Computing R&D Strategies

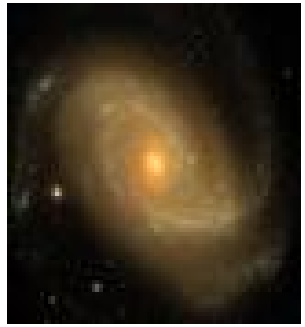
R. Tschirhart

Fermilab

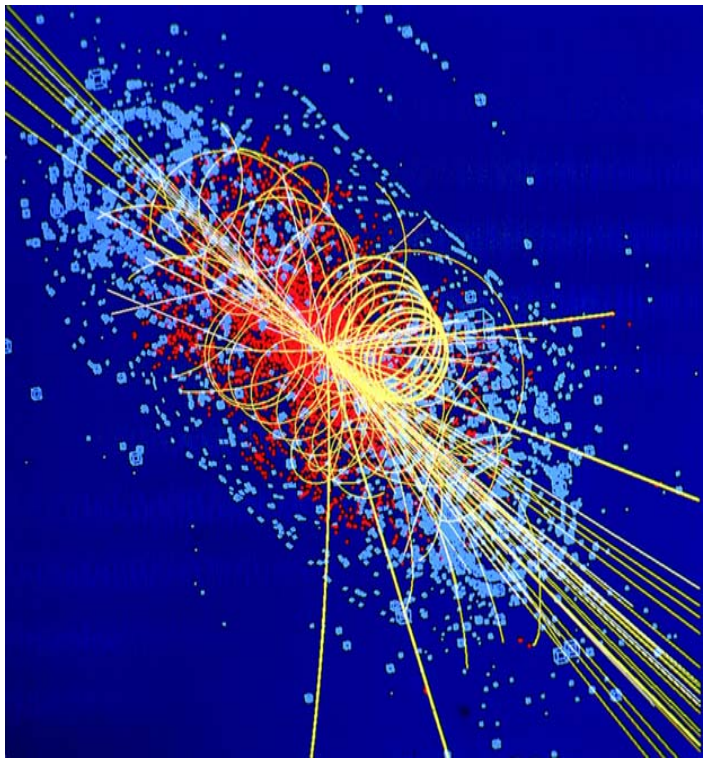
URA Visiting Committee

April 23rd, 2005

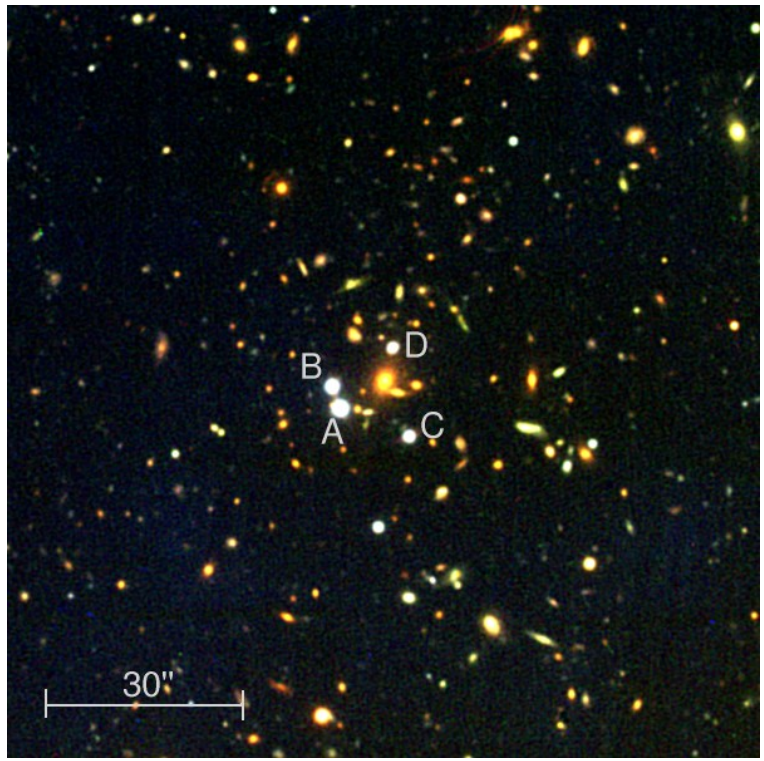
Image of the Week



Not a Discussion of Production Computing….



Simulated CMS Higgs Event



SDSS Probing G-Lensing.

Computing R&D Central to Fermilab's Future Program

- **Tightly coupled computing:**

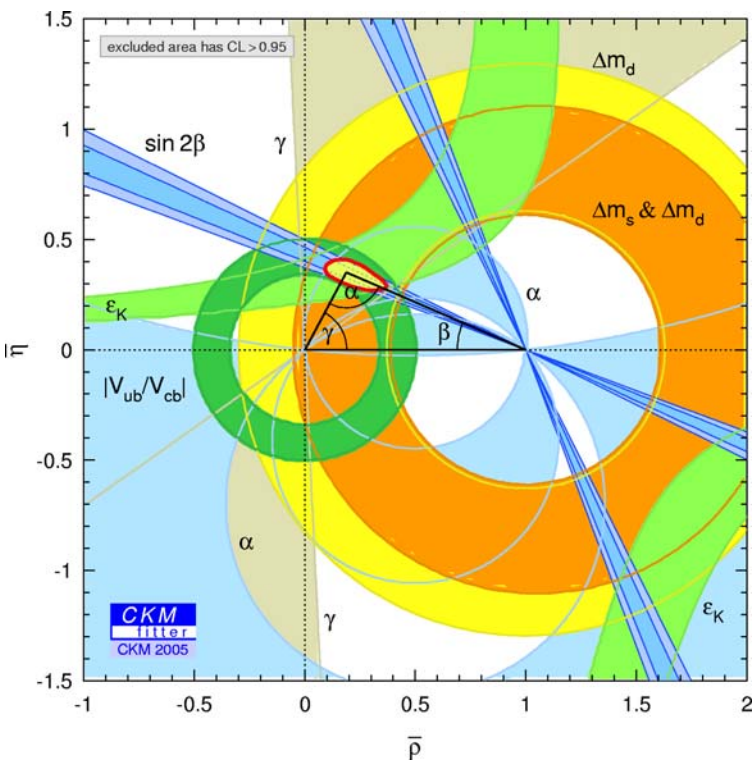
  Unraveling the present through Lattice QCD.
- **Simulating & Analyzing the present and future:**

  Adventures in Accelerator Simulation, evolution of C++, mathematical libraries for the field.
- **Enabling Technologies:**

