



Database Mini-Review - 23rd Sept. 1997

BABAR OO Databases Overall Design Issues

David R. Quarrie

Lawrence Berkeley National Laboratory

DRQuarrie@LBL.Gov



Objectivity/DB

- Technology that all the OO databases are based on
- Basis of RD45 collaboration from CERN
- Our approach has been to demonstrate that it will fulfill all roles for us, even if an alternative technology might have been as suitable for one of the several roles.
- Collaboration with RD45 can help but cannot be the complete answer
 - Very different timescales



Overall Design

- BABAR DBMS covers several *domains*
 - Event Store
 - Conditions Database
 - Online Databases
- Objectivity is underlying technology
- Domains are logically independent
 - Actually tied together by underlying Objectivity Federated Database
 - ▶ Also transaction boundaries
- Allows independent development with late binding



Data Protection & Safety

- Need to protect data against accidental corruption or deletion.
- Understanding of tools needed
- Concept of authorization levels to protect database contents
 - Domain specific
 - ▶ System, Group, User
- Anyone has read access to complete database
 - Including all subdetectors, physics groups & other users data
- Talk by Igor Gaponenko



Placement & Clustering

- Performance is gained by placement & *clustering* of data
 - Minimize wasted data access
 - If you have to access a tape, try to access multiple databases
 - If you have to open a database, try to access multiple pages
 - If you have to access a page, try to access multiple objects
 - Predictive clustering & heuristic monitoring
- Clustering cannot solve problem of doubly-dilute data
 - Cluster for small fraction of events in sample
 - Cluster for small fraction of data per event
 - Only solution to both is to duplicate relevant data/events
- Talk by Jacek Becla

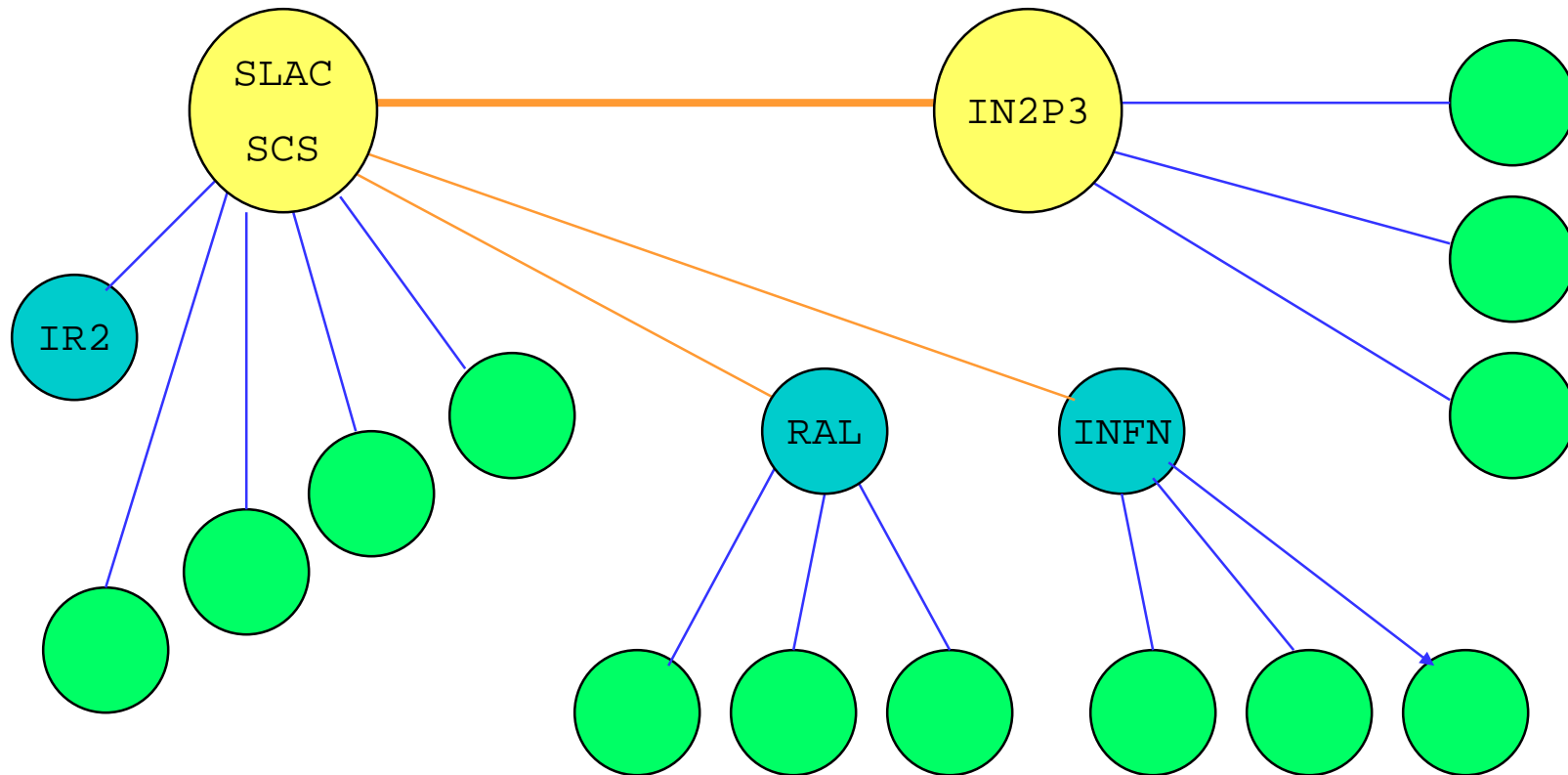


Data Distribution

- Part of placement strategy
 - Place data close to where it is being accessed
- Issues of network reliability & bandwidth
- Role of Regional Centers
 - IN2P2, RAL, INFN
- Data sharing between people doing physics analysis
- Talk by Jean-Noel Albert



Data Distribution Strategy



— Tape transfers

— Partitioning?



