An incomplete view of the future of HEP Computing

Some things I like to know for the future of HEP Computing

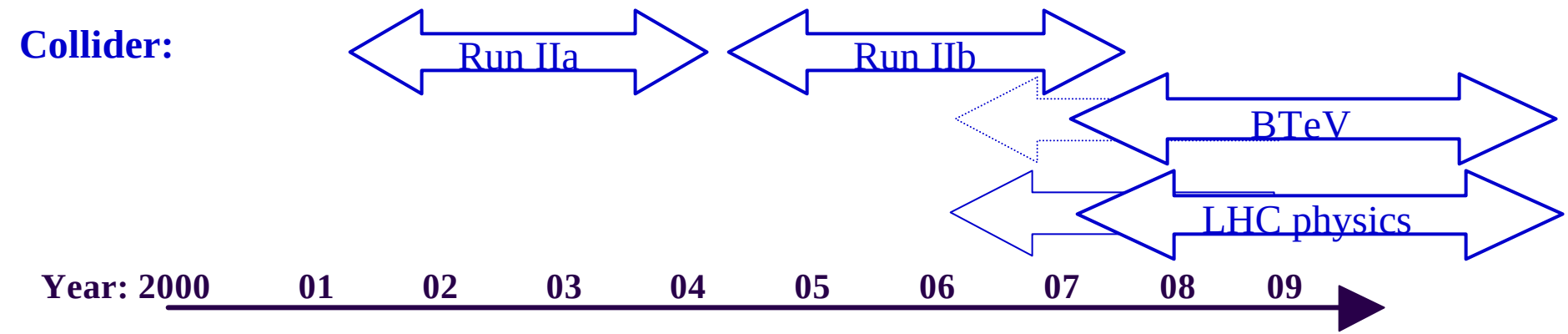
Matthias Kasemann
Fermilab

Disclaimer

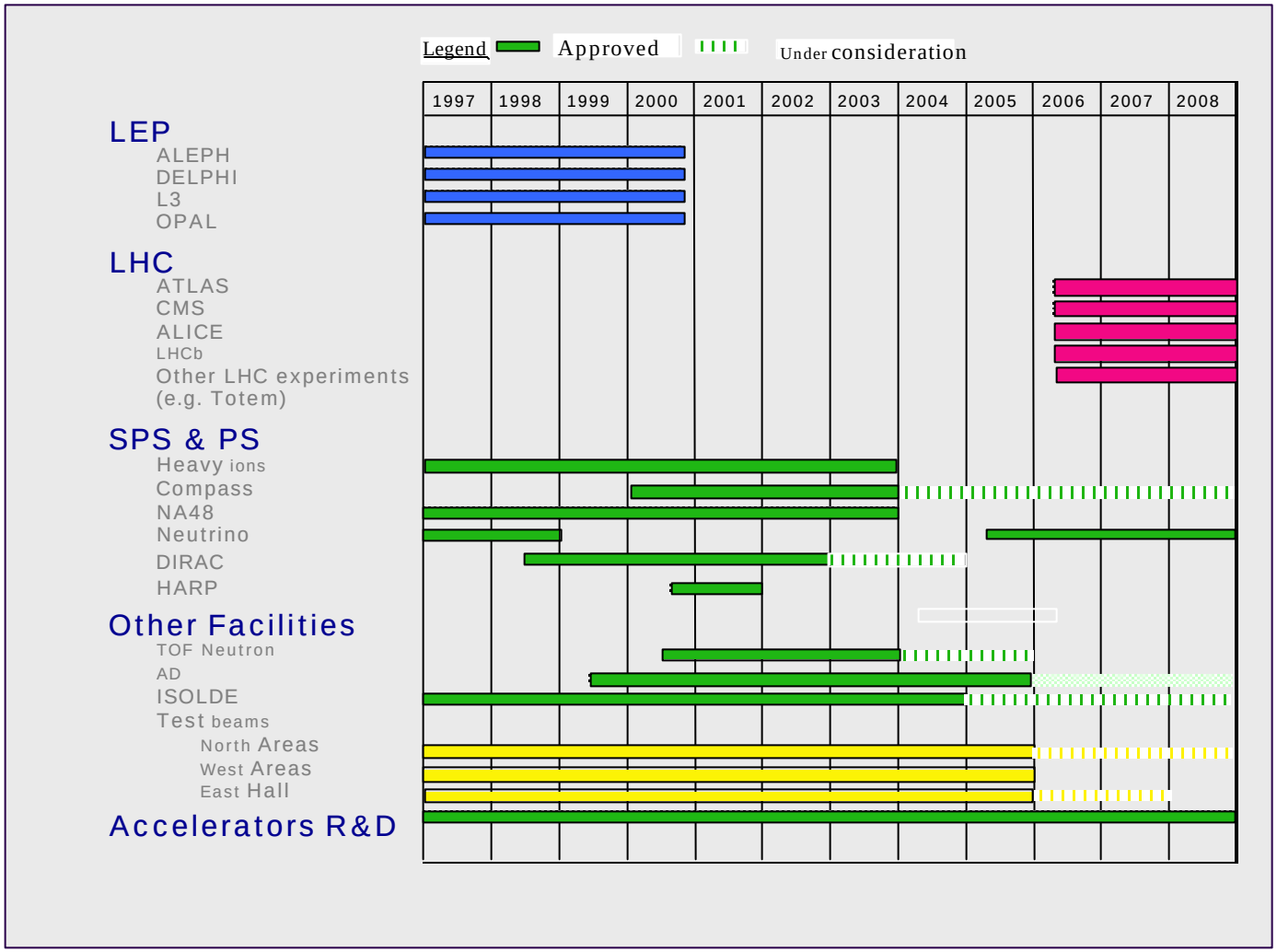
Although we are here in the Root workshop I don't present topics which are necessarily to be answered by Root in its current implementation or any derivative or future development in Root. I simply put down what worries me when I think about computing for future HEP experiments.

```
(Speaking for myself and not for US, US DOE, FNAL nor URA.)  (Product,
trade, or service marks herein belong to their respective owners.)
```

