Computing Challenges for Large Experiments

DOE Review
Brookhaven National Laboratory
July 26, 2002

Torre Wenaus, BNL/CERN
Physics Applications Software Group Leader, BNL Physics Dept
Software Manager, US ATLAS
Applications Manager, LHC Computing Grid Project
wenaus@bnl.gov

This Talk

Preface:

- Talk will be representative, not comprehensive
  - Major BNL roles in RHIC, ATLAS, D0
- Emphasis on ATLAS and STAR

Outline:

- Computing for large experiments at BNL
- Computing challenges in ATLAS
- The Data Grid
- Conclusion

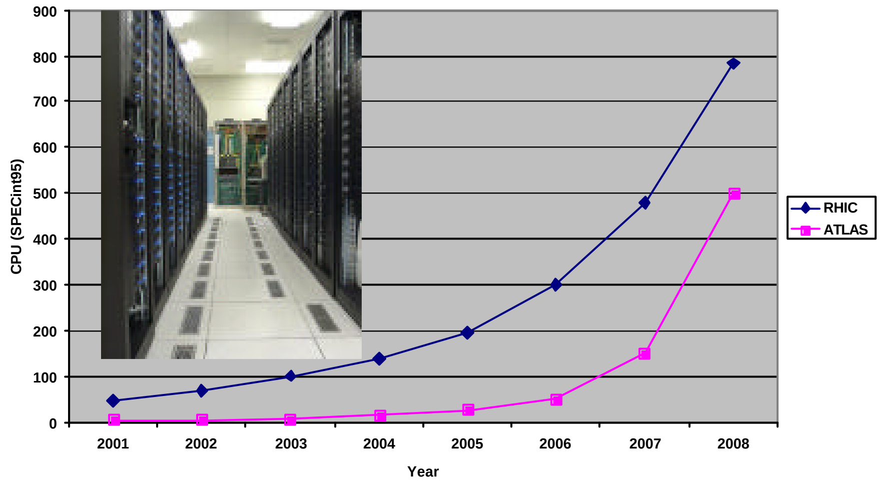
Computing for Large Experiments at BNL

- **BNL must be a leader in state-of-the-art HENP computing for success in the physics**
- Host lab for RHIC, one of the most computationally demanding HENP activities today
  - Petabyte data volumes analyzed worldwide
- U.S. host lab for ATLAS, leading Tier 1 Facility and software projects
  - Ensure the computing for a leading U.S. physics role is in place
  - Small relative to RHIC today – physics in 2007
    - LHC computing scale will be ~5x that of RHIC

BNL Facilities

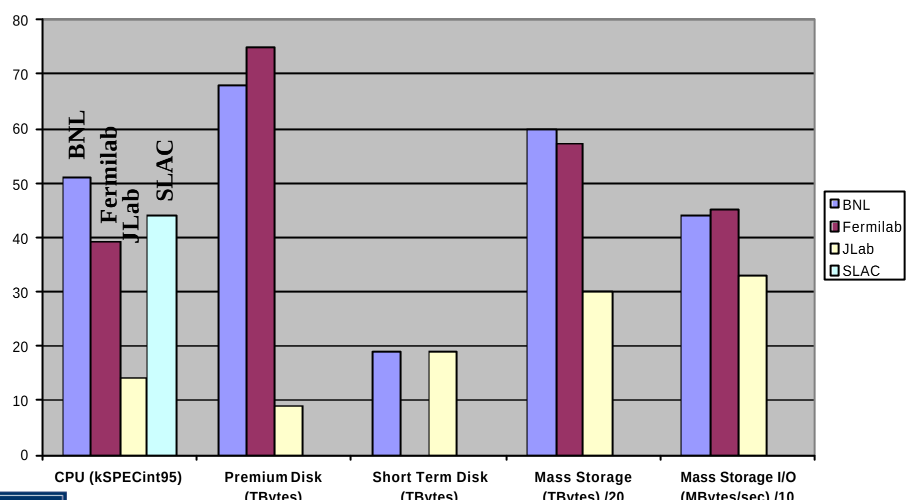
- RHIC Computing Facility (RCF) and ATLAS Computing Facility (ACF) are closely allied
- Strong synergy: almost identical requirements and configurations
- Managed sharing of resources (human and material) and expertise benefit both programs
  - e.g. ATLAS runs on quiet STAR CPUs & vice versa
- Facility is at the frontier:
  - In deployed capacities of CPU, Disk, Mass Storage, & Network
  - In utilization of emerging **Data Grid** technology
    - Bringing practical benefits to both RHIC and ATLAS in distributed data and production

