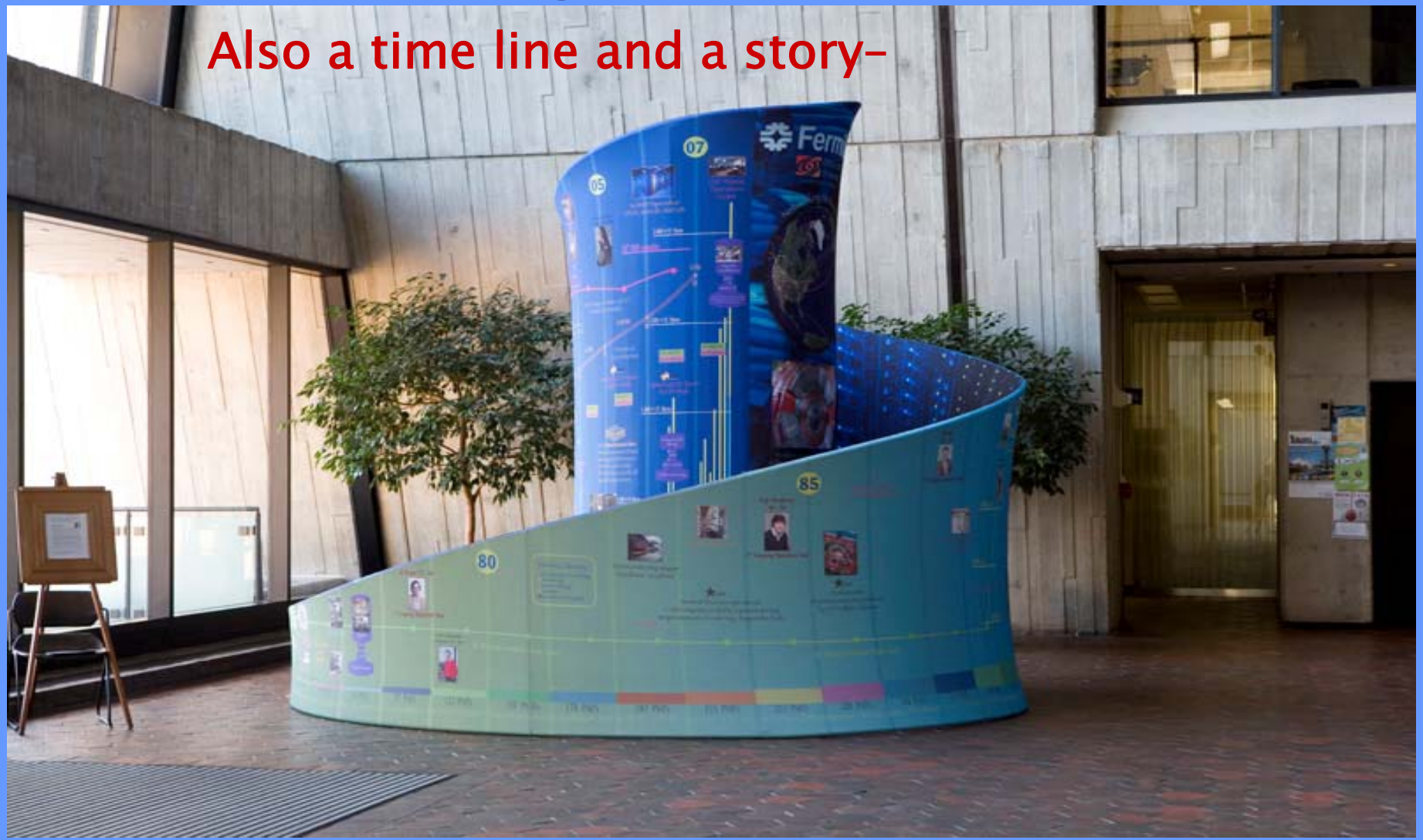
The CDF Experiment



The CMS Experiment at the LHC

More than ~900 experiments have been proposed as of this date.

Computing At Fermilab is…



Computing At Fermilab is…

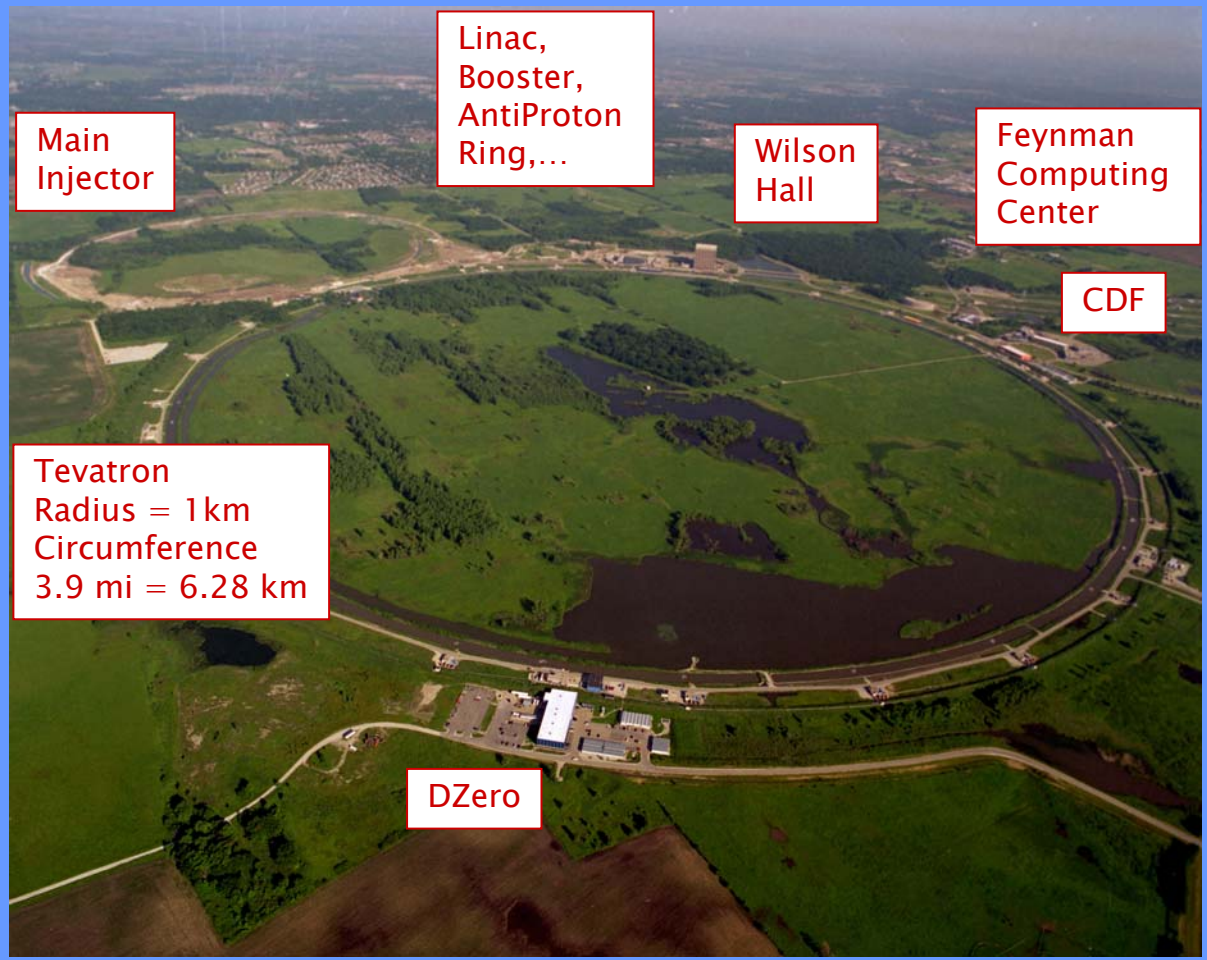
Part of a multi-year story that…

- Began in ~1967…
- Documented on a unique spiral in the Atrium
- Set up especially for you

