PROOF in PHOBOS

Maarten Ballintijn / MIT

maartenb@mit.edu

ROOT Workshop September 2005

Geneva Switzerland

Outline

- The PHOBOS Experiment at RHIC

- Data Analysis at RCF

- PROOF at RCF

- Analysis Examples

- PROOF Developments at MIT

- Future Work

Relativistic Heavy Ion Collider (RHIC)
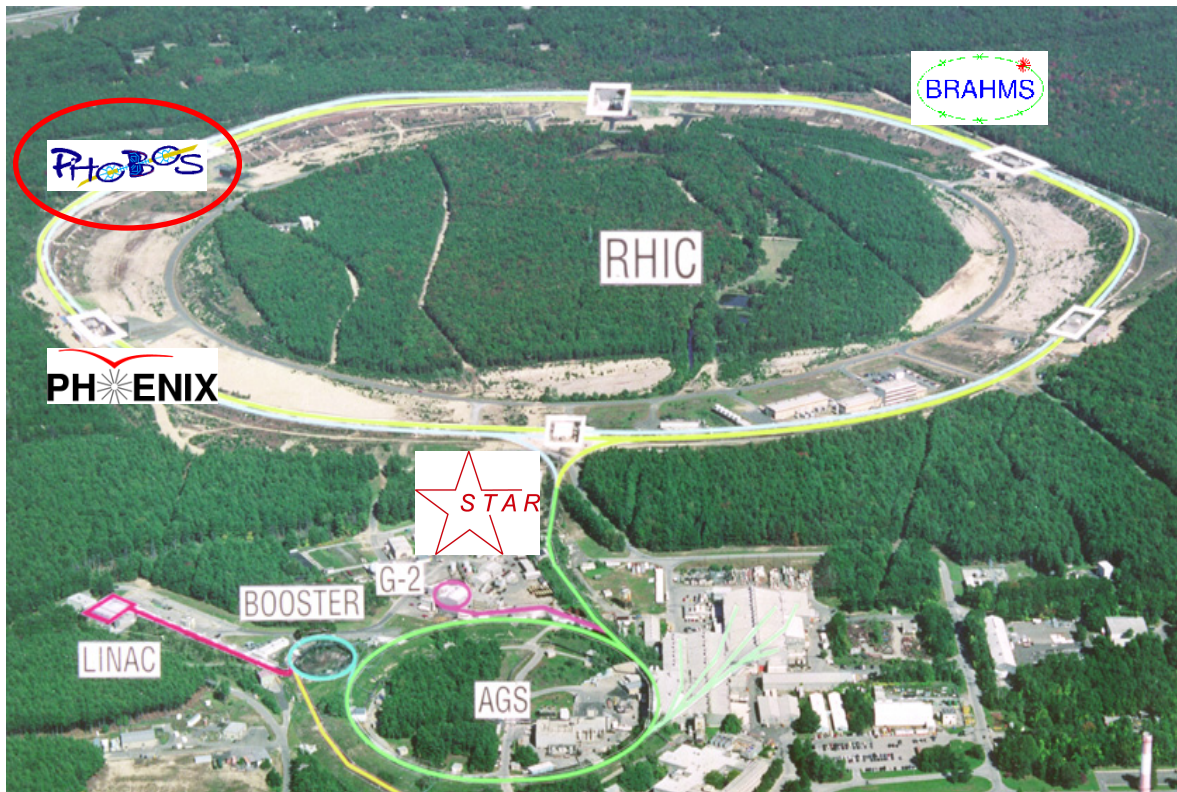
First Physics in '00