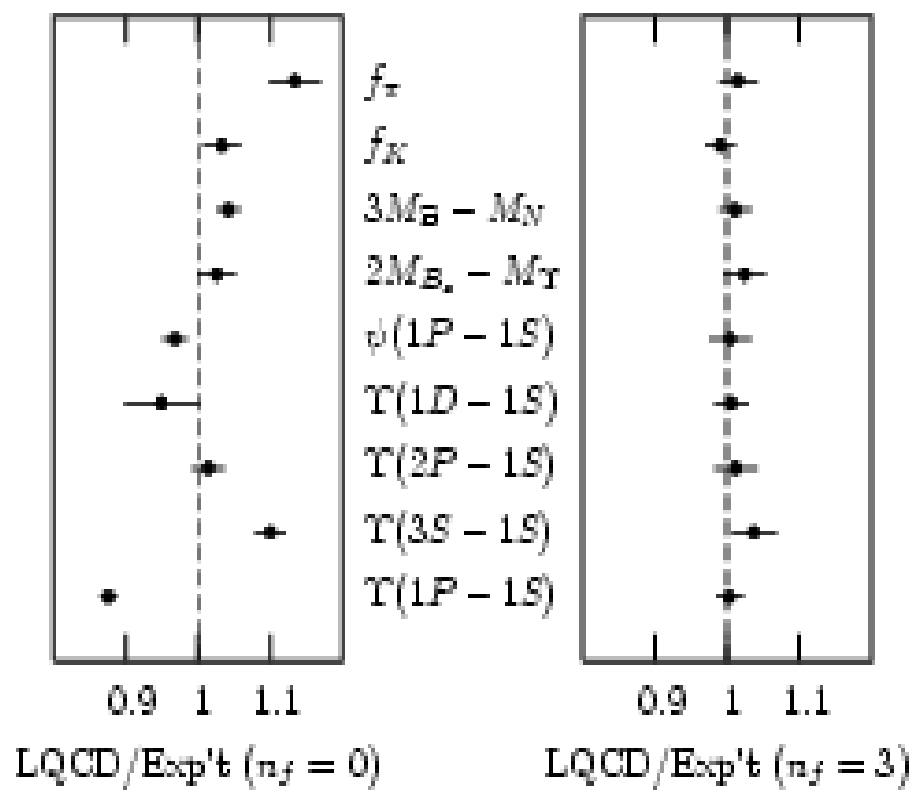
  Grid Computing, Networks, Mass Storage, Compute-Intensive Electrical Engineering.

Lattice QCD Drivers & Successes

Enabling access to the CKM matrix



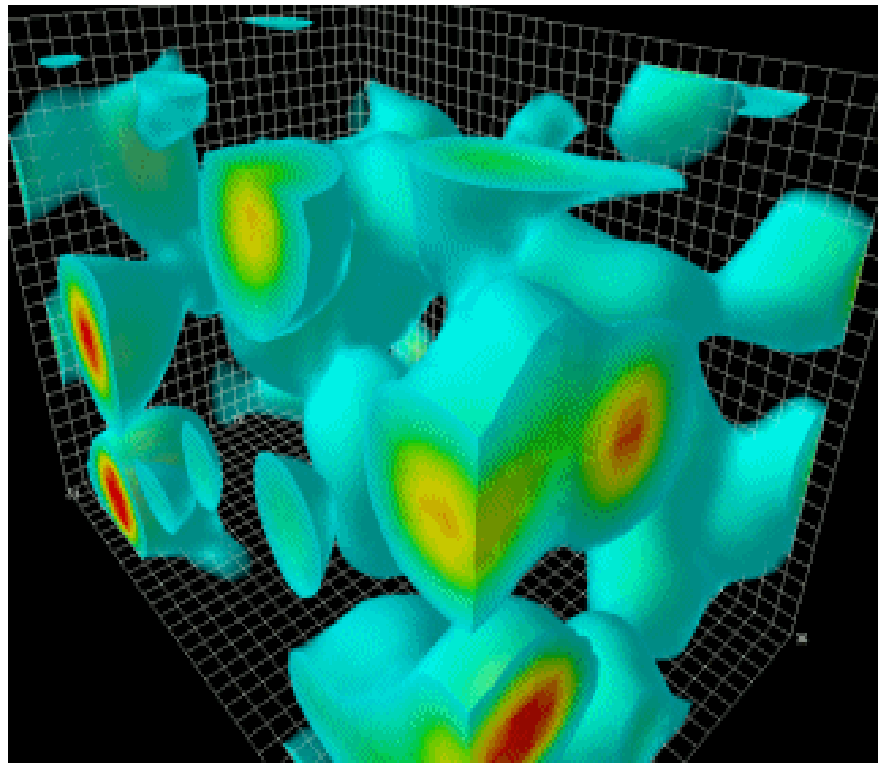
Probing QCD directly.



Realizing the QCD Action on the Lattice

- Lattice techniques key to connect measurements to underlying quark & gluon interactions.

- Two enabling technologies supported by the DOE: Tightly coupled commodity PCs and tightly coupled custom processors.



Gluon action visualization courtesy: Pr. D. Leinweber CSSM, University of Adelaide

Custom Processors optimized for QCD calculations.

office of high energy physics

Office of Nuclear Physics
Exploring Nuclear Matter - Quarks to Stars

SciDAC
Scientific Discovery through Advanced Computing
DOE Office of Science ASCR BER BES FE HENP MICS

Commodity PCs clustered with high performance networks.

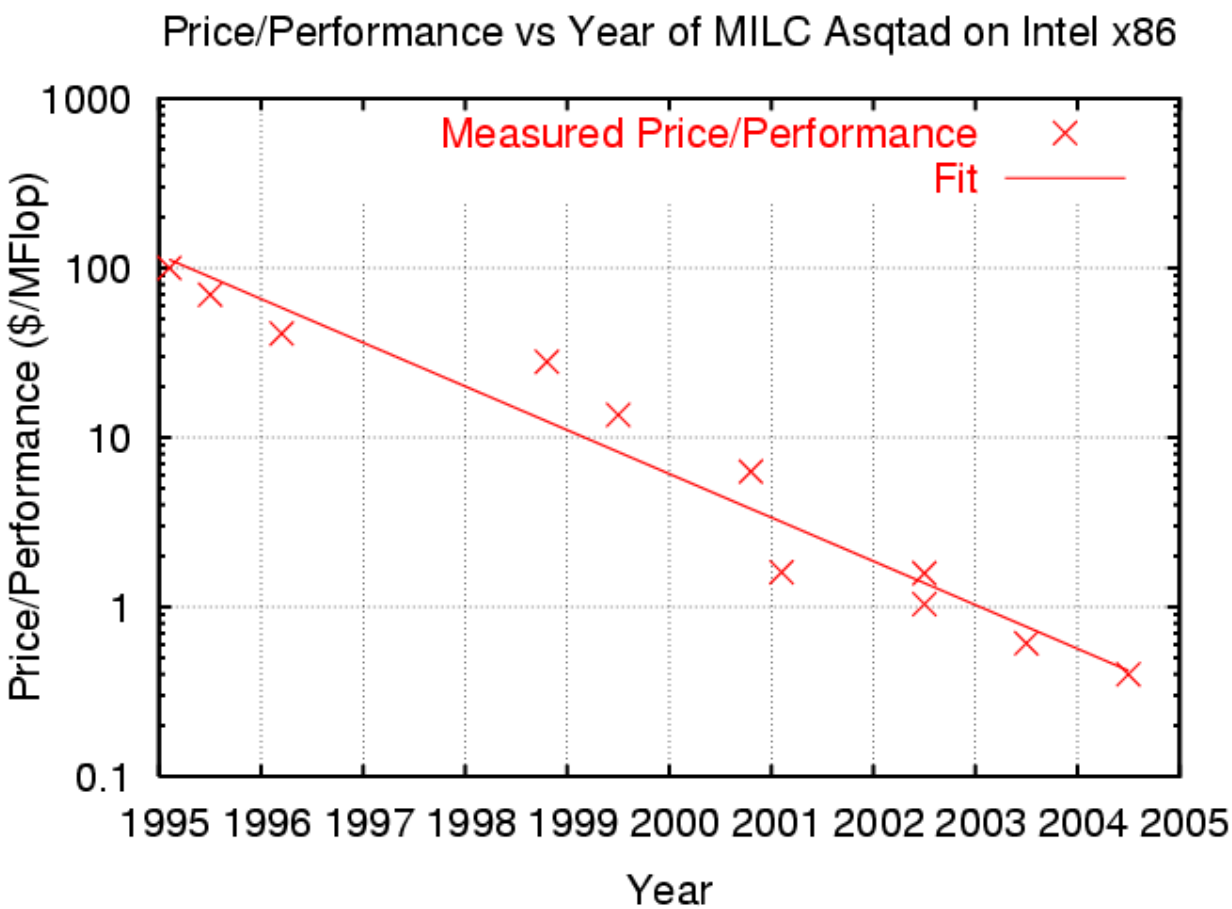
SciDAC Prototype Clusters

- Jefferson Lab (http://lqcd.jlab.org)
  - 256 node single 2.66 GHz Xeon, 3-D gigE mesh (9/2003)
    - Pr. Fodor's gigE mesh machines have had a huge impact on recent designs
  - 384 NODE SINGLE 2.8 GHz Xeon, 4-D gigE mesh (now)
    - $1700/node including networking
- Fermilab (http://lqcd.fnal.gov)
  - 128 dual 2.4 GHz Xeon, Myrinet (1/2003)
  - 32 dual 2.0 GHz Xeon, Infiniband (7/2004)
  - 128 single 2.8 GHz P4, Myrinet (7/2004)
    - $900/node, reused Myrinet from 2000
  - 260 single 3.2 GHz P4, Infiniband (2/2005)
    - ~ $900/node + Infiniband ($880/node)

