ATLAS Computing Resource Requirements for Data-taking in 2013-2015

3rd of March 2013

Introduction

This document presents the ATLAS Computing Resource Requirements for 2013-2015. Requirements are estimated using an updated model incorporating adjustments based upon operational experience from Run 1 (2010-2012) as well as the LHC schedule and expected operational parameters in 2015. The principal updates result from:

- Estimates of LHC beam parameters and beam energy, which together will lead to substantially higher luminosity and increased event pileup in 2015.
- Re-evaluation of the quantity of simulated data needed by ATLAS analyses, based upon 2010-2012 experience.
- Re-evaluation of the group and user analysis needs, extrapolated from the 2012 experience and incorporating evolution of the ATLAS Computing model planned for the 2013/2014 long shutdown (LS1).

The preliminary ATLAS computing resource requirements, which were presented to the April and October 2012 RRBs for the RRB year 2013, already reflected the change in the LHC schedule to continue and extend the operations in 2012 and to begin the first long shutdown in 2013.

The recently presented LHC plans for the running conditions in 2015 at the higher centre-of-mass energy of 13 TeV and for even more intense colliding bunches will increase the CPU consumption of event processing and the event sizes of all recorded formats. In addition, due to the increased beam energy, revised detector configuration and physics modelling improvements in the Monte-Carlo generators, the bulk of our simulated data samples must be regenerated at the new centre-of-mass energy. For these reasons, ATLAS computing resource requirements for 2013-2015 naturally grow beyond the the needs for 2012 despite improvements and optimization obtained after intensive work on both the ATLAS software performance and data handling. However, further improvements and adjustments are being made to the ATLAS resource model, as described in this document, coupled with ambitious ATLAS software and computing development plans during the LS1, in order to constrain requirements to approximately the level currently pledged for 2013 and 2014 and to limit the increase in the ATLAS requirements for 2015.

2012 was an exceptionally successful year for the LHC. It achieved instantaneous and integrated luminosity values that far exceeded expectations, and it provided the experiments with excellent data sets. 2012 was also an excellent year for ATLAS itself. By continual improvements to the performance of its software, the availability of beyond-pledge resources and by taking advantage of the flexibility of its distributed computing, ATLAS managed the unexpected levels of pileup and larger physics data sets, coped with higher trigger rates, and generated more simulated data sets than planned in order to support the broad range of physics analyses, thereby making possible a productive and timely physics programme. The successful experience of 2012 has inspired and informed several changes to the planned utilization of computing resources, as described below, in order to address the new challenges of operation in 2013 and beyond.

During 2012 operations, ATLAS profited from data-taking at rates higher than its historically planned average output trigger rate of 200 Hz. The increased rate allowed more inclusive triggers to be run for longer as LHC luminosity ramped up, affording lower trigger thresholds, more open trigger selections for data-driven background estimates, enhanced samples for trigger studies, *etc.*, all of which, in turn, enhanced the breadth, quality, quantity, and timeliness of ATLAS physics output. Successful running at trigger rates of 400 Hz had been demonstrated already in 2011, and consequently the computing resource requirements for 2012 were based upon an average 400 Hz trigger output rate for prompt reconstruction. In addition, dedicated streams involving triggers tuned to special physics interests (such as b-quark physics) with a combined rate up to 150 Hz were recorded as 'delayed' streams in 2012, meaning that the RAW data was stored on tape in Tier-0 and one Tier-1 to be reconstructed only in 2013.

In 2015, LHC will run at a centre-of-mass energy of ~13 TeV. At these centre-of-mass energies the cross sections for interesting physics processes will be at least a factor of two higher than at 8 TeV. For example, Higgs cross sections increase by a factor of 2.4 or more. In order to profit from the increased cross sections for the interesting processes, it is proposed to increase the output rate of the ATLAS High Level Trigger (HLT) to ~1 kHz. The increased cross sections will also stress our system at Level 1, but it is expected that improvements will allow the Level 1 trigger rate to be increased from its Run 1 value of 65 kHz rate to 100 kHz while somewhat improving the selectivity of Level 1. With these improvements, HLT selections tuned to have an output rate of 1 kHz would have acceptance and signal-to-background similar to that of our current selections at 8 TeV.

In 2015 ATLAS will again be confronted with much higher event pileup than previously planned. The increase in centre-of-mass energy to ~13 TeV and plans to squeeze the beams as tightly as possible will give rise to increased luminosity. Operation with 50-ns bunch spacing will then lead to pileup beyond any design levels and could achieve values of peak number of interactions per beam crossing (μ) of μ=80 or equivalently an average value of $\mu\approx 50$ (more precisely μ=56). ATLAS has in 2012 presented a complete set of arguments that the 50-ns mode of LHC operations at high pileup would cause several issues, not least a very substantial increase in the computing resources required, thus the assumption in this document is that there will be no extended physics data taking at 50-ns and high pileup values and that LHC will quickly move to 25-ns bunch spacing giving more moderate values of pileup. The working assumption in this document is thus that the LHC might be operating for two months of data taking at 50-ns and average pileup of $\mu\approx 50$ before transitioning to 25-ns bunch spacing, gradually increasing its luminosity. Consequently, for the remainder of 2015, LHC is assumed to achieve stable operation at the average $\mu\approx 25$ for the luminosity of 10^{34} at 25-ns bunch spacing. For the purpose of the 2015 resource estimation, the LHC is assumed to deliver 5 million live seconds of physics collisions.

Higher pileup has substantial repercussions for the mean reconstruction time per event and for the event size in all formats. Current best estimates are that both the reconstruction time per event and event size will, compared to the 2012 values, increase at 13 TeV by a factor of two for μ=25 with 25-ns bunch spacing and factor more than three for μ=50 with 50-ns bunch spacing.

To address these challenges, ATLAS is establishing an ambitious development programme during the LS1 in both ATLAS software and ATLAS distributed computing, which is being incorporated into an updated ATLAS computing model for 2015 and beyond. At present the related ATLAS activities are in the prototyping and testing stages; consequently, the updated computing model is not yet completely specified. The target for the updated model is October 2013, when it is planned to be included in a document for the LHCC together with the other LHC experiments. Nonetheless, the principal aims and planned activities are specified below.

