

FNAL-SAM Data Grid

Rodney Walker

Beauty 2002 GRID

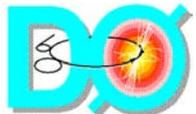
June 18, 2002



Overview



- The SAM Data Handling System
- SAM as a Data Grid
- Experience in the Dzero Experiment
- CDF Evaluation
- Future SAM-Grid architecture





What is SAM?



- SAM is Sequential data Access via Meta-data
- Project started in 1997 to handle D0's needs for Run II data system.
- Current SAM team includes:
 - ◆ Andrew Baranovski, Lauri Loebel-Carpenter, Lee Lueking*, Carmenita Moore, Igor Terekhov, Julie Trumbo, Sinisa Veseli, Matthew Vranicar, Stephen P. White. (*project leader)
 - ◆ Emeritus: Vicky White
- <http://d0db.fnal.gov/sam>





The SAM Architecture



- SAM is a vertically integrated system: its components span all the layers of standard Grid architecture
- SAM is a Grid-enabled system
- Employs a central Oracle Database to store all file metadata, replica catalog, processing information, group and user registration



Client Applications

Web

Command line

D0 Framework C++ codes

Python codes, Java codes

Collective Services

Request Formulator and Planner

Request Manager

Cache Manager

Job Manager

Storage Manager

“Dataset Editor”

“Project Master”

“Station Master”

“Station Master”

“File Storage Server”

SAM Resource Management

Batch Systems - LSF, FBS, PBS, Condor

Job Services

Data Mover

“Optimiser”

“Stager”

Significant Event Logger

Naming Service

Catalog Manager

Database Manager

CORBA

UDP

Catalog protocols

File transfer protocols - ftp, bbftp, rcv

GridFTP

Mass Storage systems protocols e.g. encp, hpss

Connectivity and Resource

SAM-specific user, group, node, station registration

GSI

Bbftp ‘cookie’

Authentication and Security

Fabric

Tape Storage Elements

Disk Storage Elements

Compute Elements

LANs and WANs

Code Repository

Resource and Services Catalog

Replica Catalog

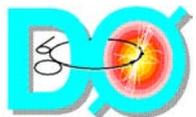
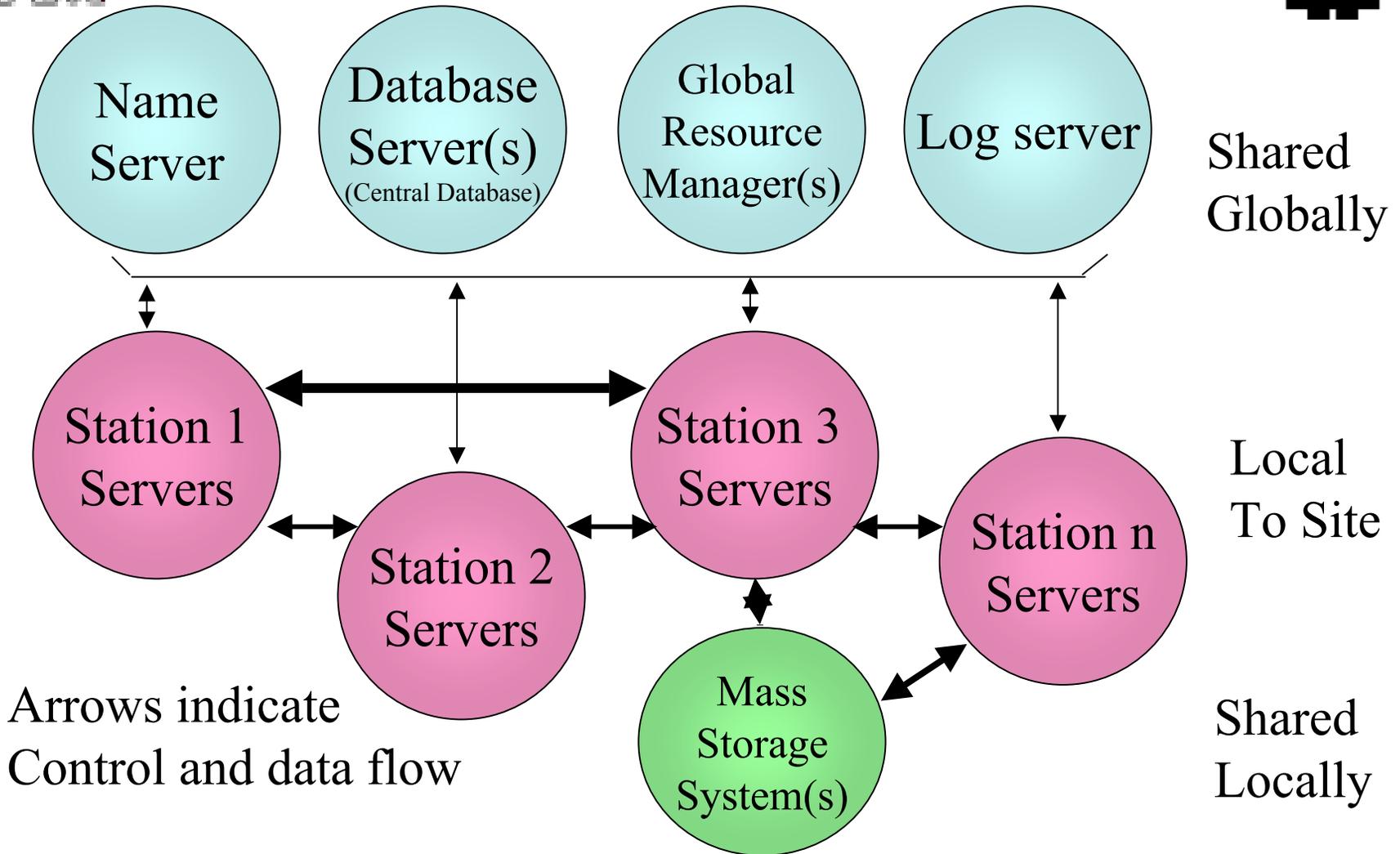
Meta-data Catalog

Indicates component that will be replaced  enhanced  or added  using PPDG and Grid tools

