ATLAS Resource Requirements: 2013, 2014 and 2015

Hans von der Schmitt
Borut Kersevan

Comp. Resource Usage in 2013-2015

- Brief outline of our resource planning guidelines for 2013-2015:
  - **In 2013:**
    - there will be one full reprocessing of 2010-2012 data and MC to further improve the quality of our reconstruction and simulation.
    - (More) new MC for analysis will be produced.
    - Very active group/user analysis.
  - **In 2014:**
    - Largish MC samples for high energy running will be produced and related physics group/user analysis.
    - The final full reprocessing of 2010-2012 data and MC, foreseen to use the evolved event formatting/data model/data distribution prepared for 2015 high-energy data taking.
  - **In 2015:**
    - Processing and reprocessing of new high energy data.
    - Related production of MC samples matching the data.
    - Increased group/user activity.

Resource Documents Update

- We have sent the mid-term Resource Usage Report, as well as our revised Resources Requests for 2012-2014 to the Computing Resources Scrutiny Group for the October Resources Review Board.
- We have revised the requirements for 2012-2014, in view of longer LHC running time in 2012.
- The assumptions in the LHC schedule and performance are:
  - The LHC running time increases from 21 to 30 weeks.
  - The LHC efficiency is changed from 30% to about 40%, due to the longer period of stable beams.
  - The extension of LHC running into 2013 leads to a delay in the resumption of LHC physics running until RRB year 2015. Consequently, it is now assumed that there will be no physics data taking in the RRB year 2014 (i.e. including the first 3 months of 2015).
- **In order to keep our disk space needs under control we have had to 'tune' down our disk usage (replication policies) and we are making plans to further optimize our Computing Model for 2015.**

Data Placement

- A table summarizing the updated replication policy:

Tier-1 disk policy (sum Tier-1s)	2012	2013	2014
Sim RDO disk copies	0	0	0
Sim ESD disk copies (current)	0	0	0
Sim ESD disk copies (previous)	0	0	0
Sim AOD disk copies (current)	2	2	2
Sim AOD disk copies (previous)	0	0	0
Real RAW disk copies	1	1	1
Real ESD disk copies (current)	~~0.2~~ 0.13	~~0.2~~ 0.05	~~0.2~~ 0.05
Real ESD disk copies (previous)	0	0	0
Real AOD disk copies (current)	2	2	2
Real AOD disk copies (previous)	0	0	0
Real DESD disk copies	~~2~~ 1	~~2~~ 1	~~2~~ 1
Tier-1 processing			
Number of reprocessing/year	~~1.5~~ 1	1	1
Tier-2 disk policy (sum of T2's)			
Sim RDO disk copies	0.05	0.05	0.05
Sim AOD disk copies (current)	2	2	2
Sim AOD disk copies (previous)	2	2	2
Sim ESD disk copies (current)	0.2	0.1	0.1
Sim ESD disk copies (previous)	0.2	0.1	0.1
Real RAW disk copies	0	0	0
Real AOD disk copies (current)	2	2	2
Real AOD disk copies (previous)	2	2	2
Real DESD disk copies	~~2~~ 1	~~2~~ 1	~~2~~ 1
Real DESD disk copies (previous)	~~2~~ 1	~~2~~ 1	~~2~~ 1

Extra copies!

Popular data (group production, AODs..) replicated dynamically by PD2P

The crossed out values describe the changes to our model in view of more data in 2012
'Current' describes 2012 data and MC
'Previous' denotes older processings as well as last year's data and MC