Think of a year >~1972

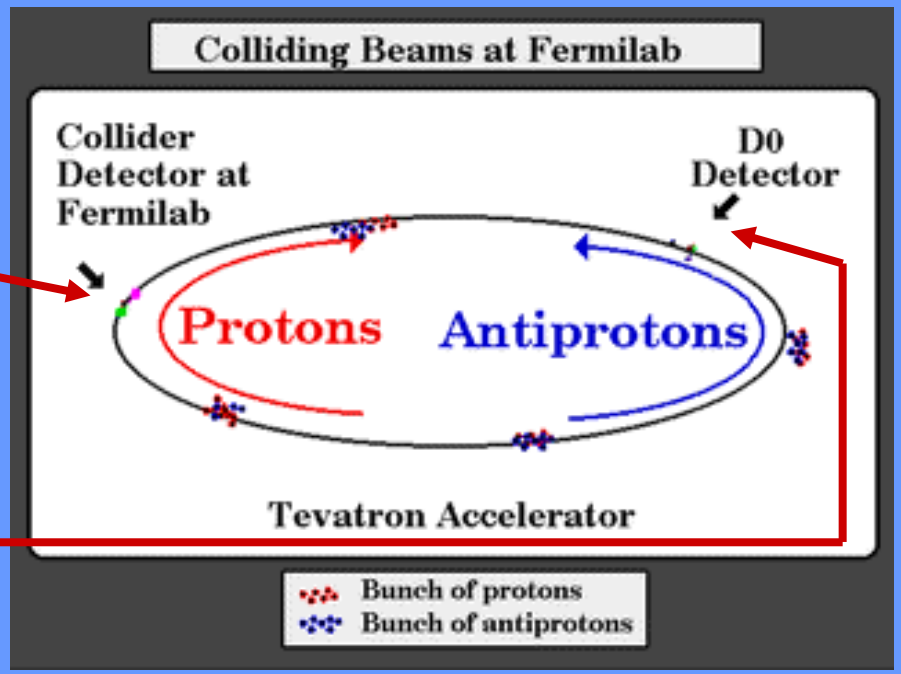
- Join me after the talk at the Spiral in the Atrium
- Learn what was happening in Computing at Fermilab at that time.

The Accelerators
(and other points of interest)

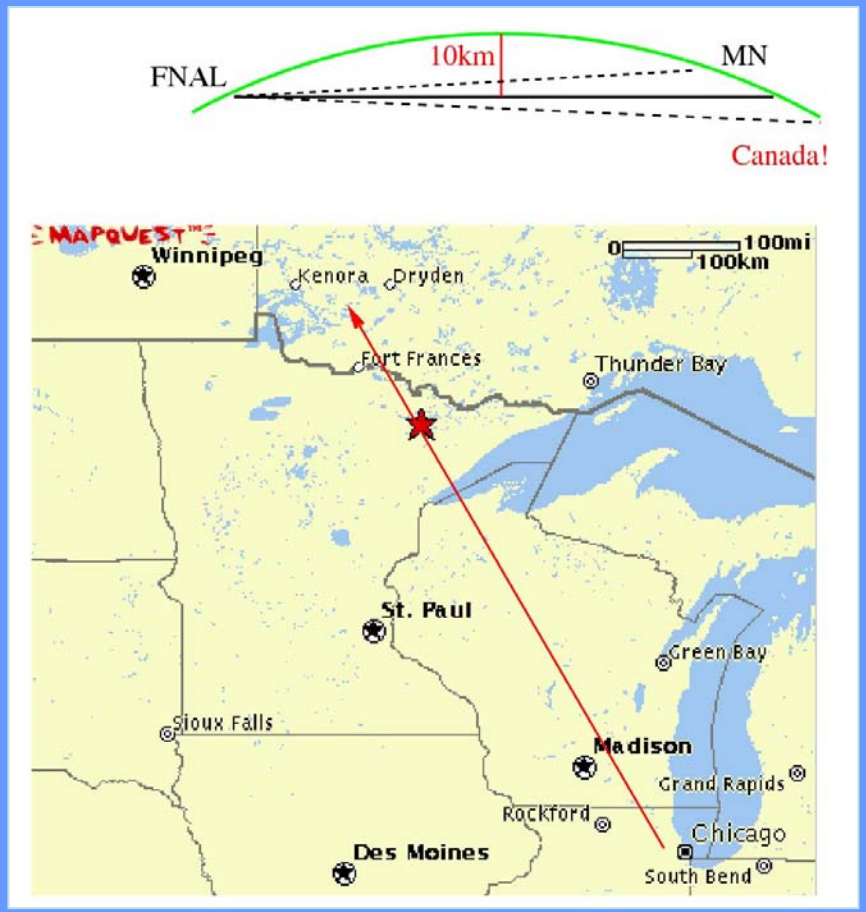


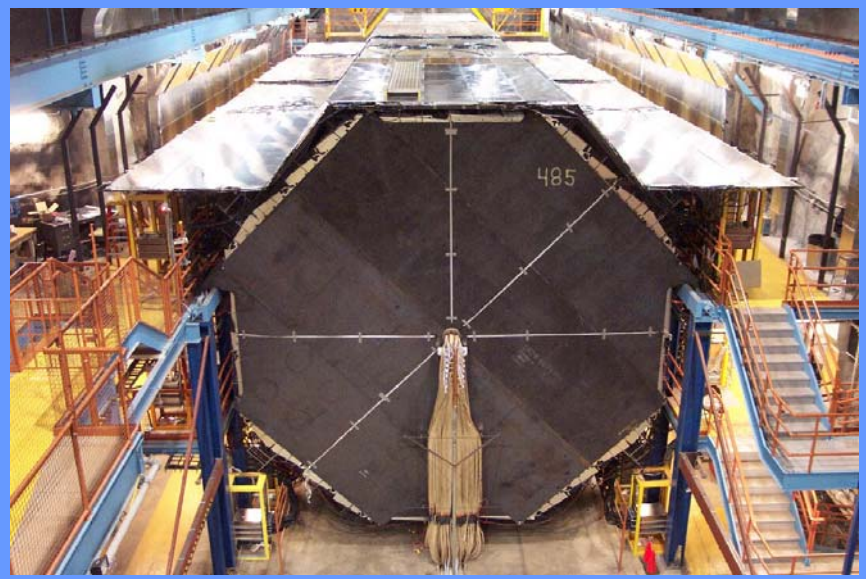
The Accelerators

(which way do those protons and antiprotons go?)