QCDOC

- "QCD on a Chip" (http://phys.columbia.edu/~cqft/)
- Designed by Columbia University with IBM
  - in many ways, a precursor to Blue Gene/L
- Based on special Power PC core:
  - ~ 500 MHz G4 with 1 GFlops double precision
  - 4 MB of embedded DRAM
  - 12 bidirectional, 1 Gbit/sec serial links
  - 2.6 GB/sec external memory interface
  - fast ethernet
- Architecture:
  - PPC cores connected in 6-D torus
  - up to 20K processors
  - up to 50% of peak (~ 5 TFlops from 10K cpus)

Performance Trends – Single Node



Processors used:

- Pentium Pro, 66 MHz FSB
- Pentium II, 100 MHz FSB
- Pentium III, 100/133 FSB
- P4, 400/533/800 FSB
- Xeon, 400 MHz FSB
- P4E, 800 MHz FSB

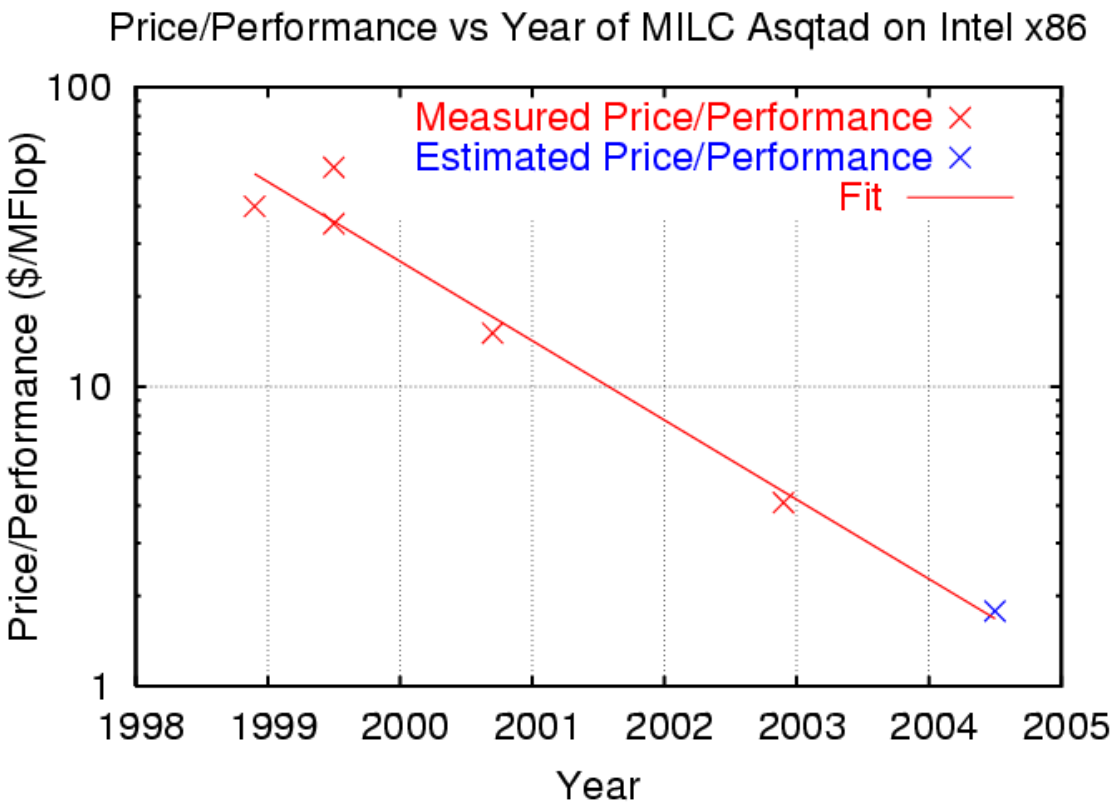
Performance range:

- 48 to 1600 MFlop/sec
- measured at 12^4

Doubling times:

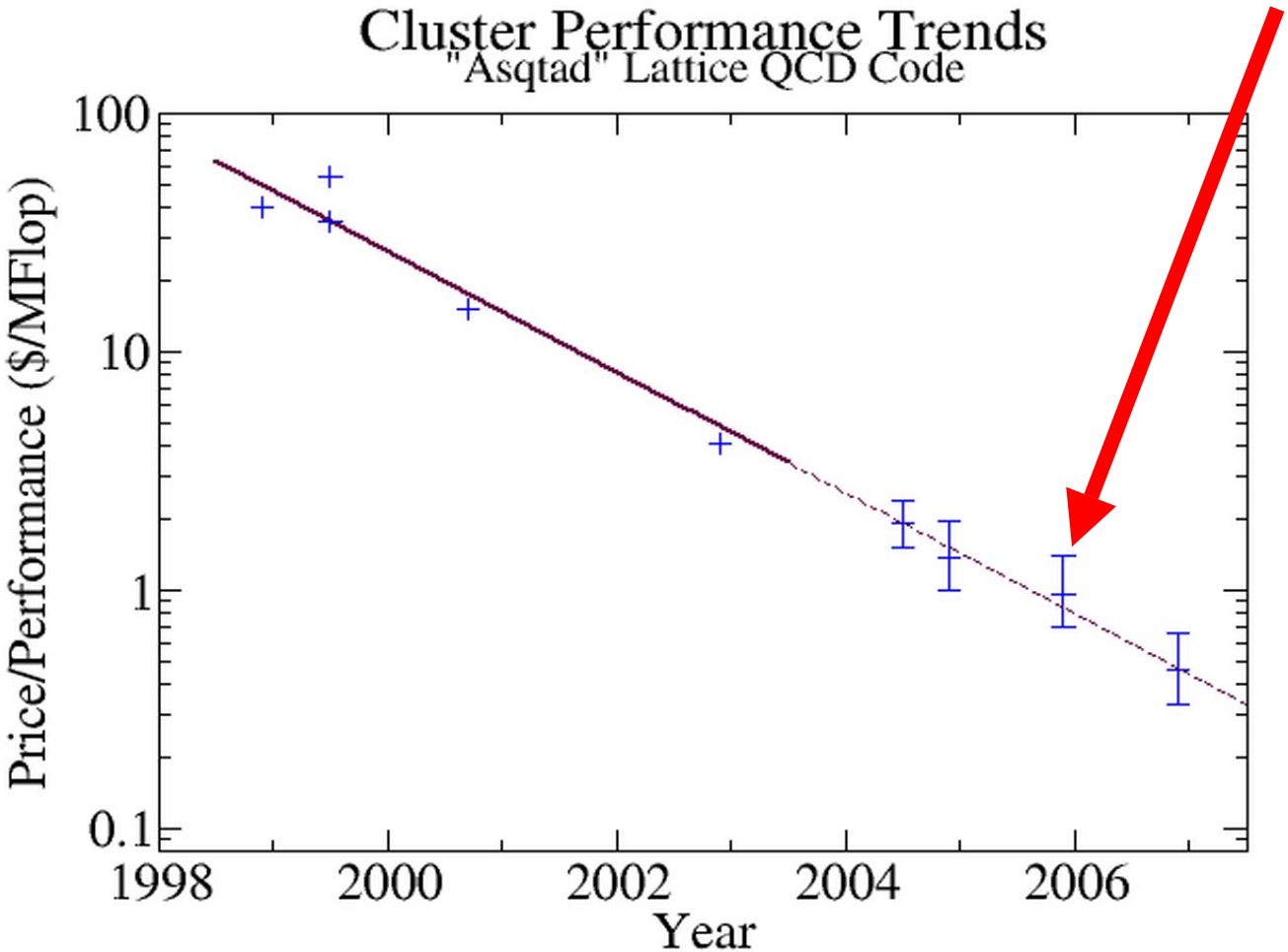
- **Performance: 1.88 years**
- **Price/Perf.: 1.19 years !!**

