Remote Production and Regional Analysis Centres

Iain Bertram

5 June 2002

Lancaster University

Overview

- Definitions
  - What is Remote Analysis
- MC Production
  - Status
  - Future
- Remote Analysis
- Conclusions

What is Remote?

- We have defined it as follows:
  Local: d0mino, FNAL Farms, Online
  Remote: clued0, club, non-FNAL.
- 76 Institutions - 670 people on Masthead
  Most of these work remotely

Goals

- Major Goal: Produce High Quality Physics Results!
- Run II: Major Problem, reducing systematic uncertainties --- Need Excellent simulation of detector.
- Regional Production Facilities will provide a significant fraction of the processing power (with storage) for the DØ collaboration.
- Production Tasks are:
  - Monte Carlo (MC) Production
  - Secondary reprocessing of the data.
  - CPU intensive user Analysis Jobs.
  - Data storage
- MC Production
  - Run IIB Data Rate ~ 50 Hz
  - Aim MC rate of 25 Hz

Monte Carlo

- The generation of Monte Carlo events in DØ involves multiple stages and many executables:
  - Organized in independent packages using the standard DØ framework
  - Generation of a physical process. Many Programs to be integrated with DØ: Pythia, Herwig, Isajet, Comhep, Pompyt, Vecbos, etc.
  - Detector Simulation: Trace the particles through the DØ detector, determine energy deposition in active areas using Geant.
  - Detector Digitization: Conversion of energy deposits into RAW data format as read out by the detector.
    Simulation of noise, and additional interactions in a given proton-antiproton interaction
- Monte Carlo currently run using a macro driven interface
  - Mc_runjob – allows description of full Monte Carlo job in terms of keyword – value pairs.