Name in “quotes” is SAM-given software component name



SAM as a Distributed System

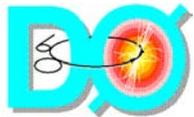
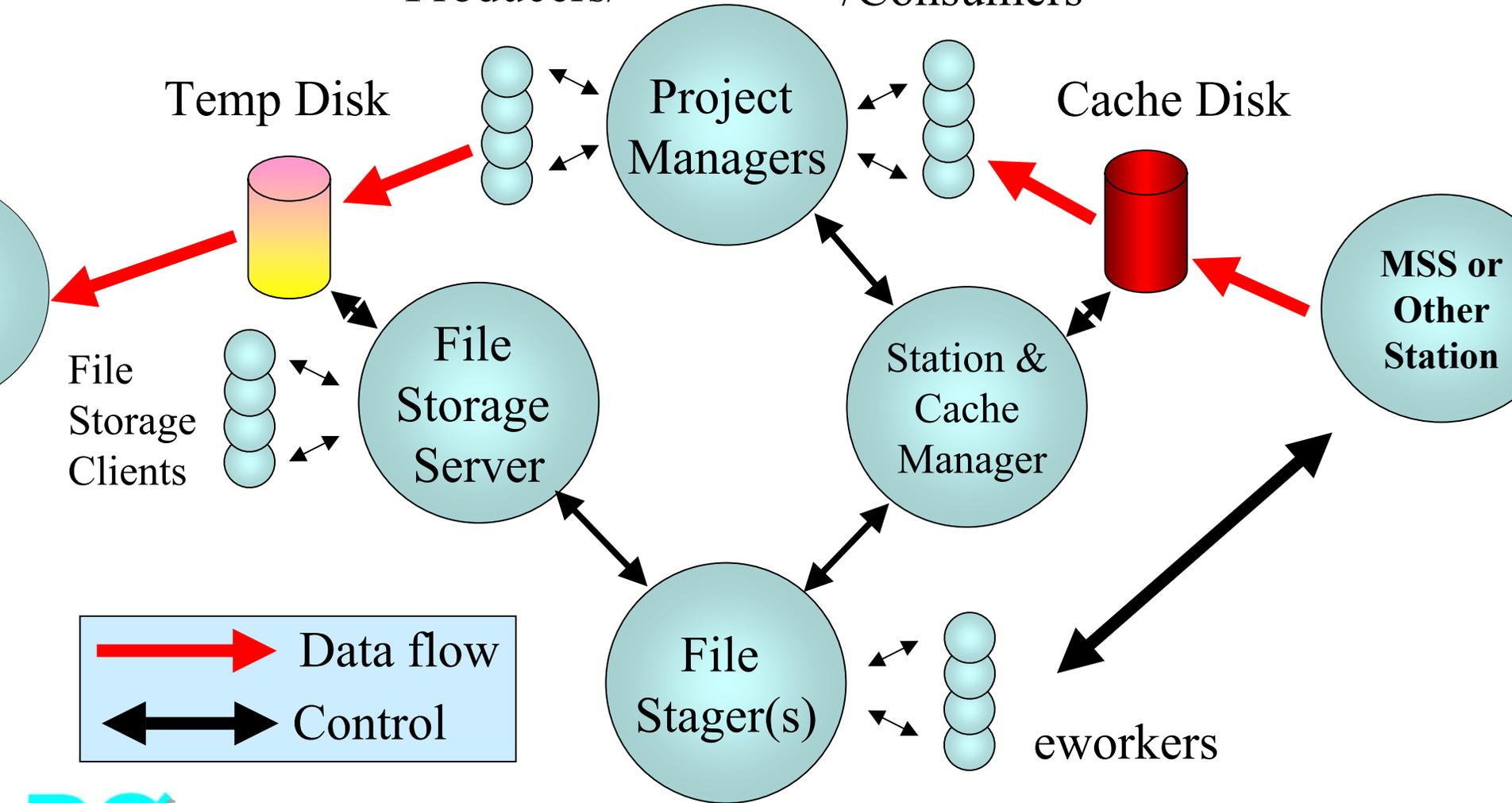