Performance Trends - Clusters



- Clusters based on:
  - Pentium II, 100 MHz FSB
  - Pentium III, 100 MHz FSB
  - Xeon, 400 MHz FSB
  - P4E (estimate), 800 FSB
- Performance range:
  - 50 to 1200 MFlop/sec/node
  - measured at 14^4 local lattice per node
- Doubling Times:
  - **Performance: 1.22 years**
  - **Price/Perf: 1.25 years**

Predictions



- Early 2006:
  - 4.0 GHz P4E
  - 1066 MHz FSB
  - PCI-Express
  - Infiniband
  - $900 + $600 (system + network per node)
  - 1.9 GFlop/node, based on faster CPU and higher memory bandwidth

The Future, USQCD: United States Lattice Gauge Theory Computational Program

- USQCD is a new initiative starting FY06.
- Fermilab is providing project management and leading roles in technical and scientific management
- Scope: $2M/yr HEP & $0.5M/yr NP.

BROOKHAVEN NATIONAL LABORATORY

Fermilab

Jefferson Lab

SciDAC
Scientific Discovery through Advanced Computing
DOE Office of Science ASCR BER BES FE HENP MICS

HEP
office of high energy physics

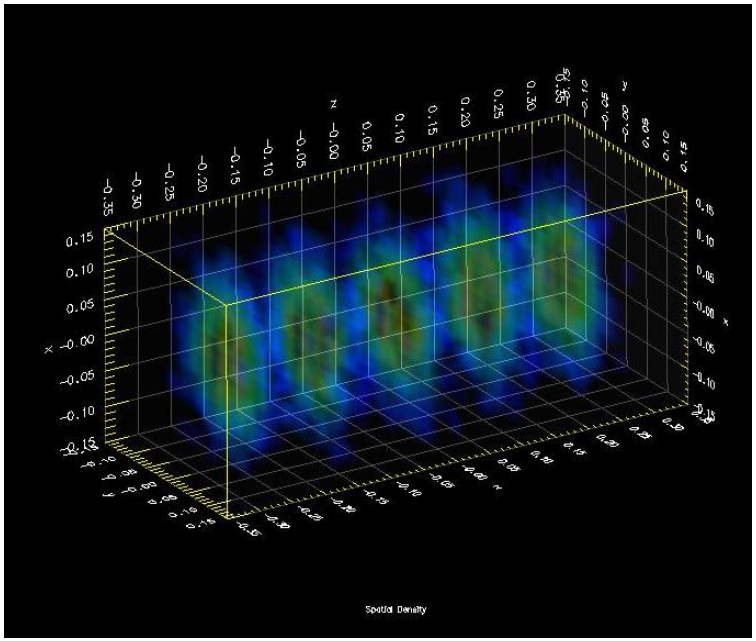
Office of Nuclear Physics
Exploring Nuclear Matter - Quarks to Stars

Συνεργεια

Simulating Present and Future Accelerators.

- Simulation of many particles in full phase space requires specialized parallel codes.
- “Synergia” is a software framework that provides a tightly coupled coding environment with accurate simulation of instrumentation.
- Evolving to full end-to-end simulations, adopting techniques from experimental HEP.

Case Study: The Fermilab Booster



Multi-bunch modeling in 3D

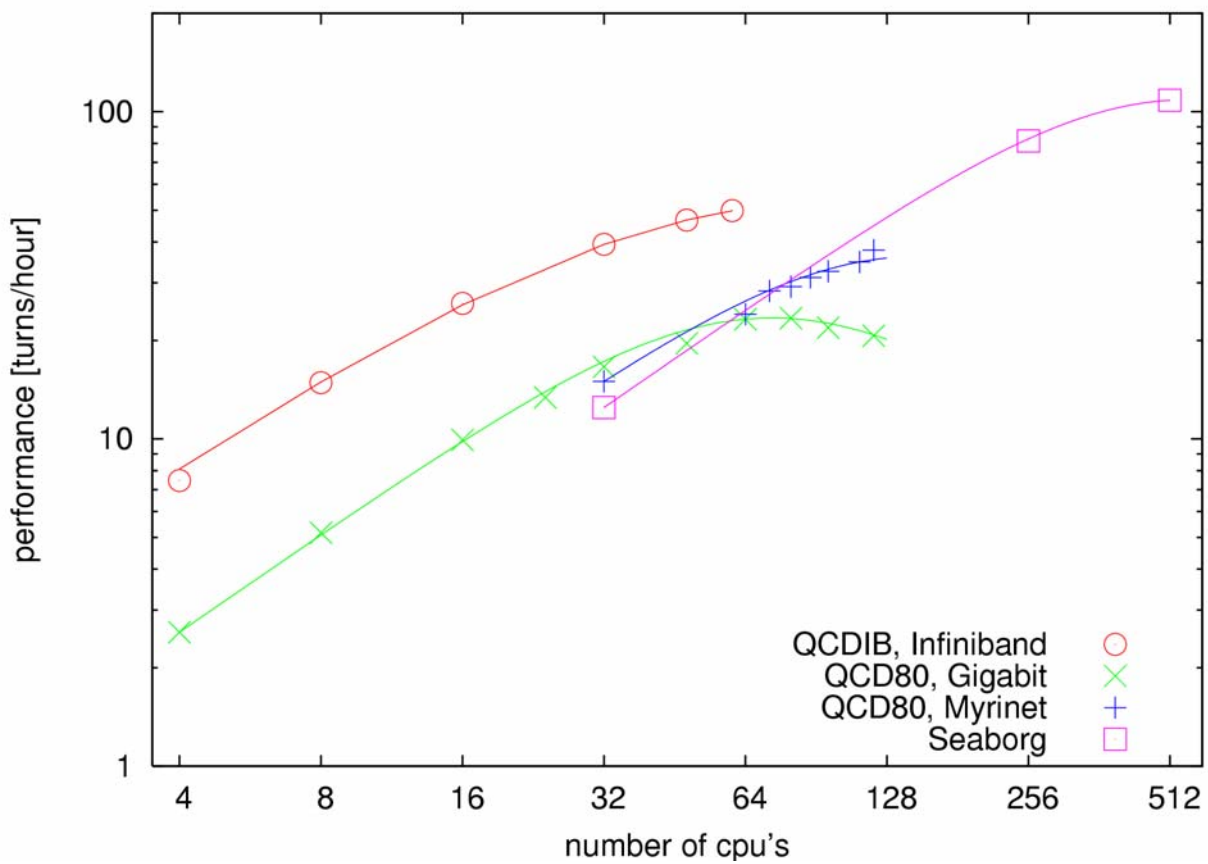
- FNAL Booster simulations follow 5 200 MHz Linac micro-bunches in a 37.8 MHz PhS slice.