Planned software improvements during LS1:

The target for data reconstruction is to achieve a speedup factor of two to three compared with the current software performance, i.e. to achieve a reconstruction time in high energy running with 25-ns bunch spacing at μ=25 close to the one in the 2012 data taking. The target speed is at present evaluated at ~180 HS06sec/event for μ=25 at 25-ns bunch spacing and ~230 HS06sec/event at μ=50 at 50-ns bunch spacing. The envisaged improvements are to be achieved by targeted algorithmic improvements as well as introducing (auto) vectorisation into as many areas of ATLAS code as possible, thereby gaining factors of speedup by utilizing the current CPU architectures in an optimal way, and introducing modern, more efficient, linear algebra libraries into the ATLAS code base. In order to reduce the memory requirement, ATLAS is introducing the multi-process event-level parallel version of the ATLAS software framework, which takes advantage of the shared memory by utilizing the copy-on-write feature of the Linux kernel. As the next step, ATLAS software will be moved to a new concurrent framework with both event-level and algorithm-level parallelism using modern threading techniques and libraries such as the Intel TBB. The effort is planned to be shared together with some other LHC experiments and should also result in ATLAS software with a low enough memory footprint to be ready for the future many-core architectures, as well as potential use at High Performance Computing facilities and IAAS resources. The work on the new framework started at the end of 2012 but it is at present still difficult to predict with confidence if we will succeed in commissioning it by 2015, or if it will be introduced somewhat later.

Similar plans are also in place for the full (Geant4) and fast (Geant4 combined with parameterized response) simulation, pileup simulation and digitization and simulation reconstruction. Speed improvements in both the full and fast simulation of on average 10% per year are expected between 2013-2015 and about 20% per year for simulation reconstruction over the same period. It has to be stressed that these are average improvements, which will come in discrete steps in the ATLAS software development and deployment cycle, and will depend also on the schedule for deploying this new software for data and simulation reprocessing, which will be driven above all by the ATLAS physics needs.

As a parallel target, the AOD event sizes should be optimized to the values of 0.25 MB/event (the 2012 value) for μ=25 (and 0.4 MB/event for μ=50) pileup conditions in 2015, which compared to the current software performance requires again a factor two to three event size reduction. The objective is to achieve this by technological improvements, reducing the physics content only as the last resort to remain within an achievable computing resource budget. Analogous improvements are planned for the other data formats, i.e. RAW, ESD, simulated HITS, *etc*.

The campaign to increase simulation speed further during LS1 is already underway. Considerable effort is being invested in developing a new simulation framework (ISF), which will enable mixing fast (parameterized) and full simulation options, selectable at the particle level, and in speeding up the pileup simulation (digitization). Another promising avenue being investigated is the possibility to use zero-bias events from real data to accurately describe the actual pileup when overlaid with simulated physics processes. The viability and speed gains of this approach in pileup digitization are at present still being investigated.

An ultra-fast simulation chain, involving special fast simulation and reconstruction by developing 'truth-assisted' algorithms is also being envisaged. These algorithms could very much reduce the total ultra-fast simulation time, and would, due to the small CPU penalty for re-runs, enable one to avoid the storage of all intermediate steps (HITS, AOD) and thus to store samples directly in 'group analysis' format only. This possibility is as present only in the prototyping stage and is not expected to be available until 2015.

Planned distributed computing improvements during LS1:

Based on the Run 1 experience, ATLAS is developing a new distributed data management system called Rucio and an upgrade to the ATLAS production system based on PanDA, introducing a new job management system called JEDI and a new workflow and task management system DEfT. The new data management system will introduce a simplified and scaleable handling of data replicas, deletion and ownership, especially for the ATLAS group and user data sets, thereby optimising disk space usage. The upgraded production system will also be better suited to the evolving workflows of group and user analysis on the Grid and will speed up the job processing sequence, thereby minimizing the time needed for the full analysis workflow. Another substantial improvement will be the handling of 'transient' data sets, which are files in the production chain that are subsequently merged into bigger files, to optimize the file sizes for data transfer and storage, before being deleted. These 'transient' data sets currently occupy a substantial fraction of both the buffering disk space on the Grid and of the network transfer bandwidth. Thus, optimizing this workflow will allow more efficient use of these resources, e.g. for additional dynamically created replicas of popular data. Further gains in optimal data placement are projected to be achieved by the use of 'federated' disk resources, e.g. XRootD federations, which is at present being evaluated by an ATLAS distributed computing taskforce.

Planned workflow and data distribution improvements during LS1:

The workflows for data processing at the Tier-0, reprocessing at Tier-1s and simulation production at Tier-1s and Tier-2s have been optimized in the Run 1 period 2010-2012 and provided a very successful and robust service. Further improvements to the system will be introduced with the new production and data management systems described above.

A major focus of LS1 improvements will be the optimization of the (physics) group and user analysis workflow and model. In the Run 1 period the ATLAS Physics Analysis and Combined Performance working groups organized their analysis workflow into processing the data and simulation samples from AODs into dedicated (group and analysis specific) data, mostly in Root ntuple format. Initially, these 'group productions' were made by dedicated user jobs and the group data produced were stored in assigned group disk spaces on the Grid. In order to provide better control over the software used and to guarantee consistency and more effective data placement, this group production activity was centralized by introducing a new Group Production working group and related Group Production workflow on the Grid at the end of 2011.

The Group Production workflow, i.e. the new centralized production of analysis formats targeted for specific sets of analyses (e.g. Standard Model Measurements, the new Higgs-like boson property measurements, SUSY searches…) has been a success in terms of consolidating separate group workflows and of common/coherent software usage. It also incorporates many features of the reprocessing improvements, to some extent reducing the need for full data and simulation reprocessing, in principle giving a net gain in Grid CPU utilization, but at the cost of some increase in disk resource needs. We realized there were opportunities for further improvement, and we are well advanced in further revising the group activities on the Grid.

In order to optimize this workflow, a taskforce was set up by the ATLAS management to provide a recommendation on the improvements of the analysis workflow both in terms of software and operation, in order to reduce needed disk space and CPU consumption. Once the list of these recommendations is final, all the appropriate developments will be incorporated into the ATLAS Software and Computing plans for LS1. Certain aspects of the findings are already being incorporated, e.g. into the new distributed data management and production systems, as already described. Until the new analysis workflow and infrastructure is commissioned, which is anticipated for mid-2014, ATLAS will need to continue the present mode of Group Production operation, which motivates some upward revision in the disk and CPU requirements for 2013 and 2014.

Run conditions and performance parameters in 2013-2015

LHC operation plans and ATLAS trigger rates, amount of simulated data, event sizes, and processing times comprise the parameters that drive computing resource requirements, and therefore are input parameters to the resource model. Such parameters are shown in Table 1. Principal parameter values and changes are discussed below.

Table 1. Event sizes, samples, processing times and other input parameters for resource calculations.