Components of a SAM Station



Producers/-----/Consumers





SAM Station: Dzero Distributed Reconstruction Farm

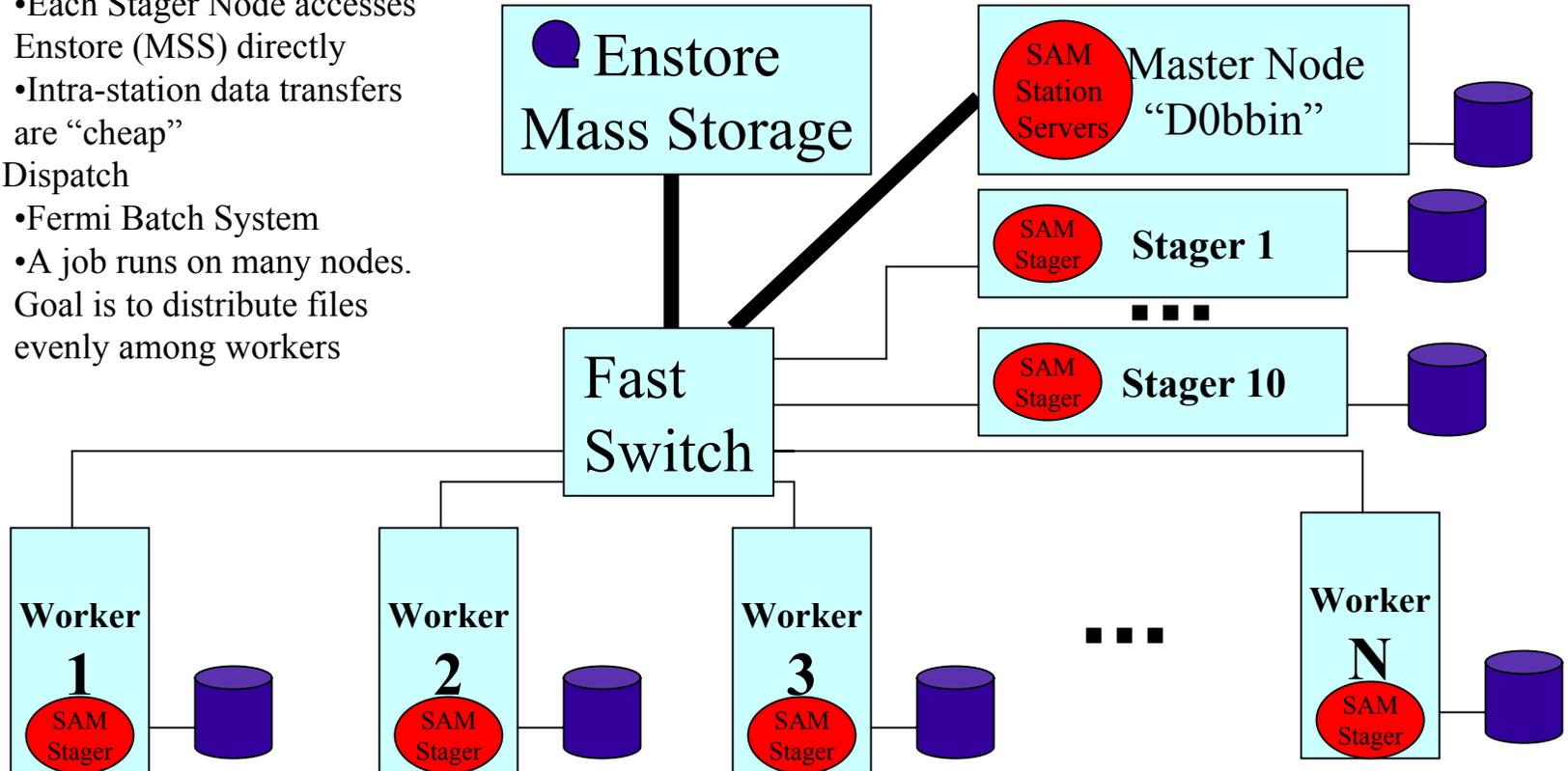


•Network

- Each Stager Node accesses Enstore (MSS) directly
- Intra-station data transfers are “cheap”

• Job Dispatch

- Fermi Batch System
- A job runs on many nodes. Goal is to distribute files evenly among workers

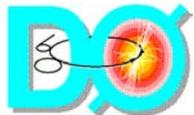
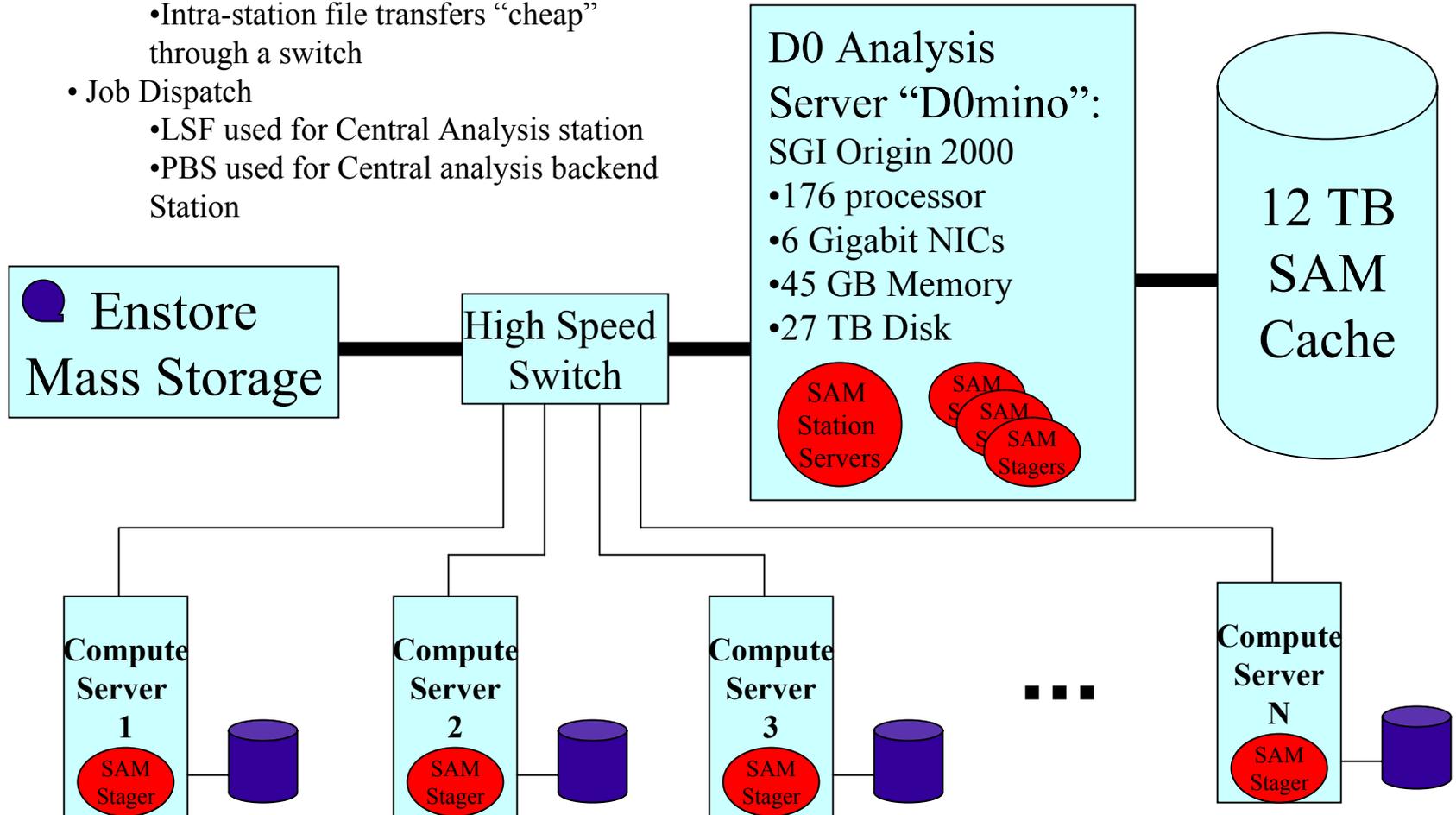




SAM Stations: Dzero Central Analysis and Central Analysis Backend



- Network
 - Access to Enstore is through D0mino
 - Intra-station file transfers “cheap” through a switch
- Job Dispatch
 - LSF used for Central Analysis station
 - PBS used for Central analysis backend Station