Comp. Model Changes for 2012-2014

- **The increase in the volume of the data to be recorded in 2012 motivated the following changes in the resource model in order to keep the increase in required disk space at a modest level:**
  - The lifetime of the current data ESD rolling buffer is reduced from 6 to 4 weeks, thus the ESD volume on disk is reduced proportionally.
  - The number of data reprocessing campaigns in RRB year 2012 is changed from 1.5 to one campaign.
  - The assumed volume of data ESDs on disk after reprocessings in years 2013 and 2014 is reduced from 20% to 5% of the total data volume, which would contain a representative sample and data sets of special interest.
  - The number of dESD copies of data and simulation is reduced from two to one at both Tier-1 and Tier-2. In view of the need for optimization, this change in replication policy and reduction in accessibility of dESDs is estimated to be the reduction that will have the least impact on ATLAS physics performance.
  - The increase in the needed group space in 2014 is reduced from 4 PB to 2 PB because of no data being taken in 2014, while group activity on 7 TeV and 8 TeV data should consolidate by then.
  - The simulation of 13-TeV collisions in 2014 in preparation for 2015 running will now be produced with full statistics in one campaign rather than two. The second campaign, which will correct the simulation parameters (beam spot, pileup calibration, etc.) based upon physics data, will be performed when physics running resumes in 2015.
  - All CERN computing resources, including Tier-0, will be available throughout 2014 for simulation. They were previously assumed to be available only half of the year due to data taking at the end of the year.

Current Estimates in a Compact Form

LHC and ATLAS parameters for 2012		Actual p-p (Mar–Aug 2012)	Revised model 2012 p-p	Model 2012 HI	Model 2013 p-p	Revised model 2014 p-p
Trigger Rate	events/sec	350 (prompt) + 150 (delayed)	400 (prompt) + 200 (delayed)	200	0	~~1000~~ 0
Live Time	Msec total	3.8	~~3.8~~ 7.3	0.7	0	~~1.6~~ 0
Average pileup μ	events	20	25			
Real data (prompt only)	Bevents total	1.3	~~1.5~~ 2.9	0.14	0	~~1.6~~ 0
Full Simulation	Bevents total	2.4	1.9	0.01	4	2.5
Fast Simulation	Bevents total	1.4	2	0	4.6	4
Real RAW	MB/event	0.79	0.8	5	0.8	0.8
Real ESD	MB/event	2.4	2.5	3	2.5	2.5
Real AOD	MB/event	0.24	0.35	1	0.35	0.35
Simulated HITS	MB/event	0.9	1	5	1	1.6
Simulated ESD	MB/event	3.3	3.5	3	3.5	3.5
Simulated AOD	MB/event	0.4	0.5	1	0.5	0.5
Full Simulation	HS06sec/event	3100	3300	48000	3300	5400
Fast Simulation	HS06sec/event	260	310	-	310	500
Real Reconstruction	HS06sec/event	210	230	480	230	230
Simulation Reconstruction	HS06sec/event	770	830	1200	830	830
Group analysis	HS06sec/event		20	20	20	20
User analysis	HS06sec/event		0.4	0.2	0.4	0.4

Current Usage and Plans

- Our predictions turned out to be accurate enough, modulo the LHC run extension.
  - Due to the efforts of sites we are getting much more CPU than approved by CRSG.
  - The disk space will be tight in 2012, we have asked for early deployment of 2013 pledges and we will get them wherever possible.

Current (2012) usage and plans

	Location	Predicted	Actual	Revised model
CPU [kHS06]	CERN	111	111	111
	Tier-1	289	436	295
	Tier-2	319	612	319
Disk [PB]	CERN	9	6	11
	Tier-1	27	27	29
	Tier-2	47	30	48
Tape [PB]	CERN	18	20 (incl. ESD)	21
	Tier-1	32	24	35

2012 data types beakdown, to CERN tape

Total 7 PB to date
ESD 4.3 PB
RAW 1.7 PB
AOD 0.5 PB
Other 0.5 PB

Longer view up to 2014:
Numbers in brackets are the current pledged numbers.

CPU [kHS06]	2012	2013	2014
CERN	111 (111)	111 (111)	111 ⇒ 281
Tier-1	289 295 (290)	303 319 (304)	405 373
Tier-2	319 (336)	355 (365)	444 408
Disk [PB]			
CERN	9 11 (9)	9 11 (10)	11
Tier-1	27 29 (31)	33 35 (33)	40 36
Tier-2	47 48 (45)	51 53 (51)	60 56
Tape [PB]			
CERN	18 21 (18)	23 27 (18)	26 31
Tier-1	32 35 (39)	40 43 (43)	53 53

CPU grouped by activities:

ATLAS CPU requirements [kHS06]	2012	2013	2014
Re-processing	23 29	29 45	29 45
Simulation production	271	521	511
Simulation reconstruction	164	181	240 212
Group (+user)	150	170	190
Total	607 613	901 917	971 958
Total without CERN resources (or HLT)	607 613	658 674	849 781