Versatile machine

- **Au+Au ('00-'02)**
    - 19.6 GeV
    - 56 GeV
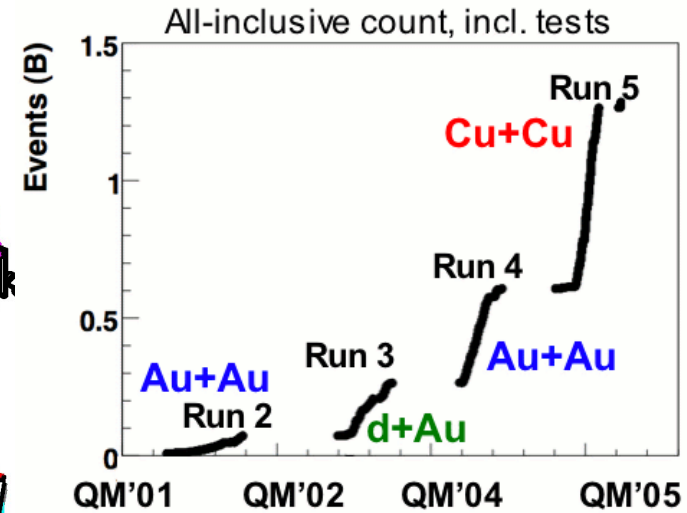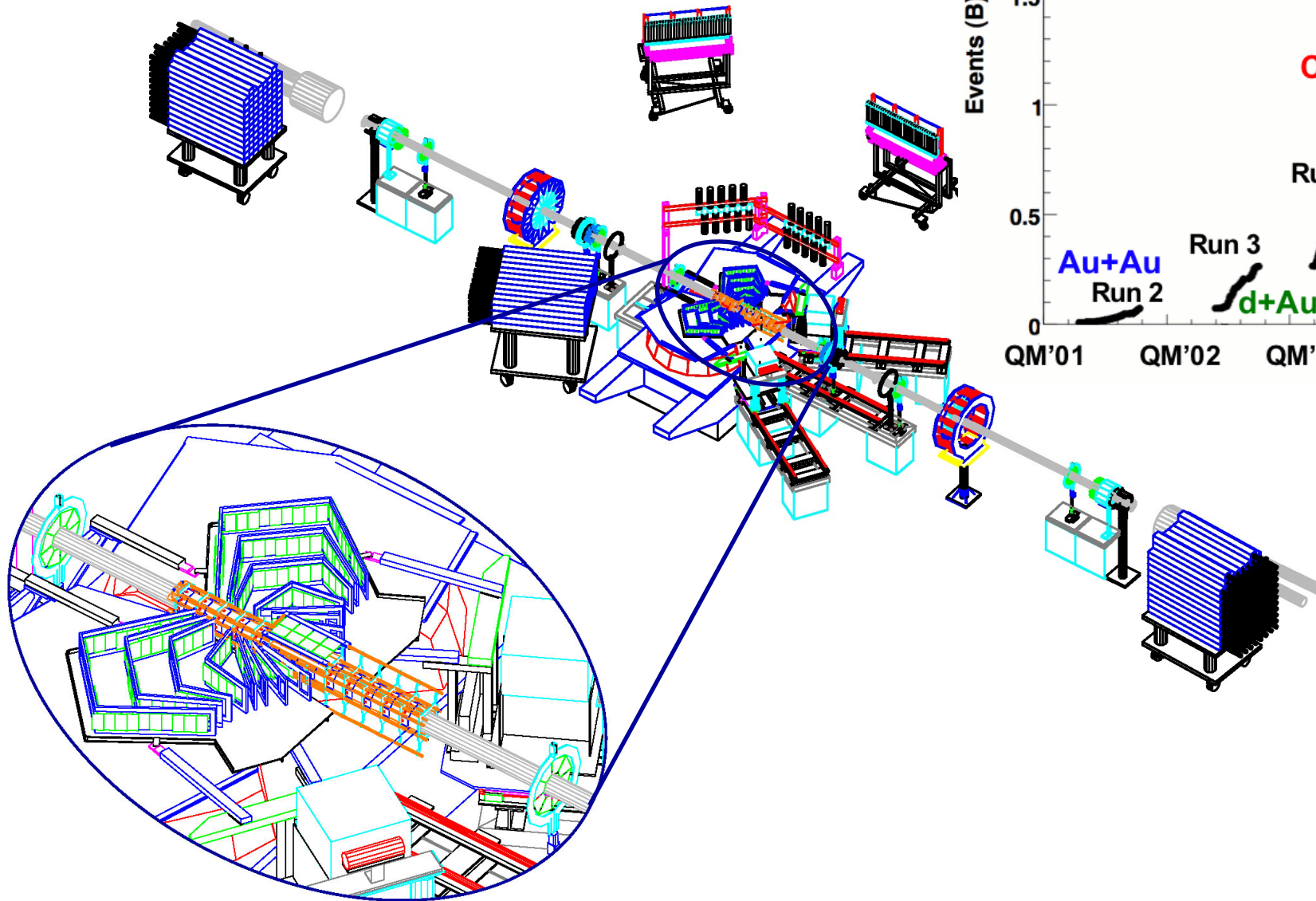   - 130 GeV
    - 200 GeV
- **p+p ('02,'03)**
    - 200 GeV polarized
- **d+Au ('03)**
    - 200 GeV