LHC and data taking parameters		**2012 pp actual**	**2013 pp**	**2014 pp**	**2015 pp**
Rate [Hz]	Hz	400 + 150 (delayed)	0	0	1000
Time [sec]	MSeconds	6.6	0	0	5.0
Real data	B Events	3.0 + 0.9 (delayed)	0	0	5.0
Full Simulation	B Events	2.6 (8 TeV) + 0.8 (7 TeV)	0.4 (2010 MC) + 0.5 (2011 MC) +2.6 (2012 MC)	1.0 (2012 MC) + 2.0 (13 TeV MC)	1.0 (50 ns MC) + 2.0 (25 ns MC)
Fast Simulation	B Events	1.9 (8TeV) + 1.0 (7 TeV)	0.6 (2010 MC) +1.0 (2011 MC) +4.4 (2012 MC)	2.0 (2012 MC) + 2.0 (13 TeV MC)	2.0 (50 ns MC) + 3.0 (25 ns MC)
Event sizes					
Real RAW	MB	0.8	0.8	0.8	1.1 (50 ns) 1. (25 ns)
Real ESD	MB	2.4	1.1 (2010) 1.1 (2011) 2.4 (2012)	1.1 (2010) 1.1 (2011) 2.4 (2012)	2.5 (50 ns) 2.5 (25 ns)
Real AOD	MB	0.24	0.16 (2010) 0.16 (2011) 0.24 (2012)	0.16 (2010) 0.16 (2011) 0.24 (2012)	0.35 (50 ns) 0.25 (25 ns)
Sim HITS	MB	0.9	0.8 (2010 MC) 0.8 (2011 MC) 0.9 (2012 MC)	0.9 (2012 MC) 1.2 (13 TeV MC)	1.2 (50 ns MC) 1.2 (25 ns MC)
Sim ESD	MB	3.3	1.9 (2010 MC) 1.9 (2011 MC) 3.3 (2012 MC)	3.3 (2012 MC) 3.5 (13 TeV MC)	3.7 (50 ns MC) 3.5 (25 ns MC)
Sim AOD	MB	0.4	0.26 (2010 MC) 0.26 (2011 MC) 0.4 (2012 MC)	0.4 (2012 MC) 0.5 (13 TeV MC)	0.55 (50 ns MC) 0.5 (25 ns MC)
CPU times per event					
Full sim	HS06 sec	3100	2700 (2010 MC) 2700 (2011 MC) 2790 (2012 MC)	2511 (2012 MC) 3500 (13 TeV MC)	3500 (50 ns MC) 3500 (25 ns MC)
Fast sim	HS06 sec	260	250 (2010 MC) 250 (2011 MC) 234 (2012 MC)	211 (2012 MC) 250 (13 TeV MC)	250 (50 ns MC) 250 (25 ns MC)
Real recon	HS06 sec	190	108 (2010) 108 (2011) 190 (2012)	108 (2010) 108 (2011) 150 (2012)	230 (50 ns) 180 (25 ns)
Sim recon	HS06 sec	770	200 (2010 MC) 300 (2011 MC) 616 (2012 MC)	493 (2012 MC) 500 (13 TeV MC)	560 (50 ns MC) 500 (25 ns MC)

LHC and data-taking parameters, data reprocessing and simulated data samples

Run 1 Data reprocessing

In 2012, one major reprocessing campaign was conducted to reprocess all data at the end of the calendar year 2012, along with some smaller reprocessings earlier in 2012. The high quality of the reprocessed data and corresponding simulation, as well as short-term software improvements in terms of physics performance result in no envisaged need for full data and simulation reprocessing in 2013. In March 2013 the first processing of the delayed streams (0.9 billion events) will take place on the Grid. Smaller targeted reprocessings are foreseen in 2013 for a fraction of Run 1 data and corresponding simulation. A full data and simulation reprocessing is still foreseen for the year 2014, at minimum to guarantee appropriate data preservation as well as to enable consistent use of Run 1 data together with the Run 2 data in view of expected evolution of ATLAS software environment, including a possible evolution of the persistent event formats. (See Table 2.)

LHC and data-taking parameters

The working assumption is that the LHC operations will provide 5 million seconds of proton-proton collisions in the RRB year 2015, mostly at 25-ns bunch spacing, with two months of initial running at 50-ns bunch spacing, as described in the introduction. This amounts to 27 weeks of proton-proton running with stable LHC beams 30% of the time. As discussed in the Introduction, average prompt trigger output rate for 2015 is specified as 1000 Hz.

During LS1 in 2013-2014, ATLAS will periodically take a small amount of cosmic data, inline with the ATLAS detector re-commissioning schedule.

Simulated data samples

The excellent LHC performance in 2011 and 2012 resulted in the ability to address a much wider range of physics analyses, with a higher level of precision, surpassing the most optimistic expectations. In addition, detailed physics studies established that the simulation is of unprecedented quality compared to previous generations of experiments, describing the data very well in most analyses; this quality opened-up more ‘safe’ applications for the simulated data. These two facts together significantly enhanced ATLAS physics output in 2011 and 2012, and they motivated production of larger than foreseen simulation statistics, as reported in the ATLAS Computing Resources Usage document. In particular, the unique opportunities afforded by the 2011 (7 TeV) data, with lower pileup conditions, meant there were many ongoing analyses of that dataset in the year 2012, resulting also in a bigger simulation volume for 7 TeV data analyses than initially anticipated.

The numbers of simulated events budgeted for 2013-2015 are derived from the number and size of data sets generated during simulation campaigns in 2011 and 2012, including both full and fast simulation. These estimates are considered modest in terms of physics requirements, particularly considering the ratio of the recorded data in 2010-2012 to anticipated data volume in 2015. If we are successful in decreasing CPU requirements and event sizes more than we predict, additional simulation statistics and samples would give further significant benefits to the physics programme.

Fast simulation was commissioned and shown to be successful for specific analysis purposes, particularly for simulation-intensive activities such as scans in the parameter space of new physics and various systematic error studies. Moreover, the ATLAS simulation team is working to extend the range of use of fast simulation. Consequently, plans include a larger fraction of fast simulation in the coming years.

The plan for the 2013-2014 period is to expand upon the existing simulated data sets. With the analyses of Run 1 data increasingly leading to very precise measurements, high statistics simulated background samples including several systematic variations of both the background and signal simulations are needed. Examples include the use of different and improved Monte Carlo generators with the inclusion of next-to-leading-order contributions, multi-leg matrix elements and improved techniques for parton-shower and matrix element matching, as well as different MC tunes, systematic variations in the detector material description, *etc*. This programme in 2013 includes a reprocessing of a limited volume of the 2010-2012 simulation, corresponding to the targeted partial data reprocessings mentioned above.

In addition, in 2014, an initial simulation volume describing 13 TeV data will be created, amounting to 2 billion fast and 2 billion fully simulated events for the analysis preparations for high-energy data taking. These samples will also be used in commissioning ('analysis challenge') activities of the new analysis model and the new distributed computing components. The Grid load of these commissioning activities is at present hard to quantify precisely, but lies within the uncertainty of the 2014 simulation budget.

In 2015, once data taking starts, a second ~13 TeV campaign will simulate the actual LHC configuration and ATLAS detector response with samples combining to a total volume of 5B fast and 3B fully simulated events in proportion to both lower and higher luminosity (pileup) scenarios.