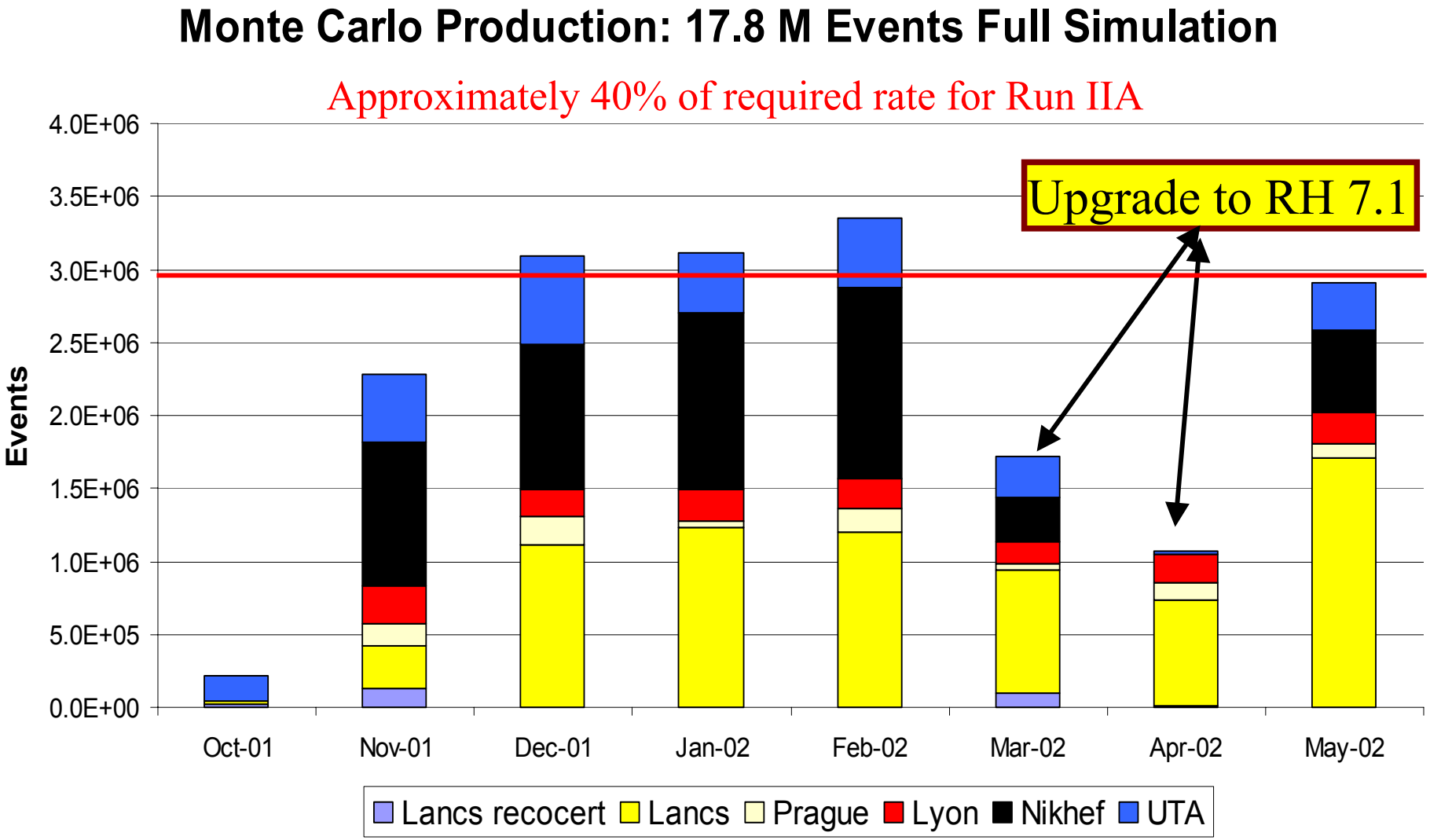
Monte Carlo Production - Current

➔All Production Monte Carlo Running Off-Site

- Run the complete analysis chain:
- Fully functional SAM station.
- MC_Runjob allows fully parameterised running of MC using macros
- Redhat Linux 6.2 or 7.1
- Tarballs created from Official Releases

- In process of implementing integrated request system using SAM
  - Allows complex search on data files
  - First step in automated MC processing chain

Monte Carlo Production Rates



MC Requirements RunIIB

➔Requirements for 25 Hz Monte Carlo Rate

- 25% Event full Geant - Rest Parameterised MC
- Time Per Geant Event on 500 MHz PIII 216 seconds

Process/Time per Event	*Generation*	*Detector Simulation*	*Digitization*	*Reconstruction*	*Analyze*	*Total*
0.5 Events Overlaid, Plate Level Geant (sec/event)						
WW inclusive	0.8	280	20	19	4.5	325
Technirho	0.8	300	20	21	5	345

- **Digitize, Reconstruct and Analyze each event 6 times**
- **Time per Geant event: 550 seconds**
- **Meet 25 Hz Goal requires 5k 500 MHz CPU's**

MC Production Cost requirements

➔Same Cost Assumptions as Fermi Farms

- Dual CPU Node: $2,500
- One I/O Node per 100 node's
- No remote mass storage costs included
- Three Year Purchase Profile
- Time Per Event = 0.25*550 seconds + 0.75*2 seconds

FY05 Target Fraction:		20%		30%		50%		Total
Time Event	500 MHz CPUs	FY03, 3GHz Nodes	Cost	FY04, 4GHz Nodes	Cost	FY05, 6GHz Nodes	Cost	Cost
100	3571	60	172,500	67	190,000	74	210,000	201	572,500
140	5000	83	207,500	94	232,500	104	260,000	281	700,000
180	6429	107	267,500	121	300,000	134	332,500	362	900,000

Software Requirements

➔Operating Systems:

- Current RH 7.1
- Future: DØ Fermi Redhat Linux or a similar Redhat release
- DØ will support official builds for the Fermi Redhat Releases
- Remote production sites could be shared facilities,
  I.e. LHC experiments, etc.
- Will not be able to upgrade operating systems at remote facilities purely to meet DØ software requirements.