Fermilab HEP Program



The CERN Scientific Programme



HEP computing: The next 5 years…(1)

- Data analysis for completed experiments continues
  - Challenges:
    - No major change to analysis model, code or infrastructure
    - Operation, continuity, maintaining expertise and effort
- Data collection and analysis for ongoing experiments
  - Challenges:
    - Data volume, compute resources, software organization
    - Operation, continuity, maintaining expertise and effort

HEP computing: The next 5 years…(2)

- Starting experiments:
  - Challenges:
    - Completion and verification of data and analysis model,
    - Data volume, compute resources, software organization, $$'s
    - Operation, continuity, maintaining expertise and effort

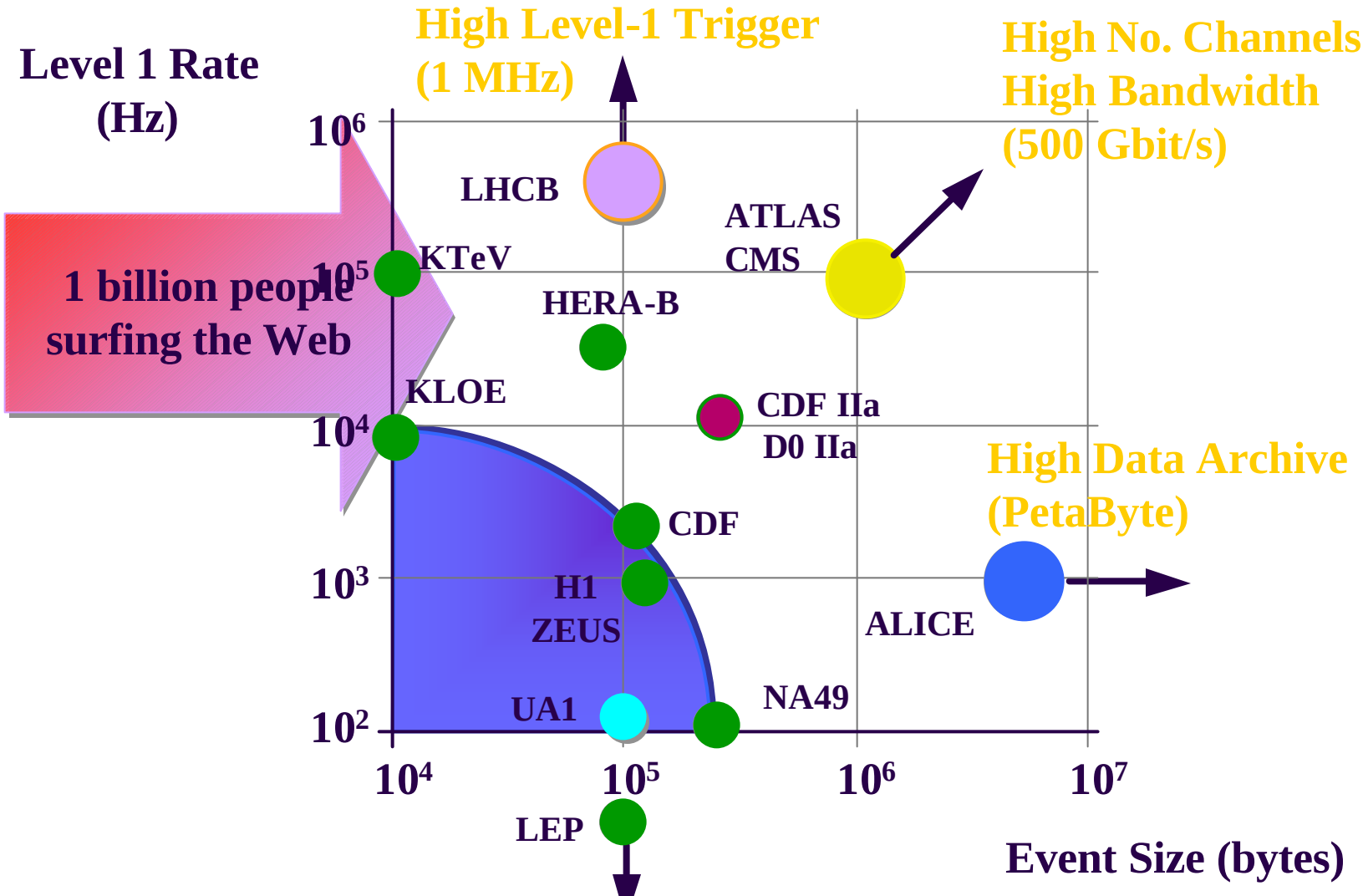
- Experiments in preparation:
  - Challenges:
    - Definition and implementation of data and analysis model,
    - data volume, compute resources, software organization, $$'s
    - continuity, getting and maintaining expertise and effort
    - Build for change: applications, data models…
    - Build compute models which are adaptable to different local environments

