

DØ RAC Working Group Report

DØRACE Meeting

June 6, 2002

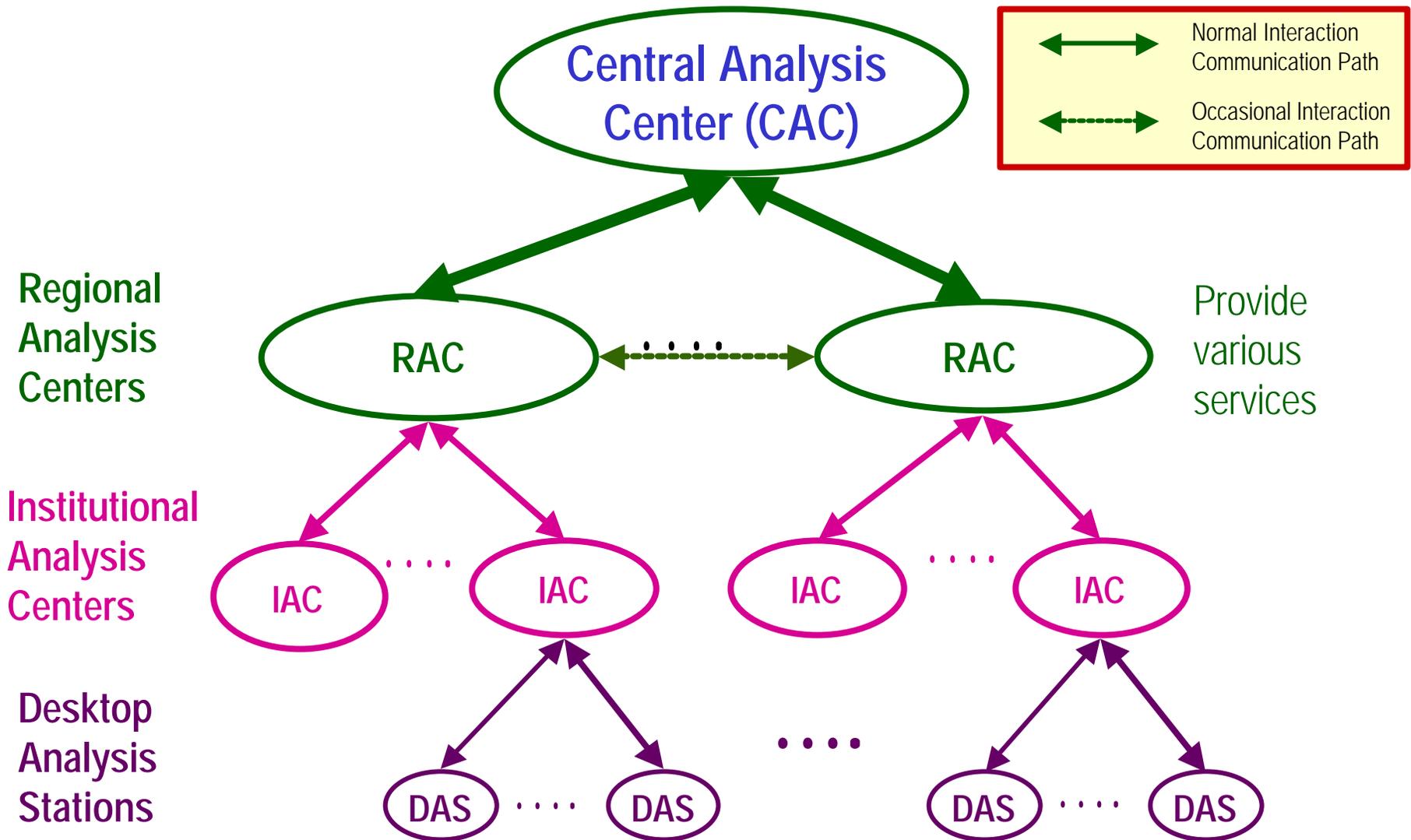
Jae Yu

- Progress
- Definition of an RAC
- Services provided by an RAC
- Requirements of RAC
- Pilot RAC program
- Open Issues

