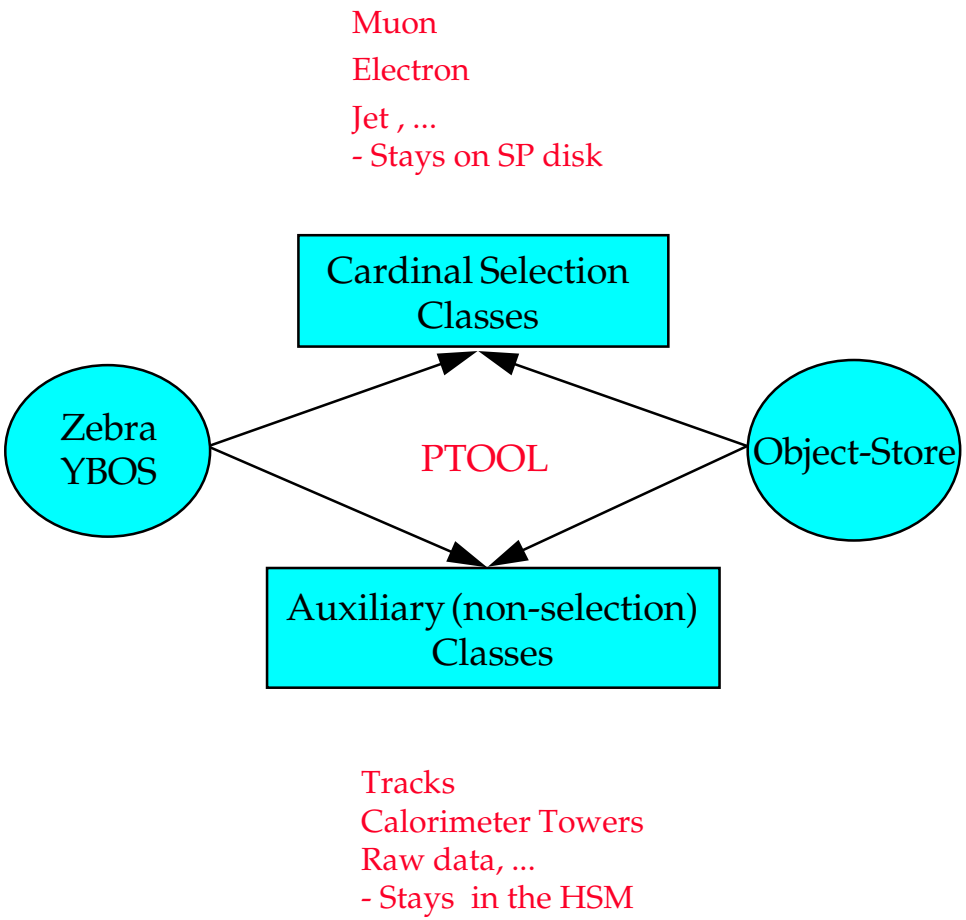
CAP Data Model

Objects

and

Access

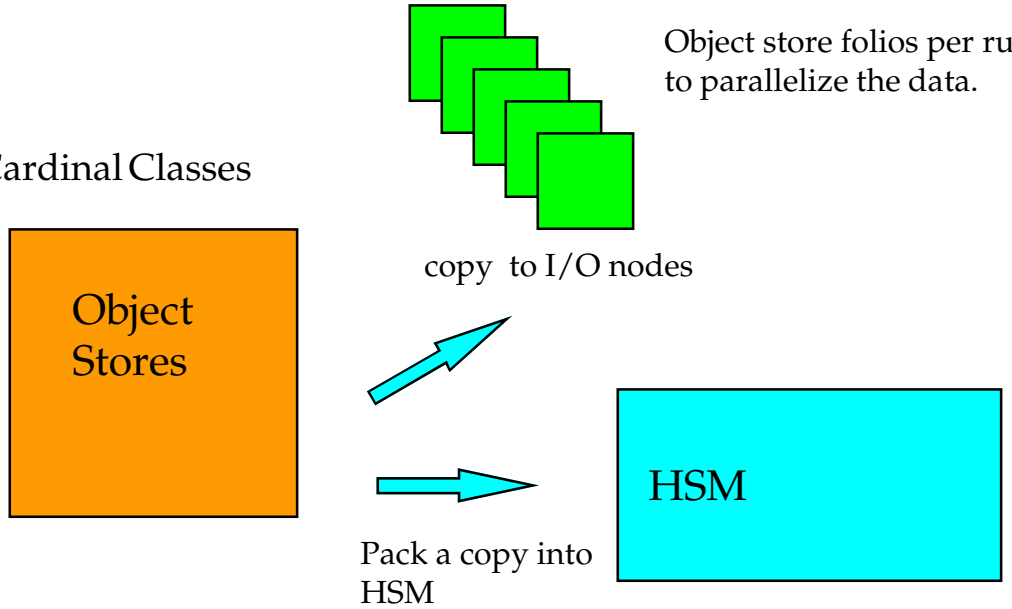
Data View



- ptool - high performance object manager - provides object storage and access, clustering of data
- ptool can support Vesta-like access or a “client-server” model

Save Object Stores Onto HSM