MINOS Near and Far

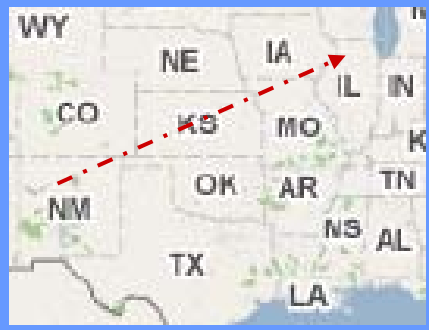




The Far Detector

Sloan Digital Sky Survey

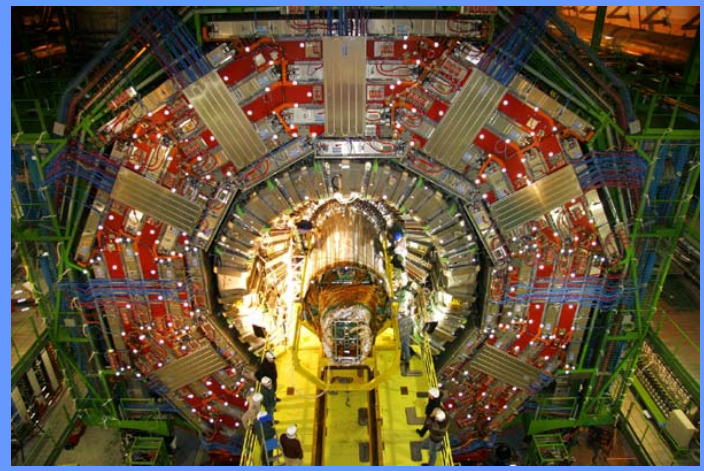
Apache Point
New Mexico



CMS Experiment At The LHC

LHC
Circumference
17 Miles = 27 Kilometers

Compare to Fermilab Tevatron 3.9 mi = 6.28 km



At the Large Hadron Collider (LHC) at CERN in Geneva, Switzerland

How Do You Do Science? (elementary school version)

1. Observe **phenomena**.
2. Develop a ***hypothesis***.
3. Use hypothesis to make ***predictions***.
4. Devise ***experiment*** to look for predictions.
5. Obtain ***results*** demonstrating (or not) predictions.
6. Draw ***conclusions*** about correctness of hypothesis.
7. If correct, add to accumulated hypotheses which are the **theory.** If not, revise hypothesis and …

Repeat

How Do You Do Science?
(Large Scale Science Version)

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan and Design **Experiment**
- Build / Install Equipment
- Acquire / Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**

Repeat

Computing Needs
Phenomena and Hypothesis

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**

Need: access previous results

- PCs (Windows, Linux), Macs
- Printers – color and B/W
- World-wide networking
- MSWord, TeX, Acrobat
- Web preprint repositories
  - Arxiv.org at Cornell U.
  - Spires at Stanford U.

Repeat

Computing Needs
Organize Collaboration

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**

Need: work together

- PCs (Windows, Linux), Macs
- World-wide networking
- E-mail
- Web Pages
  - DZero "Top Group"
  - MINOS (For Scientists)

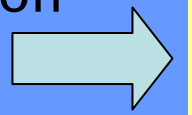
Repeat

Work compatibly and productively with >~400 colleagues from around the world

Computing Needs
Propose Experiment

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**



Need: show proposal feasibility

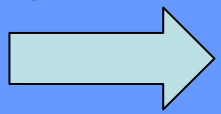
- PCs (Windows, Linux), Macs
- World-wide networking
- Web pages
  - The Dark Energy Survey
  - Minerva Proposal
- Experiment simulations

Repeat

Computing Needs
Get Approved and Funded

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Acquire / Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**



Need: Project Management

Funds come from DOE, NSF, NASA, Congress, Taxpayer—YOU.

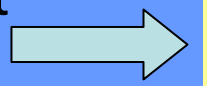
- PC's (Windows, Linux, Macs)
- Printers – color and B/W
- World-wide networking
- MSProject, Power Point…
- Proposal review…
  - MINERvA Review

Repeat

Computing Needs
Plan / Design, Build / Install

Each stage of scientific process has different computing needs.

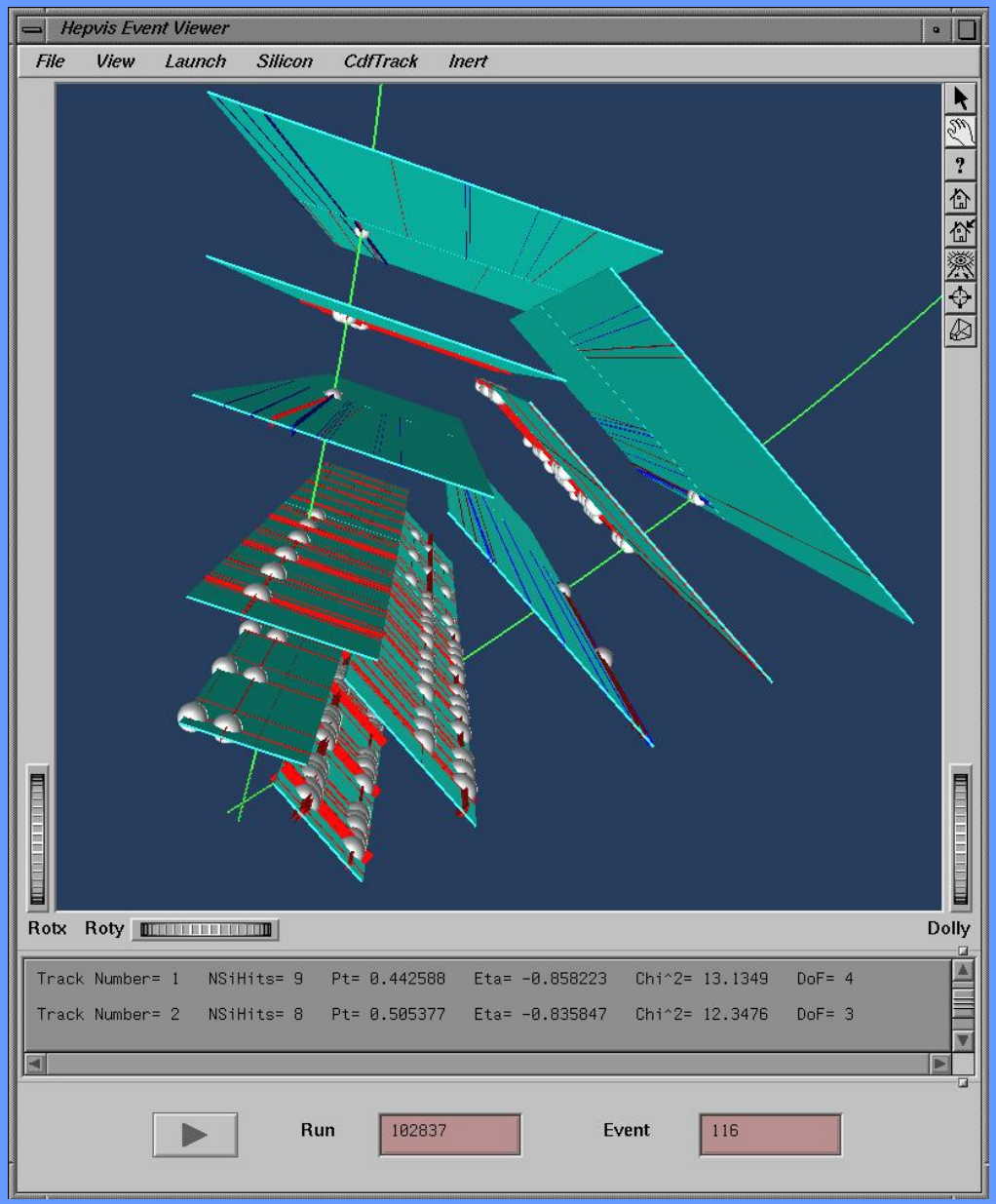
- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Acquire / Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**



Need: Plan and Design S/W

- Experiment simulation s/w,
  - Engineering design software
- Unusual sensors…
- Custom electronics
  - Integrated circuits,

