



Database Consolidation in 2011

Dario Barberis

Genoa University/INFN



Databases and services

- Conditions DB + Frontier/Squid
- TAG DB + ELSSI
- LFC



CondDB + Frontier Deployment

- Simulation and reconstruction of simulated data, as well as real data re-processing, need only DB releases.
- Analysis accesses conditions through Squids and Frontier servers at the Tier-1s.
 - Squid caches reduce access time for the jobs and load on the databases
 - also for local jobs
 - Currently 6 Tier-1s (BNL, TRIUMF, RAL, KIT, PIC, IN2P3) with a Frontier server plus CERN
 - The Tier-1s at ASGC, SARA, NDGF and CNAF are providing Squids
 - Streaming of CondDB to these sites is being stopped
- Existing Frontier servers will be consolidated and the sites need to give support in manpower and hardware for a robust service
 - PIC will discontinue support for CondDB and Frontier and concentrate on the TAG DB instead

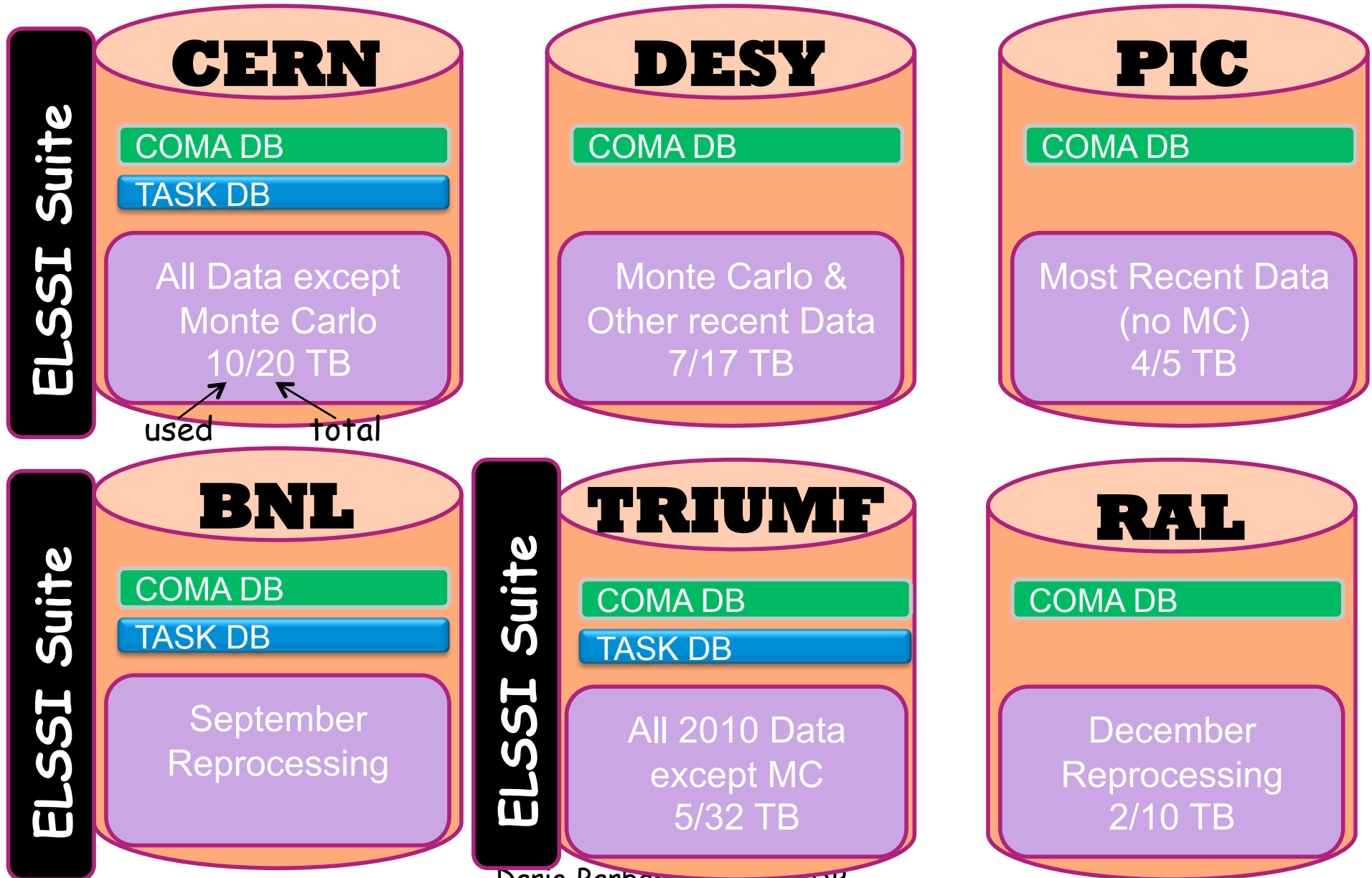


Overview of the TAG Info System

- The TAG Information System is composed of several databases scattered across Europe and America and a suite of services used to access the data
- The event data is composed of POOL Collections, where a run from a collection directly maps to a merged TAG dataset produced at Tier-0 and Tier-1s.
- The TAG catalog maintains the information of what data is available where
 - It is accessed by the ELSSI Suite services, at each application site, presently CERN, BNL and TRIUMF.
 - The TAG catalog is replicated from CERN to the other sites. It presently resides in the ATLR database.
- The TAG catalog will very soon include a «Service catalog» which will keep updated the state of all the services in the TAG Information systems
 - This will enable to make decisions at run time about failover and load balancing. For this need, replicas are needed to other sites
- We propose to replicate the TAG catalog from ATLR to the Tier-1 databases (small size ~30 MB)



Current TAG distribution





Consolidation of TAG DB and services

- We need to keep in TAG databases 2 (and only 2) copies of the same data at different sites
 - For obvious reasons of robustness against site down times
 - Load on the servers is not much
 - So far difficult to measure but tools in preparation