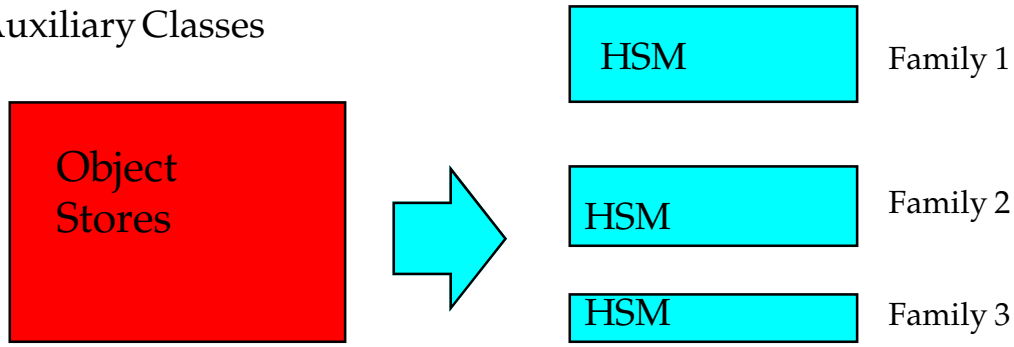
Cardinal Classes



Object store folios per run are spread across the I/O nodes to parallelize the data.

Cardinal classes are stored for recovery and I/O node changes. Data is packed for more efficient retrieval from HSM. Packing is organized by grouping several runs together.