Event sizes for real and simulated data

For 2013, event sizes will remain unchanged because no change in the event data model and storage technologies will be made. In 2014, with the commissioning of the new software and analysis model for 2015 data taking, event sizes will change. They will become lower for the reprocessed Run 1 data. For simulated Run 2 data, the software improvements described in the introduction will moderate growth with respect to Run 1; however, higher energy and anticipated higher pileup will still result in some growth.

In 2015 the event sizes listed in Table 1 assume that targets of the LS1 plans for optimizing event sizes, described in the introduction, are achieved, Thus, they have significant uncertainties. Constant effort will continue to be invested in reducing event sizes as much as possible.

Reconstruction of real data

The reconstruction time in 2013 for real Run 1 data is estimated to remain the same as in 2012, because no reprocessing and the coupled software change is envisaged. For 2014, with the planned full reprocessing using improved software, the LS1 work program described in the Introduction should substantially reduce the reconstruction time of Run 1 data.

Based upon expected running conditions and average pileup levels in 2015, as well as upon success of ATLAS LS1 improvement plans, the reconstruction times are projected to be ~180

HS06sec/event with 25-ns bunch spacing at μ=25 and ~230 HS06sec/event for μ=50 with 50-ns bunch spacing.

Reconstruction of simulated data

Reconstruction of simulated data includes the step of pileup simulation and detector response (digitization), as well as the simulation of the trigger response. Digitization improvements during LS1 are expected to result in approximately 20% speedup in 2013, while the time for reconstruction *per se* is not expected to change until the 2014 reprocessing using the new software, as with the real data. In 2014 and 2015, the new software is expected to result in reconstruction gains proportional to those for the real data reconstruction, along with expected improvements in pileup digitization and trigger simulation. Thus, Table 1 reflects improvements in simulation reconstruction times in 2013, and additional improvements in 2014. The time for the digitization and reconstruction steps, together referred to as simulation reconstruction, is the same for full and fast simulation.

CPU time for simulation

During 2012 the CPU time required for full GEANT-4 based simulation was improved by an additional 10%. It is assumed that this trend will continue during LS1 for both full and fast simulation, leading to the evolution of simulation CPU times shown in Table 1. By the end of 2012, investigation of more optimized mathematical libraries and faster random number generation confirmed the 10% gain envisaged for 2013.

Data distribution plan for 2013-2015

The significant increase in data volume due to pileup (coupled with the physics-driven increase in 2012 trigger rate) presented a substantial challenge to fitting the data within the pledged, and often already installed, 2013 computing resources. To address this challenge, ATLAS has conducted a careful review of the use of the different data formats in order to optimize their placement and accessibility.

The success of using a rotating disk-resident buffer of real ESD in Tier-1s in 2011 and 2012 gives confidence that the same procedure can be used in 2015. In 2013 and 2014, the ESD volume of Run 1 data on Tier-1 disks will be reduced to 5% of the total volume, consisting of a representative subset of Run 1 data and small special streams for trigger studies. This fraction will revert to 20% of the year's data when the rotating disk-resident buffer of Run 2 data ESD will be re-introduced.

The single copy of RAW data retained on disk at Tier-1s proved to be essential for trigger and other studies involving sparse event picking procedures in 2011 and 2012. At the end of 2012 we reduced the volume of Run 1 data RAW to 5% of the total volume, and we intend to limit it to this fraction for 2013 and 2014. In 2015, a full single copy of the new RAW data will be kept on Tier-1 disk as was done in the Run 1 data taking years.

The evolution of ATLAS distributed computing in 2011 and 2012 has demonstrated that group and user analysis can run successfully both in Tier-1 and Tier-2 centres. Performance metrics for data placement have been put in place to guarantee good data accessibility, which steers the "PanDA Dynamic Data Placement" ("PD2P"), the automated, popularity-based replication mechanism for both AODs and derived formats (group data). This approach has demonstrated its ability to optimize group/user analysis throughput using a low number of pre-placed replicas, improved CPU usage and decreased the overall network bandwidth and disk space demands. The same mechanism will therefore remain in use in the 2013-2015 period.

In the updated model foreseen for 2013-2015, the number of pre-placed AOD replicas kept in Tier-1 and Tier-2 disks for both real and simulated data will remain the same as in 2012. AODs from the most recent (real data) reprocessing and the corresponding simulation will be pre-placed in two copies, and the corresponding real data DESDs in one copy in Tier-1 disks. All AODs from data (re)processings and simulation production that remain relevant for analysis will be kept in two pre-placed disk copies in the Tier-2s, as will the real data DESDs in one copy. The number of copies of AODs corresponding to different (re)processings and simulation productions in Tier-1s will be decreased from two to zero according to their relevance and popularity for physics analysis. While in 2011 and 2012 this was done via a human decision process within the ATLAS Computing Resources Management, in 2013 and thereafter auxiliary automated mechanisms based on popularity will be introduced to provide an additional dynamic component in reducing the number of pre-placed AOD replicas in Tier-1 disks.

In addition, the simulated ESD and RDO (where RDO is the simulation equivalent of RAW), which are produced only upon explicit request for specific simulated samples, and which have proven to be essential to the combined performance and trigger groups, will be kept at Tier-2s in one copy in 2013-2015. The total simulated ESD volume is estimated to correspond to 5% of the total simulation statistics in 2012 (20% was assumed in the 2012

Resource Request), and is expected to remain at this level in 2013 and 2014, after which it is expected to increase to 10% in 2015. The RDO fraction is assumed to remain constant at 5% of the total statistics.

The HITS from fast simulation in 2012-2014 will be saved to tape, as is the case for full simulation HITS. This practice is in order to save on the CPU resources that would otherwise be needed to redo the fast simulation and in order to achieve an optimal grid load distribution with a reasonably small additional tape usage penalty. In 2015, assuming further speedup in the fast simulation and anticipating disk space and tape usage constraints, the fast simulation HITS will no longer be written to tape and the fast simulation production chain will write out only AODs.

The new data distribution plan is summarized in Table 2.

Our data distribution plan depends critically on allocation of adequate disk space for buffers for dynamically-placed, transient copies of data sets at both Tier-1 and Tier-2. Accounting for this required space has been complicated and confusing in the past, because ATLAS exploited the difference between actual ATLAS disk utilization efficiency, which is close to 90%, and the WLCG efficiency factor of 70% to provide the required buffer space. **It was agreed by the WLCG Management Board and the C-RSG that for the current request document the WLCG efficiency factor of 70% is to be omitted in the 2013-2015 projections, while still to be used for 2012 projections and usage.** Consequently, the required buffer space remains implicit in the 2012 values reported here, whereas for the 2013-2015 projections, the projected disk space requirements include pre-placed data, required buffer space for dynamically-placed data, and considerations of the practically utilizable fraction of the disk.

