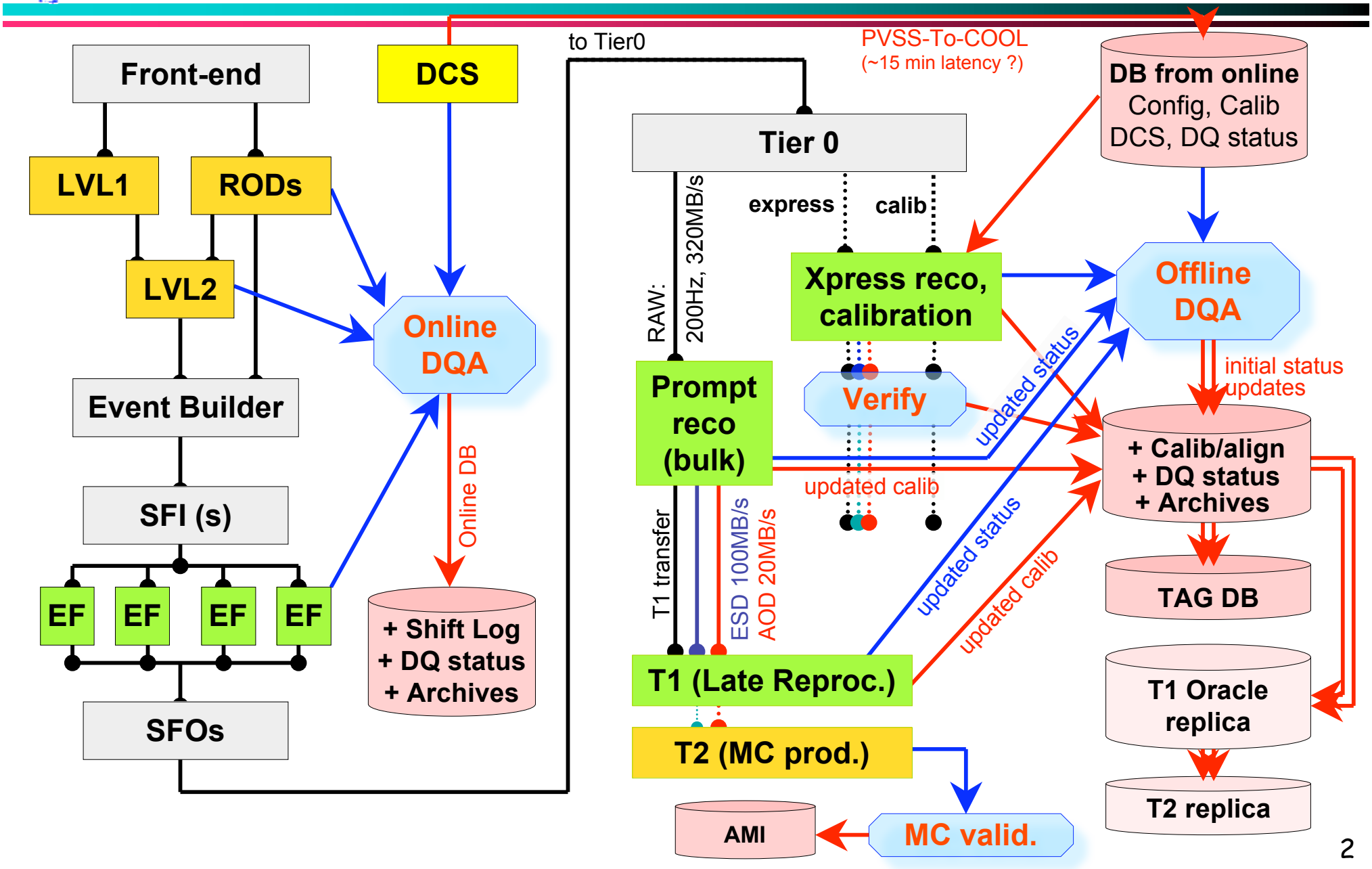
ATLAS Computing Model in a Nutshell

Dario Barberis

CERN & Genoa University/INFN

Data flow from DAQ to offline



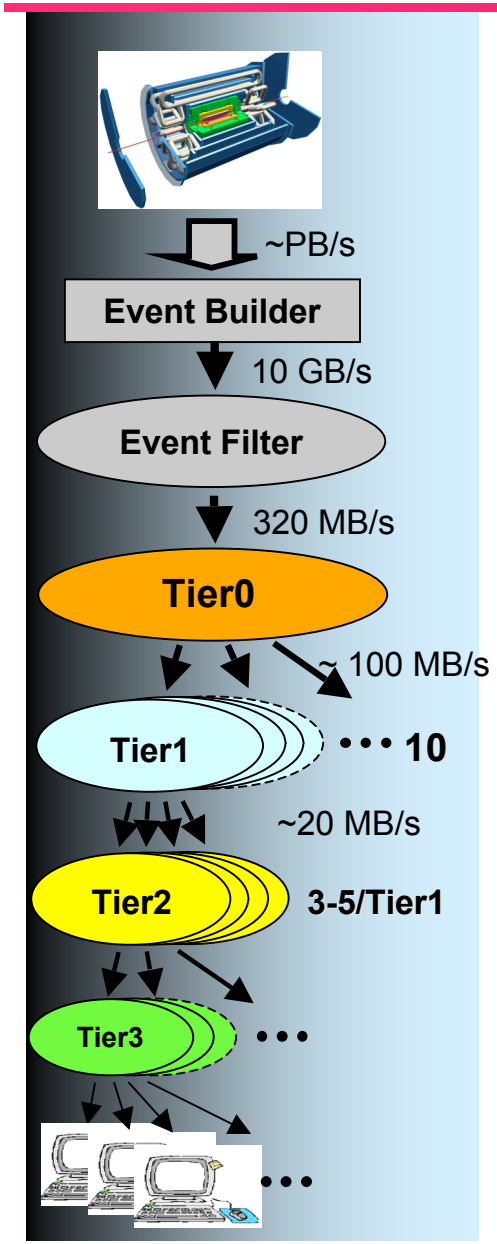
Event Data Model

- RAW:
  - "ByteStream" format, ~1.6 MB/event
- ESD (Event Summary Data):
  - Full output of reconstruction in object (POOL/ROOT) format:
    - Tracks (and their hits), Calo Clusters, Calo Cells, combined reconstruction objects etc.
  - Nominal size 1 MB/event initially, to decrease as the understanding of the detector improves
    - Compromise between "being able to do everything on the ESD" and "not enough disk space to store too large events"
- AOD (Analysis Object Data):
  - Summary of event reconstruction with "physics" (POOL/ROOT) objects:
    - electrons, muons, jets, etc.
  - Nominal size 100 kB/event (now 200 kB/event including MC truth)
- DPD (Derived Physics Data):
  - Skimmed/slimmed/thinned events + other useful "user" data derived from AODs and conditions data
  - Nominally 10 kB/event on average
    - Large variations depending on physics channels
- TAG:
  - Database (or ROOT files) used to quickly select events in AOD and/or ESD files

Computing Model: central operations

- Tier-0:
  - Copy RAW data to CERN Castor Mass Storage System tape for archival
  - Copy RAW data to Tier-1s for storage and subsequent reprocessing
  - Run first-pass calibration/alignment (within 24 hrs)