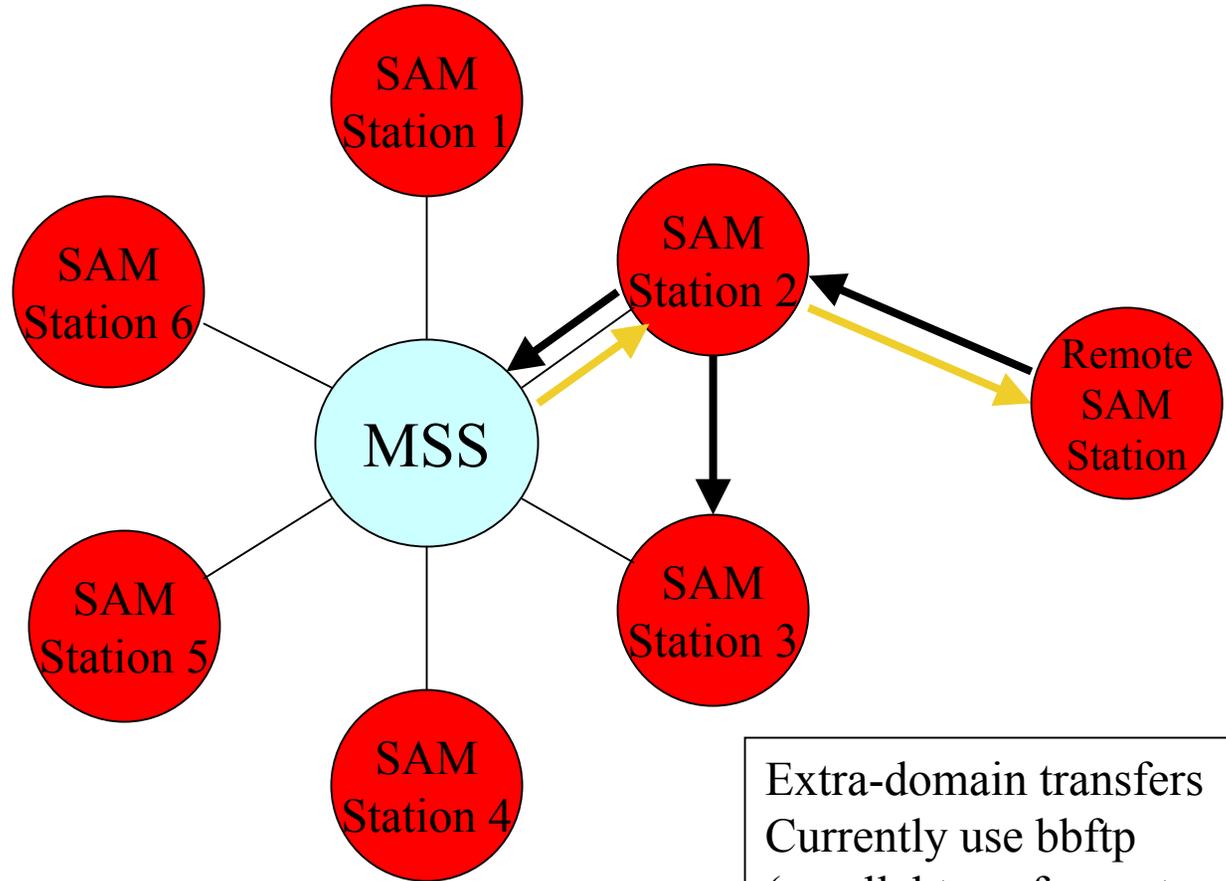


Data to and from Remote Sites



Station Configuration

- Replica location
 - Prefer
 - Avoid
- Forwarding
 - File stores can be forwarded through other stations
- Routing
 - Routes for file transfers is configurable

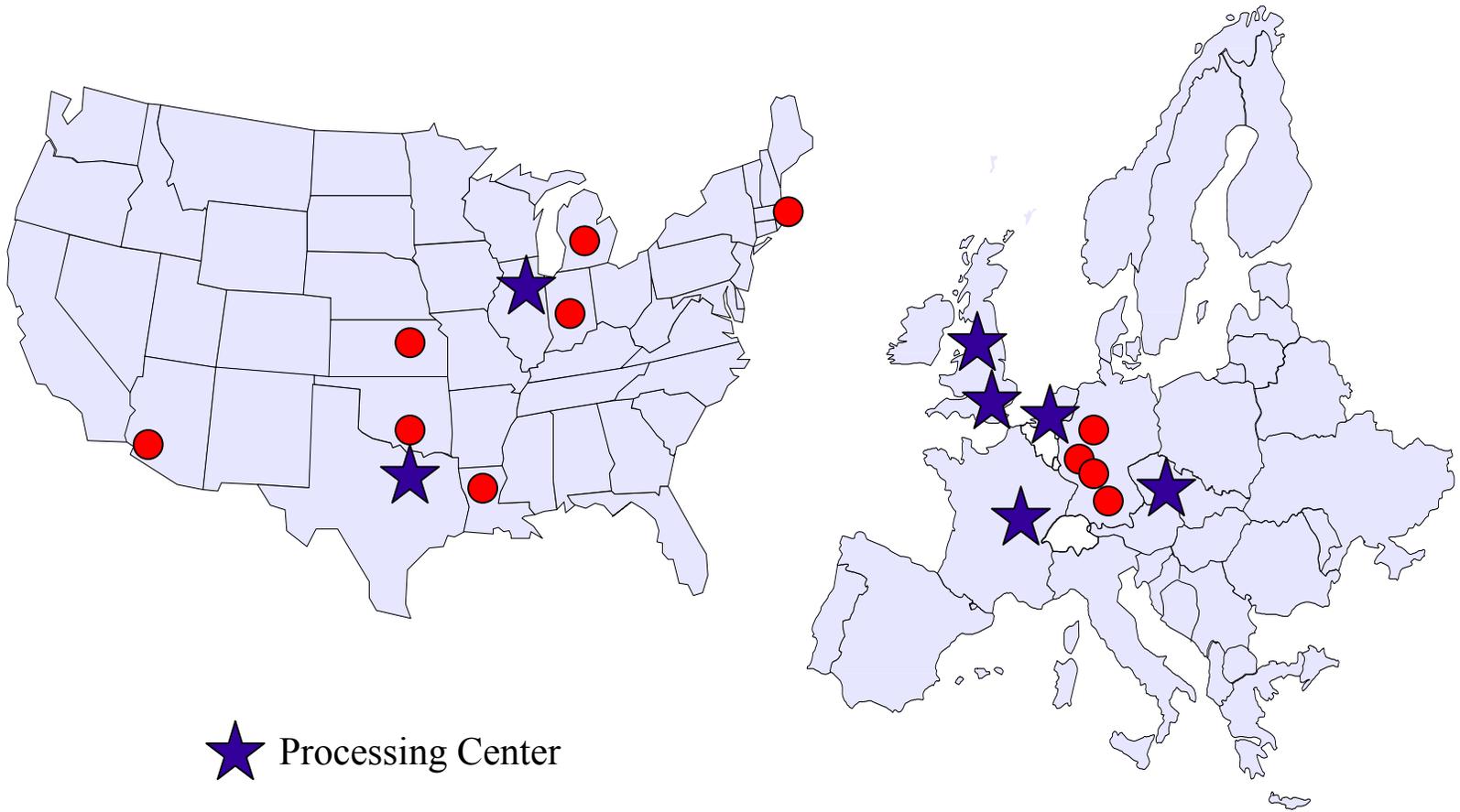


Extra-domain transfers
Currently use bbftp
(parallel transfer protocol)



