DØ Computing at UTA

Jae Yu

DoE Site Visit, Nov. 13-14, 2001

Dr. Michael Procario

- Introduction
- UTA DØ MC Farm
- Local Computing Resources
- DØ Remote Analysis Coordination Effort (RACE)
- UTA DØ Grid Effort
- Plans for the near future
- Conclusions

Introduction

- UTA was the largest offsite farm in Run I
- Playing a leading roll in RunII MC production
- Switched over to MC production network event delivery
- Begin to set up analysis servers for remote-analysis
- Recent appointment for DØ Remote Analysis Coordination
  - Setting up seamless software distribution
  - Efficient resource management
  - Transparent data delivery
- Strong participation to DØ Grid effort ➔ Working together with ATLAS

MC Farm Responsibility at DØ

- Has been a main offsite facility for MC generation and production for DØ in both Run I and II
- Increased capacity to produce about 30k events/day in two independent Linux farm clusters
  - Collaboration with UTA Computing Science Engineering Department and Academic Computing Services
  - Will double the capacity when ACS farm integrated in
- Has a very efficient home grown MC farm job control software and monitoring systems

UTA DØ Monte Carlo Farm

- Largest offsite computing facility in Run I (European institutions has taken the title away in Run II)
- Current UTA HEP MC Farm system:
  - 5 dual 566MHz + 18 dual 866MHz Linux PCs
  - 0.5GB RAM per machine
  - 0.61 TB total disk storage

- Use Home grown job and resource management software ➔ Preserve high efficiency
- To date over 4.3 million events generated in 8 Mo. for Run II
- Second farm of five dual 866MHz Linux cpu in CSE recently added
  - Promotes inter-departmental collaboration
  - UTA CSE interested in GRID development
- Working with the ACS to add 16 dual 866MHz Lunux farm
- Switched from Tape to Network event delivery system

Local Analysis Computing

- New analysis server added (Mark Sosebee)
  - 600GB disk space
  - 100MBit/sec network card
  - Dual 1GHz cpu
  - 1GByte of ram
  - Linux OS installed
  - Ready for DØ software deployment
- Supplemented with separate mail and web servers
- Other Desk top machines for desk top analysis
- HEP MC Farm machines + ATLAS Test machines
- Network bandwidth improvements

Remote Analysis Coordination

- Computing hardware is rather inexpensive
  - CPU and storage media are inexpensive
  - Small institutions can afford to own reasonable size computing facilities
- DØ collaboration is larger and more international
  - Most the collaborating institutions are remote
  - Code development can occur at remote stations
    - Exploit available man-hours for much needed software development
    - Give ownership to collaborators from remote institutes
  - Optimal and efficient access to data is of utmost importance to expedite analyses
  - Minimize travel around the globe for data access
  - Exploit existing but scattered computing resources
  - Sociological issue of HEP people at the home institutions
- It is quite certain that sharing a 15-20fb^{-1} worth of raw and reconstructed data (~5-7PByte?) efficiently will be a big issue
- **Primary goal is empowering individual desktop users**

What do we need?

- Remote DØ software development environment
  - Allow remote participation for code development which might soon be a bottleneck in expediting physics results
  - Allow remote analysis for histogram production
  - Allow remote reconstruction or production environment
- Optimized resource management tools
  - Allow to maximally utilize local resources
  - Allow to tap into available computing resources in other locations
  - Allow participation of remote resources for global reconstruction or production
- Efficient and transparent data delivery and sharing
  - Allow location independent access to data
  - Allow quicker access to sufficient statistics for initial analyses
  - Allow data sharing throughout the entire network of collaboration
  - Minimize central data storage dependence
  - Alleviate load for central data storage and servers

Remote System Software Usage

- Three usage categories of remote DØ code development system
  - Minimal executables and necessary configuration files: MC farms, releases done through very light mini-tar
  - Accepting binary only done via tar files, heavier than mini-tar. For local executable running ➔ Run-time environment effort will help this case
  - Full blown code releases; binaries via tar files and sources via ups/upd ➔ UTA wants this mode
- Most the institutions want the entire software