  - Run first-pass reconstruction (within 48 hrs)
  - Distribute reconstruction output (ESDs, AODs & TAGS) to Tier-1s
- Tier-1s:
  - Store and take care of a fraction of RAW data (forever)
  - Run "slow" calibration/alignment procedures
  - Rerun reconstruction with better calib/align and/or algorithms
  - Distribute reconstruction output to Tier-2s
  - Keep current versions of ESDs and AODs on disk for analysis
- Tier-2s:
  - Run simulation (and calibration/alignment when appropriate)
  - Keep current versions of AODs on disk for analysis

Data replication and distribution



In order to provide a reasonable level of data access for analysis, it is necessary to replicate the ESD, AOD and TAGs to Tier-1s and Tier-2s.

RAW:

- Original data at Tier-0
- Complete replica distributed among all Tier-1
  - Randomized datasets to make reprocessing more efficient

ESD:

- ESDs produced by primary reconstruction reside at Tier-0 and are exported to 2 Tier-1s
- Subsequent versions of ESDs, produced at Tier-1s (each one processing its own RAW), are stored locally and replicated to another Tier-1, to have globally 2 copies on disk

AOD:

- Completely replicated at each Tier-1
- Partially replicated to Tier-2s (~1/3 – 1/4 in each Tier-2) so as to have at least a complete set in the Tier-2s associated to each Tier-1
  - Every Tier-2 specifies which datasets are most interesting for their reference community; the rest are distributed according to capacity