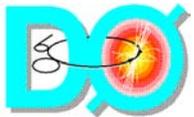


Dzero SAM Deployment Map



★ Processing Center

● Analysis site





SAM usage statistics for DZero

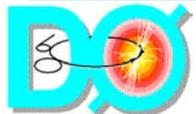
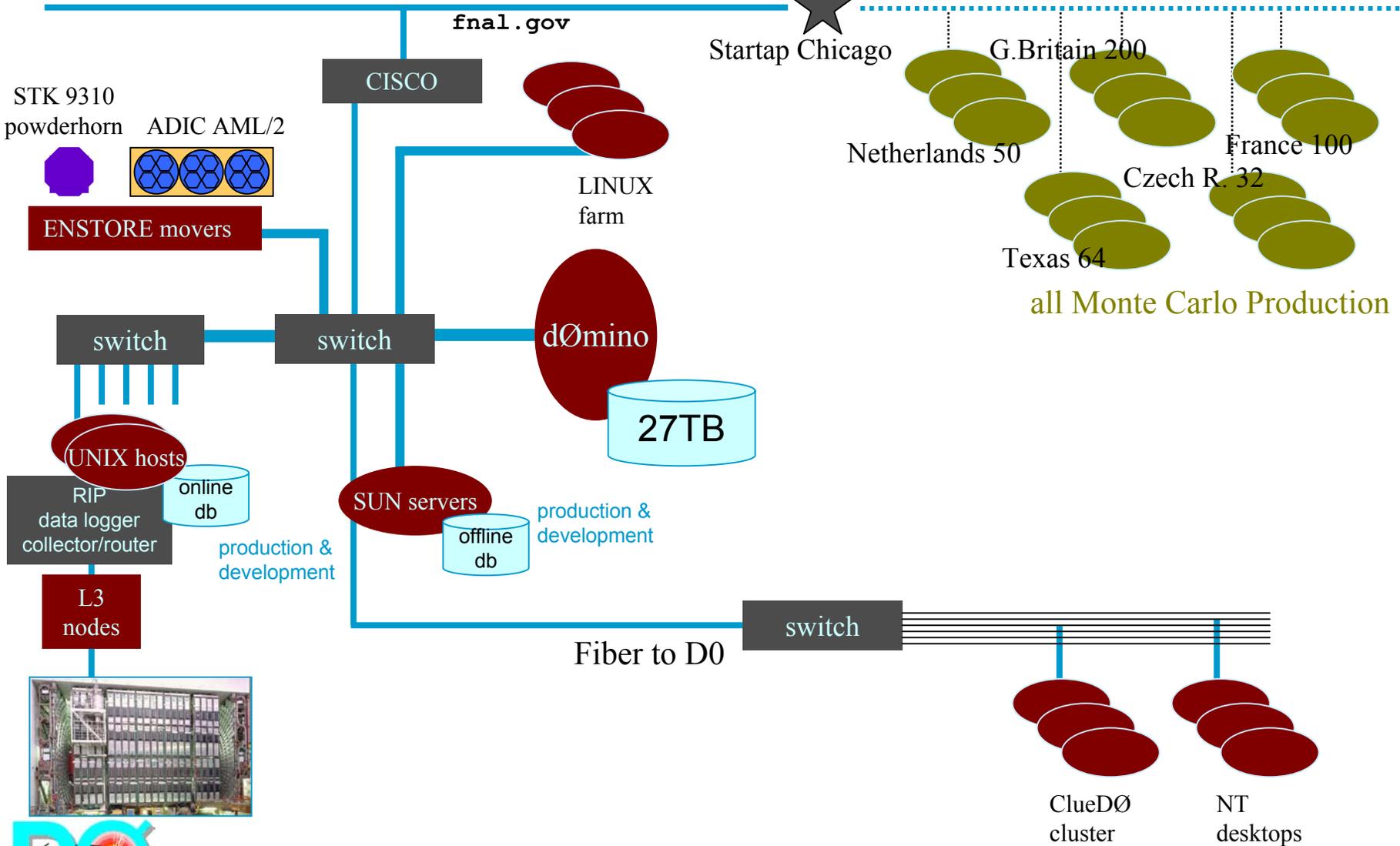


- 497 registered SAM users in production
 - 360 of them have at some time run at least one SAM project
 - 132 of them have run more than 100 SAM projects
 - 323 of them have run a SAM project at some time in the past year
 - 195 of them have run a SAM project in the past 2 months
- 340 registered nodes
- 48 registered stations, over two dozen vary active.
- 57,698 cached files on disk somewhere
- 668,490 physical and virtual data files known to SAM
- 409,087 physical files
 - 88,597 raw file
 - 181,414 reconstructed files
 - 86,733 root-tuple files
 - 293,041 montecarlo related files