Run 2 Data Volumes

Category	Parameter	D0	CDF
DAQ rates	Peak rate	53 Hz	75 Hz
	Avg. evt. Size	250 KB	250 KB
	Level 2 output	1000 Hz	300 Hz
	maximum log rate	Scalable	80 BM/s
Data storage	# of events	600M/year	900 M/year
	RAW data	150 TB/year	250 TB/year
	Reconstructed data	75 TB/year	135 TB/year

- First Run 2b costs estimates based on scaling arguments
  - Use predicted luminosity profile
  - Assume technology advance (Moore's law)
  - CPU and data storage requirements both scale with data volume stored
- Data volume depends on physics selection in trigger
  - Can vary between 1 – 8 PB (Run 2a: 1 PB) per experiment
- Have to start preparation by 2002/2003

How Much Data is Involved?

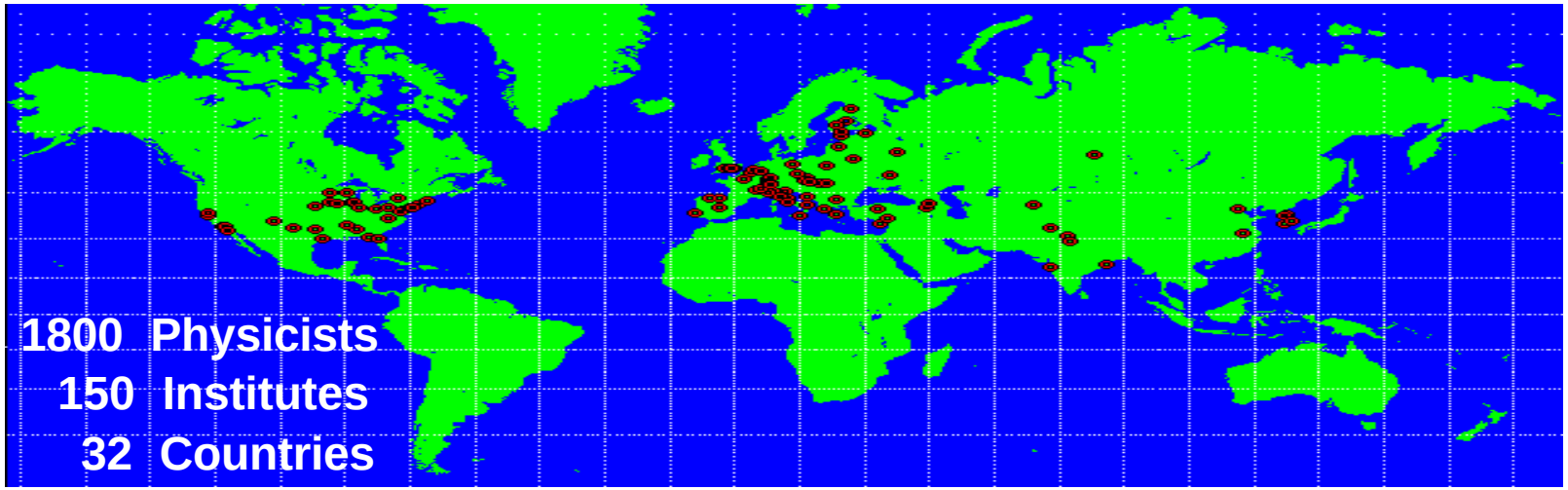


HEP computing: The next 5 years…(3)

