DB activities

◆ Richard's and Torre's 2-year term as DB project leaders ended recently. Both stay in the database activities.

◆ The overall scope of the ATLAS database activities is unchanged. Emphasis shifts from development to commissioning+operation.

◆ Two activities shifted to other areas - which are GRID and Computing Operations related:

   1. Assign responsibility for the development of the Grid-related databases and data management system (ProdDB and DDM/DQ2) to the Grid Tools & Services area

   2. Define a high-profile "Database Deployment and Operations" activity within the Offline Computing Operations group

DB activities

◆ The database project itself covers:

a) event persistency (POOL)

b) calibration/alignment/conditions databases (COOL and related work)

c) detector geometry, and DB releases

d) online usage of databases - trigger, DAQ, DCS, detectors ...

Emphasis is on co-ordination of several contributing projects - e.g. OKS in TDAQ, DCS mostly in TDAQ except interface to COOL, COOL/CORAL in DB project, detector work, ...

e) TC databases (racks, cabling, MTF)

f) non-event metadata (AMI etc.)

This area is new to the project and will be influenced by the outcome of the metadata task force.

Non-event-related metadata will have a strong interplay with event-related metadata, dealt with mainly in activity a), as well as with DDM metadata, dealt with in activity 1) on previous page.

DB Co-ordination

◆ A newly established DB co-ordination group covers all of the activities a-f, with liaison to 1-2

◆ Composition (could still vary a bit)
  ◆ event persistency: David Malon
  ◆ calib/align/conditions: Richard Hawkings
  ◆ geometry/releases: Vakho Tsulaia
  ◆ online: Larry Price, Johannes Haller, Igor Soloview, HvdS
  ◆ TC: Pascal Perrodo
  ◆ metadata: Stephane Jezequel
  ◆ production/DDM: -tbd-
  ◆ deployment/operations: Sasha Vanyashin

DB testing in 2006

◆ Several occasions in 2006

  ◆ Tier0 scaling tests: RAW data to ESD/AOD/TAG (i.e. around Tier0), mainly with repeated data samples; also with replication of some of the data to Tier1s ("Oracle streams" used) started on June 19th, for 3 weeks

  ◆ Within CSC: Similar with much refined data samples, RAW data streams, lumi blocks - so interesting data in TAG
  in September 2006

  ◆ Large Scale Tests - focus on configuring fast enough from configuration DBs, DCS, conditions, geometry DB - involves multiple layers of replication/caching

  ◆ Second focus on conditions writing

  ◆ Need to test the MDT calibration scheme (external calibration centres => results to Tier0) once the scheme is in place

◆ Involve some 100 ... 1200 nodes

◆ Some functional aspects not finalized:
  - ◆ Replication or caching to (200) rack servers
    - ◆ MySQL or sqlite replication
    - ◆ FroNtier or DbProxy caching (SLAC)
  - ◆ Caching stage below this
    - ◆ per ROD crate?
    - ◆ for HLT: within a box - cache or share?
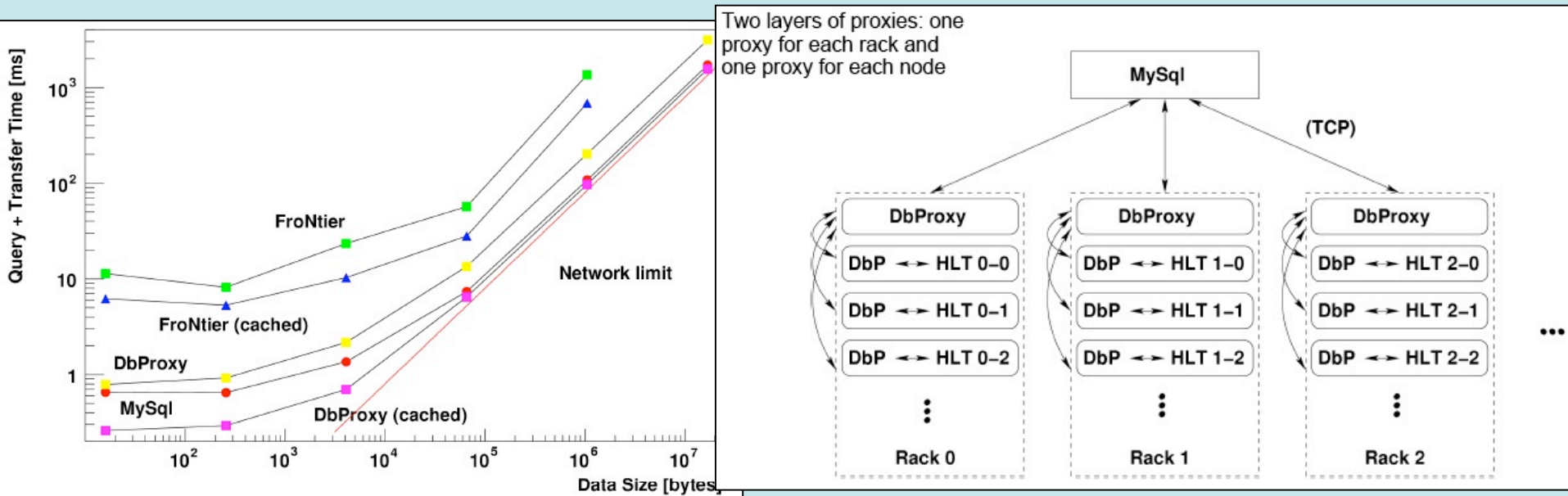
Sharing of process-constant data