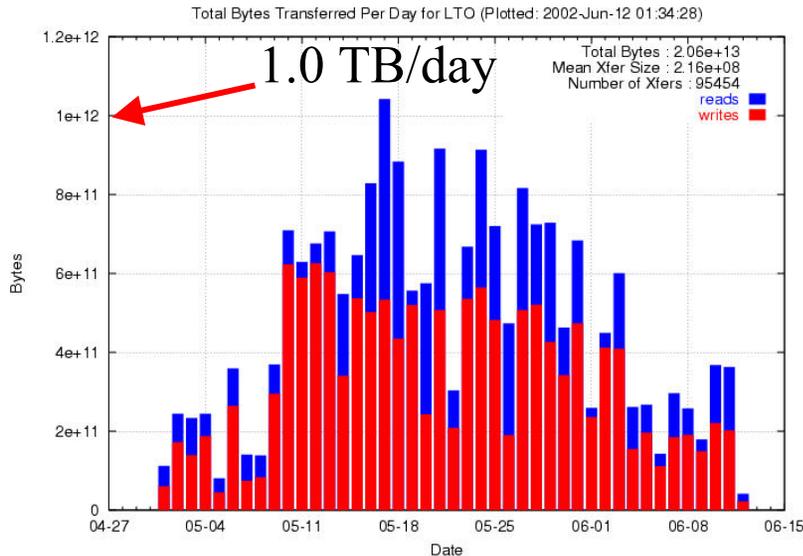
DØ data handling/database system architecture



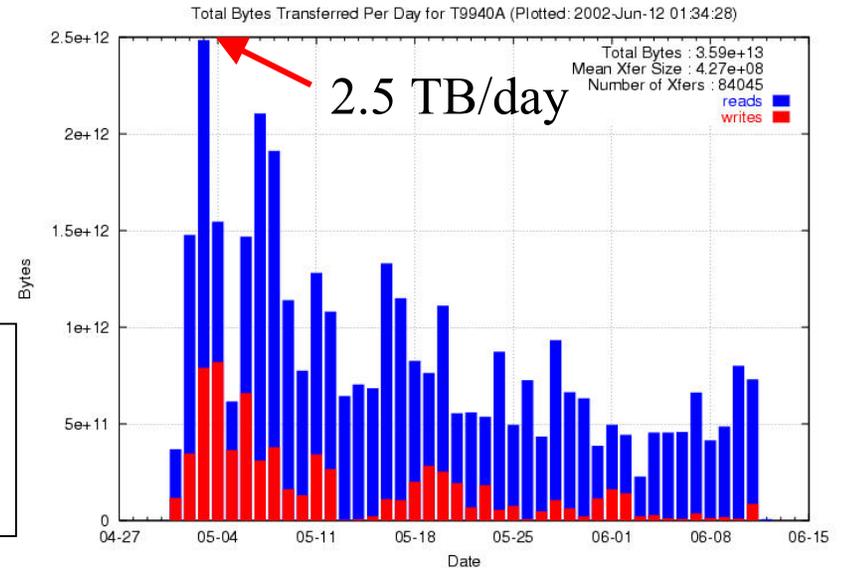
d0db.fnal.gov/sam



Typical Tape Robot Data Rates



ADIC Robot/LTO Drives



STK Robot/ 9940 Drives



Legend:

Read (blue)

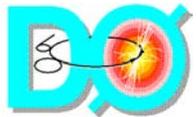
Write (red)

Total data: 115 TB





The commissioning of SAM for CDF





The Commissioning Project

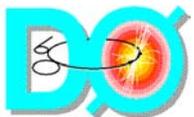


■ History

- ◆ CDF started the evaluation of SAM near the end of 2001
- ◆ During the last three months additional CD/DZero/CDF effort has been added.

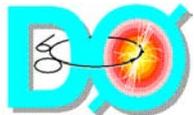
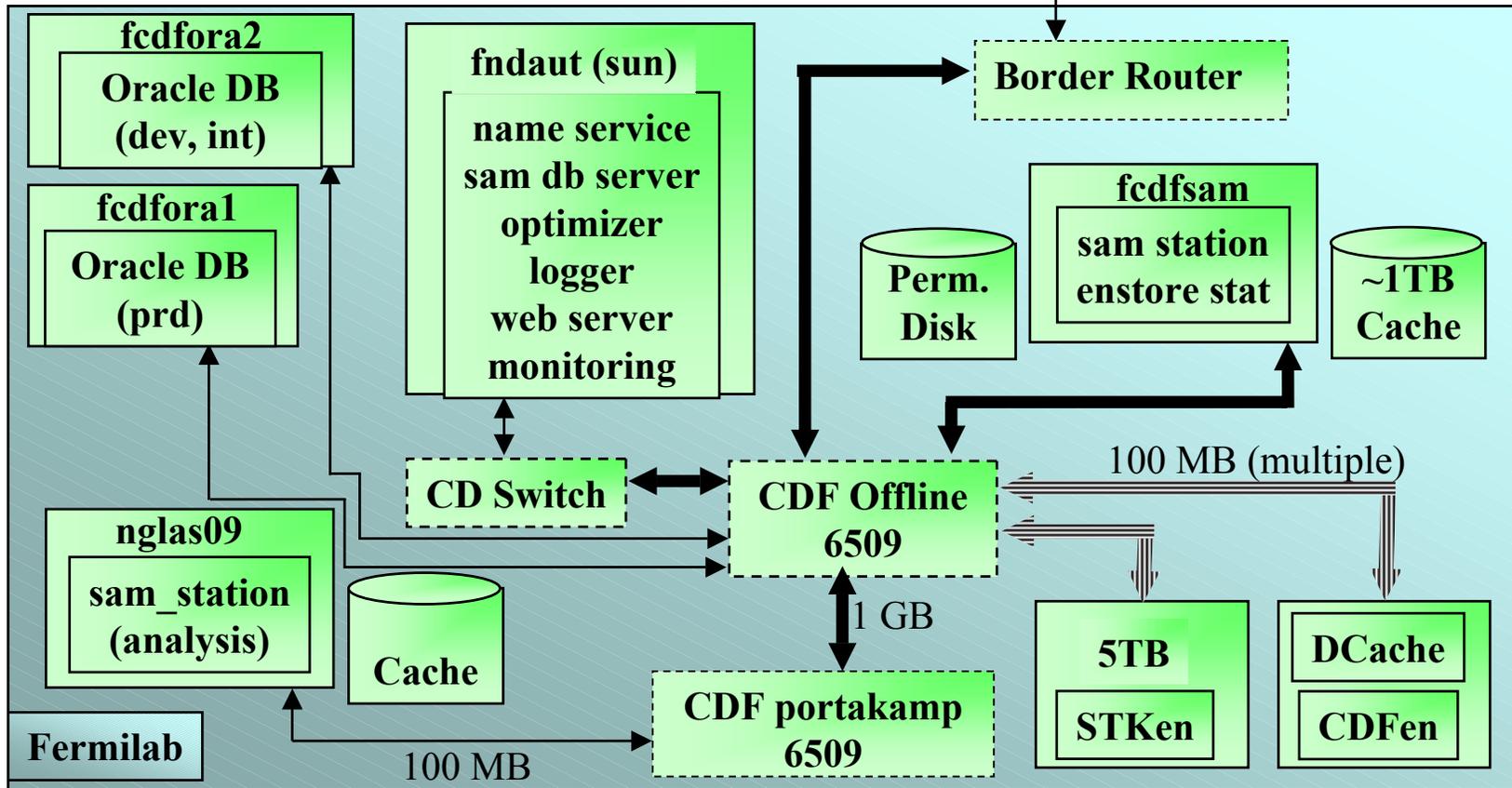
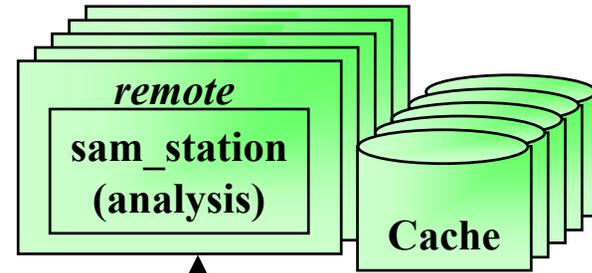
■ Goals of the Commissioning Project