- Challenges in big collaborations
  - Long and difficult planning process
  - More formal procedure required to commit resources
  - Long lifetime, need flexible solutions which allow for change
    - Any state of experiment longer than typical PhD or postdoc time
    - Need for professional IT participation and support
- Challenges in smaller collaborations
  - Limited in resources
  - Adapt and implement available solutions (“b-b-s”)

CMS Computing Challenges

- Experiment in preparation at CERN/Switzerland
- Strong US participation: ~20%
- Startup: by 2005/2006, will run for 15+ years



Major challenges associated with:
Communication and collaboration at a distance
Distributed computing resources
Remote software development and physics analysis
R&D: New Forms of Distributed Systems

Role of computer networking (1)

- State-of-the-art computer networking enables large international collaborations
  - needed for all aspects of collaborative work
    - to write the proposal,
    - produce and agree on the designs of the components and systems,
    - collaborate on overall planning and integration of the detector, confer on all aspects of the device, including the final physics results, and
    - provide information to collaborators and to the physics community and general public
  - Data from the experiment lives more-and-more on the network
    - All levels: raw, dst, aod, ntuple, draft-paper, paper

Role of computer networking (2)

- HEP developed its own national network in the early 1980s
- National research network backbones generally provide adequate support to HEP and other sciences.
- Specific network connections are used where HEP has found it necessary to support special capabilities that could not be supplied efficiently or capably enough through more general networks.
  - US-CERN, several HEP links in Europe…
- Dedicated HEP links are needed in special cases because
  - HEP requirements can be large and can overwhelm those of researchers in other fields
  - because regional networks do not give top priority to interregional connections

Data analysis in international collaborations: past

- In the past analysis was centered at the experimental site
  - a few major external centers were used.
  - Up the mid 90s bulk data were transferred by shipping tapes, networks were used for programs and conditions data.
  - External analysis centers served the local/national users only.
  - Often staff (and equipment) from the external center being placed at the experimental site to ensure the flow of tapes.
  - The external analysis often was significantly disconnected from the collaboration mainstream.

Data analysis in international collaborations: truly distributed

- Why?
  - For one experiment looking ahead for a few years only centralized resources may be most cost effective, but:
  - national and local interests leads to massive national and local investments
  - For BaBar:
    - The total annual value of foreign centers to the US-based program is greatly in excess of the estimated cost to the US of creating the required high-speed paths from SLAC to the landing points of lines WAN funded by foreign collaborators
  - Future world-scale experimental programs must be planned with explicit support for a collaborative environment that allows many nations to be full participants in the challenges of data analysis.

Distributed computing:

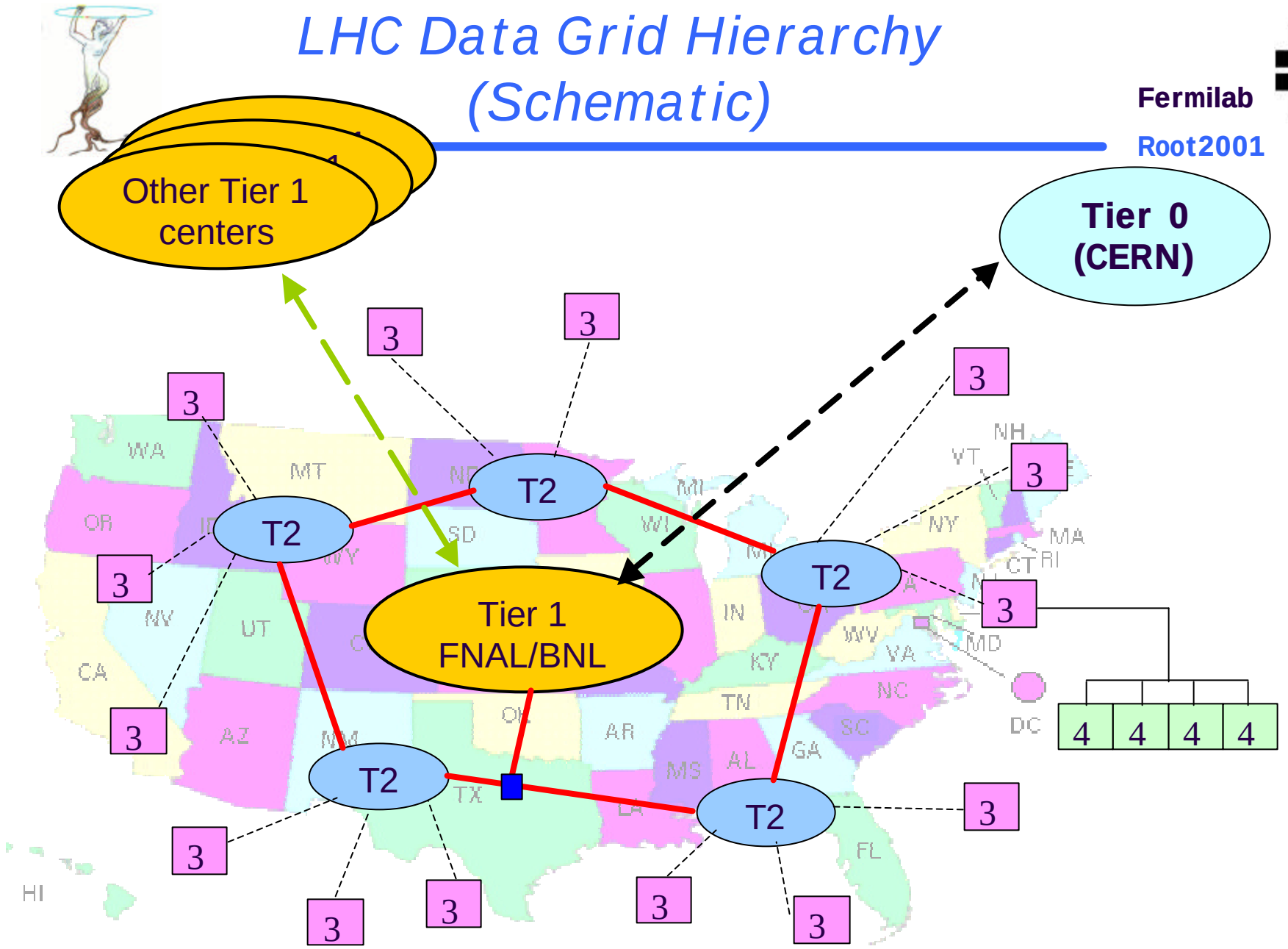
- Networking is an expensive resource, should be minimized
- Pre-emptive transfers can be used to improve responsiveness at the cost of some extra network traffic.
- Multi-tiered architecture must become more general and flexible
  - to accommodate the very large uncertainties in the relative costs of CPU, storage and networking
  - To enable physicists to work effectively in the face of data having unprecedented volume and complexity
- Aim for transparency and location independence of data access
  - the need for individual physicists to understand and manipulate all the underlying transport and task-management systems would be too complex

Distributed Computing

6/13/01: The New York Times ON THE WEB

"It turns out that distributed computing is really hard," *said Eric Schmidt, the* ***chairman of Google****, the Internet search engine company.*
"It's much harder than it looks. It has to work across different networks with different kinds of security, or otherwise it ends up being a single-vendor solution, which is not what the industry wants."

LHC Data Grid Hierarchy (Schematic)



Many more technical questions to answer (1)

- Operating system:
  - UNIX seems to be favored for data handling and analysis,
  - LINUX is most cost effective
- Mainframe vs. commodity computing:
  - commodity computing can provide many solutions
  - Only affordable solution for future requirements
  - How to operate several thousand nodes?
  - How to write applications to benefit from several thousand nodes?
- Data access and formats:
  - Metadata databases, event storage

Many more technical questions to answer (2)

- Commercial vs. custom software, public domain
- Programming languages:
  - Compiled languages for CPU intensive parts
  - Scripting languages provide excellent frameworks
- How to handle and control big numbers in big detectors:
  - Number of channels, modules improves (several millions of channels, hundreds of modules
  - Need new automatic tools to calibrate, monitor and align channels

Some more thoughts

- Computing for HEP experiments is costly
  - In $$'s, people and time
  - Need R&D, prototyping and test-beds to develop solutions and validate choices
- Improving the engineering aspect of computing for HEP experiments is essential
  - Treat computing and software as a project (see www.pmi.org):
    - Project lifecycles, milestones, resource estimates, reviews
- Documenting conditions and work performed is essential for success
  - Track detector building for 20 years
  - Log data taking and processing conditions
  - Analysis steps, algorithms, cuts

As transparent and automatic as possible