- **4 Experiments**
    - **2 big and 2 small**
- **Complementary capabilities**

PHOBOS Experiment

HI Collisions & Quark-Gluon Plasma

Initial Stage

Hadronization

Hot (thermalized?) Stage

Detector
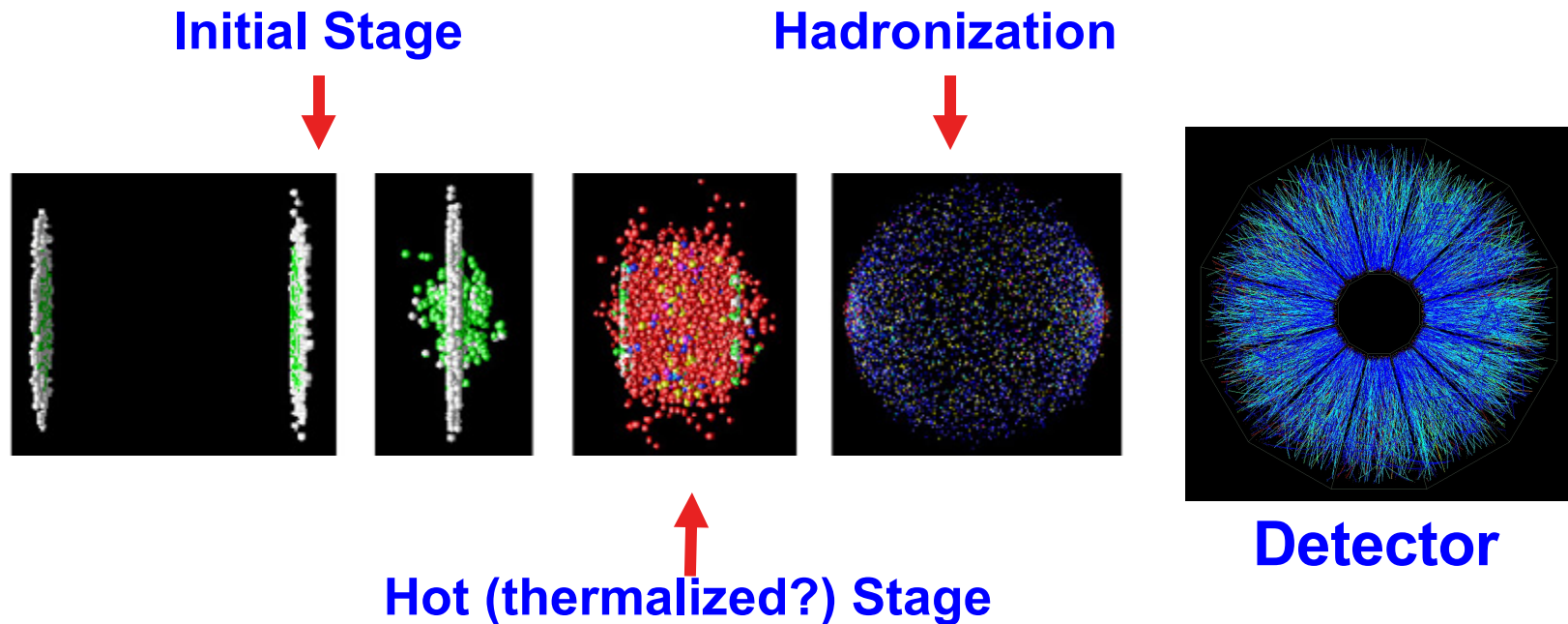
Au+Au collision @RHIC: 30 TeV deposited in collision zone