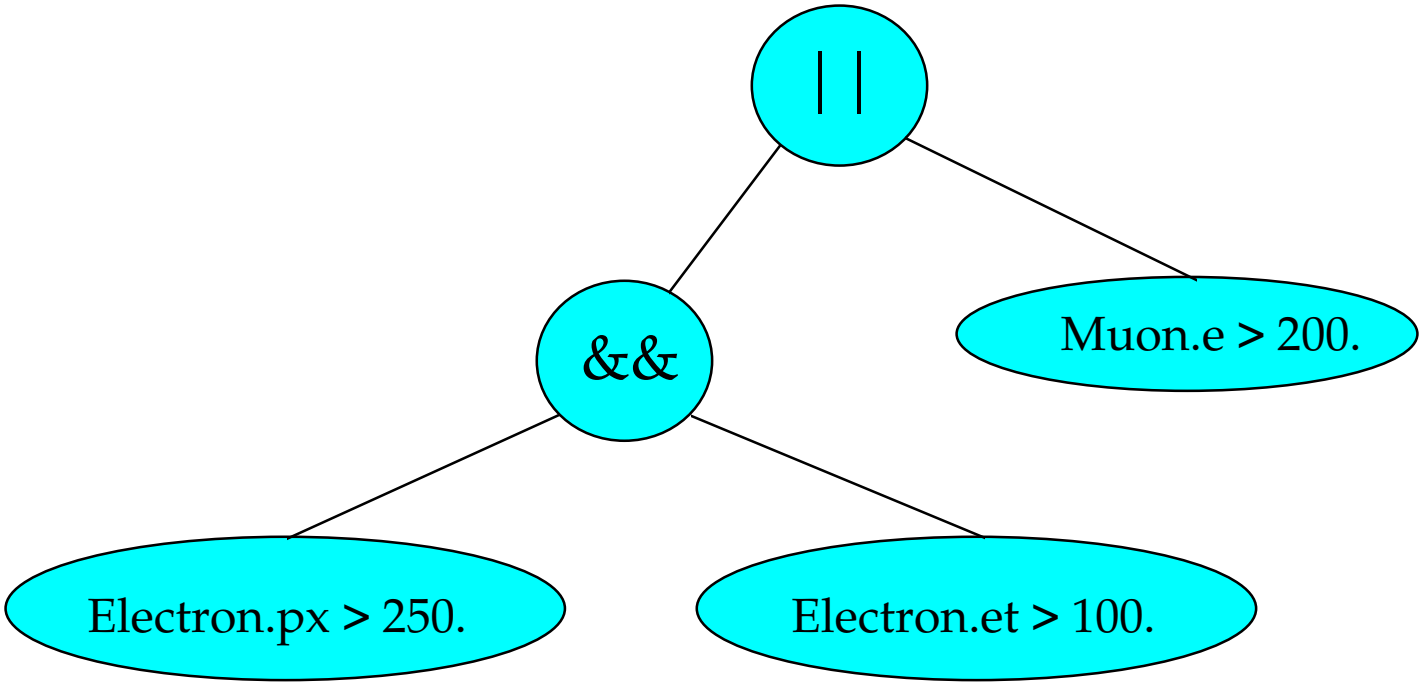
Auxiliary Classes



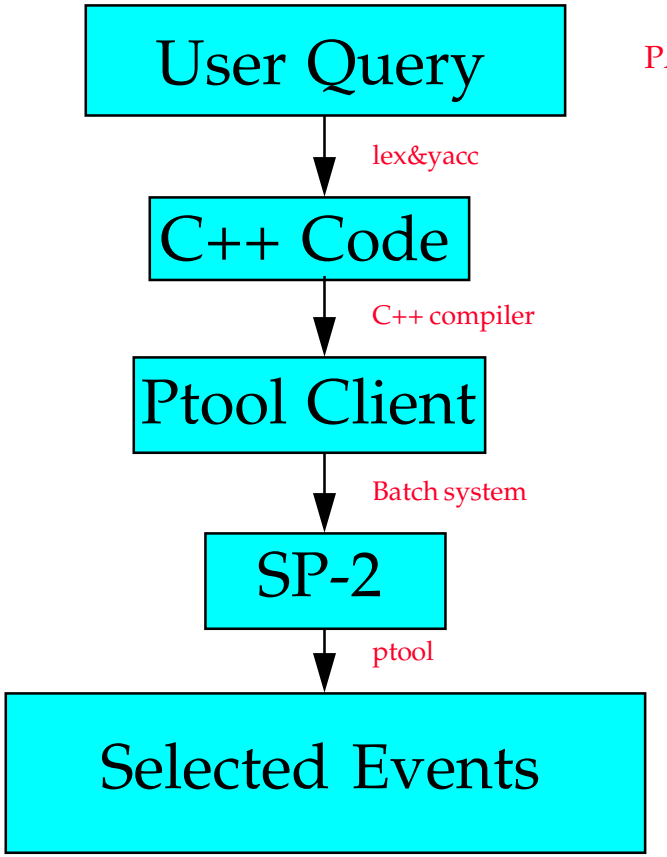
Data is not packed since folios will be requested in a random fashion. Data is organized by run. Data is migrated to tape in run groups by using families and controlling the migration times. Recall of data will be queued and ordered based on how the data was organized in the HSM ATL

Finding data: the query language

Electron.px > 250. && Electron.et > 100. || Muon.e > 200.