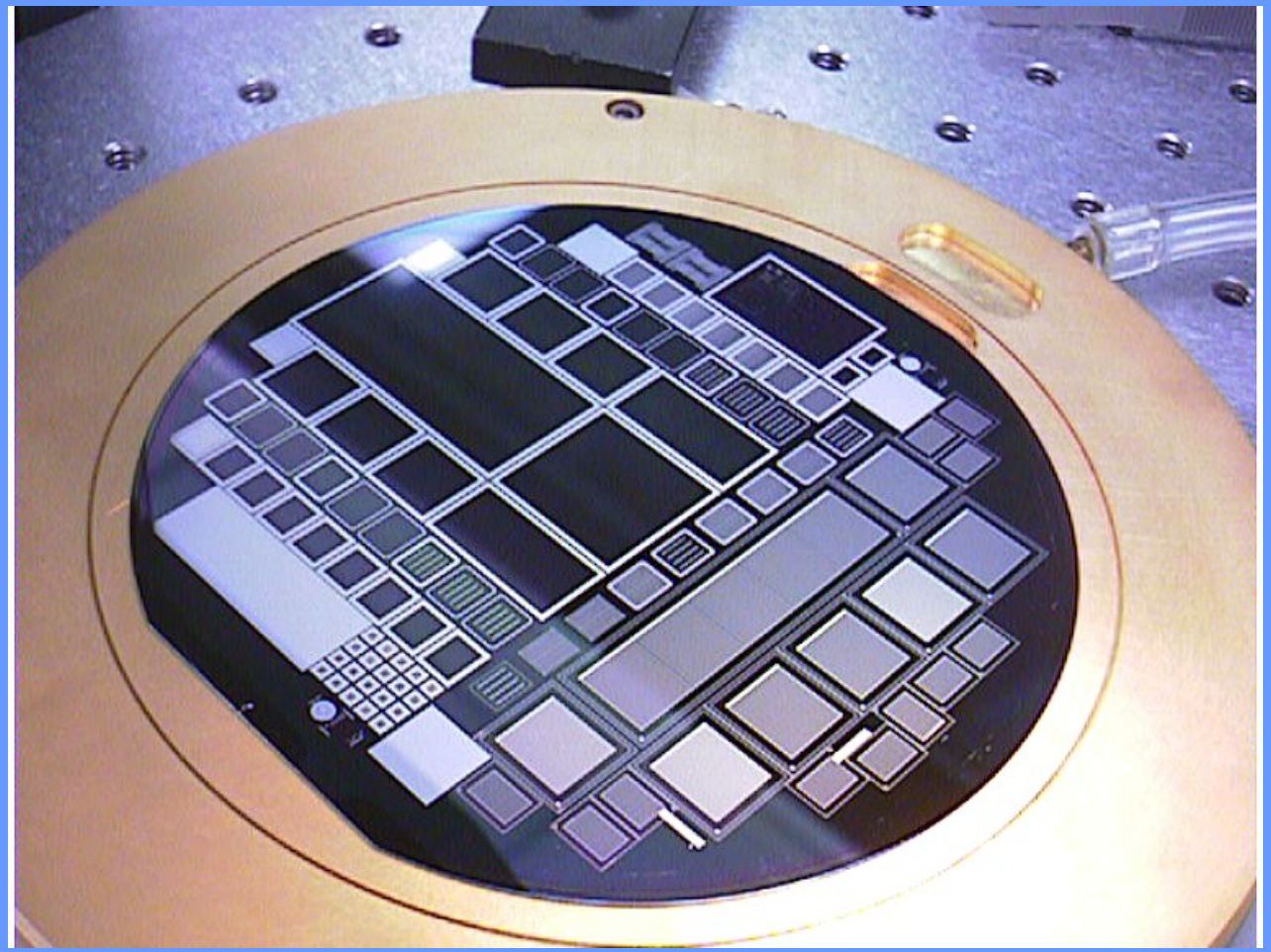
Repeat

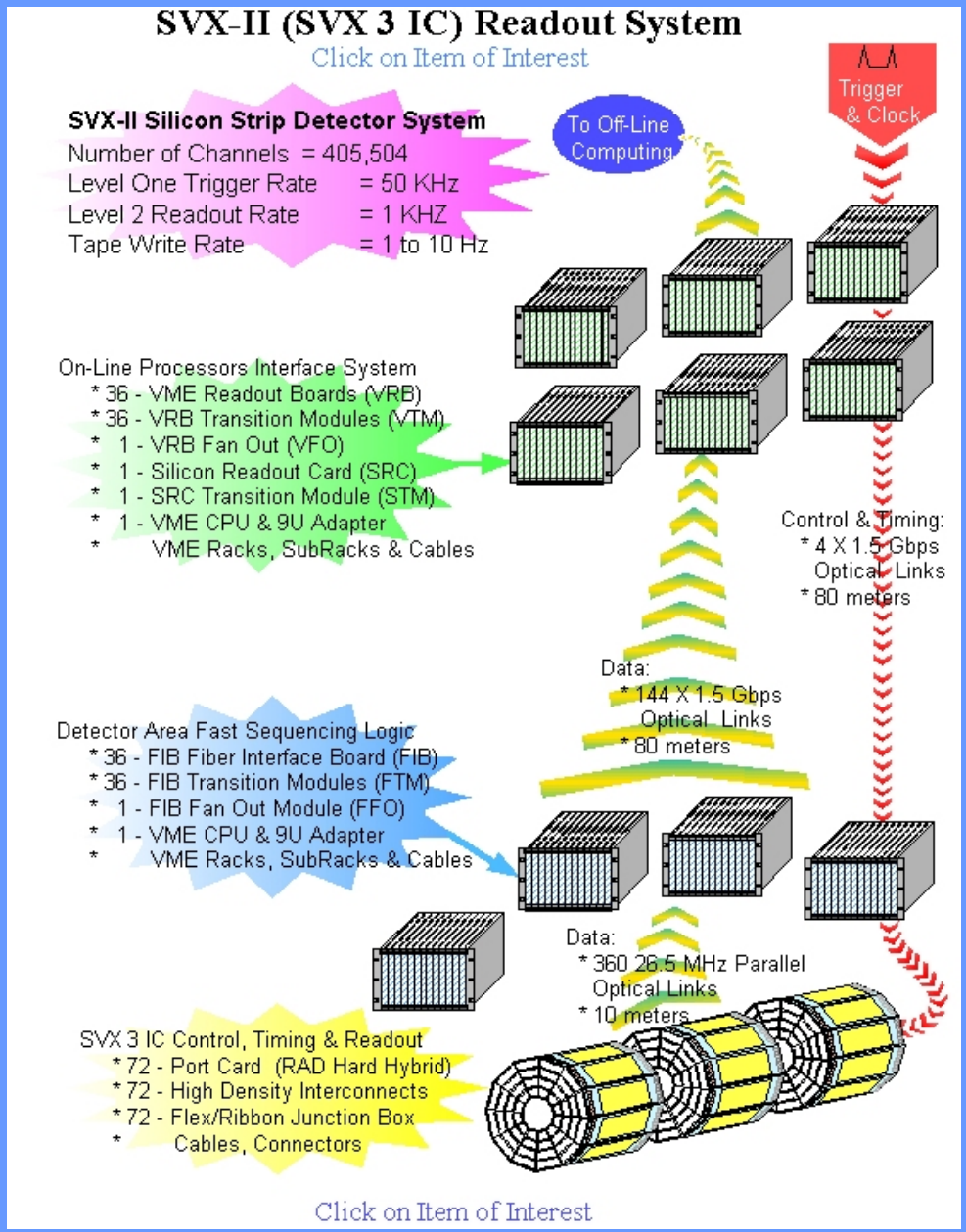


Experiment simulation S/W

A 3D simulation program models particles going through the sensors in the CDF Silicon Vertex Detector.