Sufficient energy density to create Quark-Gluon Plasma

Find rare fluctuations in 500M event data sets: PROOF

Data Analysis at RCF

- RHIC Computing Facility
  - Large centrally managed facility
  - Sub cluster for each experiment
- Components of the Facility
- Batch System
- Data Storage and Data Handling

Components of the Facility

- ## 25 Interactive Nodes
- ## 425 Compute nodes w/ distributed disk
  - ### 100 TB disk space
  - ### Mix of 100Mb and 1Gbit Ethernet
- ## HPSS tape robot / Mass Storage System
- ## Centralized disk space
  - ### NFS (0.9 TB) – home directories, software
  - ### Panasas (3.8 TB) – data, proof work directories

Batch Systems

- ## LFS batch system
  - ### Phased out in favor of Condor because of cost
- ## Condor
  - ### Kerberos authentication
  - ### Computing on Demand (COD) enabled for use with PROOF
  - ### Complex configuration – 4 priority levels (queues) – 4 x larger # VMs

Data Storage and Handling

- **NFS servers**
  - Home directories, Software, some Data
- **Panasas**
  - Replacing NFS for data
  - Used for PROOF directories
- **Distributed Disk and rootd**
  - Highly scalable
  - Cost effective
  - Needs management software: CatWeb

CatWeb Catalog and Data Manager

- PHOBOS File Catalog
  - All reconstructed and DST data (in HPSS mass storage system)
- Data management
  - Web based user interface
  - Database back-end and daemons for pools
- Storage pools – scatter data to avoid hot spots
- FileSets
  - Global and per user
  - Unit of data management and file access

PHOBOS Software Environment

- ## PhAT – Phobos Analysis Toolkit
  - ROOT based analysis environment
  - Collection of modules implementing reconstruction, calibration and analysis
  - AnT – Analysis Tree – DST supporting all PHOBOS Analysis efforts
  - PAR files available for all the modules
 - TGrid Interface to CatWeb

PROOF At RCF

- Direct (xinetd) vs. Condor COD
- Intelligent proofserv wrapper
  - Multiple versions
  - No root access
  - Debug support (e.g. run with valgrind)
- Proof server configuration
  - Per user config files - deprecated
  - Global predefined config files
  - Need a resource broker / scheduler

Analysis Examples

- These analysis examples were prepared using PROOF

- Presented at the Quark Matter 2005 conference in Budapest, Hungary

- For detailed information see:

  - Structure and Fine Structure of Hadron Production at RHIC, Gunther Roland, QM05 proceedings (to be published)

Rare high multiplicity event search
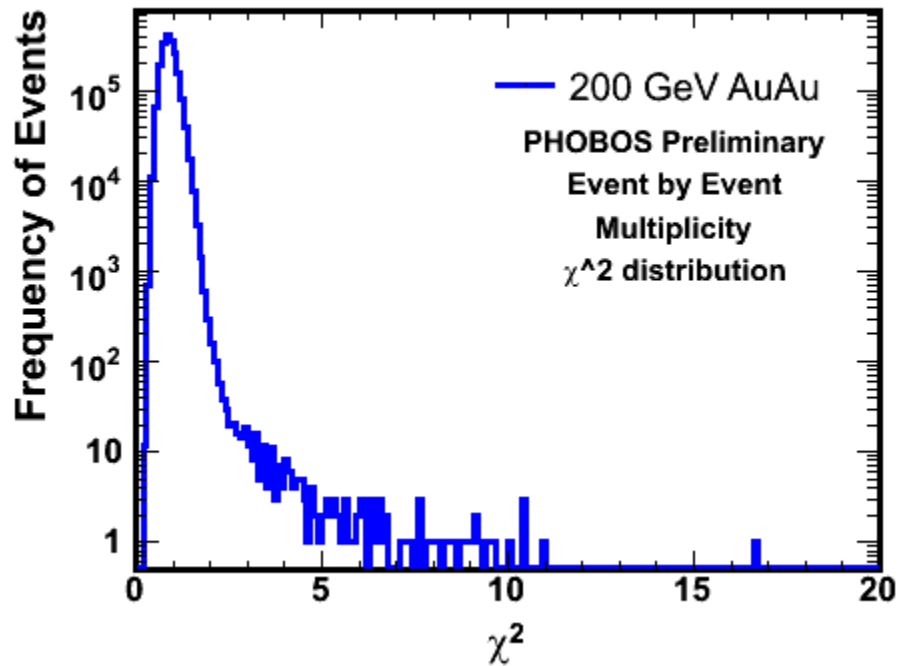
- Burak Alver
- Dataset: 11k files, 4.5 TB
- 150 slaves, ~1 hour