Projected CPU Capacity Growth



Fully integrated support staff of (currently) 25

Comparison to Other Major NP/HEP Facilities



BNL Software: From RHIC to the LHC

- RHIC software successful in enabling a strong physics program
  - Physics results presented ~1 month after RHIC startup
  - Good performance as data volumes have grown
- Proven RHIC approaches RHIC being propagated to LHC
- My roles:
  - STAR Software and Computing Leader 1997-2000
    - Led core software team at BNL
  - Physics Applications Software (PAS) Group Leader
    - STAR team migrating to ATLAS and expanding
    - Transmitting RHIC experience to LHC
  - Software Manager for US ATLAS

RHIC (STAR) Software for the LHC

- 'Hybrid' object-oriented event store combining files with an open source database: a community first
  - Recently adopted by ATLAS and the LHC
- LDRD project 'NOVA' for distributed analysis tools; precursor to Grid, based on STAR experience
  - Basis for several tools adopted by ATLAS
    - Distributed Data Manager; Virtual Data Catalog; Parameter Database
- ATLAS distributed production system based on STAR
- STAR deployed Grand Challenge 'STACS' mass store management software
  - Evolved into PPDG SciDAC project (Shoshani LBNL)

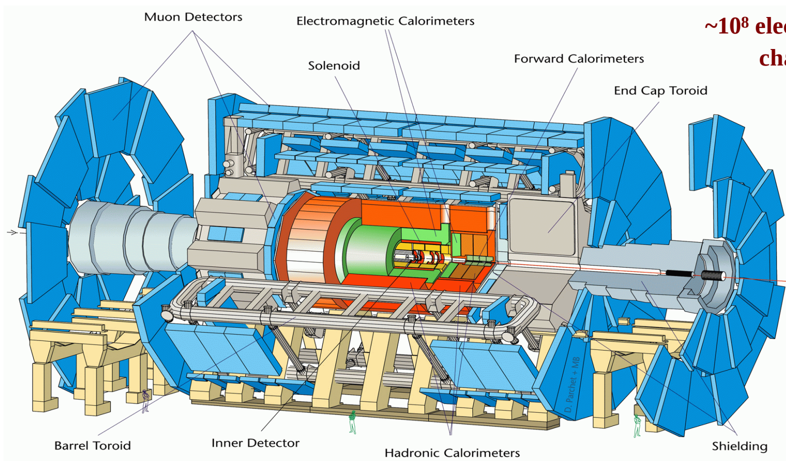
ATLAS Software at BNL

- PAS group activities:
  - Migrating successful RHIC approaches to the LHC
  - Software support for the U.S. ATLAS community
  - SciDAC-sponsored Particle Physics Data Grid, collaborating with computer scientists
    - Developed distributed data management system now used in production throughout ATLAS
    - Extending program to distributed processing and transparent grid-based analysis
  - Support and participation in detector software efforts

BNL and the LCG

- LHC Computing Grid Project established in March to prepare and deploy the LHC computing environment
- Active in
  - Common software for physics applications
    - Data management, simulation, analysis tools,…
    - ~20 people today, growing to 40-50
  - Computing for the LHC
    - Facilities, grid middleware and deployment
- I lead the Physics Applications Software activity
  - BNL in a central, high visibility role
  - Strong coupling to our local software work