Table 2: Input parameters for Tier-1 and Tier-2 resource calculations.

Tier-1 disk policy (sum Tier-1s)	**2012**	**2013**	**2014**	**2015**
Sim RDO disk copies	0	0	0	0
Sim ESD disk copies	0	0	0	0
Sim AOD disk copies	2	2	2	2
Real RAW disk copies	1	0.05	0.05	1
Real ESD disk copies	0.13	0.05	0.05	0.2
Real AOD disk copies	2	2	2	2
Real DESD disk copies	1	1	1	1
Tier-1 processing				
Number of reprocessing/year	1	0.22	1.06	1
Tier-2 disk policy (sum of T2's)				
Sim RDO disk copies	0.05	0.05	0.05	0.05
Sim AOD disk copies	2	2	2	2
Sim ESD disk copies	0.2	0.05	0.05	0.1
Real RAW disk copies	0	0	0	0
Real AOD disk copies	2	2	2	2
Real DESD disk copies	1	1	1	1

Summary of resource requirements in 2013-2015

The table below summarizes requirements for CERN, Tier-1, and Tier-2 centres. These requirements are broken down in subsequent tables. In the tables, the numbers for 2012 come from the resource requirements document update of October 2012, and from the actual usage reported at the end of RRB year (February 2013). The numbers in the 2013-2015 columns are current estimates of resource requirements for those years.

Table 3: ATLAS resource needs for CERN, Tier-1's and Tier-2's for 2012-2015. The values in square brackets [] give the value from the ATLAS October 2012 resource request [1] and the values in round brackets () represent the total existing pledge obtained from [2].

CPU [kHS06]	2012 requested	2012 actual	2013	2014	2015
CERN	*111 (111)*	*111*	*111 [111] (111)*	*111 [111] (111)*	*111 -> 240*
Tier-1	*295 (259)*	*420*	*316 [319] (333)*	*385 [373] (326)*	*478 [502]*
Tier-2	*319 (332)*	*634*	*360 [355] (395)*	*412 [408] (398)*	*522 [540]*
Disk [PB]	WLCG factor 0.7		WLCG factor 1.0		
CERN	*11 (9)*	*10*	*10 [11] (10)*	*12 [11] (10)*	*15*
Tier-1	*29 (30)*	*47*	*35 [35] (36)*	*35 [36] (33)*	*47 [51]*
Tier-2	*48 (45)*	*52*	*51 [52] (49)*	*56 [56] (46)*	*65 [69]*
Tape [PB]					
CERN	*21 (18)*	*21(and 9 ESD)*	*25 [27] (27)*	*29 [31] (31)*	*38*
Tier-1	*31 (38)*	*31*	*42 [43] (41)*	*55 [53] (53)*	*74 [77]*

In 2013 and 2014, when there is no beam data, it is envisaged that all the CERN resources (Tier-0 and CAF, 111 kHS06) will be used for processing of the delayed streams and data reprocessing, calibration, simulation and group activities.

In 2015 it is anticipated that the CERN CPU capacity will asymptotically increase to 240 kHS06, as the Tier-0 grows to accommodate the high luminosity, high energy running and 1-kHz data taking. This growth is indicated in the table by the range of values at Tier-0 in 2015.

In addition, CPU resources from the High Level Trigger (HLT) will be used in 2013 and 2014 for simulation production and reconstruction. Due to the specific nature of the HLT cluster, it is currently hard to estimate how much throughput can be achieved with its integration into the ATLAS Grid and what fraction of it can be used. For planning, 60% of the nominal HLT cluster capacity is assumed to be available and to be used effectively, resulting in 132 kHS06 in 2013 and for half of 2014 (i.e. 66 kHS06-year). The HLT contribution is not broken out explicitly in the table below.

The ATLAS computing model has proven to be very flexible in 2010-2012, as ATLAS has varied its data replication and distribution policies and various group and user activities, and simulation activities have been performed at both Tier-1s and Tier-2s. For this reason, the sum of Tier-1 and Tier-2 CPU requirements and the sum of Tier-1 and Tier-2 disk storage requirements is for many purposes more relevant than the detailed division of requirements between the two tiers. For example, total ATLAS CPU requirements grouped by activities are presented in Table 4.

ATLAS is now developing practices and policies that will allow clouds to partition resources between tiers, within certain constraints, as best suits the particular characteristics of the centres and funding specific to the cloud. Constraints upon the use of Tier-2 centres must account for such additional factors as disk and tape availability and accessibility as well as network connectivity.

ATLAS has benefitted from the availability of substantial beyond-pledge and opportunistic CPU resources, as evident from the actual CPU usage of 2012 in Table 3 and as detailed in the Computing Resource Usage document. These additional resources proved extremely valuable in 2012, allowing ATLAS to pursue an even richer and more precise set of physics results than would otherwise have been possible in the same time frame. Our resource planning continues to be based upon the physics programme that can be accomplished within pledged resources, while we hope that our centres and FAs will continue to provide ATLAS with the invaluable resources beyond those pledged that will allow us to accomplish an optimal research programme and physics productivity.

Table 4: ATLAS CPU needs grouped by activities for 2013-2015

ATLAS CPU requirements [kHS06]	2012	2013	2014	2015
Re-processing	29	8	31	39
Simulation production	271	436	410	461
Simulation reconstruction	164	195	253	250
Group+user	166	280	280	273
Total	**629**	**919**	**974**	**1023**
Total without CERN resources	**613**	**676**	**797**	**1000**

CERN Resources

Table 5: CERN CPU

CERN CPU (kHS06)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
CERN CPU Total	***111***	***111***	***111***	***111***	***240***
Tier-0 subtotal	***77***	***111***	***0***	***0***	***185***
T0: Full reconstruction	64				161
T0: Partial processing and validation	5				12
T0:Merging and monitoring	2				5
T0: Automatic calibration	5				5
T0: Servers	2				2
CAF subtotal	***35***				***56***
CAF: Partial reconstruction, debugging and monitoring	6				16
CAF: Non-automatic calibrations	4				4
CAF: Group activities	12				18
CAF: User activities	4				5
CAF: Servers	12				12
Simulation production and reconstruction	0	0	111	111	0
HLT: Simulation production and reconstruction	0	0	132	66	0

Table 6: CERN Disk

	WLCG factor 0.7				
CERN Disk (PB)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
CERN Disk Total	***10.8***	***10.1***	***9.6***	***11.7***	***15.3***
Tier-0 Disk Subtotal	***2.70***	***4.55***	***2.80***	***2.80***	***3.10***
Buffer for RAW and processed data	2.30	3.86	2.70	2.70	2.70
Buffers for merging	0.30	0.59	0.00	0.00	0.30
Tape buffer	0.10	0.10	0.10	0.10	0.10
CAF Total	***8.1***	***5.5***	***6.7***	***8.9***	***12.2***
CAF: Calibration and alignment	0.7	0.7	0	0	0.5
CAF: Derived detector data	3.2	1.7	1.2	1.2	3.3
CAF: Derived simulated data	2.8	1.3	4.0	6.2	6.8
CAF: Group data	1.0	1.0	1.0	1.0	1.0
CAF: User data	0.4	0.8	0.5	0.5	0.5