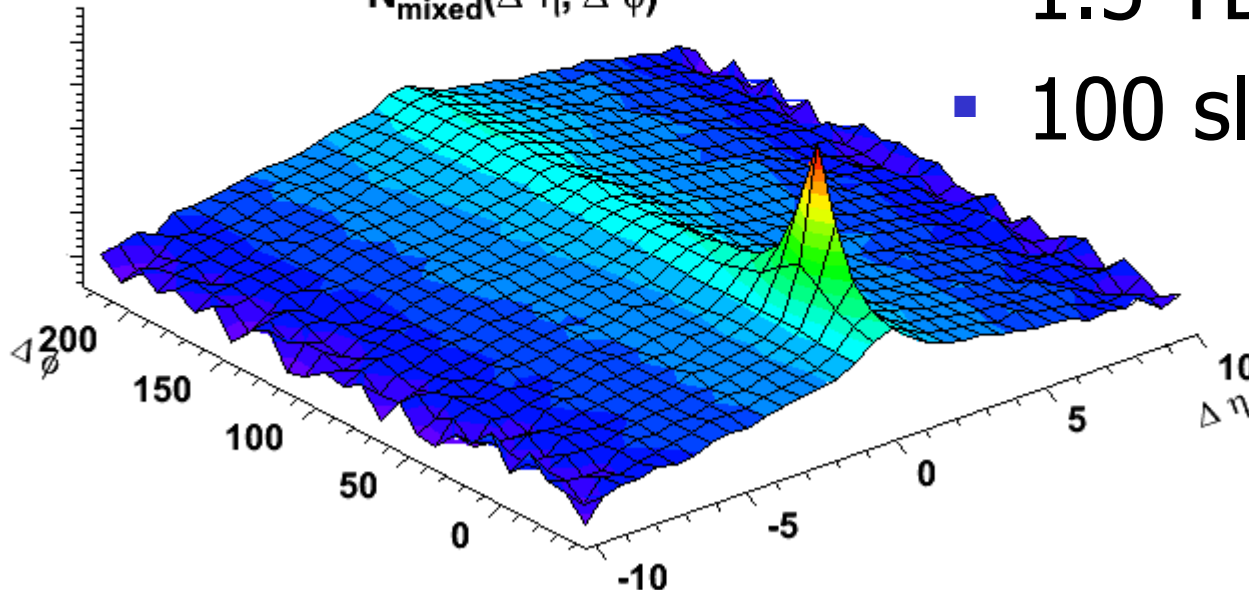# Cu and Au Transverse Momentum Spectra
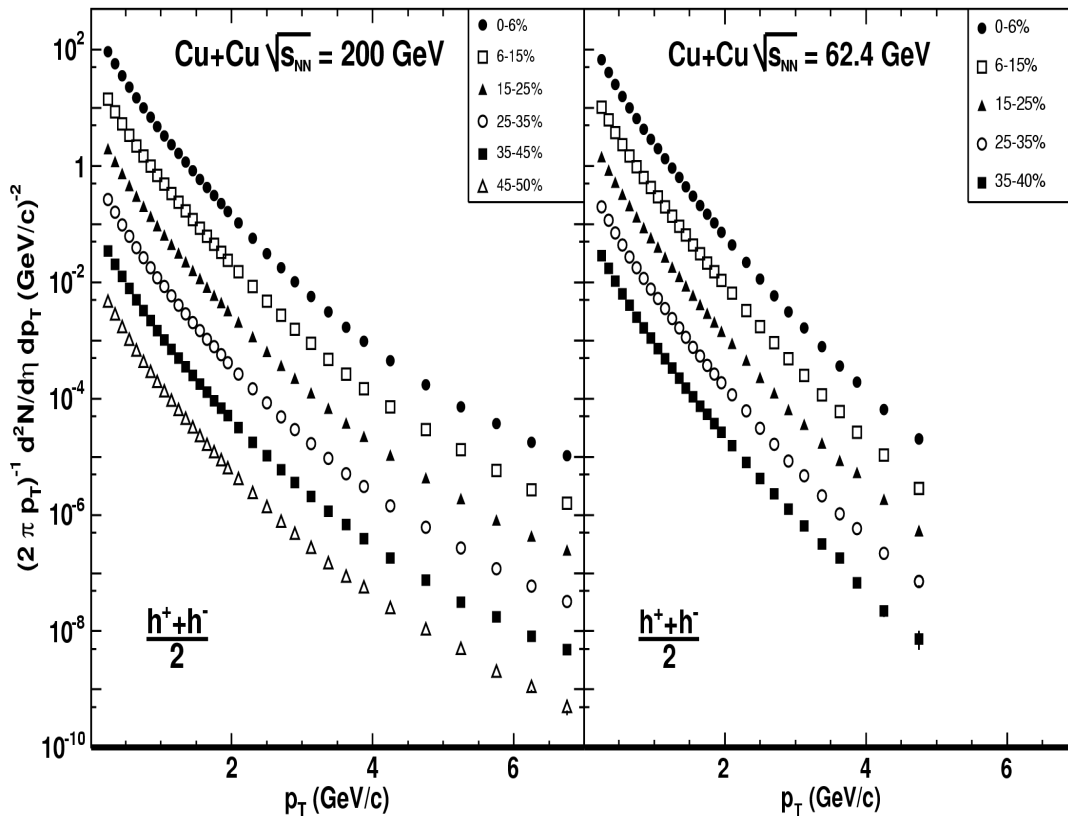
- Edward Wenger
- Dataset: 40k files, 13.5 TB
- 100 slaves, 45 min
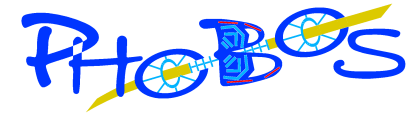
Two Particle Correlations @ 200GeV