First Look at 2015-2016

- We made a quick projection so far on what resources we could need in 2015 RRB year (until April 2016), keeping the current Computing Model.
  - The event sizes & CPU times are set to be equal to 2012 values or taken from ATLAS upgrade MC samples @ 14 TeV.
  - 21 weeks and 30% LHC efficiency assumed.
  - **Our goal is data taking @ 1kHz trigger rate - big advantage for Physics, but we need to work on our software and computing to be able to achieve it within reasonable resource constraints.**
  - relatively modest amount of simulation w.r.t. data taken assumed.

LHC and data taking parameters		2012 pp	2013 pp	2014 pp	2015 pp
Rate [Hz]	Hz	400	0	0	1000
Time [sec]	MSeconds	7.3	0	0	3.8
Real data	B Events	2.9	0	0	3.8
Full Simulation	B Events	1.9	4	2.5	2
Fast Simulation	B Events	2	4.6	4	5
Event sizes					
Real RAW	MB	0.8	0.8	0.8	0.8
Real ESD	MB	2.5	2.5	2.5	2.5
Real AOD	MB	0.35	0.35	0.35	0.35
Sim HITS	MB	1	1	1.6	1.6
Sim ESD	MB	3.5	3.5	3.5	3.5
Sim AOD	MB	0.5	0.5	0.5	0.5
CPU times per event					
Full sim	HS06 sec	3300	3300	5400	5400
Fast sim	HS06 sec	310	310	500	500
Real recon	HS06 sec	230	230	230	230
Sim recon	HS06 sec	830	830	830	830
Group analysis	HS06 sec	20	20	20	20
User analysis	HS06 sec	0.4	0.4	0.4	0.4

CPU [kHS06]	2011	2012	2013	2014	2015
CERN	74	111	111	111	228
Tier-1	202	295	319	373	502
Tier-2	275	319	355	408	540
Disk [PB]					
CERN	7	11	11	11	14
Tier-1	22	29	35	36	51
Tier-2	35	48	53	56	69
Tape [PB]					
CERN	14	21	27	31	45
Tier-1	16	35	43	53	78

- **The CPU demand most drastic at T0 due to the trigger rate.**
- **The increases in CPU and disk at T1, T2 'not too bad' but still rather high, so:**
  - **Quite some work will be needed to match (and improve) on the event sizes, reco times etc..**
  - **We need to start investing effort and discuss event sizes, EDM, CPU times and Computing model changes to improve the predictions.**

Plans for LS1

- **A selection of topics:**

- **Improvements/simplification of the Event Data Model (EDM)**
  - Some of the C++ classes designed at the time are not well suited for modern CPU architectures
  - Data components scattered over memory: bad for memory cache usage
  - And bad for vector-unit usage: compilers can't auto-vectorize
  - ⇒ Design simplified EDM, for usage within algorithms; start with tracking algorithms

- **Concurrency: usage of many-core designs with small memory footprint**
  - Multi-core used since a while with multi-process software (not multi-threaded); memory footprint reduced to ~1/3, still too big for many-core
  - Multi-threaded designs being tackled algorithm by algorithm
  - Need some support in the software framework: underway
- **benefitting from collaboration with IT/OpenLab, PH-SFT and CMS in these and related fields**

- **New Distributed Data Management system (DDM) being implemented**
- **WAN data access and data caching (file level, event level)**
- **Applications-driven usage of networks**
- **New MC production system**