Remote Analysis

- LHC and Grid Developments lead to many new large Computing Resources.
- Need to access data software, and database information remote from FNAL
- Integrate Software with new GRID Tools

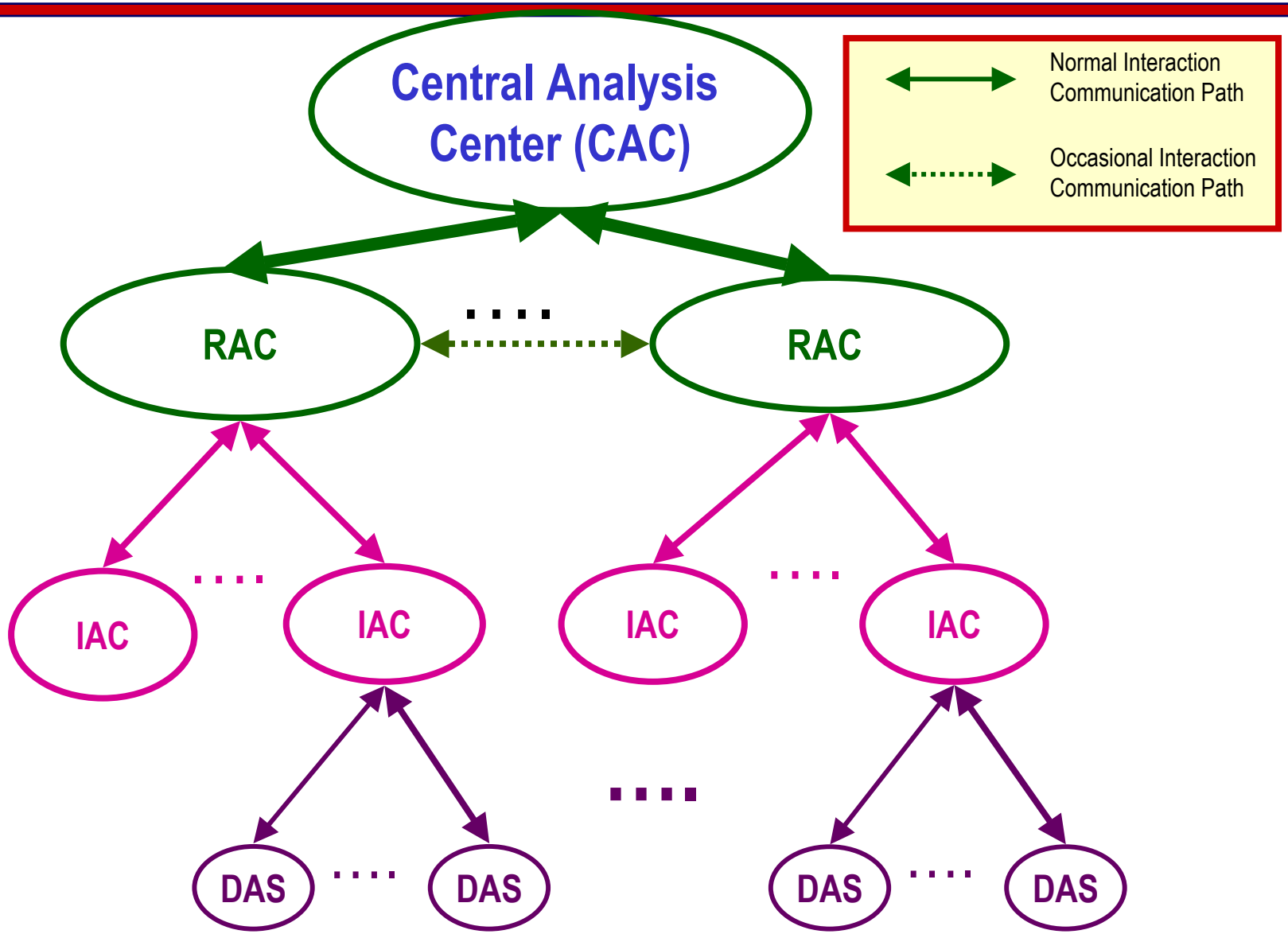
Propose to set up Regional Analysis Centres (RAC)

A series of locations that offer centralized access points for data analysis for remote users.