Two particle correlation function of minbias dAu 200Gev

$$C(\Delta\eta, \Delta\phi) = \frac{N_{real}(\Delta\eta, \Delta\phi)}{N_{mixed}(\Delta\eta, \Delta\phi)}$$
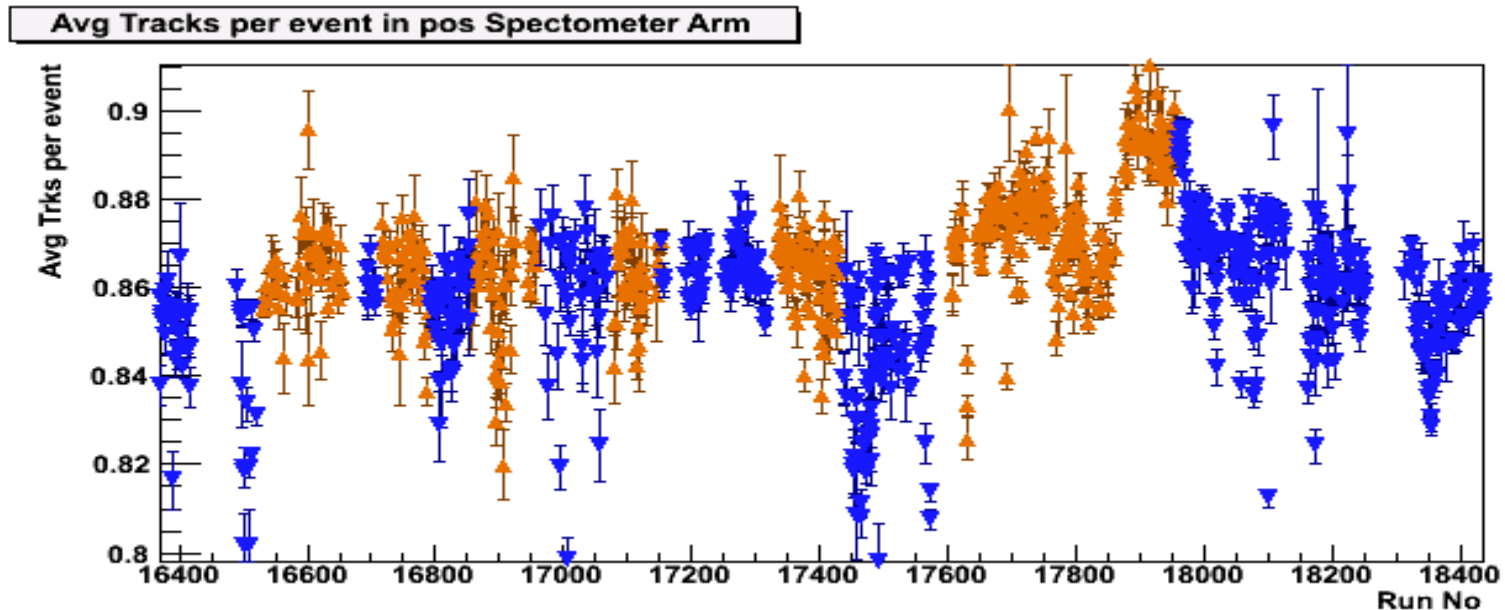