◆ Now: 4 CPU cores per box: 2 chips * 2 cores => 8 cores per box
◆ Typically one Athena process (LVL2: one thread) per core
◆ Multihreading: naturally assign sharable data to global thread, i.e. have only one copy for all threads
◆ Geometry, field maps, ...
and more types of data which are constant throughout process

◆ Use also for single-threaded environment (event filter, offline)

◆ Advantage:
  ◆ Save DB accesses (volume, #connections)
  ◆ Save physical memory

◆ Athena not there yet...
  ◆ Change within Gaudi
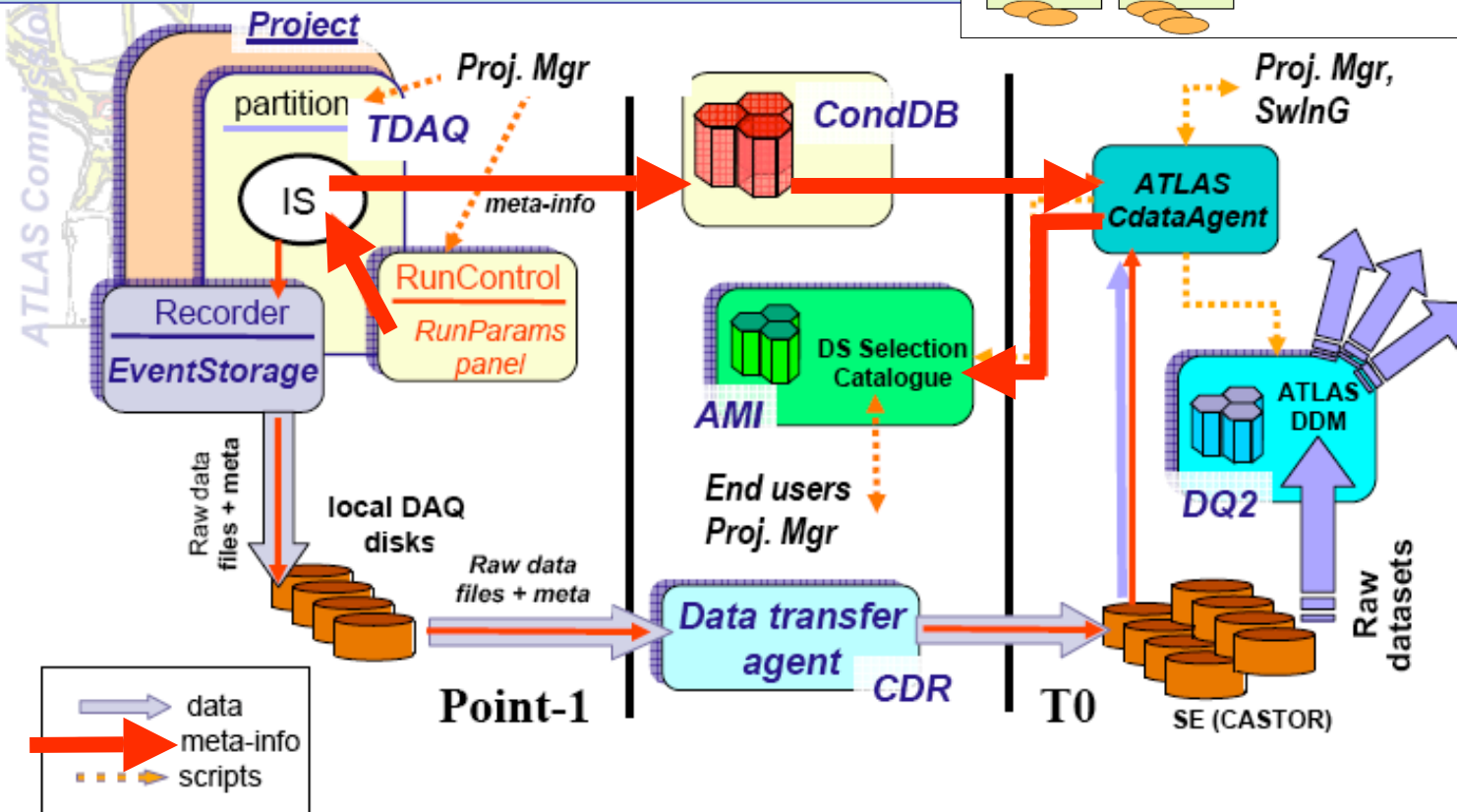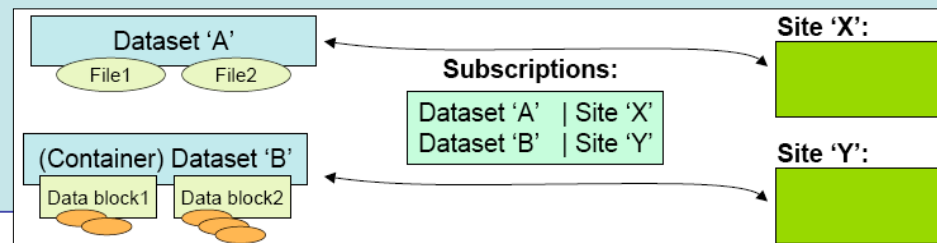 ◆ Keep constants constant
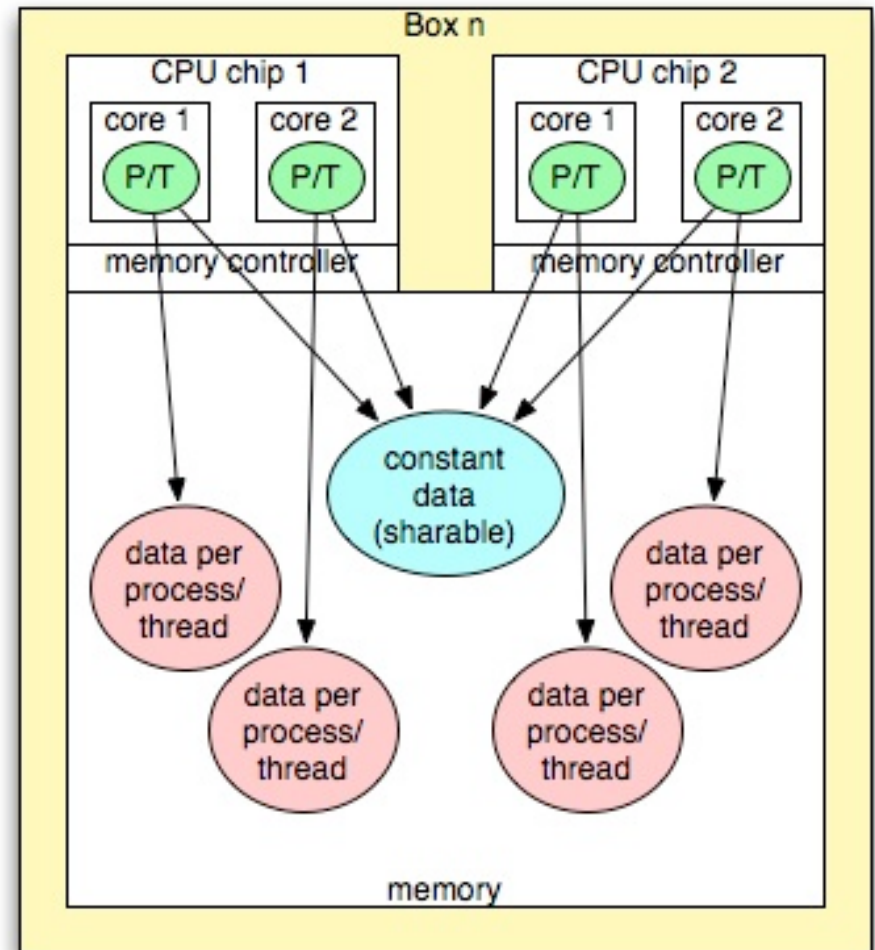
Metadata and DDM

Flow of Data and Metadata from the Cavern via Tier0 worldwide: **Distributed Data Management**

Primary metadata originate in RunControl and proceed via COOL to cataloging in DQ2 (AMI: metadata DB)

DQ2 replicates sets of data files (datasets) to the remote sites which can subscribe to datasets

Options under study:

CDR or DQ2 for data transfer from Point1 to Tier0