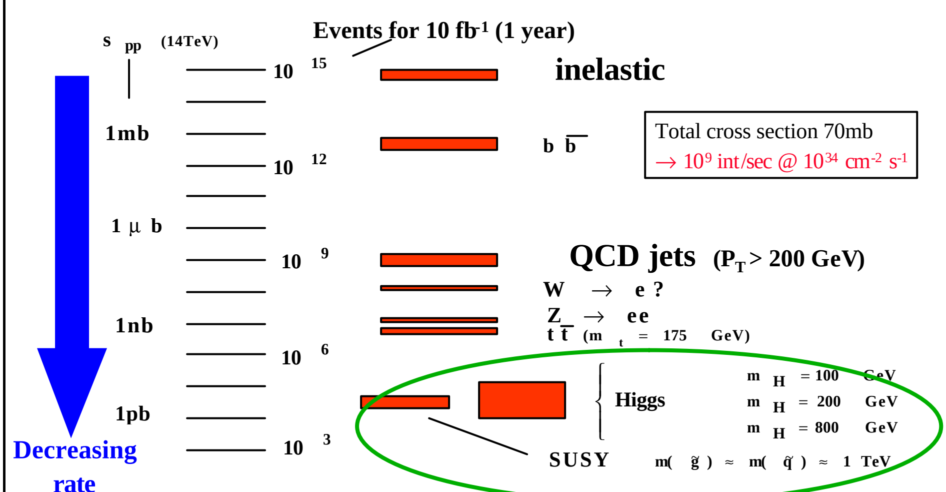
ATLAS Detector at the LHC



~10^8 electronics channels

33 U.S. Institutes

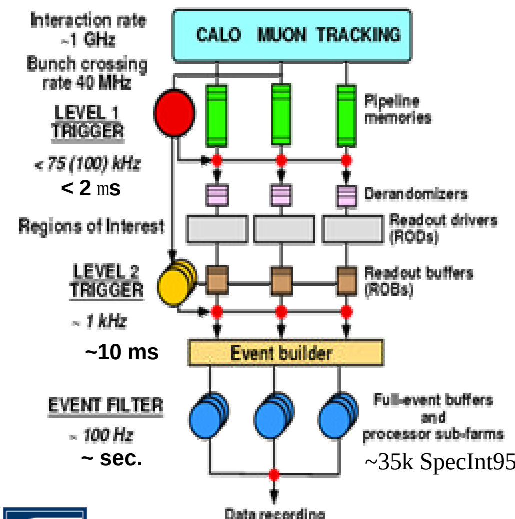
PP Cross Section and Event Production Rates



Total cross section 70mb
→ 10^9 int/sec @ 10^{34} cm^{-2} s^{-1}

Low rates for new physics demand
➔ Online: Highly selective trigger
➔ Offline: Efficient extraction of sparse samples from huge datasets

ATLAS Online System



Physics selection of the 100 'best' events/sec:

40 MHz, 1 PB/sec

Level 1: Coarse calorimeter data and muon trigger chambers

75 kHz, 75 GB/sec

Level 2: Full information from all detectors in regions of interest

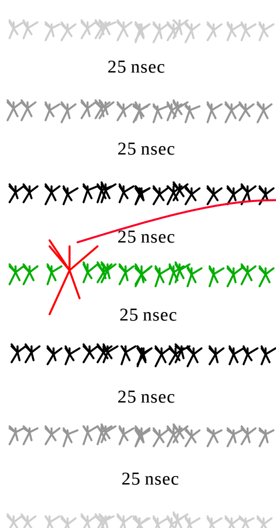
1 kHz, 1 GB/sec

Event Filter: Reconstruction of complete event using latest alignment and calibration data

100 Hz, ~100 MB/sec

~10 TB/day, 1 Petabyte/year of recorded raw data

Events in ATLAS



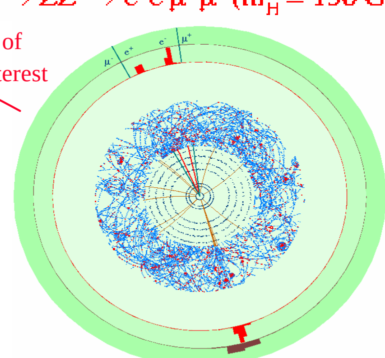
Bunch crossing every 25 nsec

Each with ~20 interactions

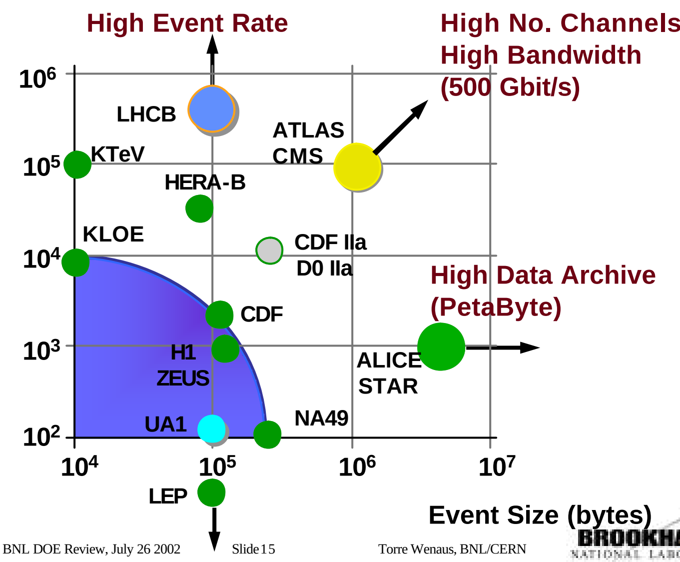
Event contains the products of many interactions

$H \rightarrow ZZ^{*} \rightarrow e^{+}e^{-}\mu^{+}\mu^{-}$ ($m_H = 130$ GeV)