- Wei Li, Constantin Loizides
- Dataset: 4.5k files, 1.5 TB
- 100 slaves, 75 min

Identified Particle ratios using the Spectrometer

- Vasundhara Chetluru
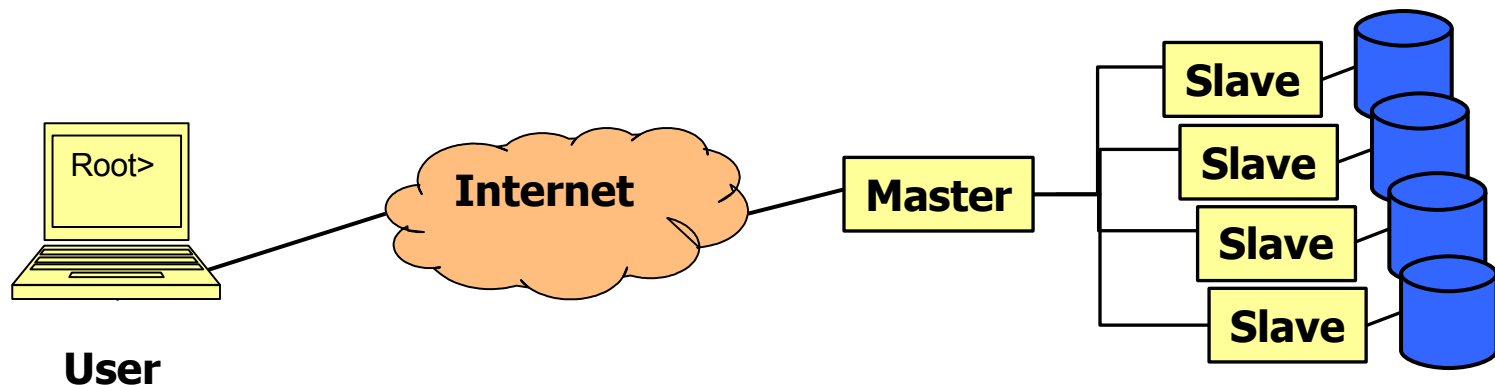- Dataset 40k files, 13.5 TB
- 100 slaves, 90 min