Initial Infrastructure Set Up

- Have established a listserver distribution list, d0-remote-analysis
  - Consists of 50 subscribers who are institutional contacts so far
  - Provide more restricted forum to share experiences and information
- One institutional Contact requested per each institution responsible for (People with more permanent position for continuity)
  - Setting up remote analysis stations
  - Keeping up releases
  - Answer institutional user inquiries
  - Channel through the inquiries that can't be answered locally to larger and more experienced crowd, d0-remote-analysis, to share experiences
  - Provide institutional supports for analysis or code development efforts
  - Participate in development of remote-analysis tools
  - Participate in testing and evaluating the tools
  - Establish necessary infrastructure for institutions (network, disk space, etc)

- Sent out Survey to gather information
  - How many were established
  - What is the depth of software download and installation (binaries, source)
  - Preference in Pull vs Push release system?
  - Biggest difficulties in establishing remote-sites?
  - What can be improved?
  - What can institutions offer?
  - Tasks for efficient remote analysis establishments?
  - Attend remote-analysis workshops?
  - Topics to be discussed in the workshops?
- 53/76 institutions (30/34US, 19/30 European, 5/7SA, 1/5A) assigned institutional contacts (45)
- Will organize a series of workshops, first one in Feb. at FNAL

Workshop Goals

- Primary goals:
  - Identify available resources within the collaboration
  - Sharing experiences
  - Understand the current status ➔ Exchange ideas
  - Identify missing or anticipated to be missing pieces for exploiting remote resources
  - Identify items that need to be prepared for expediting data analysis at remote sites
  - Identify necessary tools to empower desktop users
  - Set common goals, task lists and schedules
    - Distribute tasks
- Establishing clear road map for the future

Biggest Difficulties

- Having hard time setting up initially
  - Lack of updated documentation
  - Rather complicated set up procedure
  - Lack of experience➔ No forum to share experiences
- OS version differences (RH6.2 vs 7.1), let alone OS
- Most the established sites have easier time updating releases
- Network problems affecting successful completion of large size releases (4GB) takes a couple of hours (SA)
- No specific responsible persons to ask questions
- Availability of all necessary software via ups/upd
- Time difference between continents affecting efficiencies

Plans for DØ Remote Analysis

- Make the initial set up simpler and easier
  - Provide updated document and maintain them in regular bases
  - Provide initial set up script that needs only a push of a button
- Establish automatic “release-ready” notification system using the distribution d0-remote-analysis
  - Need to agree on a definition of “release-ready”
  - My suggestion is let the release managers define this
  - But provide sufficient information on the release
- Improve and provide tools for simplified (preferably one button web operation) “pull” based download and installation

- Limit other OS supports (IRIX, LINUX + minimal OSF)
  - Remote stations are responsible for local build for unsupported OS
- Start bi-weekly (every on-week) remote-analysis meetings to share information and experiences
  - Video conferencing >30 institutions at the same time is the first huddle
  - Pick a few sites and get them through the set up process ➔ Refine the documents, tools, and procedures
  - Goal is to get all other institutions that want remote-analysis release system ready within 6 months or by next summer
- In the mean time, will prepare for workshop in Feb.

- Offer for help
  - Some institutions volunteer for testing
    - Univ. of Wuppertal
    - Mainz
    - UTA
    - KSU
    - BU
  - Some institutions offer manpower in various forms
    - UTA
    - Langston
    - UC Fresno
    - LA Tech
  - Some offer specific tool developments
    - Build error information (Washington)
    - Pick-n-choose download and installation (UC Riverside)
    - Run time environment (Imperial College)
    - SAM & Condor batch submission (Imperial College)
    - PACMAN development (Michigan)