Interaction of interest

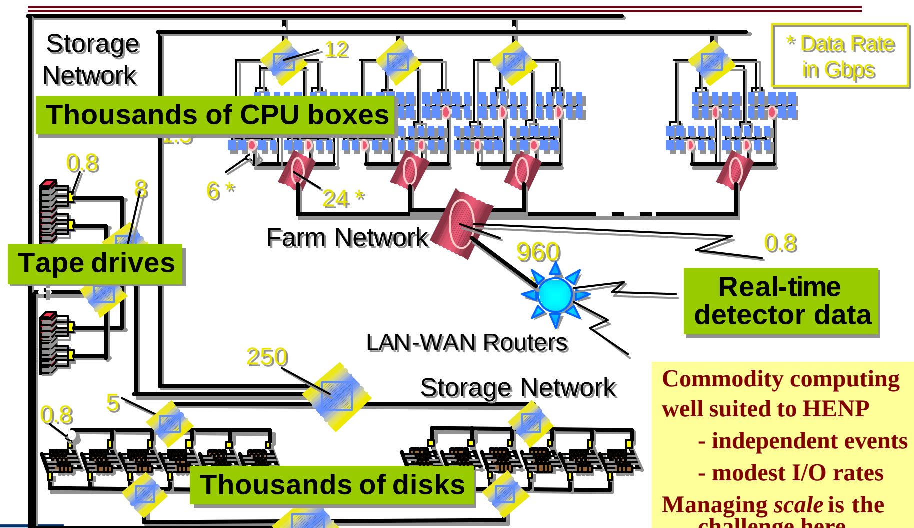


Data Handling in HENP Experiments



STAR:
20MB, 1 kHz
ATLAS:
1MB, 75 kHz

LHC-Era Computing Farm



Commodity computing well suited to HENP
- independent events
- modest I/O rates

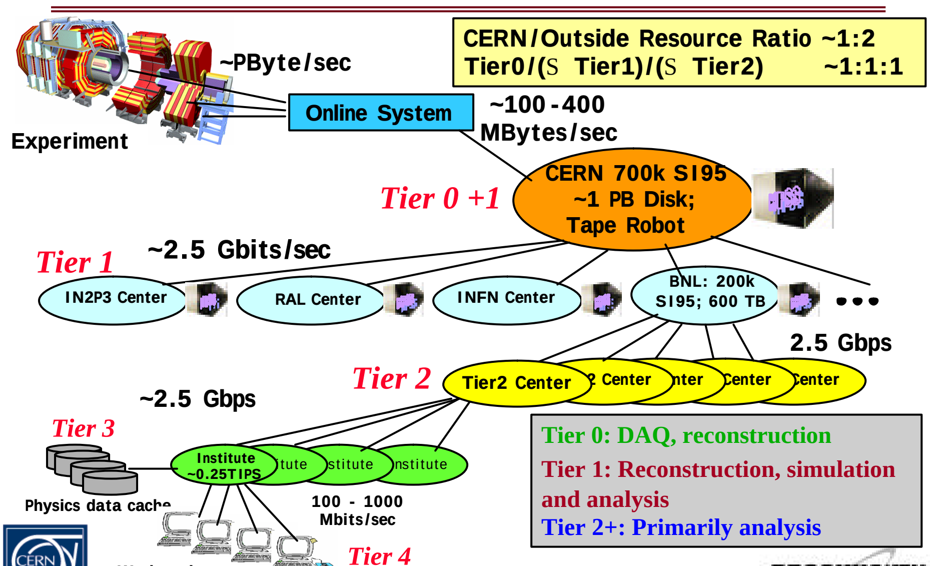
Managing *scale* is the challenge here

Beyond the Farm: Distributed Computing

- Central *requirement*: enabling remote physicists in analysis
  - Location independence of data access and analysis
- Central *reality*: needed capacity can only be achieved by employing widely distributed resources
  - Limited central resources; national and local interests