TAG:

- TAG databases are replicated to all Tier-1s (Oracle and ROOT files)
- Partial replicas of the TAG will be distributed to Tier-2 as ROOT files
  - Each Tier-2 will have at least all ROOT files of the TAGs that correspond to the AODs stored there

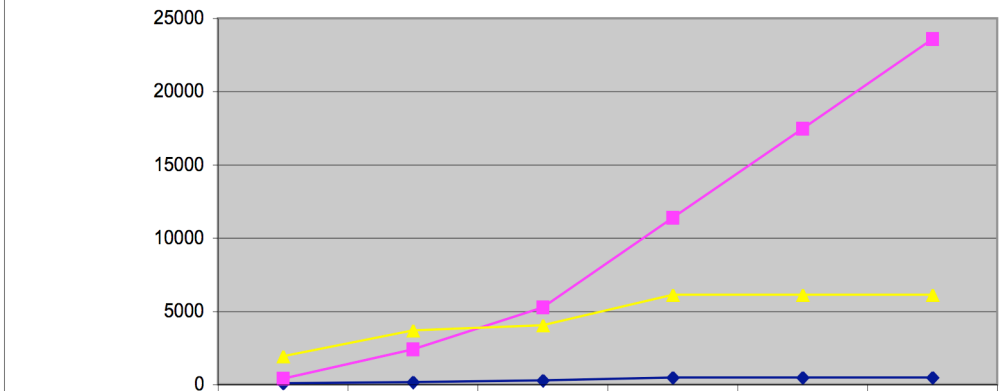
Samples of events of all types can be stored anywhere, compatibly with available disk capacity, for particular analysis studies or for software (algorithm) development.

(Simplified) Analysis Action Sequence

- Access the metadata catalogue (AMI) and find the datasets of interest
  - Based on physics trigger signatures, time range, detector status etc.
- (Optional) Use the TAG data (in Oracle DB or ROOT format) to build a list of interesting events to analyse further
- Use Distributed Analysis tools (e.g. Ganga) to submit jobs running on AOD data at Tier-2s (or on ESD at Tier-1s for larger-scale group-level analysis tasks)
  - Accessing only the selected datasets
  - (Optionally) taking the event list from the TAG selection as input
  - Producing DPD (Derived Physics Data) samples as output
    - Selected events in AOD format (skimming)
    - "Thinned/Slimmed" events in AOD format (selected event contents)
    - Any other simpler format (e.g. ntuples) for subsequent interactive analysis
  - Storing DPD on the Grid for group access or on local resources for interactive access
- DPD production can be also a group activity in case they can be used by several analyses
  - In this case DPDs must be stored on Tier-2s for global access
- Finish with interactive analysis (typically using ROOT) on the DPD files
  - Producing histograms and physics results