- ◆ Supporting 5 groups for data analysis
- ◆ Enabling access to datasets of interest
- ◆ Production availability of the systems
- ◆ Limited impact on CDF Enstore effort (CDF has already converted to use Enstore)





Infrastructure for CDF-SAM





Status of CDF-SAM project



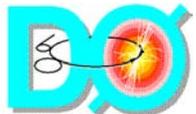
- The Hardware and Software infrastructure are in place
- Translation of the CDF Data File Catalog in production on June 3.
- Developed AC++ (CDF analysis package) interfaces to SAM to retrieve and analyze files.
- Enabled access to dCache (Enstore + DESY)
- Deploying to test sites to sort out configuration issues.
- Test users in UK, Italy and US are starting to use SAM to do physics.





SAM and the Grid

SAM-Grid

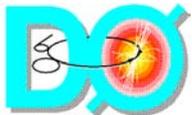




What is SAM-Grid?



- Project to include Job and Information Management with the SAM Data Management System
- Project started in 2001 as part of the PPDG collaboration to handle D0's expanded needs.
- Recently included CDF
- Current SAM-Grid team includes:
 - ◆ Andrew Baranovski, Gabriele Garzoglio, Lee Lueking, Dane Skow, Igor Terekhov, Rod Walker (Imperial College), Jae Yu (U. Texas Arlington)
 - ◆ Collaboration with U. Wisconsin Condor team.
- **<http://www-d0.fnal.gov/computing/grid>**





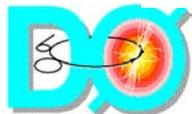
The Goal



- Enable fully distributed computing for the DZero and CDF, by enhancing SAM and incorporating standard Grid tools and protocols. Developing new solutions for Grid computing in a secure and accountable environment.
- The SAM grid-ification is funded by PPDG and GridPP. The collaborators we are working with include the Condor Team (via PPDG) and Imperial College (via GridPP)
- We are communicating with other groups working on Grid technologies as well (EDG among them).
- Regular CDF/DZero joint grid meetings
- We promote interoperability and code reuse



Condor
High Throughput Computing



d0db.fnal.gov/sam



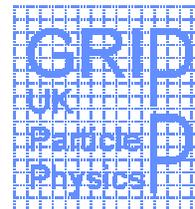
Major Components



- **Job Definition and Management:** The preliminary job management architecture is aggressively based on the Condor technology provided by through our collaboration with University of Wisconsin CS Group.
- **Monitoring and Information Services:** We assign a critical role to this part of the system and widen the boundaries of this component to include all services that provide, or receive, information relevant for job and data management.
- **Data Handling:** The existing SAM Data Handling system, when properly abstracted, plays a principal role in the overall architecture and has direct effects on the Job Management services.