- Fully 3D space-charge
- For high-precision simulations use 33x33x257 grid and ~1,000,000 particles
- Periodic boundary conditions – periodic/long beam
- Multi-turn injection
- 6-D PhS matched beam generation utilities

Computational Performance

- parallel computing required
- Code ported to supercomputers (NERSC) and parallel PC clusters
- for a FNAL Booster simulation:

~100 turns/hr on 512 procs



Performance model: $t = \frac{A}{N} + B\,N$

The FNAL Booster

The Village Crier

national accelerator laboratory

Operated by Universities Research Association Inc.
Under Contract with the United States Atomic Energy Commission

VOL. 5, NO. 45 DECEMBER 20, 1973

THE BOOSTER ACCELERATOR

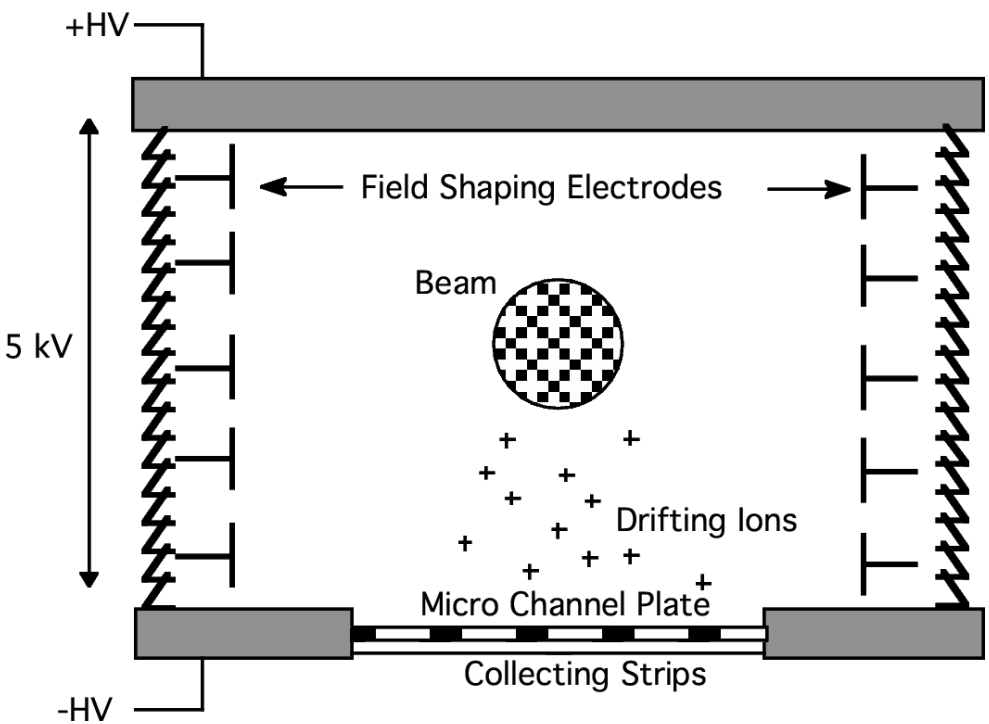
Since May, 1971, the Booster Accelerator has taken 200 million electron volt protons from the Linac and "boosted" their energy to 8 billion electron volts. They are then injected into the Main Acceelerator. The Booster synchrotron is about 500 feet in diameter. Twelve "bunches" of Booster protons are required to fill the Main Accelerator.

In December, 1970, the last Booster magnet was installed in the completed tunnel begun in 1969. Staff members continue to improve performance of the Booster.

- Rapid cycling, 15 Hz
  400 MeV → 8 GeV
- 24 FOFDOOD cells,
  total length 474.2 m
- RF 37.7 → 52.1 Mhz
- Injection/capture ~ 2 ms
- Multiturn injection, typically
  12x0.035 mA = 0.42 mA
- Here focus on first few hundred turns

1973: "Staff members continue to improve Booster Performance."
2005…….

Some Instrumentation: Booster Ion Profile Monitor (IPM) Detector

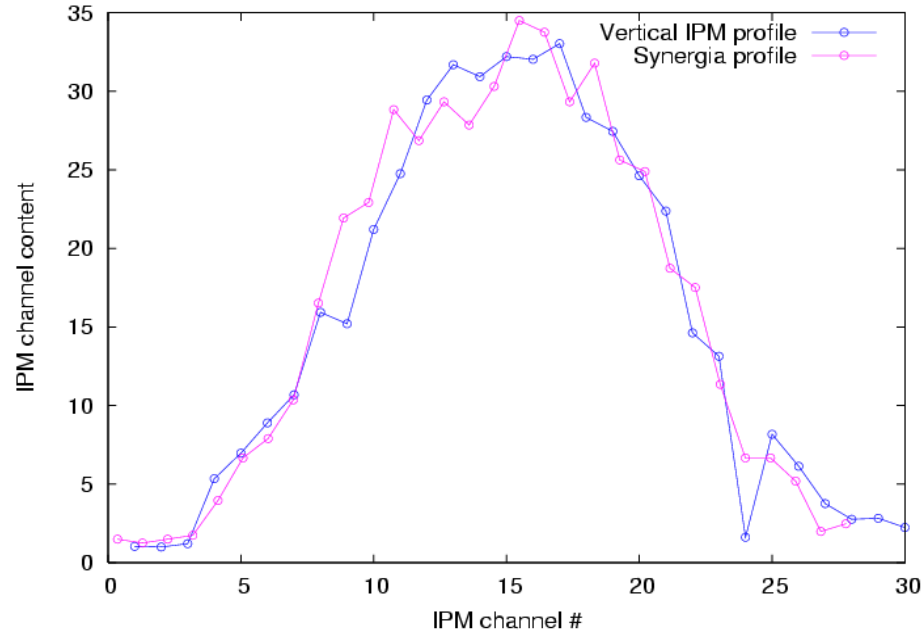


- Turn-by-turn beam profiles in horizontal and vertical planes
- Response affected by beam charge
- We have done detailed theoretical and experimental work on calibration *Phys. Rev ST AB* ***6****:102801, 2003*

Field generated by beam charge smears measured beam profiles

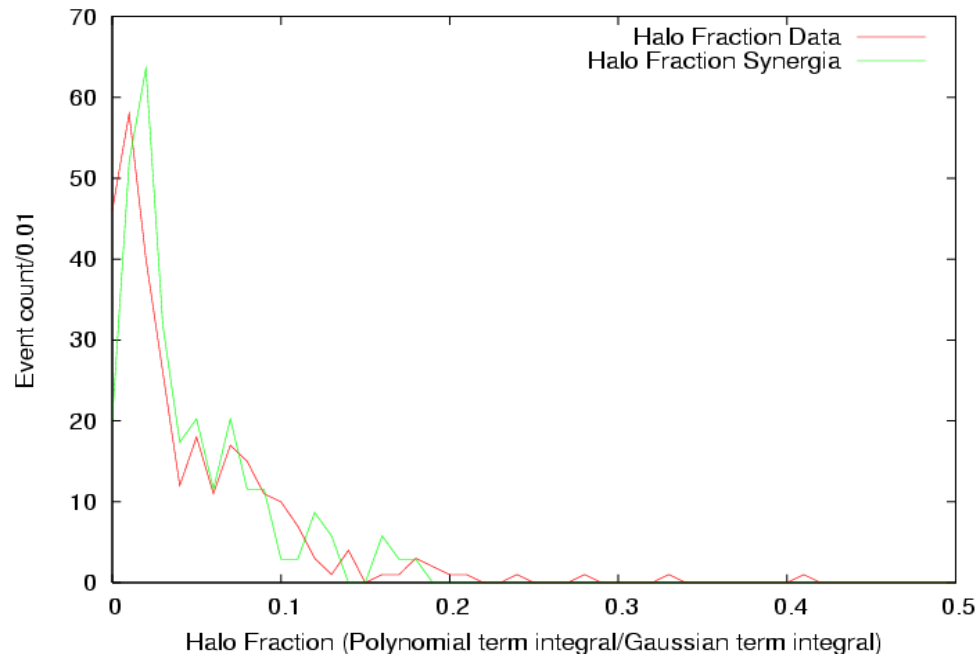
Συνεργεια

Measured IPM Profile well modelled by Synergia Profile.