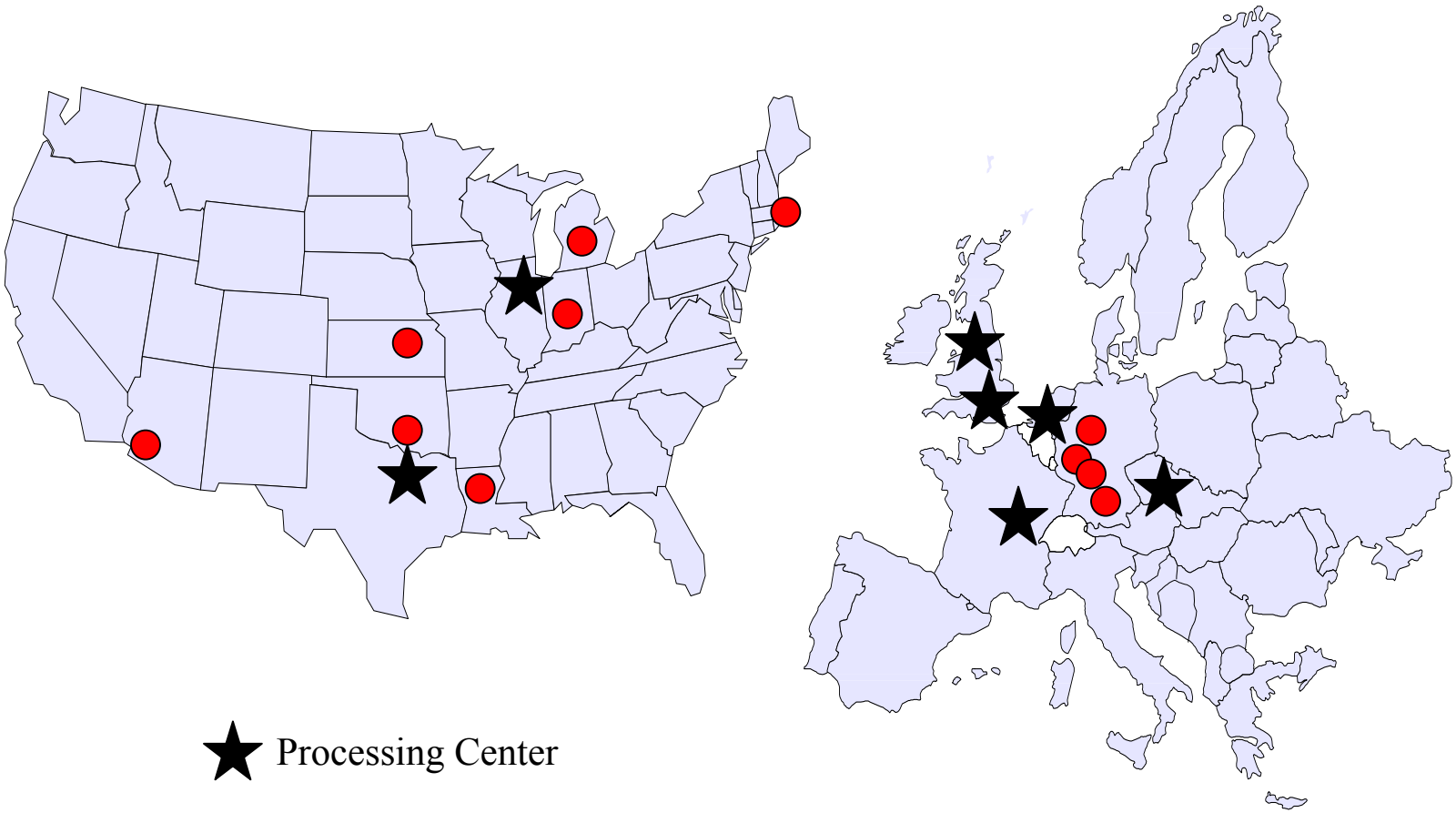
Regional Analysis Centre Requirements

- RAC are expected to provide
  - Significant CPU Power
  - Access to DØ Thumbnails and some DSTs
  - User Support for the Facility
- An institute with large concentrated and available computing resources
  - 100s of CPUs
  - 10s of TBs of disk cache
  - 100s Mbytes of network bandwidth
  - Possibly equipped with Mass Storage
- Costs Based on Lancaster Farm
  - Year 00: 200 CPU, 2TB disks, 30 TB expandable Mass Storage Cost 600 k$.
- Each RAC ~ 1M$.....
  - Required disk storage is about 50TB for Run IIa alone.