- **Simulation: Integrated Simulation Framework (ISF)**
  - Selection of full/fast/parametrized mode per subevent
  - Concurrency within event
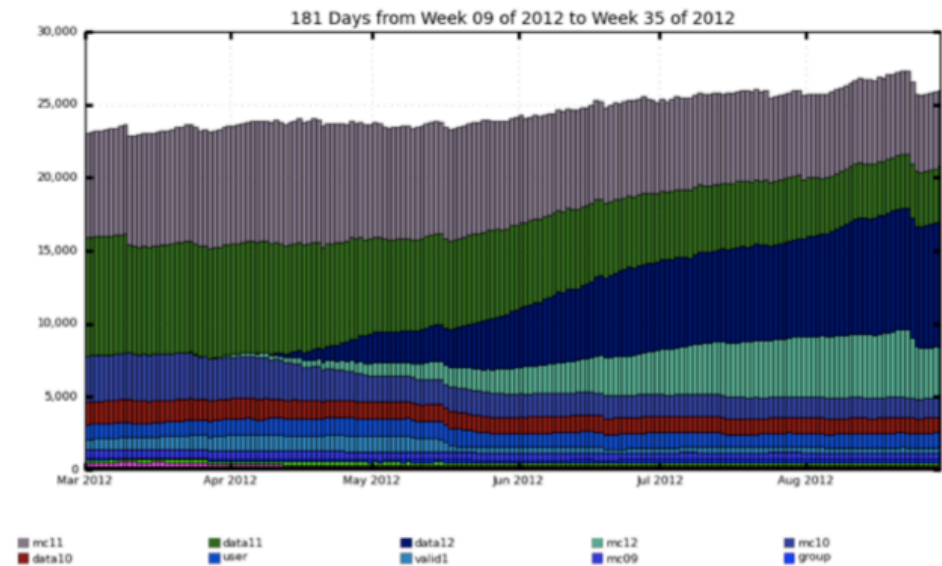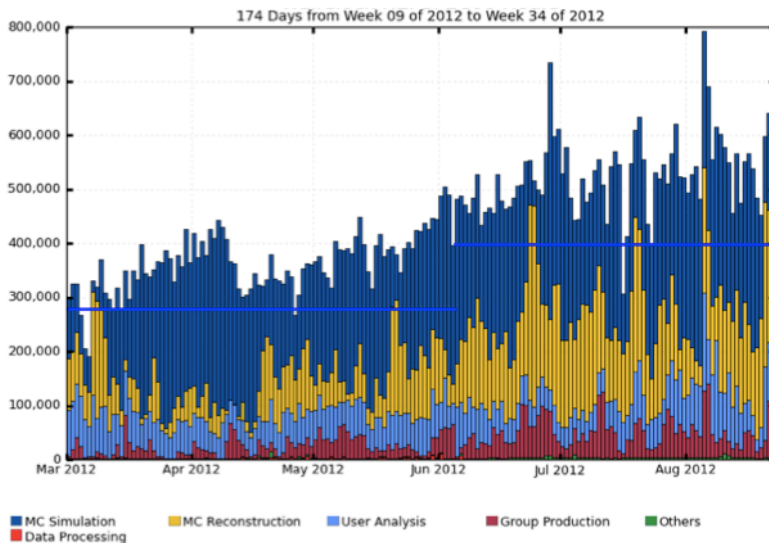 - Plus benefit from reconstruction speedup

Current Tier-1s Status

- ATLAS Grid ctivities at Tier-1s:
    - As stated, we have more CPU as well as (but not that much) disk, we need to be careful!
    - Performing very well!
    - There are two upcoming data reprocessings:
        - 8 TeV pp data,
        - Heavy ion data from 2011.

Tier-1 Disk [PB]	Predicted	Actual	Revised model
RAW data	1.9	1.4	3.5
Real ESD+AOD+DPD	3.8	4.5	4.5
Simulated data	8	7.1	8
Calibration and alignment outputs	0.4	0.4	0.4
Group data	6	3.2	6
User data (scratch)	2	2	2
Cosmic ray data	0.2	0.2	0.2
Processing and I/O buffers	4.3		4.3
Dynamic Data Buffers	Included in WLCG scaling factor	8	Included in WLCG scaling factor
Total	27	26.8	29

Tier-1 CPU [kHS06]	Predicted	Actual	Revised model
Reprocessing	23	1	29
Simulation Production	71	248	71
Simulation Reconstruction	134	109	134
Group+User activities	60	78	60
Total	289	436	295

174 Days from Week 09 of 2012 to Week 34 of 2012

MC Simulation · MC Reconstruction · User Analysis · Group Production · Others · Data Processing