Avg Tracks per event in pos Spectometer Arm
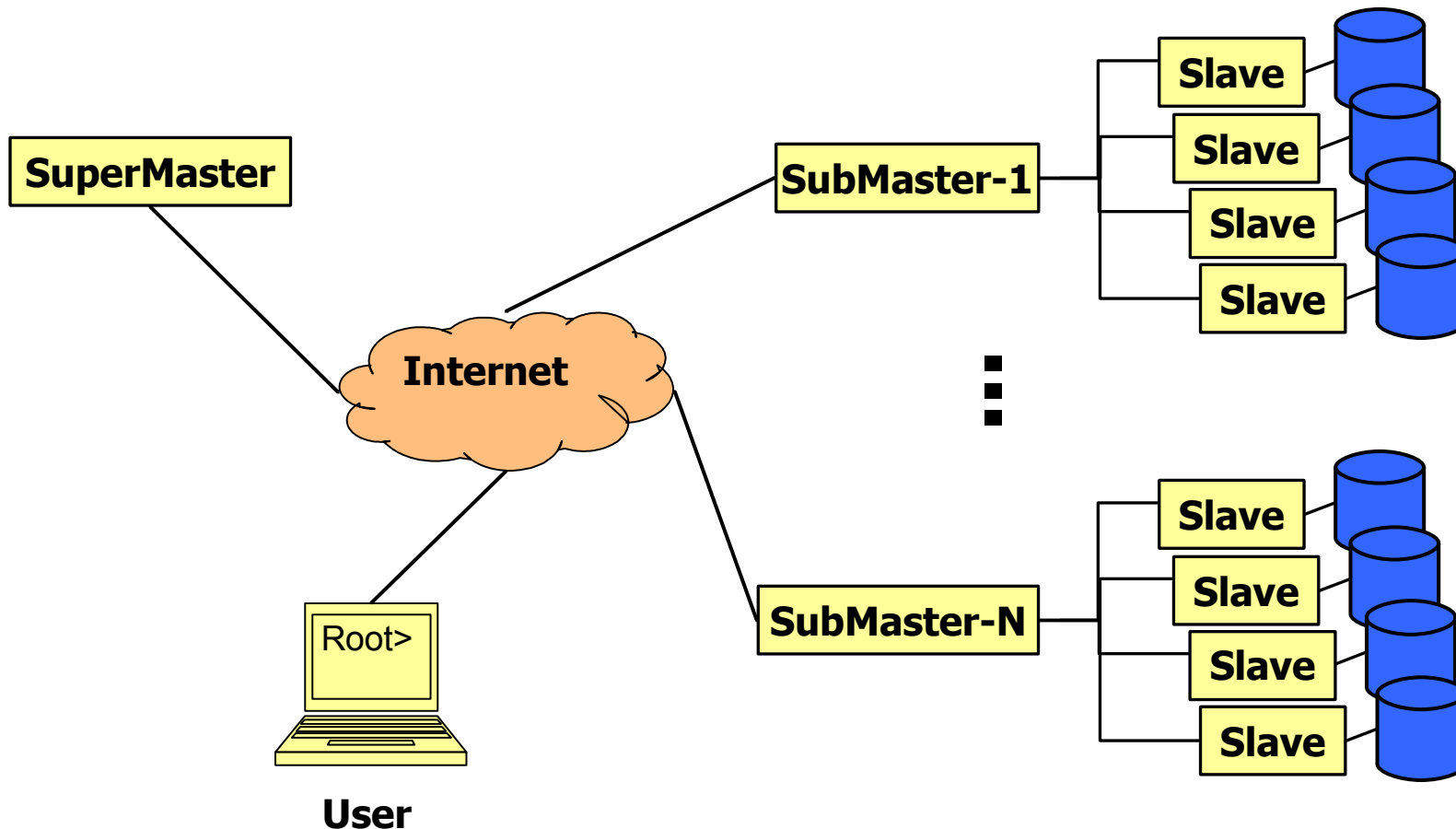
PROOF Developments at MIT

- ## Multi Level Master Configurations
  - ### Kris Gulbrandsen

- ## Progressive Packetizer
  - ### Zev Benjamin

Multi Level Master configurations

- Default 3-tier configuration
  - client – master: Low bandwitdth / high latency
  - master – slaves: High bandwidth / low latency

Multi Level Master configurations

Multi Level Master Configurations

- **Geographically distributed setup (Grid)**
  - Distributed data
  - Distributed computing power (replicated data)
- **Scalability in large cluster**
  - Parallel merging
  - Local network topology
- **Static super packetizer based on mass storage domain tag**

Progressive Packetizer

- The packetizer divides the work over the slaves

- Previous packetizers opened all files to determine number of entries
  - Not optimal for performance
  - Bad interaction with MSS

- Progressive Packetizer processes directly
  - Continuously Estimate total number of events
  - Can order files based on availability

Future Work

- Integrate PROOF into CMS Tier-2 Facilities
  - Starting with the Tier-2 at MIT
- Extend multi-level master functionality
- Improve integration with schedulers and batch systems