➔ Distributed computing is an essential reality

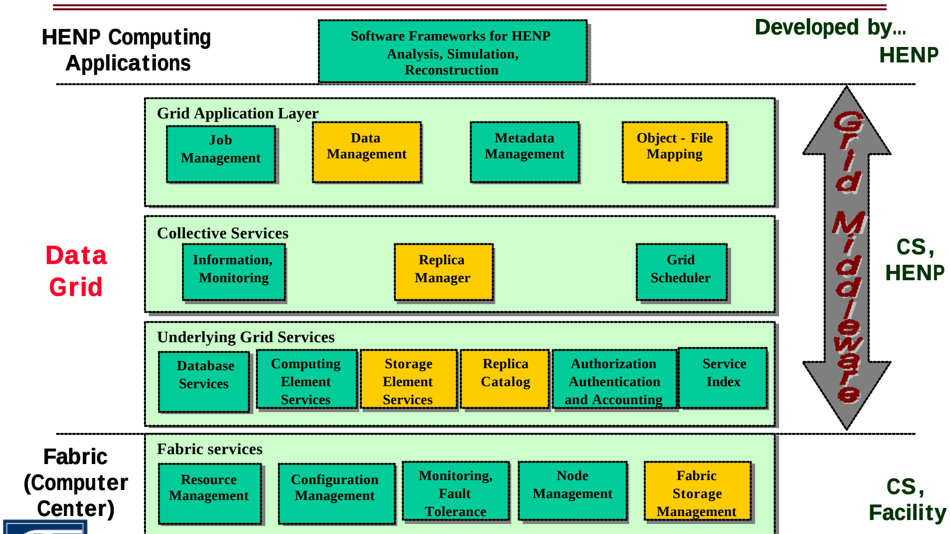
LHC Distributed Computing Model



The Data Grid

- Distributed computing brings major new challenges:
  - Large scale distributed data management
  - Management of distributed processing resources
  - Distributed monitoring and error recovery
  - Global authentication and security
  - Efficient, fault tolerant use of networks
- These are more **computer science** than HENP
  - The complex enabling software required doesn't exist
- The solution: collaborate with the CS community
  - Develop the **Data Grid** as the enabling software for HENP computing

Data Grid Architecture – Highlighting Data Management



What the Data Grid will do for the physicist

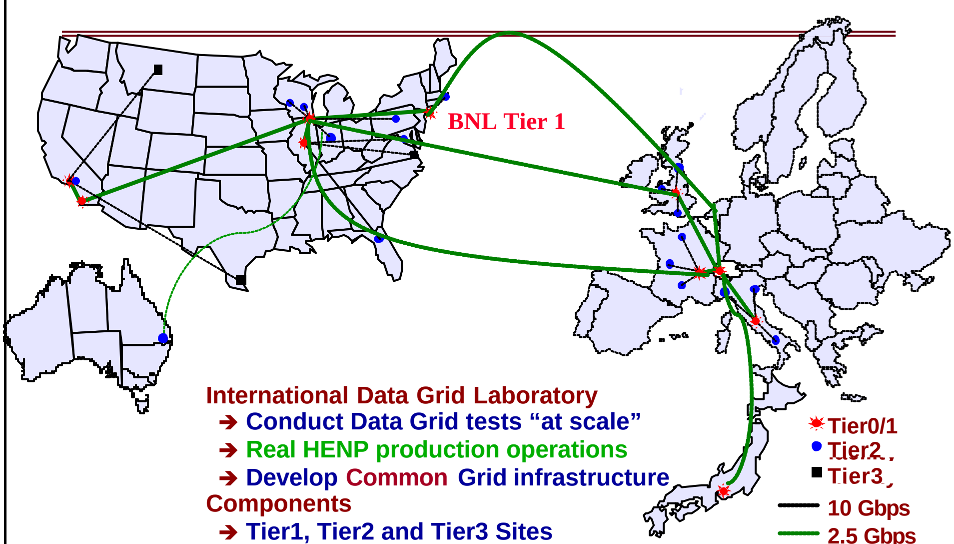
Physicist submits work, and the Grid…

- Finds convenient places for it to be run
  - Optimizing distributed resource usage
  - Providing efficient access to data
- Authenticates to sites in use
- Allocates needed CPU, storage resources
- Runs the jobs
- Monitors progress
- Recovers from problems

… and notifies user of completion.

Early versions of components exist; integration underway

Grid Infrastructure Circa 2003



International Data Grid Laboratory

- ➔ **Conduct Data Grid tests "at scale"**
- ➔ **Real HENP production operations**
- ➔ **Develop Common Grid infrastructure Components**
- ➔ **Tier1, Tier2 and Tier3 Sites**
- ➔ **0.6 - 10 Gbps networks**

Conclusions

- Greatest challenges: overcoming the *scale* and *distributed nature* of the task to provide *full physics participation* to global collaborators
- Our 'specialized computing platform' is the **Data Grid**
  - Close collaboration with computer scientists
    - New computing domain with broad applications
  - HENP and BNL are a development/deployment environment
    - First productive applications; no longer just R&D
- BNL is at leading edge of HENP computing
  - Not a luxury: essential to physics program and responsibilities to users
  - Must be consolidated and grow with support of the Lab and its funders