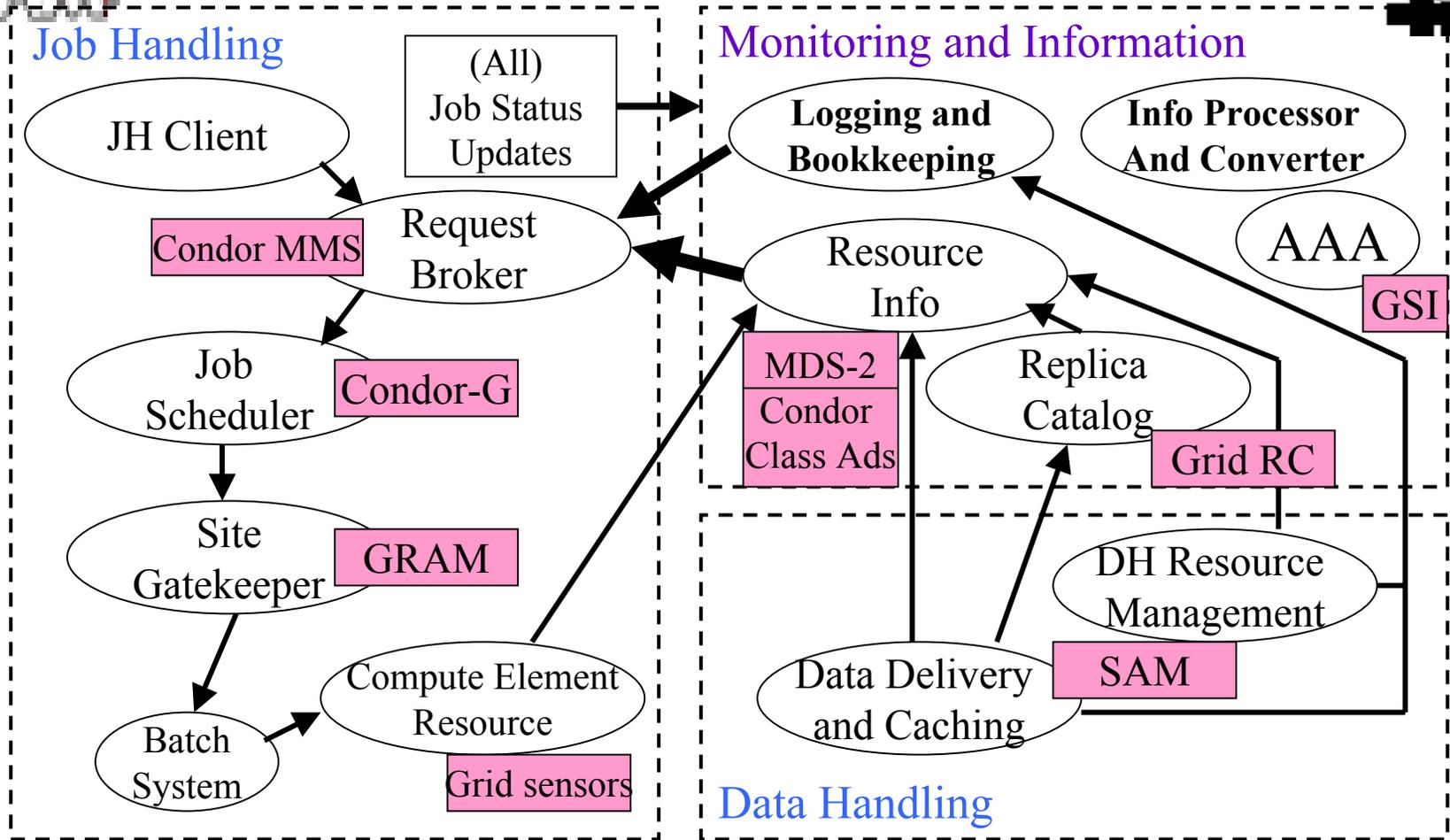


d0db.fnal.gov/sam





SAM-Grid Architecture



Principal Component

Service

Implementation Or Library

Information



d0db.fnal.gov/sam

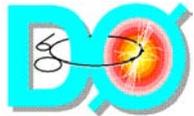
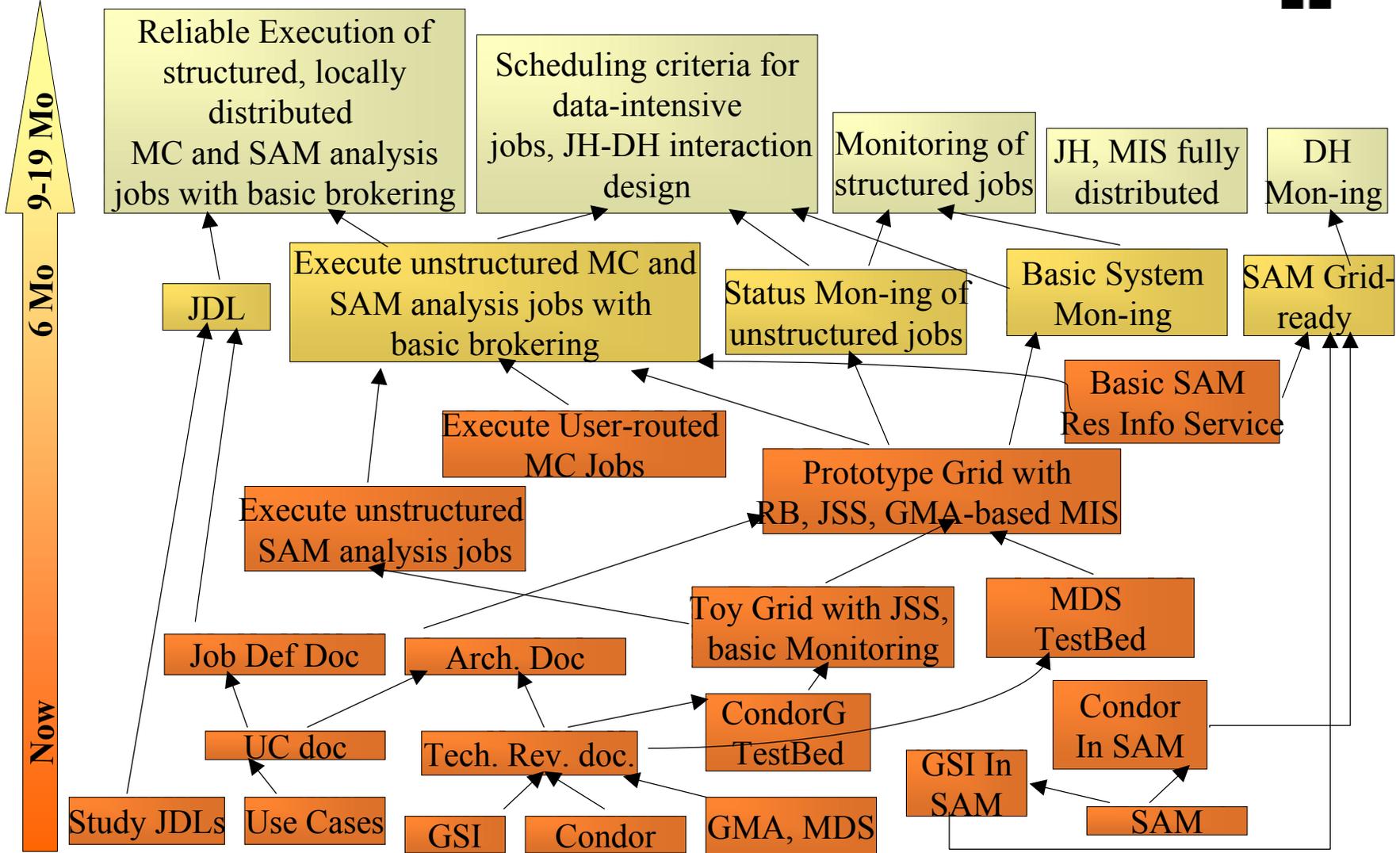


Condor
High Throughput Computing





The Milestone Dependencies





Conclusion



- The Dzero data handling system is used heavily, and has been very reliable. The distributed data handling model is proven sound and we expect to scale to meet Dzero's needs.
- SAM is being used to manage several kinds of compute hardware for various processing needs including processing farms and analysis clusters. This will continue and be expanded. The system has a worldwide deployment within the Dzero collaboration.
- CDF has evaluated SAM and is moving toward commissioning their own system.
- SAM-Grid is an architecture for a computing grid that is being developed and will enable easy and efficient use of compute resources around the world. This system is being built with “standard Grid Middleware” including Globus toolkit and Condor technology.
- This work is done in conjunction with Particle Physics Data Grid (PPDG) collaboration, and GridPP.



d0db.fnal.gov/sam