Table 7: CERN Tape

CERN Tape (PB)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
Total	***21***	***21***	***25***	***29***	***38***

Tier-1 Resources

Table 8: Tier-1 CPU

Tier-1 CPU (kHS06)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
Re-processing	29	18	8	31	39
Simulation production	71	189	80	80	142
Simulation reconstruction	134	101	108	154	207
Group (+user) activities	60	112	120	120	90
Total	***295***	***420***	***316 [319]***	***385 [373]***	***478 [502]***

Table 9: Tier-1 Disk

	WLCG factor 0.7				
Tier-1 Disk (PB)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
Current RAW data	3.5	1	0.5	0.5	5.2
Real ESD+AOD+DPD data	4.5	14	9	5	11
Simulated RAW+ESD+AOD+DPD data	8.0	12	9	13	14
Calibration and alignment outputs	0.4	0.4	0	0	0.3
Group data	6	15	12	12	12
User data (scratch)	2	2	1.4	1.4	1.4
Cosmics	0.2	0.2	0.2	0.2	0.2
Processing and I/O buffers	4.3	3	3	3	3
Total	***29***	***47***	***35 [35]***	***35 [36]***	***47 [51]***

Table 10: Tier-1 Tape

Tier-1 Tape (PB)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
Real RAW+AOD+DPD data	15	10	11	13	21
Cosmics and other data	4	4	4	4	4
Group + User	3	0.3	4	5	6
Simulated HITS+AOD data	12	17	24	33	43
Total	***35***	***31***	***42 [43]***	***55 [53]***	***74 [77]***

Tier-2 Resources

Table 11: Tier-2 CPU

Tier-2 CPU (kHS06)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
Simulation production	199	314	158	197	319
Simulation reconstruction	29	95	42	54	43
Group + User activities	90	224	160	160	160
Total	***319***	***634***	***360 [355]***	***412 [408]***	***522 [540]***

Table 12: Tier-2 Disk

	WLCG factor 0.7				
Tier-2 Disk (PB)	***2012 requested***	***2012 actual***	***2013***	***2014***	***2015***
Real AOD+DPD data	13	9	7	11	15
Simulated HITS+RDO+ESD+AOD	22	18	21	21	26
Calibration and alignment outputs	0.3	0.3	0.2	0.2	0.2
Group data	10	19	20	20	20
User data	2	2	1.4	1.4	1.4
Processing and I/O buffers	1	3	2	2	2
Total	***48***	***52***	***51 [52]***	***56 [56]***	***65 [69]***

References

[1] All related documents can be found on the ATLAS Computing Model twiki:
https://twiki.cern.ch/twiki/bin/view/Atlas/ComputingModel

[2] The WLCG REBUS monitoring pages:
http://wlcg-rebus.cern.ch/apps/pledges/summary/

Appendix 1

This appendix provides additional explanation of values in the detailed breakdown Tables 5-12. The values labelled 'actual' in the tables are taken from the ATLAS usage document for 2012 [1] and give the values obtained in the RRB year of 2012 (March 2012-March 2013).
The values in square brackets [] give the value from the ATLAS October 2012 resource request [1].

<u>Table 5: CERN CPU</u>

Tier-0 CPU: Full reconstruction

This is a product of the average input trigger rate, the average reconstruction time, and a factor that represents the fraction of events that must be processed during the live time of the LHC (and corresponds roughly to the maximum LHC live time that can be sustained over a period of several days). The values for these quantities in 2012 were taken to be:

Average input trigger rate = 400 events/s
Average reconstruction time = 230 HS06-sec/event
Maximum LHC live time = 70%

From March of 2015 up to March 2016 (i.e. the WLCG accounting year) one can assume LHC will already take a significant quantity of data, as described earlier in the document. The assumed values for 13 TeV centre-of-mass energy are:

Average input trigger rate = 1000 events/s
Average reconstruction time = 230 HS06-sec/event
Maximum LHC live time = 70%

Tier-0 CPU: Partial processing and validation

It is assumed we will be validating live data by partially reconstructing 20% of the real data as it comes into the T0 at the trigger rate (1000 Hz in 2015). The processing here is assumed to use 60 HS-sec per event, which is constant throughout the Run 1 years. It is not used in 2013-2014.

Tier-0 CPU: Merging and monitoring

This describes the merging of smaller files into large ones suitable for tape and network distribution. The method was substantially improved in the beginning of 2011 and is half the previous value, amounting to 3% of the reconstruction CPU power. It is not used in 2013-2014.

Tier-0 CPU: Automatic calibration

This is calibration production run promptly at the T0. It is assumed that a sample of 480M events is processed at a CPU power of 60 HS06-sec/event. This sample is run over a period of 80 days, and it is assumed there is an 85% CPU usage efficiency. This requires 5 kHS06 to run. This is done in all years with data collection ; it is not used in 2013-2014.

Tier-0 CPU: Servers

This is a fixed number and handles data serving needs on the T0. This is based on 2010-2012 experience. In 2013-2014, it is included in the total Grid usage.

CAF CPU: Partial, reconstruction, debugging and monitoring

This denotes special studies on the data (e.g. the express stream) as well as special simulation studies (e.g. HLT tuning studies). For 2015, it is assumed to be 10% of the full reconstruction CPU, based on Run 1 experience. It is not used in 2013-2014.

CAF CPU: Non-automatic calibrations

Based on Run 1 experience this requires 4 kHS06 for data-taking years.

CAF CPU: Group activities

We estimate this by assuming the group usage on the CAF is about 20% of the CPU activity of groups on the sum of the Tier-1s. For 2015, this gives 18 kHS06. In 2013 and 2014, the groups will do this work on the ATLAS Grid resources.

CAF CPU: User activities

In general, users do not run on the CAF, but we have special cases for prompt studies. We assume based on our Run 1 running that there are about 30% of users running jobs similar to those run by the groups, so they require 0.30 of the CAF group CPU = 5 kHS06 in 2015. In 2013 and 2014, they will instead use the ATLAS Grid resources.

CAF CPU: Servers

This is a fixed number and describes the data serving needs on the CAF. This is based on Run 1 experience.

Tier-0+CAF+HLT: Simulation production and reconstruction

As described in the document, in 2013 and 2014 the CERN resources will be integrated in the Grid for simulation production and reconstruction, involving the Tier-0, CAF and the ATLAS HLT cluster. In 2013 and 2014, the whole Tier-0 and CAF will be available, amounting to 111 kHS06 and the HLT cluster with partial availability adds 132 kHS06 in 2014 and half of that, 66 kHS06, in 2014.