Remote Analysis Centres



Current remote SAM sites



Remote Analysis Requirements

- Off Site Data Reconstruction
  - Access the appropriate DØ calibration, and luminosity databases.
  - Transfer of relevant tier of data
  - Well synchronized reconstruction executable
- Generic User Analysis Jobs
  - The user will be required to meet several mandatory conditions:
    - The software must be built using DØ RunTime Environment.
    - Any input/output data must be stored in SAM.
    - That the jobs can be fully described and submitted to a SAM queue.
- Grid Enhancements
  - It is assumed that Remote Production Jobs will make full use of the current DØGrid project.
  - Use of the Grid is not specific to remote analysis, but rather is a coherent part of the overall computing plan.

Open Issues

- Remote Mass Storage Costs.
- Network Requirements
  - Depend on development of grid
    (Minimum requirements for one Regional Analysis Center connected to FNAL: ~14MB/sec for Run IIa.)
  - Data to processors or program to data

- Available Resources:
  Incomplete List Not Easy to Plan
  - Expect minimum 600 processors.

Pilot DØRAC Program

- Regional Analysis Centers (RAC) Pilot sites:
  - We wish to set up a pilot RAC over the next six months
- What should we accomplish?
  - Transfer TMB files as they get produced
  - Network monitoring tools
- Allow Institutes users to access the Thumbnails from a RAC.
  - Monitor Usage of System
  - Optimize operating conditions

Conclusions

Off Site Computing is an essential part of meeting DØ's analysis goals for Run II.

Monte Carlo Production

Secondary Reprocessing of Data

User Analysis Jobs

Monte Carlo Production Running Off Site

First Prototype Regional Analysis Centres starting up.