Data archiving

- Deal with >150TB per year
 - Won't fit on affordable disk
 - Integrate disk & tape
- Many database files (>50,000)
 - Part of clustering/placement strategy
- Hierarchical mass store HPSS
- Need to integrate Objectivity with HPSS
 - Transparent caching & migration of database files to & from tape
- Talk by Andy Hanushevsky

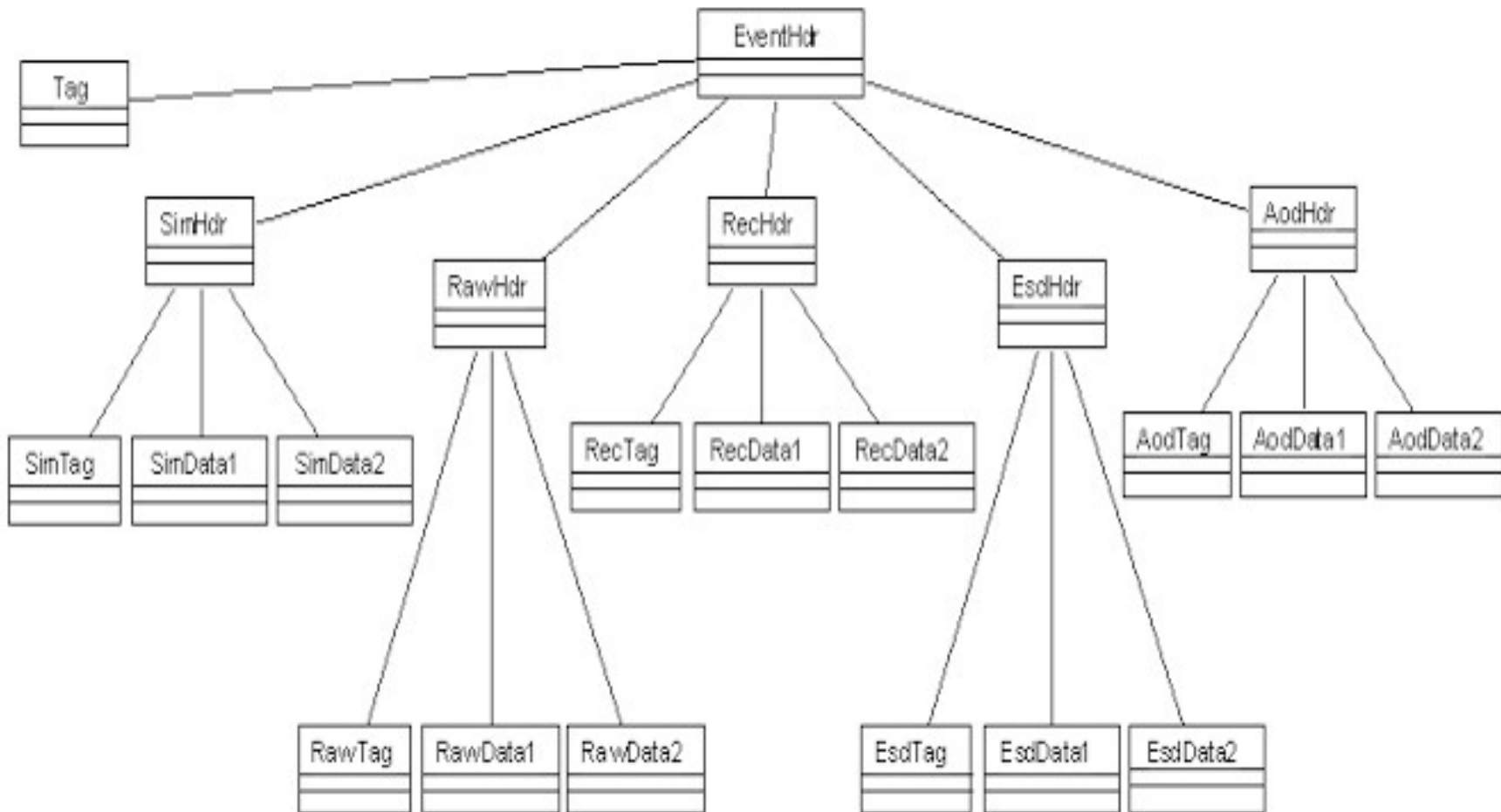


Event Structure

- Compact
 - Minimize database overhead
- Hierarchical
 - Allow access to portions of the event with minimal overhead
- Extensible
 - Allow us to add new data as our understanding improves with minimal impact
- Talk by Ray Cowan



Event Hierarchy





Collections, Dictionaries & Registry

- Event collections “hold” event samples
 - Contain references to
- One dictionary per user, group and system
 - Contain references to named event collections
 - ▶ Names must be unique within a dictionary
 - ▶ Names may have aliases (not yet implemented)
 - Mechanism by which a collection may be located and used as a source of events
 - ▶ Application locates the desired dictionary & then the required event collection
- Event Registry
 - Set of all dictionaries & other management information



Event Collections & Access

- Different Event collections by default reference the same event headers for the same event
- Have flexibility
 - Copy the Event Header for dilute samples
 - ▶ Enhanced performance
 - Copy the Event Tag
 - Copy a portion of the event hierarchy
- Garbage collection is an issue
 - Delete those portions of the event that have been copied, but *not* the original portions



Event API

- Exposure to database for reconstruction programmer?
- Decision following June Workshop
 - Transient Interface
 - Maintain our existing reconstruction code
- First generation prototype demonstrated feasibility
- Second generation underway
- Talk by Steve Gowdy (David Quarrie)