Table 6: CERN Disk

Tier-0 Disk: Buffer for RAW and processed data

The buffer for RAW and processed data was increased to 3.86 PB (2.7 PB without the WLCG factor) to accommodate the increased trigger rate and event sizes in 2012. Need continues for 2013-2015.

Tier-0 Disk: Buffers for merging

Buffers for merging increased to 0.59 PB (0.3 PB without the WLCG factor) to accommodate the increased trigger rate and event sizes in 2012. They will not be used in 2013-2014.

Tier-0 Disk: Tape buffer

The tape buffer amounts to 0.1 PB.

All the numbers for the above Tier-0 buffers come from operational experience during the Run 1 running.

CAF Disk: Calibration and alignment

This describes RAW data kept for calibration studies. The estimate is based on Run 1 experience. The data is a sample used by the Inner Detector, Liquid Argon, and Muon system alignment groups. The sample size is determined by the needed statistics and contains some data from the current year and some legacy data for comparisons. In 2012, the sample in total is 0.7 PB (0.5 PB without the WLCG factor). There is no new data in 2013-2014, and the 2015 volume is expected to remain the same as in 2012.

CAF Disk: Derived detector data

This corresponds to a fraction of real AOD and ESD as well as specialized formats (DRAW, DESD, HIST, TAG) produced in Tier-0 reconstruction. In addition, the subsample of RAW data from specialized (calibration and express) streams for detector studies, high-priority analyses and calibration done by groups is stored. With 2012 data currently stored, this is estimated to be 1.7 PB (1.2 without the WLCG factor). It is assumed this data stays useful in

2013 and 2014 until replaced with new data in 2015. The 2015 usage is estimated to be equivalent to twice the reconstructed AOD volume and 10% of total RAW data volume.

CAF Disk: Derived simulated data

The simulated AODs and RDOs from the last simulation campaign are replicated to the CAF disks for detector studies, high-priority analyses and calibration. This gives 1.3 PB (0.9 PB without the WLCG factor) for 2012 with proportional increase in 2013 and 2014 when CERN CPU resources are used for Grid production.

CAF Disk: Group data

The Group data space at CERN of 1 PB (0.7 PB without the WLCG factor) was fully utilized in 2012. The groups on the CAF are assumed to continue the present activities in 2013-2015 on the Grid while storing their input and output data on CAF and thus the needed volume is estimated at 1 PB and remains constant in 2013-15.

CAF Disk: User data

We keep a 0.8 PB scratch area open for users (0.5 PB without the WLCG factor). They are expected to remove any data sets they need to keep. This size is based on Run 1 running experience and remains valid in 2013-2015.

Table 7: CERN Tape

CERN Tape:

In February 2012 there was 29 PB of data on tape, of which ~9 PB is the data ESDs, which were put on tape as a fail-over procedure and will be deleted early in 2013. Only the residual 21 PB are listed under 2012 actual usage. Estimates for years 2013-15 thus add to the base value of 21 PB.

In 2013 and 2014 the CERN CPU resources will be used within the Grid for simulation production and analysis. Thus the CERN tape will save the HITS and AODs produced at CERN, which is projected to amount to additional 8 PB of tape used for the two years combined.

Each year of data taking, we put to CERN tape all RAW data (from beams and cosmics); all AOD and DESD from the Tier-0 processing as well as all derived formats, the volume being evaluated of the same size as the AOD volume. For 2015, the volume increase from pp collisions is thus estimated at 5.2 PB (RAW) + 3.8 AOD+DESD+derived formats = 9 PB. In addition, the opportunistic simulation that can be done on the idle CERN resources is estimated as up to 5% of total 2012 simulation production (HITS and AODs) which adds another 0.5 PB, bringing the total volume increase to 9.5 PB. Adding this to the legacy data totals to 38 PB.

Table 8: Tier-1 CPU

Tier 1 CPU: Reprocessing

This is the CPU needed to do reprocessing campaigns, averaged over each year.

As stated, we do not expect to reprocess the Run 1 data in full until after the end of 2013. Instead, a subset of the Run 1 data and corresponding simulation for the most important analyses may be reprocessed in 2013. The full reprocessing of all data and MC is foreseen in 2014.

The effective number of reprocessings is evaluated at 0.22 of the total Run 1 data volume in 2013 and 1.06 of the total Run 1 data volume in 2014. The assumed CPU usage efficiency is 85% and the overhead factor is 6% (merging of output files). The 2015 data is also expected to be reprocessed effectively once; this reprocessing will probably be made up of one or more consecutive staged reprocessings.

Tier 1 CPU: Simulation Production

Simulation production is shared between T1 and T2. This represents simulation production of the T1 portion. In the years 2013 and 2014 the CERN resources (including HLT farm) are taken into account. The assumed WLCG 85% CPU usage efficiency and 5% overhead due to merging and event generation steps are taken into account.

Tier 1 CPU: Simulation reconstruction

The Tier-1s perform the bulk of simulation reconstruction, ~80% is done at Tier-1s (rest at T2), an activity that continues through LS1. It involves simulation reprocessings and new simulation production, as detailed in the document. The assumed WLCG 85% CPU usage efficiency and 5% overhead due to merging are taken into account.

Tier 1 CPU: Group + user activities

Group activity is assumed to be present over the whole period, with increased activity before targeted physics conferences. The average CPU requirement is presented. For 2013-2015 the estimate is based on the Run 1 group and user activity, in particular on the 2012 usage when it amounted to ~112 kHS06 on the yearly average, at the end of 2012 even 30% more. As detailed earlier in the document, the group and user activity in 2013-2014 is projected to remain constant, estimated at 120 kHS06, while in 2015, this drops slightly, reflecting the higher priority again on other tasks at the Tier-1s.

Table 9: Tier-1 Disk

Tier 1 Disk: Current RAW data

As already noted in this document, ATLAS has dropped most of the ESD and decided to normally keep all raw data from the current year on disk instead. The RAW volume was reduced to ~0.5 PB at the end of 2012, which will be used for performance studies in 2013-2014. In 2015, we again expect to keep the RAW from that year on Tier-1 disks; the size is estimated at 5.5 PB, based on the number and size of events.

Tier 1 Disk: Real ESD+AOD+DESD data

In 2013-2015, as already described in the document, only the most popular data processings will be kept in the Tier-1 disks. The ESD volume will remain constant in 2013-2014, keeping 5% of the total volume for detector performance and physics studies. The rolling ESD buffer will be re-introduced in 2015 with a 20% fraction of the total ESD volume. The DESDs will be kept in one replica throughout 2013-2015 and AODs will have up to two pre-placed copies and additional volatile dynamically replicated copies. The aim is to improve and fine-tune the operational procedures during the LS1 to keep the minimal amount of pre-placed AOD datasets in the Tier-1s, which still guarantees adequate physics group and user analysis throughput.