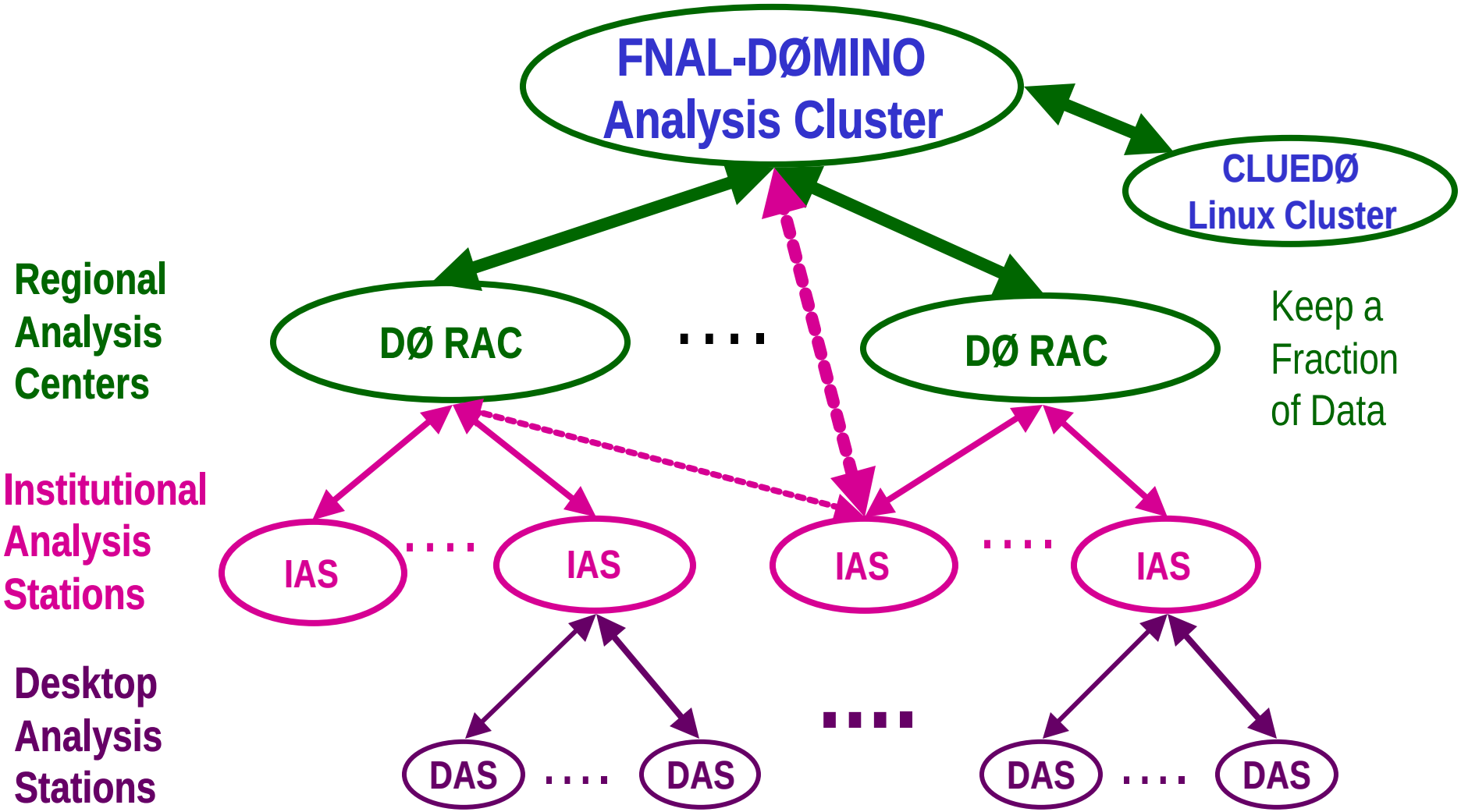
UTA DØ Grid Effort

- DØ Data delivery system (SAM) intensely used to transfer MC events to FNAL storage via network
- Installed Condor on our Two MC farms (HEP and CSE)
- Condor running on the farms
  - Submitted simple test jobs
  - Will write new control software, utilizing Condor
- Observe interactions between Condor and SAM
- Installed another Grid toolkit, Globus
- Will participate in development of job and resource management tools
  - Develop tools for DØ's use
  - Apply them to ATALS
- Use DØ as Grid test bed ➔ Must work with PPDG and FNAL-DØ

The DØ Grid Team

- Tomasz Wlodek: MC Farm Management, Liaison to DØ MC group, and DØ Lepto-quark analysis (DoE)
- Drew Meyer: MC Farm and DØ Grid design and Software Development – Lead software designer (Jae Yu's Start up)
- Karthik Gopalratnam: Physics Undergraduate, DØ Remote analysis web site, Farm daily maintenance, and initial set up of the root analysis environment (Jae Yu's Start up)
- Amruth Dattatreya/Venka Kaushik: CSE Masters Students, Grid tool installation, monitoring, benchmarking and network (Jae Yu's start up)
- Mark Sosebee: Computing system manager (0.5 UTA+0.5DoE)
- Having weekly meetings to keep up the progress
  - Weekly action items identified and checked at the next meeting