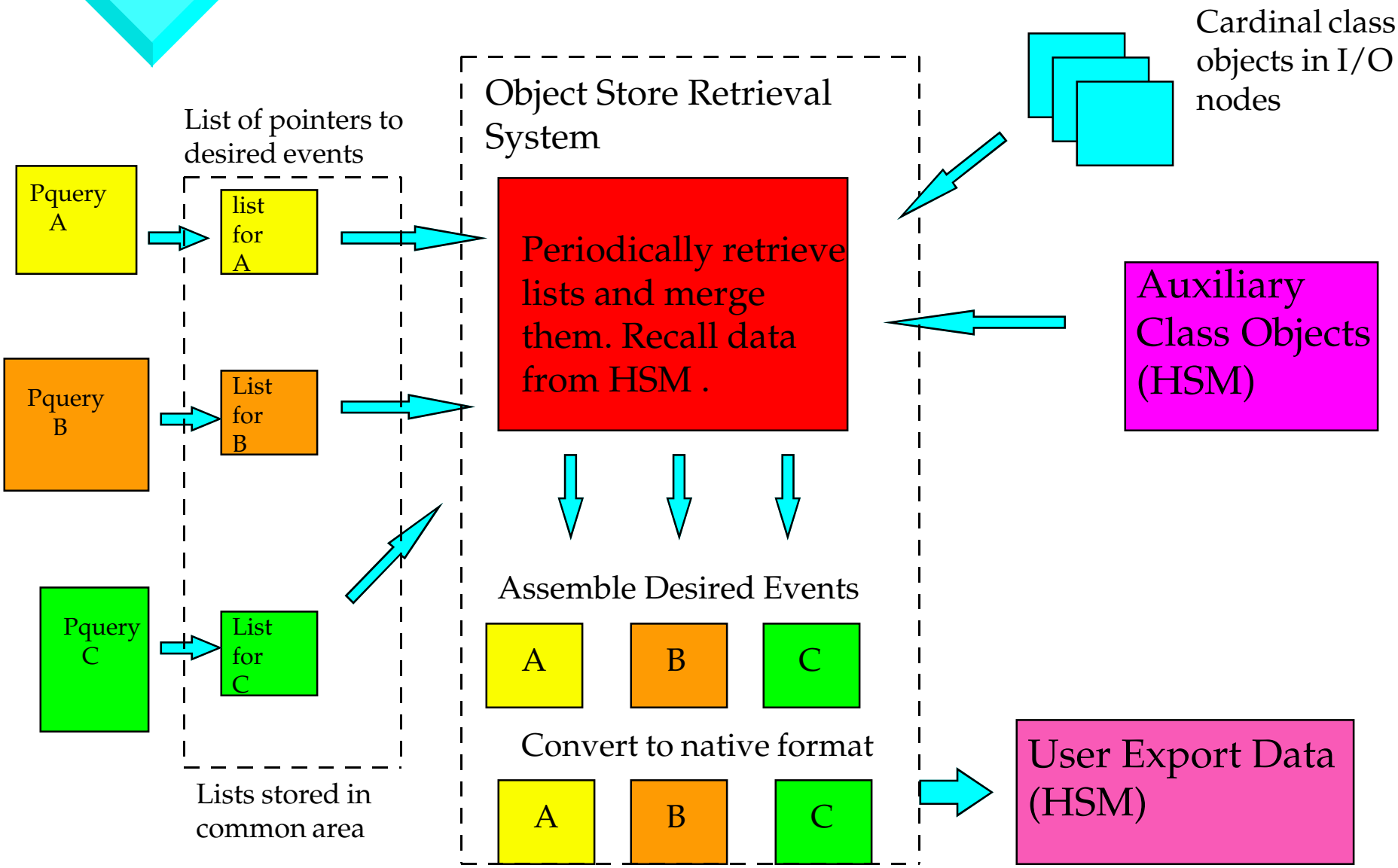
Query processing



PAW, OQL

Zebra , YBOS,
object-store, event list,
persistent-pointer list

Retrieval of selected data



HSM Roles for CAP

User Original Data
Prototype - 200 GB

Native Format

YBOS
Zebra

Needfile
(Temp space)

Generated Object Stores
Prototype - 200 GB

Cardinal Classes

STA 5%
DST 50%

Auxiliary Classes

STA 95%
DST 50%

Archive (Perm space)