DØRAC Working Group Progress

- Working group has been formed at the Feb. workshop
 - The members: I. Bertram, R. Brock, F. Filthaut, L. Lueking, P. Mattig, M. Narain, P. Lebrun, B. Thooris, J. Yu, C. Zeitnitz
 - Had many weekly meetings and a face-to-face meeting at FNAL a few weeks ago to hash out some unclear issues
- A proposal (DØ Note #3984) has been worked on
 - Specify the services and requirements for RAC's
 - Doc. At :
<http://www-hep.uta.edu/~d0race/d0rac-wg/d0rac-spec-050802-8.pdf>
 - All comments are in → Target to release after the discussion at ADM tomorrow

Proposed DØRAM Architecture



June 6, 2002

DØRAC Report
DØRACE Meeting, Jae Yu

What is a DØRAC?

- An institute with large concentrated and available computing resources
 - Many 100s of CPUs
 - Many 10s of TBs of disk cache
 - Many 100Mbytes of network bandwidth
 - Possibly equipped with HPSS
- An institute willing to provide services to a few small institutes in the region
- An institute willing to provide increased infrastructure as the data from the experiment grows
- An institute willing to provide support personnel if necessary

What services do we want a DØRAC do?

1. Provide intermediary code distribution
2. Generate and reconstruct MC data set
3. Accept and execute analysis batch job requests
4. Store data and deliver them upon requests
5. Participate in re-reconstruction of data
6. Provide database access
7. Provide manpower support for the above activities

Code Distribution Service

- Current releases: 4GB total → will grow to >8GB?
- Why needed?:
 - Downloading 8GB once every week is not a big load on network bandwidth
 - Efficiency of release update rely on Network stability
 - Exploit remote human resources
- What is needed?
 - Release synchronization must be done at all RACs every time a new release become available
 - Potentially need large disk spaces to keep releases
 - UPS/UPD deployment at RACs
 - FNAL specific
 - Interaction with other systems?
 - Need administrative support for bookkeeping
- Current DØRACE procedure works well, even for individual users
→ Do not see the need for this service at this point

Generate and Reconstruct MC data

- Currently done 100% at remote sites
- What is needed?
 - A mechanism to automate request processing
 - A Grid that can
 - Accept job requests
 - Packages the job
 - Identify and locate the necessary resources
 - Assign the job to the located institution
 - Provide status to the users
 - Deliver or keep the results
 - Database for noise and Min-bias addition
- Perhaps the most undisputable task of a DØRAC

Batch Job Processing

- Currently rely on FNAL resources (DØmino, ClueDØ, CLUBS, etc)
- What is needed?
 - Sufficient computing infrastructure to process requests
 - Network
 - CPU
 - Cache storage to hold job results till the transfer
 - Access to relevant databases
 - A Grid that can:
 - Accept job requests
 - Packages the job
 - Identify and locate the necessary resources
 - Assign the job to the located institution
 - Provide status to the users
 - Deliver or keep the results
- This task definitely needs a DØRAC
 - Bring input to the user or bring the exe to the input?

Data Caching and Delivery Service

- Currently only at FNAL (CAC)
- Why needed?
 - To make data should be readily available to the users with minimal latency
- What is needed?
 - Need to know what data and how much we want to store
 - **100% TMB**
 - **10-20% DST?** → To make up 100% of DST on the net
 - Any RAW data at all?
 - What about MC? 50% of the actual data
 - Should be on disk to minimize data caching latency
 - How much disk space? (~50TB if 100% TMB and 10% DST for RunIIa)
 - Constant shipment of data to all RACs from the CAC
 - Constant bandwidth occupation (14MB/sec for Run IIa RAW)
 - Resources from CAC needed
 - A Grid that can
 - Locate the data (SAM can do this already...)
 - Tell the requester about the extent of the request
 - Decide whether to move the data or pull the job over