181 Days from Week 09 of 2012 to Week 35 of 2012

mc11 · data11 · data12 · mc12 · mc10 · data10 · user · valid1 · mc09 · group
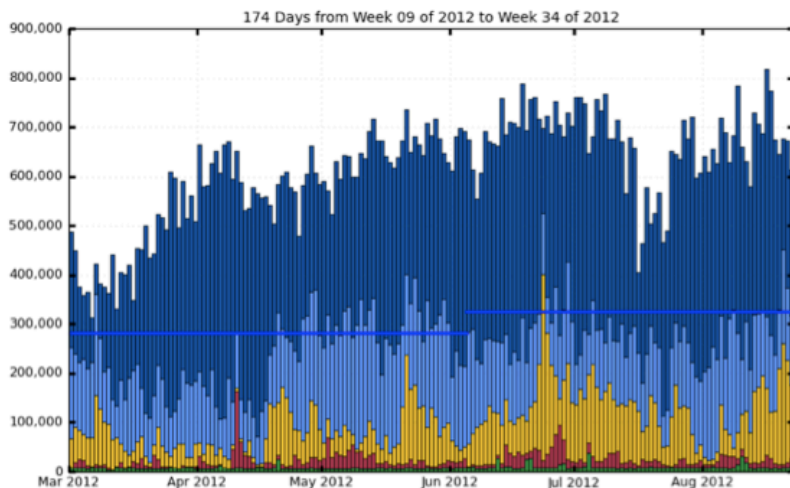
Current Tier-2s Status

- Tier-2s:
  - Successful running of digitization and reconstruction, implemented beginning of the year.
  - Now also taking a big part in group (D3PD/NTUP) production.
  - Repositories for all 'active' data and MC.
  - About 30% of disk space used for 'dynamic data buffers':
   - Popular data (group production, AODs..) replicated dynamically by PD2P.

Tier-2 CPU [kHS06]	Predicted	Actual (average Mar-Aug)	Revised model
Simulation Production	199	395	199
Simulation Reconstruction	29	79	29
Group + User activities	90	138	90
Total	*319*	*612*	*319*

Tier-2 Disk [PB]	Predicted	Actual	Revised model
Real AOD+DPD	11	5.4	13
Simulated data	22	11	22
Calibration and alignment output	0.3	0.3	0.3
Group data	10	4.8	10
User data	2	1.4	2
Processing and I/O buffers	1		1
Dynamic Data Buffers	Included in WLCG scaling factor	7.1	Included in WLCG scaling factor
Total	*47*	*30*	*48*

174 Days from Week 09 of 2012 to Week 34 of 2012

Legend: MC Simulation, User Analysis, MC Reconstruction, Group Production, Others, Data Processing
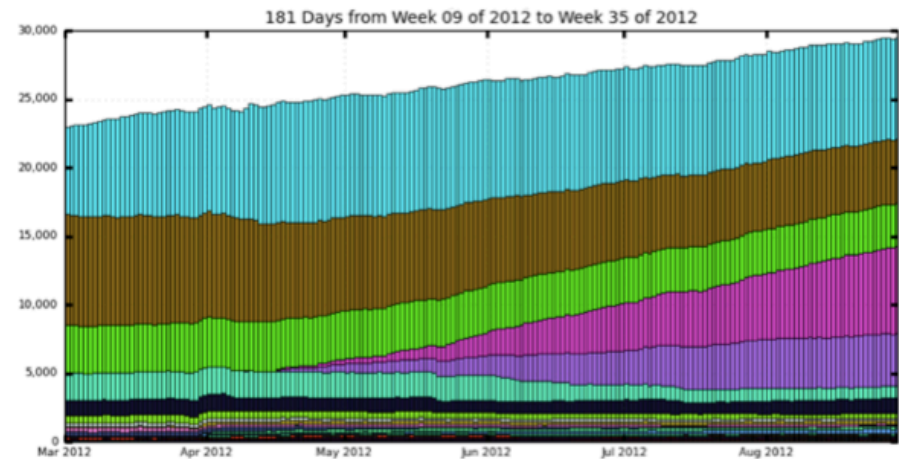
181 Days from Week 09 of 2012 to Week 35 of 2012

Legend: mc11_7tev, data11_7tev, mc10_7tev, data12_8tev, mc12_8tev, data10_7tev, user, data10_hi, data11_hi, mc10_2tev