Smeared Synergia vs data profiles

Non-Gaussian "Halo" component of IPM Profile well modelled.



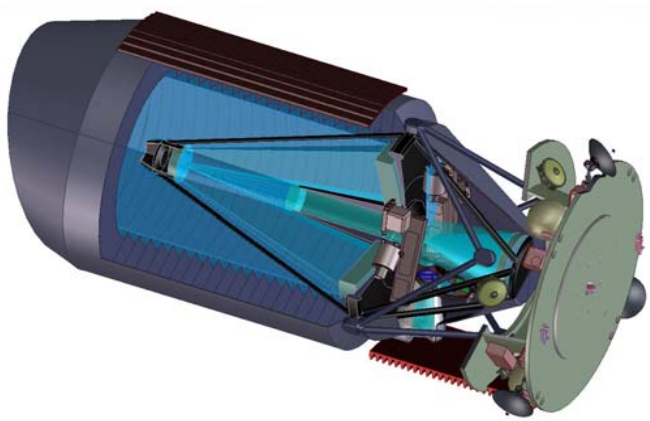
Συνεργεια

Future Directions for Synergia

- Electron cooling for the Fermilab collider complex.
- Beam-Beam effects at the ILC interaction point.
- Space charge effects in the ILC damping rings.
- Electron cloud effects in the Main Injector/Proton-Driver.

CS-Intensive Engineering: A Billion Pixels in Space

- Validation of Flash Memory designs.
- Development of novel nearly lossless data compression algorithms.
- Development of FPGA based data compression.



Enabling Technologies: Grid, Networks & Storage

Our former centralized computing model doesn't scale to future need.

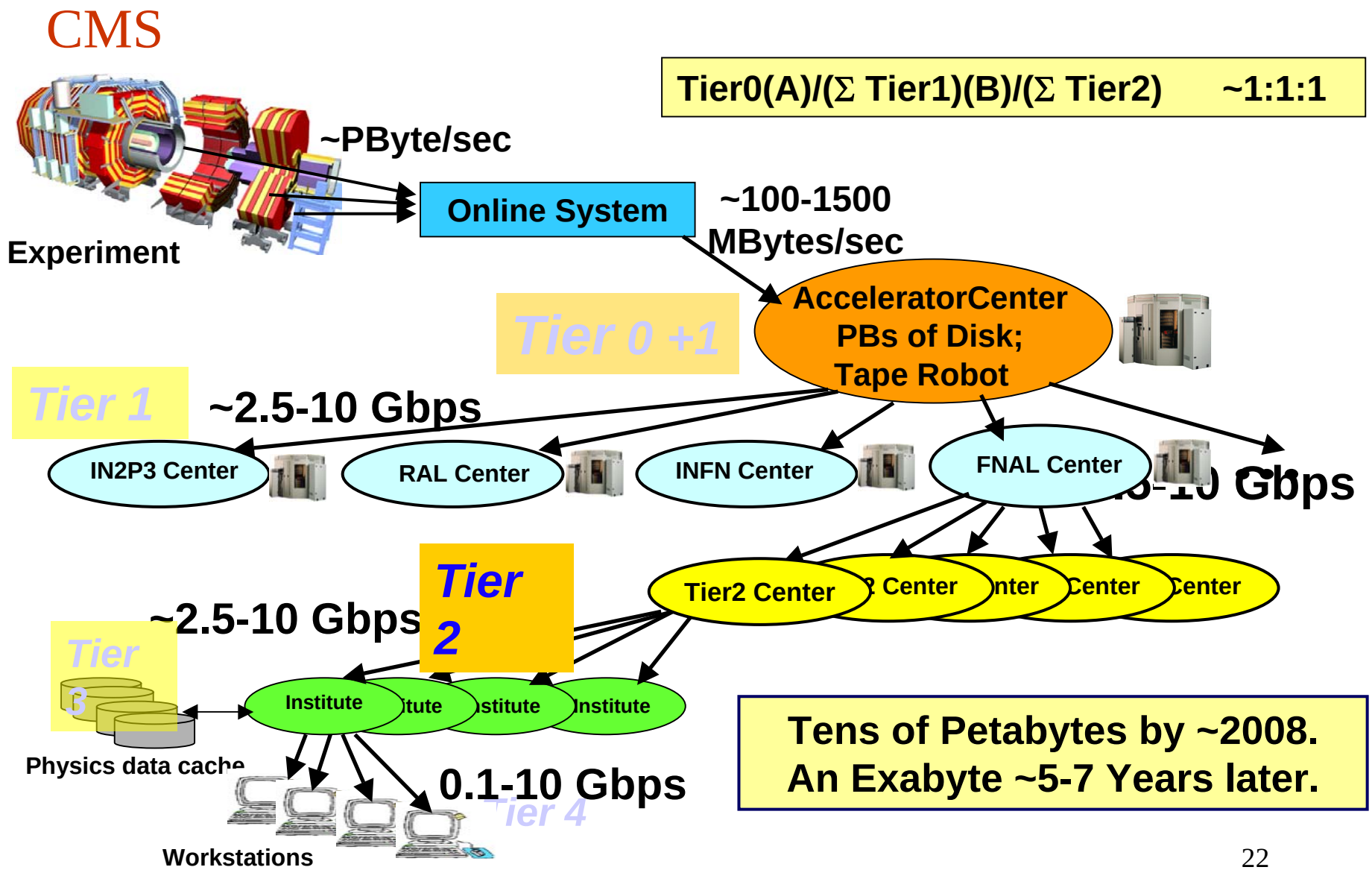
Institutions have invested in IT, leveraged among local stakeholders.

The challenge ahead is to globally leverage these resources.