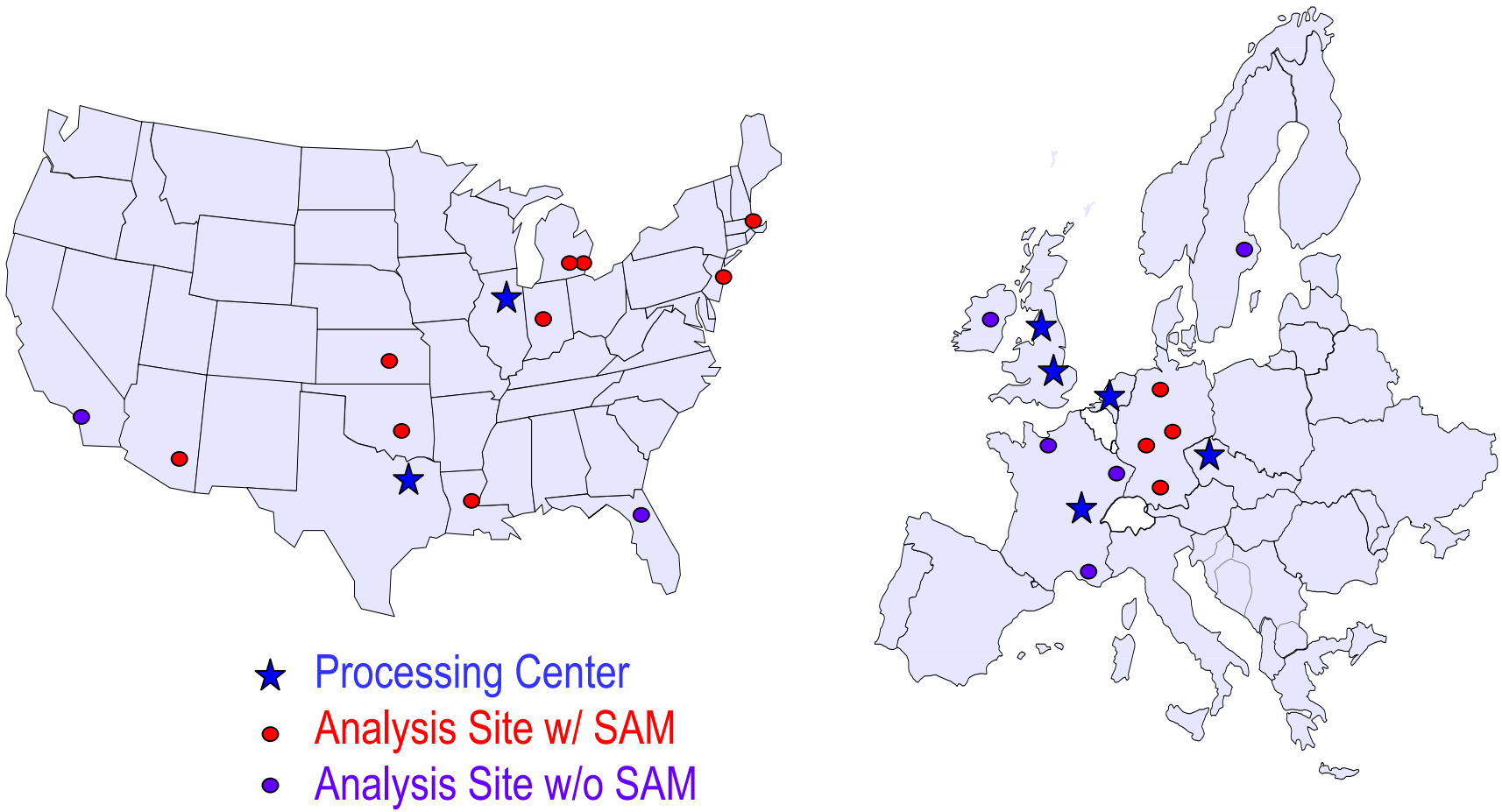
Off Site Computing is Going Strong!

Backup Slides

CLUB

- Analysis Backend to Cluedo
  - For analyzing 1TB data samples
  - Similar to MC farm's
  - Essentially a batch engine.
  - Fermilab to provide seed and infrastructure.
  - This cluster is a RAC for IAC's around Fermilab and for Fermilab physicist. Should be included in the pilot RAC program.
- Assume 200 CPU system (100 nodes 250k$)
  - 15 TB data storage (10 servers) (50k$)
  - Looks like 300k$ per CLUB type installation

DØRACE Deployment Map (US and EU only)



Remote Analysis Requirements

- Off Site Data Reconstruction
  - Access the appropriate DØ calibration, and luminosity databases. It is assumed that these will be accessed via a server interface and not require a local copy of the database.
  - Transfer of relevant tier of data
  - Well synchronized reconstruction executable
- Generic User Analysis Jobs
  - The user will be required to meet several mandatory conditions:
  - The software must be built using Dørte to ensure that the package can run without the full DØ release.
  - Any input/output data must be stored in SAM. That the jobs can be fully described and submitted to a SAM queue.
- Grid Enhancements
  - It is assumed that Remote Production Jobs will make full use of the current DØGrid project. Use of the Grid is not specific to remote analysis, but rather is a coherent part of the overall computing plan.

DØ RAC Implementation Timescale

- Implement First RAC by Oct. 1, 2002
  - CLUBs cluster at FNAL and Karlsruhe, Germany
  - Cluster associated IAC's
  - Transfer TMB (10kB/evt) data set constantly from CAC to the RACs
- Workshop on RAC in Nov., 2002
- Implement the next set of RAC by Apr. 1, 2003
- Implement and test DØGridware as they become available
- The DØGrid should work by the end of Run IIa (2004), retaining the DØRAM architecture
- The next generation DØGrid, a truly gridfied network without