Mined data for Export
Prototype - 5 GB

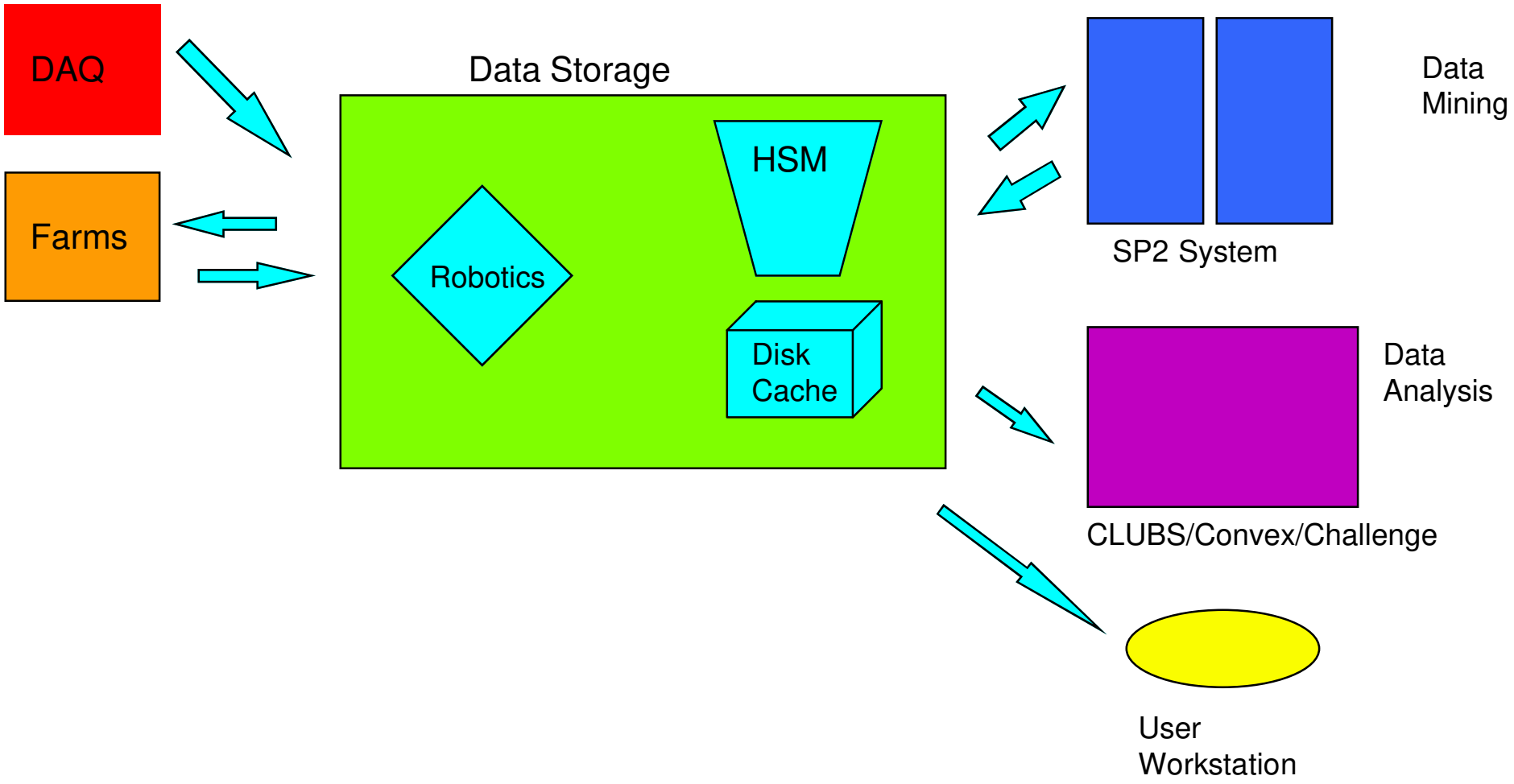
Native Format

Tcache
(Fixed retention)

Gain Experience from Prototype

- Prove data mining concept
- Determine real bandwidth needs
- Improve the HSM storage model
- Build system to handle 200 GB object store repository

Future


DAQ
Farms
Data Storage
HSM
Robotics
Disk
Cache
Data
Mining
SP2 System
Data
Analysis
CLUBS/Convex/Challenge
User
Workstation