Tier 1 Disk: Simulated HITS+RDO+ESD+AOD data

In 2013-2015, following the replication policy, only the currently active software version of AOD will be kept on Tier-1 disk. In addition, 1 PB of pileup inputs (HITS and/or data overlay) is foreseen to be distributed across Tier-1s based on 2011-2012 experience. During the LS1 the aim is to improve the operational procedures further to keep the minimal amount of pre-placed AOD datasets in the Tier-1s, which still guarantees adequate physics group and user analysis throughput.

Tier 1 Disk: Calibration and alignment outputs

In 2012 calibration and alignment jobs produced 0.4 PB of output (0.3 PB without the WLCG factor of 0.7). This volume represents three copies of this data distributed over the T1 disks to facilitate performance and physics studies. This volume will not be needed in 2013-2014 and will be again used in 2015 with data taking.

Tier 1 Disk: Group data

The volume of the widely popular group data in 2012 rose to a total of almost 11PB (15 PB including the WLCG scaling factor of 0.7). The 11 PB value includes the dynamically replicated popular data sets as well as intermediate/transient data, which are essential for the current mode of operation. These dynamic and transient buffers represent about 50% of the group volume and 20% of the total disk volume, as predicted in the 2012 resource request. For 2013-2015, the group data volume is set to 12 PB and is expected to remain constant. This volume includes the dynamically replicated popular data as well as transient data. The utilization of this volume will evolve with software and operational improvements in the group activities, as described earlier in the document.

Tier 1 Disk: User data

2 PB scratch area (1.4 PB without the WLCG factor of 0.7) is open for users running jobs at Tier-1s; they are expected to migrate any data sets they need to keep. This size was established in 2011 to optimize accessibility and prevent premature deletions. We keep this value of 1.4 PB in 2013-2015.

Tier 1 Disk: Cosmics

A sampling of cosmic data taken is kept on disk for detector performance studies. The value of 0.2 PB is based on current experience.

Tier 1 Disk: Processing and I/O buffers

Buffers are needed for all I/O and data processing, involving transient data sets to be merged or processed further, simulation generator-level inputs, *etc*. The size is based on Run 1 experience and is expected to remain constant at 3 PB (without the WLCG factor of 0.7) in 2013-2015.

Table 10: Tier-1 Tape

Tier 1 Tape: Real RAW+AOD+DESD data

In 2013-2014, the volume of real data will increase due to the planned reprocessing activities detailed earlier in this document. In 2015, a copy of the new data will be placed on the Tier-1 tapes, following the same policy as in Run 1.

Tier 1 Tape: Simulated data

The archived volume increase of simulated HITS and AODs in 2013-2015 corresponds to the predicted simulation production activities projected for these years. As already described in this document, the plan is to stop archiving the HITS from fast simulation in 2015, reducing the tape needs for 2015 and beyond.

Tier 1 Tape: Group+User

According to ATLAS publication policy, the group data formats used in analysis need to be archived to tape. The group and user quota is expected to be filled in the first half of 2013 when the groups fulfil their archival duties according to the ATLAS publication policy. The estimate of the volume increase is 1 PB per year.

Tier 1 Tape: Cosmics and other legacy data

There are 4.28 PB of legacy cosmic data on tape.

Table 11: Tier-2 CPU

Tier 2 CPU: Simulation Production

The fraction of simulated events not done in Tier-1s (or CERN resources in 2013-2014) is processed in Tier-2s. For data-taking years, this is the bulk of the simulation production.

Tier 2 CPU: Simulation Reconstruction.

The fraction of simulated events not digitized and reconstructed in Tier-1s (or CERN resources in 2013-2014) is processed in Tier-2s.

Tier 2 CPU: Group + User activities

Group activity is assumed to be present over the whole period, with increased activity before targeted physics conferences. The average CPU requirement is presented. For 2013-2015 the estimate is based on the Run 1 group and user activity, in particular the end of 2012, when the group and user activity share on all available resources exceeded 40%, as detailed earlier in the document. The group and user activity in 2013-2015 is expected to remain at the constant level of 160 kHS06, with the aim of LS1 improvements to increase the physics throughput while keeping the same CPU consumption.

Table 12: Tier-2 Disk

Tier 2 Disk: Real AOD + DESD data:

By the policy introduced in 2012 and described in this document, the data processings that are still active in physics analyses are kept in Tier-2 disks, in two copies for redundancy. The data occupying the Tier-2 disks in 2013 are the latest processings of all Run 1 data, which will in 2014 be complemented with the final reprocessed Run 1 data. In 2015 final reprocessed Run 1 data will be retained and complemented with the new high energy data.

Tier 2 Disk: Simulated HITS+RDO+ESD+AOD data

By the policy introduced in 2012 and described in this document, the simulation productions that are still active in physics analyses are kept in Tier-2 disks, in two copies for redundancy. The simulated data occupying the Tier-2 disks in 2013 are the latest simulation productions of all Run 1 data, which will in 2014 be complemented with the final reprocessed Run 1 simulated data. In 2015 the final reprocessed Run 1 simulated data, produced in 2014, will be retained on Tier-2 disks and complemented with the new high energy simulation production. In addition, 1PB of pileup inputs (HITS and/or data overlay) is foreseen to be distributed across Tier-2s based on Run 1 experience.

Tier 2 Disk: Calibration and alignment

From experience, the calibration and alignment jobs produce about 0.3 PB/yr (0.2 PB without the WLCG factor of 0.7) and this is effectively a rolling buffer which remains active for performance studies in 2013-2015 at 0.2 PB.

Tier 2 Disk: Group data

The volume of the widely popular group data in 2012 rose to the total of almost 14PB (19 PB including the WLCG scaling factor of 0.7). The value of 14 PB includes the dynamically replicated popular data sets as well as buffers of intermediate/transient data, which are essential for the current mode of operation. These dynamic and transient buffers represent about 50% of the group volume and 20% of the total disk volume, as predicted in the 2012 resource request. For 2013-2015, the group data volume is set to 20 PB and is expected to remain constant. This volume includes the dynamically replicated popular data as well as transient data. The utilization of this volume will evolve with the software and operational improvements in the group activities, as described earlier in the document.

Tier 2 Disk: User data

User data at T2 is in a scratch area that is cleaned regularly, so permanent user data must be moved off the T2. The buffers are sized by experience (1.4 PB without the WLCG factor of 0.7) and indicate to grow only slightly as the accumulated data increases.

Tier 2 Disk: Processing buffers

This is space reserved for moving data to and from T2 worker nodes for grid production jobs. Run 1 experience shows that these buffers need to be about 3 PB total (2 PB without the WLCG factor of 0.7). They remain needed and at 2 PB in 2013-2015.