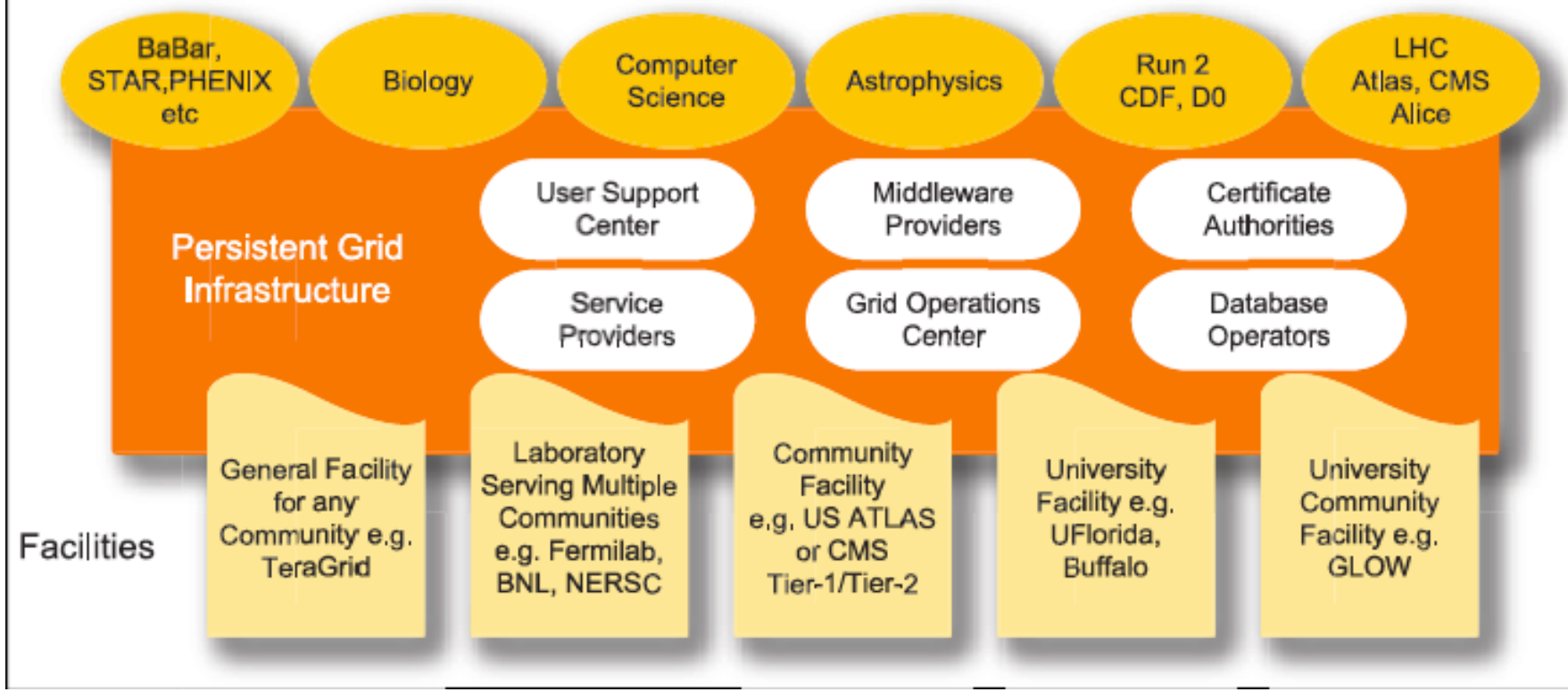
Open Science Grid

LCG

A GRID We Need…Soon!



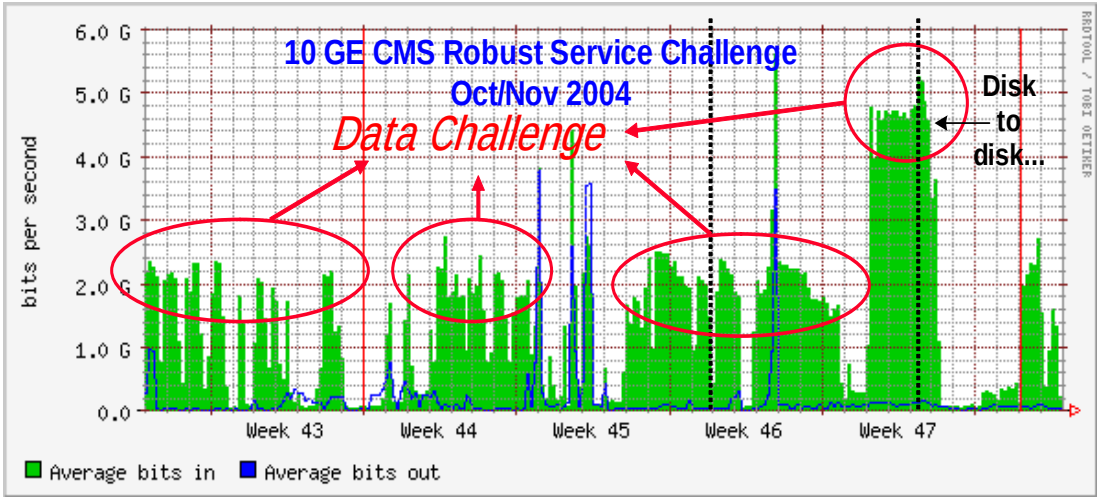
Open Science Grid



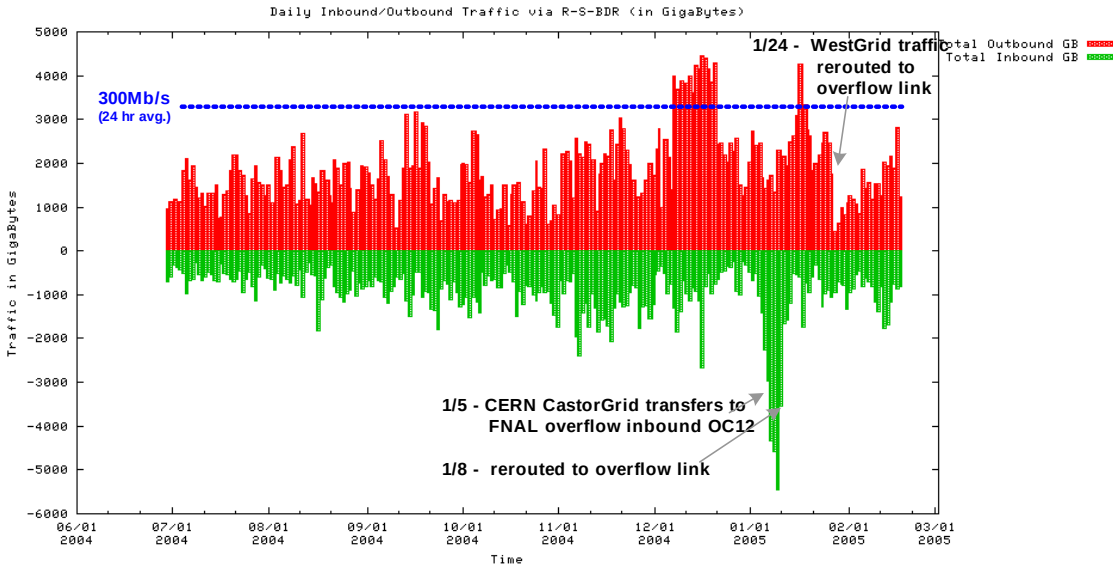
http://www.opensciencegrid.org

Enabling Technology: Network R&D on the "Starlight" Link.