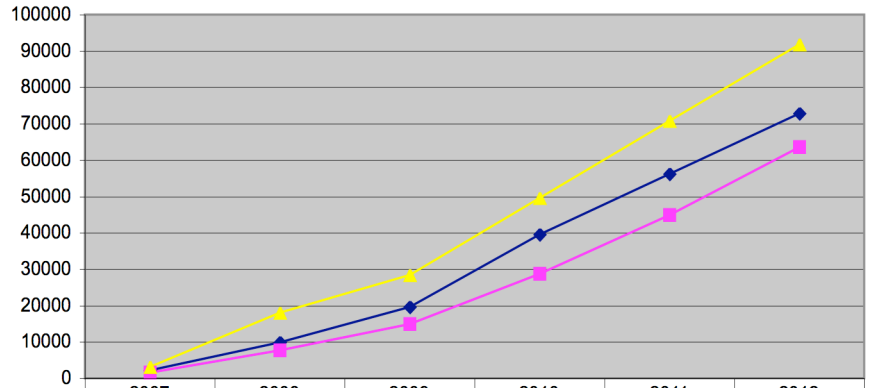
Resource Evolution

T0 Evolution



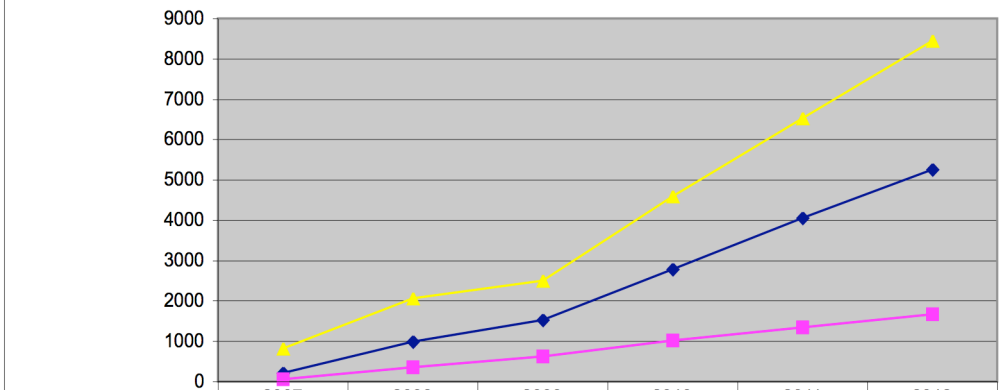
	2007	2008	2009	2010	2011	2012
Total Disk (TB)	75.14785714	152.4621429	277.3242857	472.3528571	472.3528571	472.3528571
Total Tape (TB)	381.3075	2381.711	5267.2345	11371.158	17475.0815	23579.005
Total CPU (kSI2k)	1910	3705	4058	6105	6105	6105

T1 Evolution



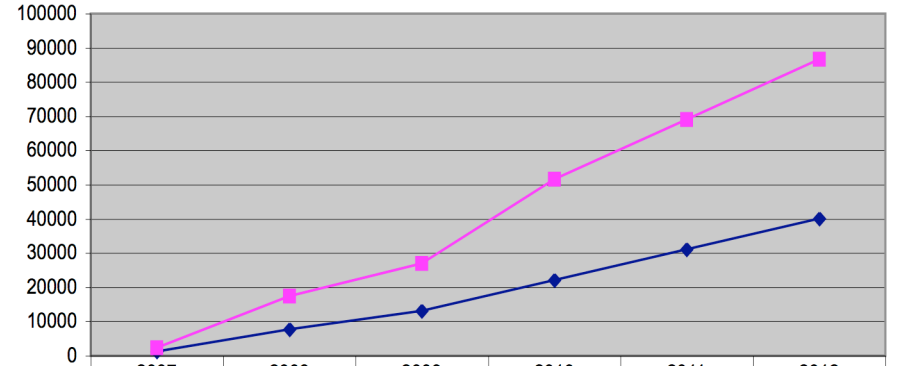
	2007	2008	2009	2010	2011	2012
Total Disk (TB)	2157.0332	9938.696929	19686.41793	39487.79764	56190.82307	72893.8485
Total Tape (TB)	1543.186667	7693.996427	14949.57676	28698.0172	44929.67775	63644.55841
Total CPU (kSI2k)	3173.323529	18122.83529	28423.02353	49573.22353	70723.42353	91873.62353

CAF Evolution



	2007	2008	2009	2010	2011	2012
Total Disk (TB)	212.2436607	986.3915464	1529.026057	2777.498914	4047.976771	5255.197486
Total Tape (TB)	57.3206625	356.5720482	625.1016482	1017.151648	1342.801648	1668.451648
Total CPU (kSI2k)	821	2069	2502	4596	6523	8450

T2 Evolution



	2007	2008	2009	2010	2011	2012
Disk (TB)	1259.040486	7744.368955	13112.03563	22132.30423	31091.45139	40050.91999
CPU (kSI2k)	2336.108333	17494.50644	26972.75589	51544.63737	69128.41886	86712.20034

Resources for Analysis (2008)

CPU share	Tier-1s	Tier-2s	CAF
Simulation	20%	33%	-
Reprocessing	20%	-	10%
Analysis	60%	67%	90%

DISK share	Tier-1s	Tier-2s	CAF
RAW	10%	1%	25%
ESD	55%	35%	30%
AOD	25%	25%	20%
DPD	10%	39%	25%

Tier-2 Data on Disk

- ~35 Tier-2 sites of very, very different size contain:
- Some of ESD and RAW
  - In 2008: 30% of RAW and 150% of ESD in Tier-2 cloud
  - In 2009 and after: 10% of RAW and 30% of ESD in Tier-2 cloud
  - This will largely be 'pre-placed' in early running
  - Recall of small samples through the group production at T1
- Additional access to ESD and RAW in CAF
  - 1/18 RAW and 10% ESD
- 10 copies of full AOD on disk
- A full set of official group DPD *(in production area)*
- Lots of small group DPD *(in production area)*
- User data
  - Access is 'on demand'