Unusual Sensors





Custom Electronics

CDF Silicon Vertex Detector Readout

Computing Needs
Acquire / Record / Store Data

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Acquire / Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**

Repeat

Need: Acquire data…

Acquire and Record Data

- Acquire data
  - The experiment detects the fragments from a proton-antiproton collision according to some specified pattern ("a possibly interesting event").
  - Electrical signals throughout the experiment are acquired as 1's and 0's and packaged into an event of whose size is ~300,000 bits or more.
  - An event rate of hundred's of Gb/second results. Many events are thrown out ("not that interesting").
- Record data
  - The *really interesting* events are recorded on disk and stored on tape for later analysis.

Computing Needs
Analyze Data

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Acquire / Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**

Need:
- Reconstruct 1's and 0's into tracks and particle identification
- Do physics analysis on resulting collections of tracks

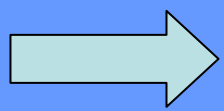
Means:
- Open Science Grid
- Disk backed up by tape with robotic tape mounts Enstore Usage.

Repeat

Computing Needs
Obtain Results

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**



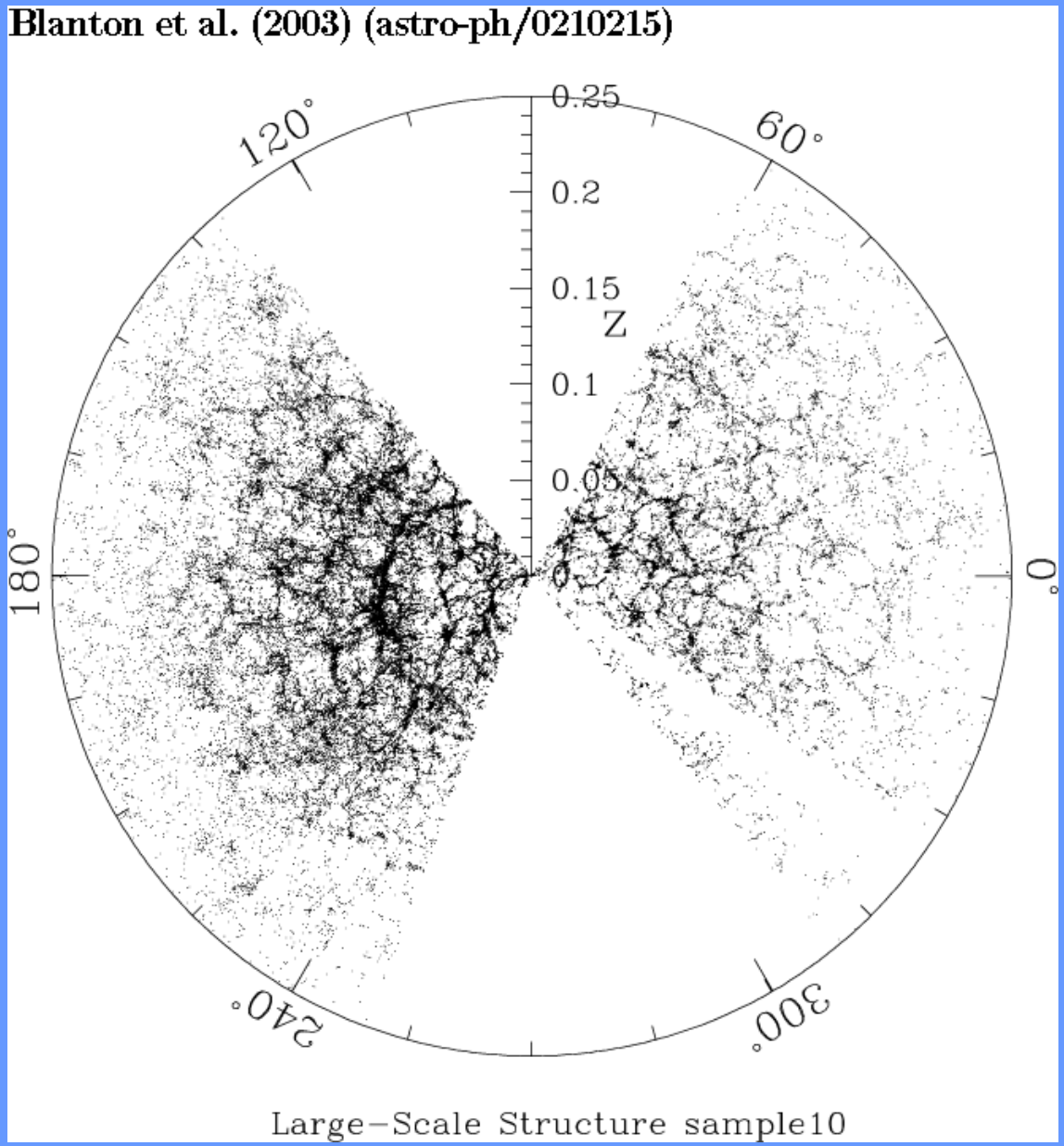
Repeat

Need: Do additional physics analysis.

Means: Use analysis farms with their disk caching capabilities to speed data analysis.

CDF Usage

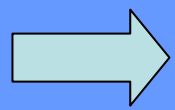
An SDSS Result



Computing Needs
Publish Conclusions

Each stage of scientific process has different computing needs.

- Identify **Phenomena**
- Develop **Hypothesis**
- Organize Collaboration
- Propose **Experiment**
- Get Approved
- Obtain Funding
- Plan / Design **Experiment**
- Build / Install Equipment
- Record / Store Data
- Analyze Data
- Obtain **Results**
- Publish **Conclusions**