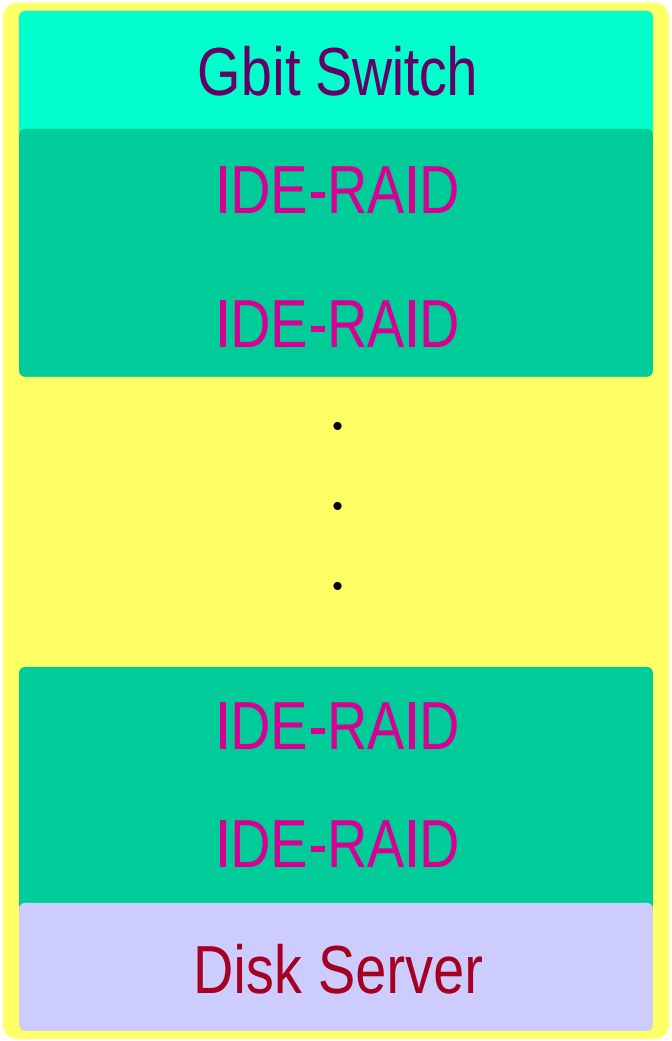
Schematics of a DØ Grid



Plans for the Near Future

- Develop Grid software for job and resource control for DØ's use
  - Testing at DØ farm (HEP, CSE, and ACS)
  - Improve and generalize the software
  - Apply for ATLAS Grid computing
- Improve Network Infrastructure
  - Improvement Swift Center bandwidth from 10Mbits/sec to 100Mbits/sec a few weeks ago by the ACS Network group of UTA
  - Plan to increase the bandwidths to 1Gbit/sec early next year
  - Internet II ready for outside the campus communication
- Establish hardware infrastructure for large scale storage
  - Need ~50TBytes to store 10% of DØ Run IIa data
  - Scalable Storage model based on Hard disk RAID array
  - Start out with a simple prototype
  - Develop and test monitoring and control software

Prototype UTA DØ Regional Center



- IDE Hard drives are ~$2.5/GByte
- Each set of IDE RAID array gives ~1.6TByte – hot swappable
- Can be configured to have up to 10-16TB in a rack
- Modest server can manage the entire system
- Gbit network switch provide high throughput transfer to outside world
- Flexible and scalable system
- Need an efficient monitoring and error recovery system
- Communication to resource management

Conclusions

- UTA has been the largest offsite farm in Run I and still and will be playing crucial roll in DØ MC production
  - The leadership is being taken away by the European institutions
- The recent appointment places UTA in a good position to play a leading roll in DØ remote analysis and Grid development effort
  - European dominated effort
  - UTA must be a flagship institution in the US effort
  - Must work together with DØ computing and PPDG groups
- UTA Plans to lead DØ Offsite Analysis System establishment effort ➔ will play a leading role in the effort for DØ GRID development
- Plans to establish UTA as a DØ Regional Center
  - Improved network bandwidths
  - Establishment of the scalable storage system
  - Resource management software
  - Storage management software
- Prepare UTA to become an ATLAS Tier 2 site