- In 2010 each reprocessing campaign produced ~3 TB of new data (this is the minimum "chunk" size)
 - Some old data (mainly MC) could be deleted in the past but it's more difficult in the future
- Each full year of running will produce at least 24 TB of additional data
 - 3 TB x 2 versions x 2 copies x 2 for MC
 - So we need increasing storage capacity but not necessarily a lot of computing power
- With site agreement, the Oracle capacity and manpower freed by CondDB can be used to increase the available storage capacity and server power for the TAG DB
 - Technical requirements have been specified by the TAG group
 - Work load for sites would not be appreciably different from now but they would provide a more useful service to the Collaboration
 - Discussions started with CNAF and SARA (both will be online by mid-2011)
 - BNL will discontinue support for their test instance of the TAG DB (that was very useful during the development and deployment period, thanks!) and concentrate their efforts on CondDB+Frontier



LFC consolidation at CERN

- Currently we have 16 LFC catalogues:
 - 1 at Tier-0 (Oracle in the LCGR database RAC)
 - 10 at each Tier-1 (9 Oracle, 1 MySQL)
 - 5 at US Tier-2s (all MySQL)
- Each of them constitutes a single point of failure for jobs running and data access in the corresponding sites
- The policy agreed at last S&C week foresees the consolidation and merger of all LFCs at CERN with the provision of a high availability service
 - In any case most of the LFC interactions are with DDM services and the Panda server @ CERN
- Technical discussion last Friday (ATLAS-ADC and CERN-IT people) resulted in the following plan:
 - J-Ph Baud develops the consistency check and merger scripts and tests them on a replica of a medium-sized LFC to measure and optimize:
 - Time needed for consistency check
 - Policy for reconciling LFCs in case of data mismatch
 - Time needed for merger
 - Total cloud downtime during merger
 - When these parameters are known (~1 month) the detailed schedule can be built
 - One cloud at a time will be merged, starting from Oracle ones
 - More scripts may have to be developed for the MySQL LFC merger
 - When a cloud is done, the LFC at the Tier-1 can be (should be) shut down