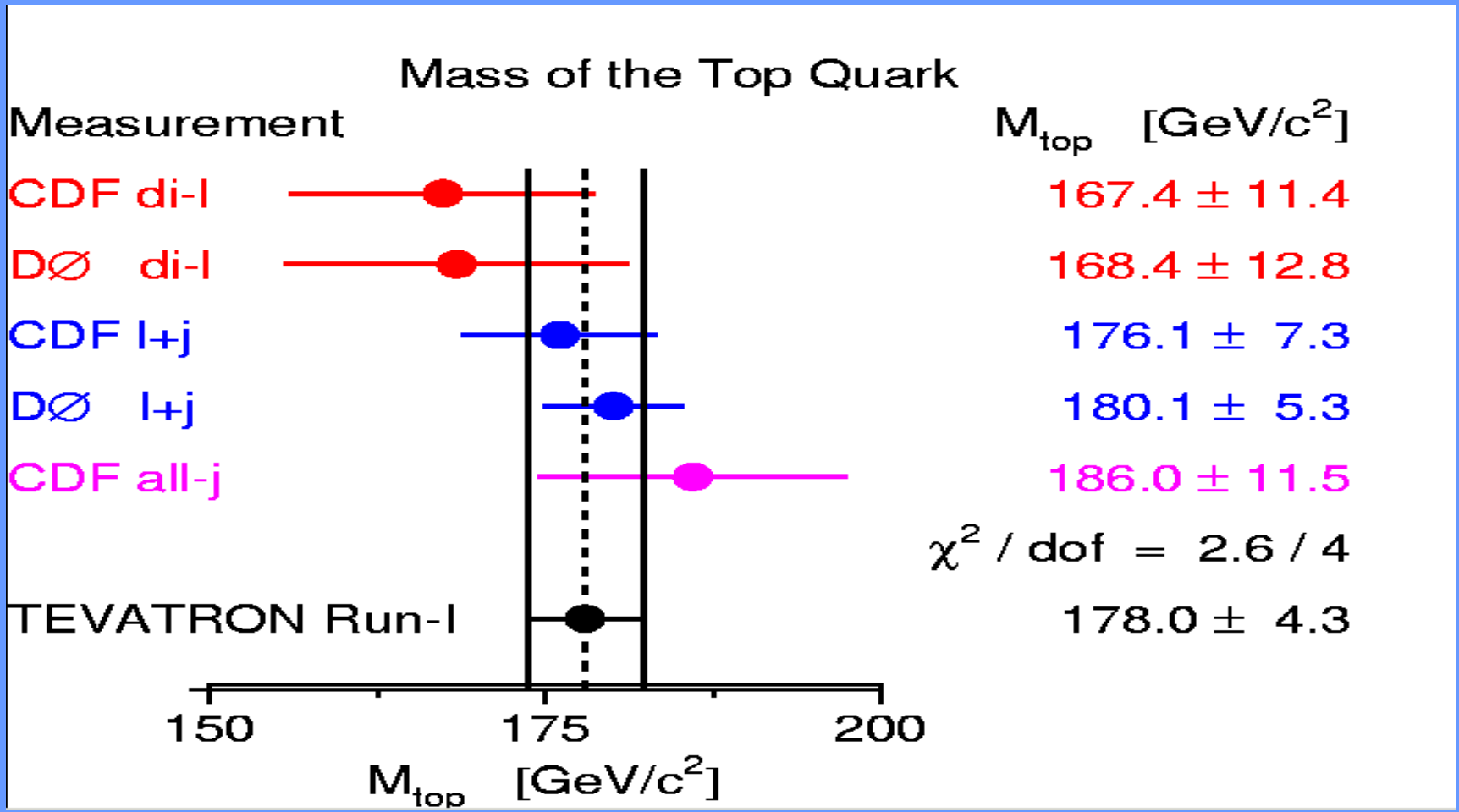
Repeat

Need: Show evidence of progress in scientific work.

Means: Publish papers and conferences.

Top Quark Mass

Mass of the Top Quark



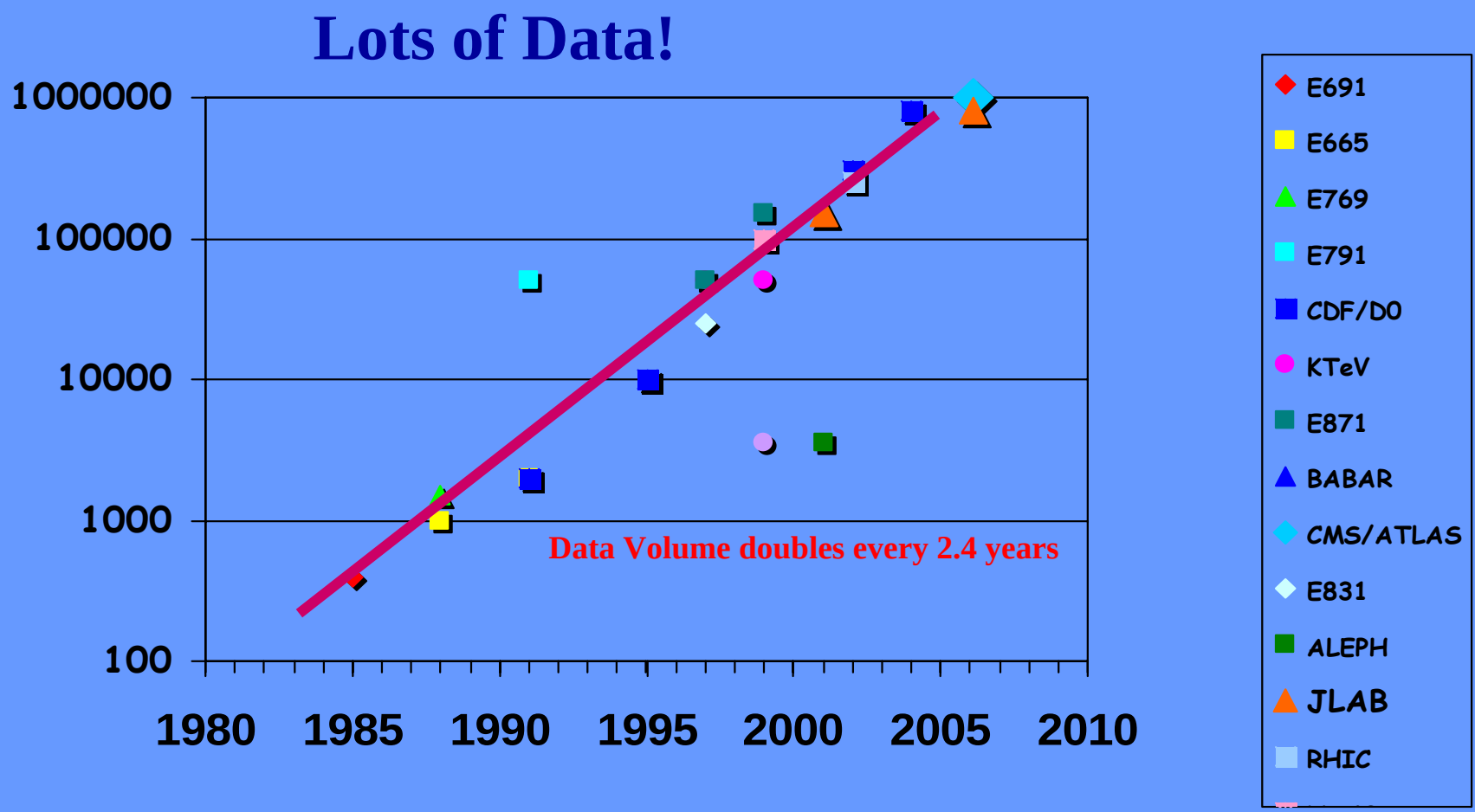
And Now On Another Topic…

Computing Facilities

Computing Facilities

- The facilities can be sub-divided into:
  - Networks (>10,000 network devices)
  - Mass Storage (> 10 PB data on tape)
  - Grid Installations (>~6,000 cores)
  - Three facilities: FCC, GCC, LCC