Data Reprocessing Services

- These include:
 - Re-reconstruction of the actual and MC data
 - From DST?
 - From RAW?
 - Re-streaming of data
 - Re-production of TMB data sets
 - Re-production of roottree
 - *ab initio* reconstruction
- Currently done only at CAC offline farm

Reprocessing Services cont'd

- What is needed?
 - Sufficiently large bandwidth to transfer necessary data or HPSS (?)
 - DSTs
 - RAC's will have 10% or so already permanently stored
 - RAW
 - Transfer should begin when need arises
 - RAC's reconstruct as data gets transferred
 - Large data storage
 - Constant data transfer from CAC to RACs as CAC reconstructs fresh data
 - Dedicated file server at CAC for data distribution to RACs
 - Constant bandwidth occupation
 - Sufficient buffer storage at CAC in case network goes down
 - Reliable and stable network
 - Access to relevant databases
 - Calibration
 - Luminosity
 - Geometry and Magnetic Field Map

Reprocessing Services cont'd

- RAC's have to transfer of new TMB to other sites
 - Since only 10% or so of DST's reside only the TMB equivalent to that portion can be regenerated
- Well synchronized reconstruction executable → run_time environment
- A grid that can
 - Identify resources on the net
 - Optimize resource allocation for most expeditious reproduction
 - Move data around if necessary
- A dedicated block of time for concentrated CPU usage if disaster strikes
- Questions
 - Do we keep copies of all data at the CAC?
 - Do we ship DSTs and TMBs back to CAC?

Database Access Service

- Currently done only at CAC
- What is needed?
 - Remote DB access software services
 - Some copy of DB at RACs
 - A substitute of Oracle DB at remote sites
 - A means of synchronizing DBs
- A possible solution is proxy server at the central location supplemented with a few replicated DB for backup

What services do we want a DØRAC do?

- Provide intermediary code distribution
- ✓ Generate and reconstruct MC data set
- ✓ Accept and execute analysis batch job requests
- ✓ Store data and deliver them upon requests
- ✓ Participate in re-reconstruction of data
- Provide database access
- ✓ Provide manpower support for the above activities

DØRAC 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

Pilot DØRAC Program

- RAC Pilot sites:
 - Karlsruhe, Germany → In the process of discussing with the management at the institute
 - CLUBS, Fermilab → Need to verify (Bring it up at the CPB?)
- What should we accomplish?
 - Transfer TMB files as they get produced
 - A File server (both hardware and software) at CAC for this job
 - Request driven or constant push?
 - Network monitoring tools
 - To observe network occupation and stability
 - » From CAC to RAC
 - » From RAC to IAC
 - Allow IAC users to access the TMB
 - Observe
 - Use of the data set
 - Accessing pattern
 - Performance of the accessing system
 - SAM system performance for creating DØRAC Replog

June 6, 2002

DØRACE Meeting, Jae Yu

16

- The user account assignment?
- Resource (CPU and Disk space) need?
- What are the needed Grid software functionality?
 - To interface with the users
 - To locate the input and necessary resources
 - To gauge the resources
 - To package the job requests

Some Open Issues

- What do we do with MC data?
 - Iain's suggestion is to keep DØStar format not DST
 - Additional storage for Min-bias event samples
 - What is the analysis scheme?
 - The only way is to re-do all the remaining process of MC chain
 - Require additional CPU resources
 - » DØSim
 - » Reconstruction
 - » Reco-analysis
 - Additional disk space to buffer intermediate files
 - Keep the DST and rootpule? Where?
- What are other questions we want answers from the pilot RAC program?
- How do we acquire sufficient funds for these resources?
- Which institutions are the candidates for RACs?
- Do we have full support from the collaboration?
- Other detailed issues covered in the proposal