



ATLAS Resource Requests 2009-2010

Dario Barberis

CERN & Genoa University/INFN



Model recap (1)

- For the purpose of these resource calculations, we assume the following data-taking profile:
 - A full cosmic-ray run in Q3-2009, which will produce approximately the same data volume as the cosmic-ray run between August and November 2008 (1.2 PB of raw data, ~500 million events);
 - 10^6 live seconds in both of Q4-2009 and Q1-2010;
 - 2×10^6 live seconds in both of Q2-2010 and Q3-2010;
- In 6×10^6 live seconds, at the nominal trigger rate of 200 Hz, we can collect 1.2×10^9 events, corresponding to almost 2 PB of raw data.
- We also plan to produce 900M fully-simulated (Geant4) events and 2200M Atlfast-II events. For fully-simulated events, we save HITS to tape at the production site (or its associated Tier-1) and keep on disk only a fraction (currently 40%) of ESDs. For both fully-simulated and Atlfast-II events, we distribute AODs and DPDs similarly to real events.



Model recap (2)

- RAW data are stored on tape at CERN and at Tier-1s (one copy distributed between all Tier-1s). Raw data are stored on disk temporarily at Tier-1s, to allow access for fast problem detection, and are flushed out after a few weeks. Samples of raw data can be moved to Tier-2s for longer investigations.
- ESDs (Event Summary Data) are stored on tape at the location of production (CERN for the first-pass processing, Tier-1s for reprocessing) and on disk in 2 Tier-1s. The disk copies are replaced at each reprocessing cycle. Samples of ESDs can be copied to Tier-2s for in-depth investigations.
- AODs (Analysis Object Data) are stored on tape at the location of production (CERN for the first-pass processing, Tier-1s for reprocessing) and on disk in each Tier-1 cloud (the Tier-1 and its associated Tier-2s). The exact storage pattern depends on the ratio between Tier-1 and Tier-2 disk capacity in each cloud. The number of disk copies of each AOD dataset will be reduced with time, depending on their physics interest and the frequency of access.
- DPDs (Derived Physics Data) are small subsets of ESDs or AODs that can be used to study detector performance or for a specific analysis channel. The sum of the sizes of all DPDs is constrained to be at most equivalent to the size of AODs. DPDs are distributed in the same way as AODs.



Note on resource calculations

- The pre-existing resource usage must not be forgotten
- Assume MC08 remains 'active' in 2009 and 2010 until replaced by equivalent new MC data
- Assume 2008 cosmics needed until late 2009
- Additional space needed for new cosmic data taking, disk re-used later
- This is the first time more realistic event sizes/processing-simulation times have been applied to a Tier 1/Tier 2 request
 - Previously fixed to the 2006 values
- Would not intend to show any numbers after 2010 at this point

Event size (MB)	Real	Simulated
RAW	1.6	2
ESD	1	1
AOD	0.2	0.25
DPD	0.2	0.25

CPU (HS06-s)	
Reconstruction	60
Full simulation	8000
Atlfast-II	400



Resource table

		OLD				NEW							
	kHS06 / PB / PB	<i>Pledge 2008 (C-RRB Fall 2007)</i>	<i>Pledge 2009 (C-RRB Fall 2008)</i>	<i>Req. 2009 (C-RRB Fall 2008)</i>	<i>Req. 2010 (C-RRB Fall 2008)</i>	Requirement 2009-Q2	Requirement 2009-Q3	Requirement 2009-Q4	Requirement 2010-Q1	Requirement 2010-Q2	Requirement 2010-Q3	Requirement 2010-Q4	Requirement 2011-Q1
Tier-0	CPU	14.1	16.2	30.4	30.4	12.7	18	30.4	30.4	30.4	30.4	30.4	30.4
	Disk	0.15	0.26	0.65	0.7	0.41	0.7	0.7	0.7	0.7	0.7	0.7	0.7
	Tape	2.8	6.2	9.7	14.6	3.8	5.1	6.4	7.8	8.9	9.9	9.9	9.9
CAF	CPU	8.6	10.2	23.2	37.4	7.6	14.2	20.8	27.4	37.4	37.4	37.4	37.4
	Disk	1.1	1.8	3.3	4.5	0.76	2	2.5	2.9	3.4	3.9	4.2	4.4
Tier-1	CPU	76.8	120.9	113.7	234.5	88.8	90	90	90	177	177	227	227
	Disk	10.9	19.9	20.9	41.3	7	19.6	21.7	23.8	27.5	32	34.3	36.7
	Tape	7.7	14.7	15.8	22.7	8.1	9.3	10.3	11.3	12.5	13.7	14.2	14.8
Tier-2	CPU	70.1	107.7	108	242.5	108	108	108	108	240	240	240	240
	Disk	6.3	10.7	13.3	24.8	6.8	9.2	11.4	13.3	17.5	21.5	23.1	24.8

1. All new requirements are lower than (or at most equal to) the old requirements throughout 2009 and 2010. The average saving is of order 20%.
2. The only problem may arise at CERN, as their pledges at the November 2008 C-RRB meeting were substantially lower than our requirements for 2009 and 2010.
3. For the sake of clarity, we have now put together the Tier-0 and CAF tape requirements, as there is no difference in the way data are stored on tape at CERN.
4. These resource requirements can only be as accurate as the input assumptions. We estimate our own intrinsic uncertainties to ~10%.