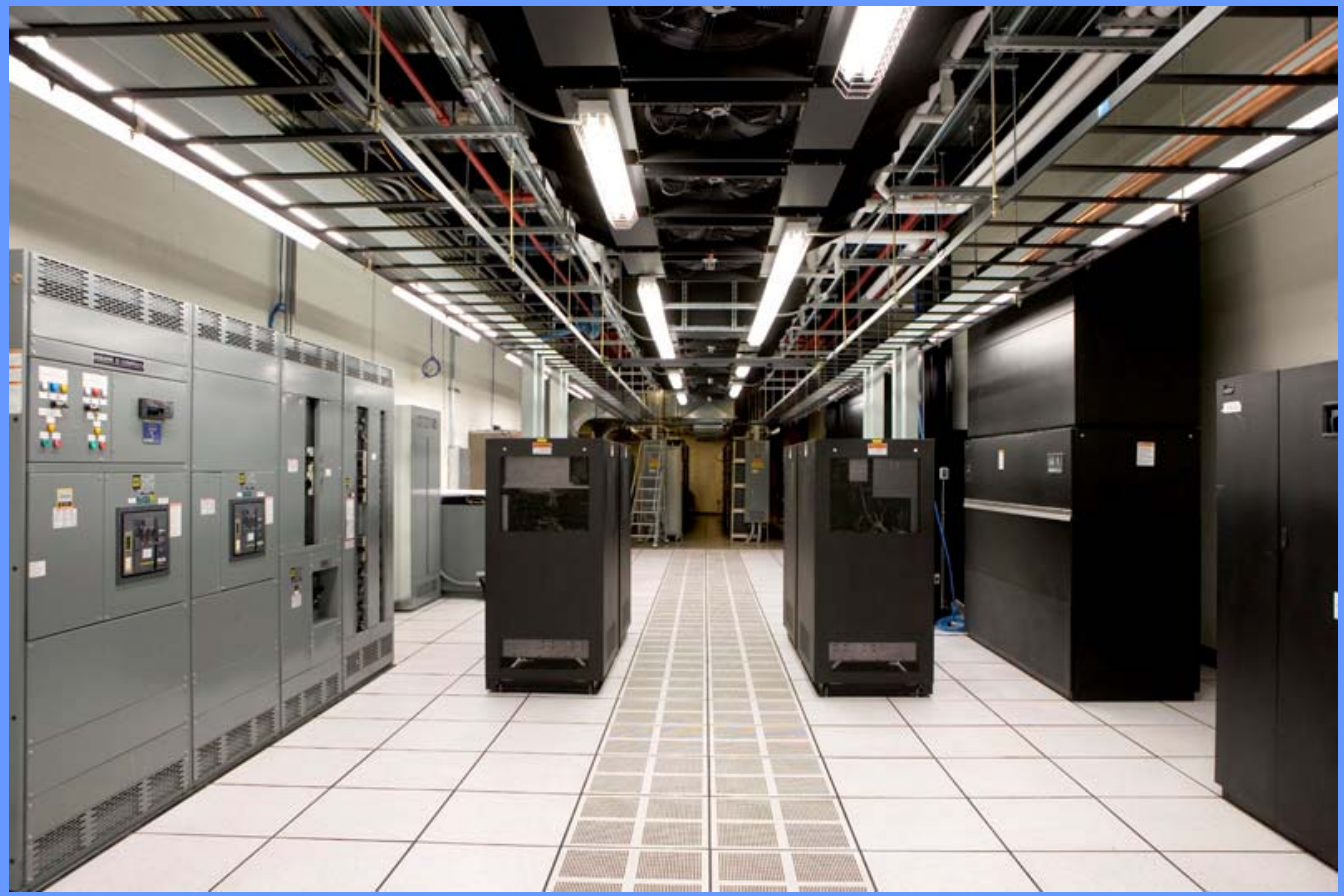
Moore's Law


Lots of Data!
1000000
100000
10000
1000
100
1980
1985
1990
1995
2000
2005
2010
Data Volume doubles every 2.4 years
E691
E665
E769
E791
CDF/D0
KTeV
E871
BABAR
CMS/ATLAS
E831
ALEPH
JLAB
RHIC

FCC

GCC

LCC
(Just before filling the facility)



And Now On Another Topic…

U. S. CMS Computing and GRID Computing

US CMS Computing

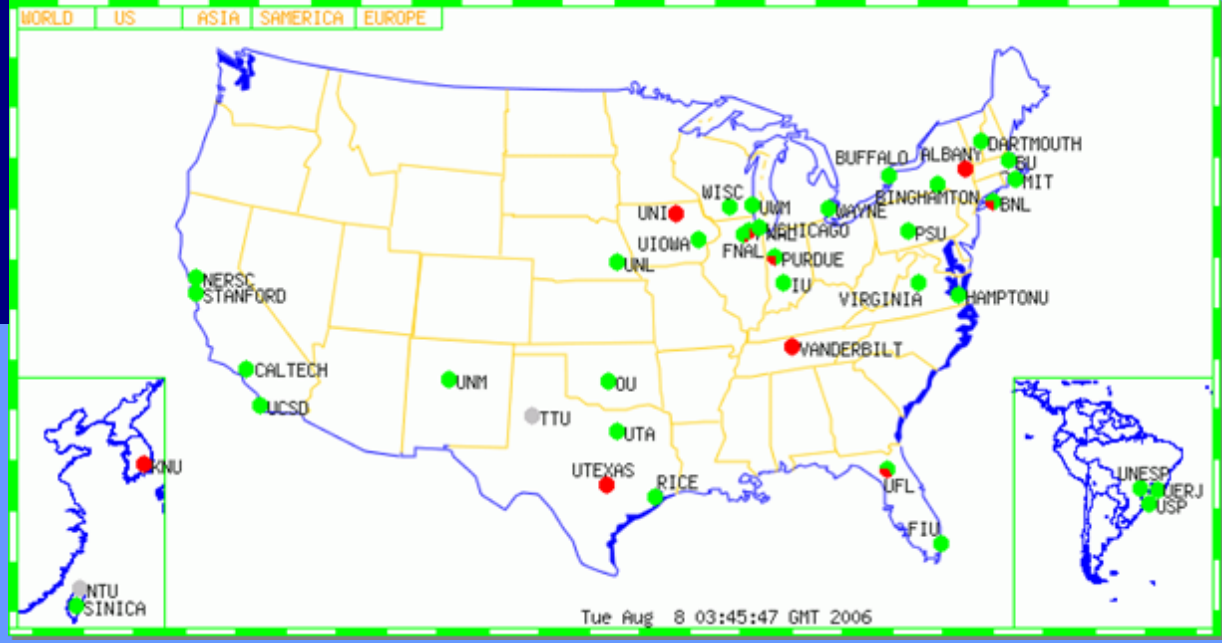
- Fermilab is the host lab of U.S. CMS experiment which will begin taking data at CERN in Geneva, Switzerland in ~2009
- Fermilab hosts the project management for the U.S. CMS Software and Computing Program in DOE
- U.S. Physicists participate in this research.