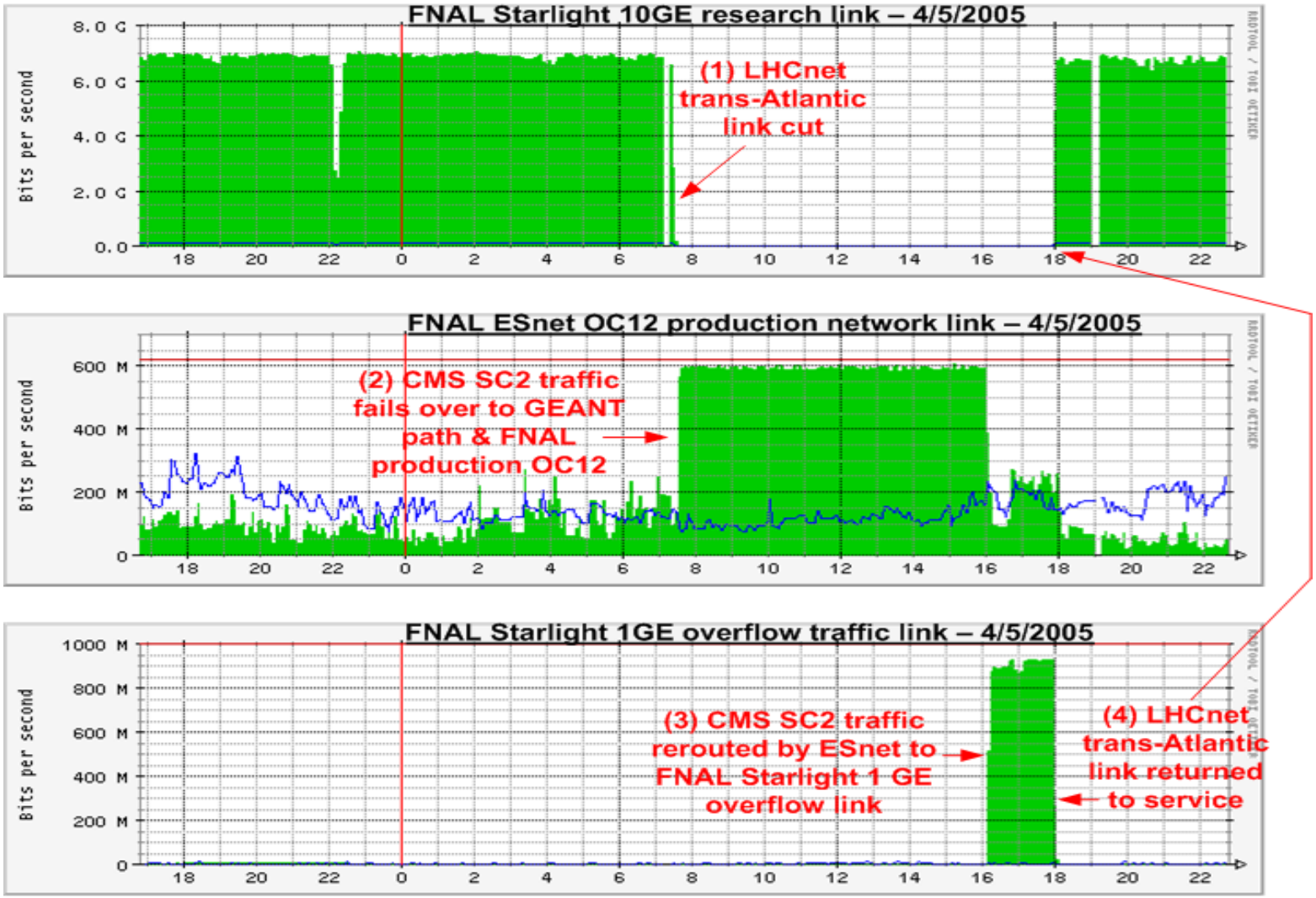
- CMS robust service challenge sustained 2.5 Gb/s for weeks
- SC2004 = 7.5 Gb/s sustained



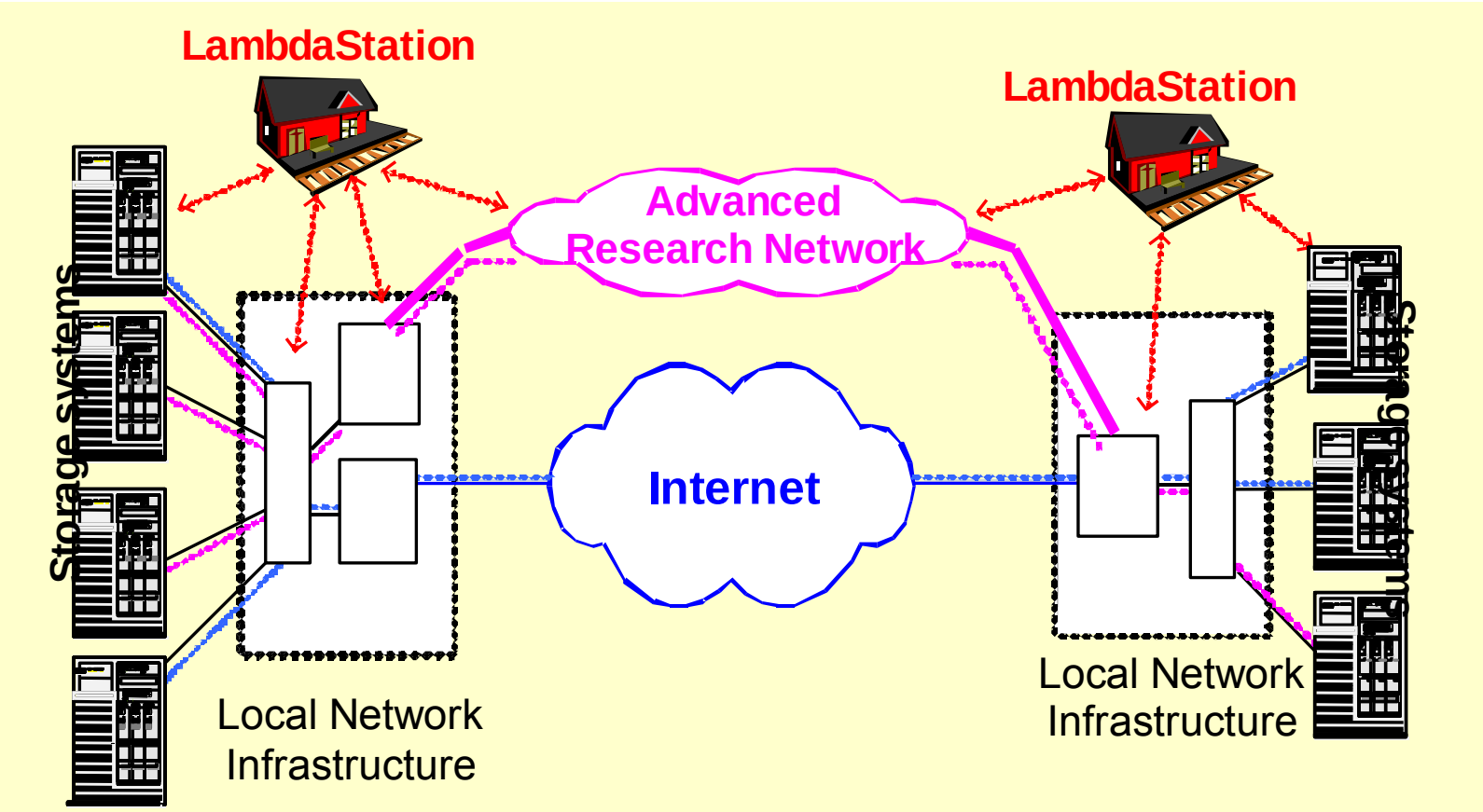
Production ES-Net OC-12 Link is saturated!



Program Need Outstrips Current Capacity…



ASCR funded R&D project. Addresses the last-mile problem of connecting resources to high bandwidth WANs.



Core Grid Middleware: Storage Resource Management (SRM)

Protocol Development. Focus on facility interoperability. Real collaborative success: LBL(Disk Resource Manager), JLAB(JLAB Storage Manager), FNAL-DESY(Dcache), International Lattice Data Grid. Grass roots effort with some key explicit SciDAC support.

Performant Data Grids for HEP as part of the broader community

- We are working with Computer Science research and development teams **(Condor, Globus)** to develop the capabilities and performance for Grids for HEP
- We are working together with Nuclear Physics, Astrophysics, and Bioinformatics scientists on a shared infrastructure.
- We are contributing to Grid standards and services through developing to our experiment needs and contributing to the common program of work for Grid Security, Storage, Operations, Workload management etc.

Open Science Grid Consortium is the Means

- Building on success of Grid3 and continued collaboration with **GriPhyn, iVDGL, PPDG**.
- Collaboration between many Lab Facilities and University IT groups, experiments, computer science groups etc.
- Collaborate and encourage Interoperability with other Grids in the US - TeraGrid, Campus Grids such as GLOW, GRASE etc.
- US LHC commitment to present their resources to OSG and to work in a common shared environment. Close working relationship with LCG and EGEE in Europe.
- Interest from Korea, Taiwan, Brazil collaborators.

Conclusions

- Lattice QCD now a precise tool. Computing Research enabled this development, and will ensure an incisive future.
- Tightly coupled accelerator simulations with integrated instrumentation modelling will be a key element in realizing future machines.
- Enabling technologies of Grid, Networking & Storage are thriving through vigorous collaboration across a broad span of stakeholders.