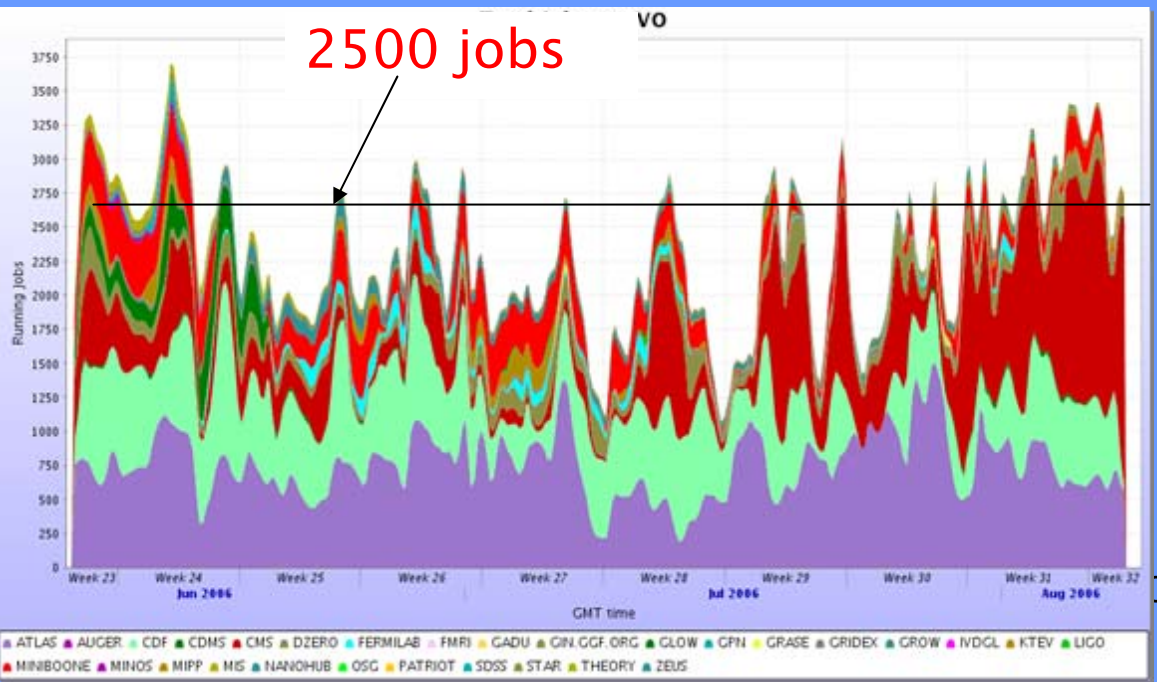
GRID Computing

- Fermilab is participating in many GRID initiatives:
  - ppdg (DOE SciDAC)
    - Particle physics data grid
    - Fermilab, SLAC, ANL, BNL, JLAB, Caltech, UCSD, Wisconsin, SDSC
  - Griphyn
    - Grid physics network
  - iVDGL
    - International virtual grid laboratory
- GRID activities have been a very natural outgrowth of distributed computing of the large collaborations.

~70 Sites Working together as a Coherent Computing Facility

Open Science Grid

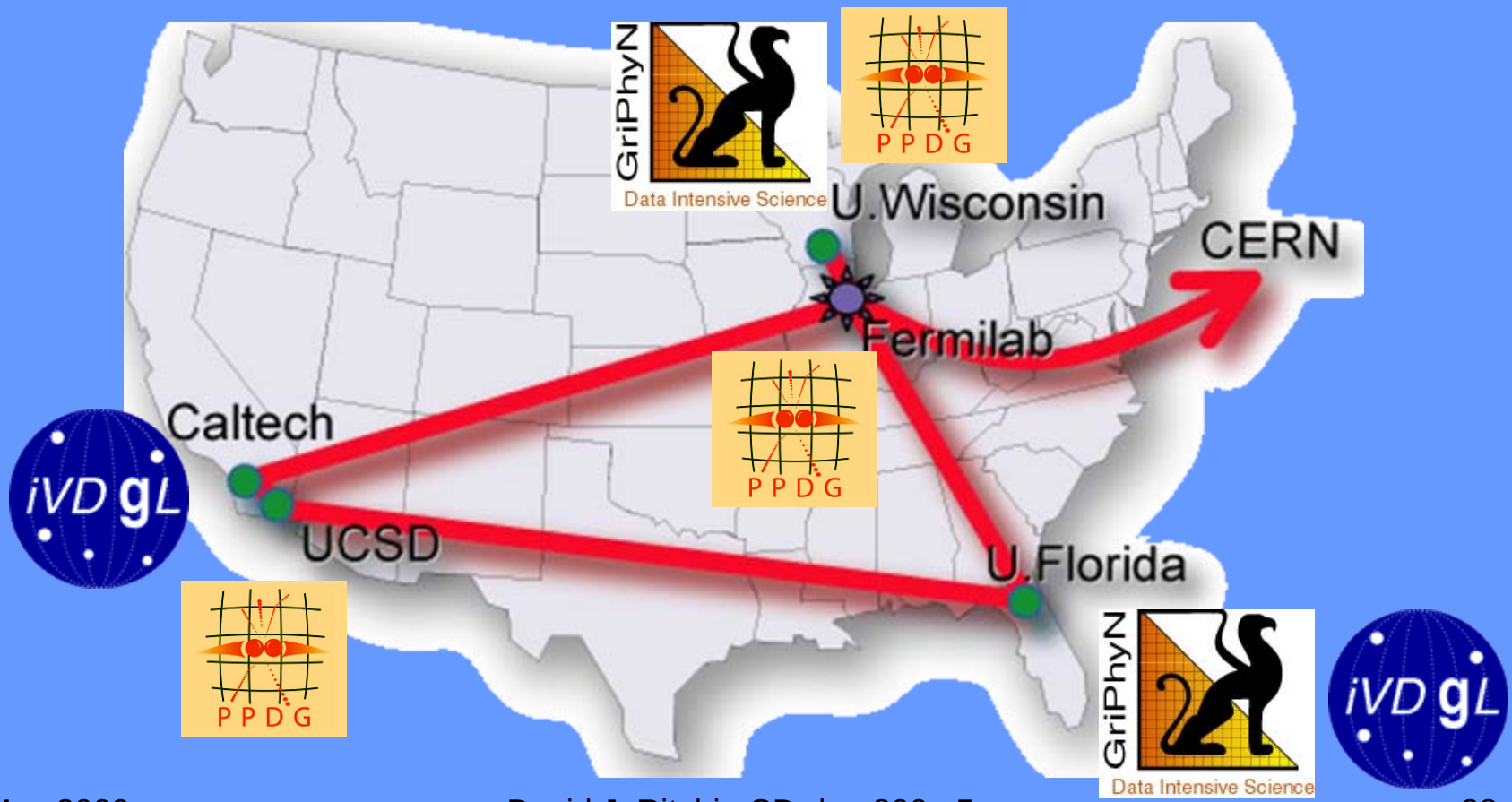




Research Groups Sharing Sites to Run Compute Intensive Jobs

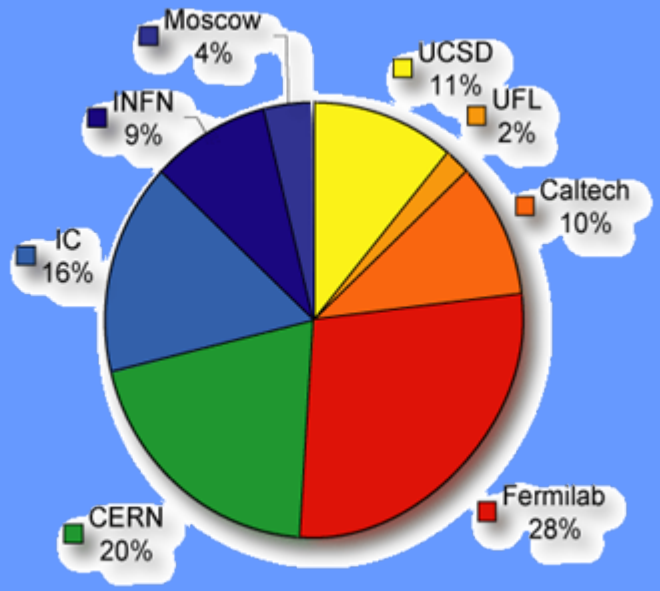
US CMS Computing

Tier 1/Tier 2 Centers and Prototypes



US CMS Computing

- Contribution to CMS is substantial
- Series of Data Challenges
  - Increasing complexity
  - >70 million events fully simulated



And Now On Another Topic…

Computer Security

Computer Security

- The Internet is now >~100 Million computers/users
  - “one in a million” events happen every day ‘somewhere on the net’
  - Only with the net ‘somewhere on the net’ is always just next door.
- We are learning how to live in a world where not everyone is friendly.

Computer Security

- This means that computer security is an ever present task—particularly, when studies have shown that a hack attempt is made once every 13 minutes.
- We use extensive automated scanning software to detect problems and isolate them typically within a few hours if not a few minutes.

Computer Security

- During your summer here, pay attention to computer security.
- Do not go to web sites not related to work.
- Follow proper password handling.
- If you violate these and other rules, your computing privileges will be withdrawn.
- A summer without access to computing is very dull so mind your p’s and q’s.

Acknowledgements

I would particularly like to thank my colleagues for contributions of slides and comments.