

# *ATLAS DP and MC plans for LS1*

L. Fiorini  
(U. Valencia and IFIC)

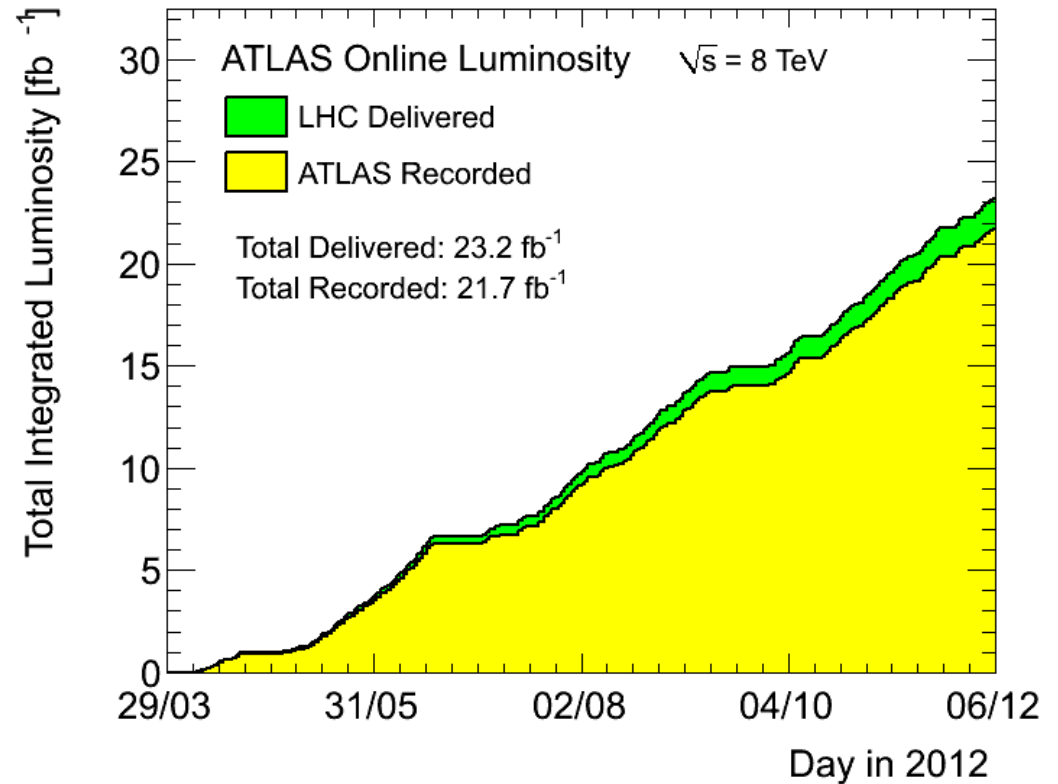
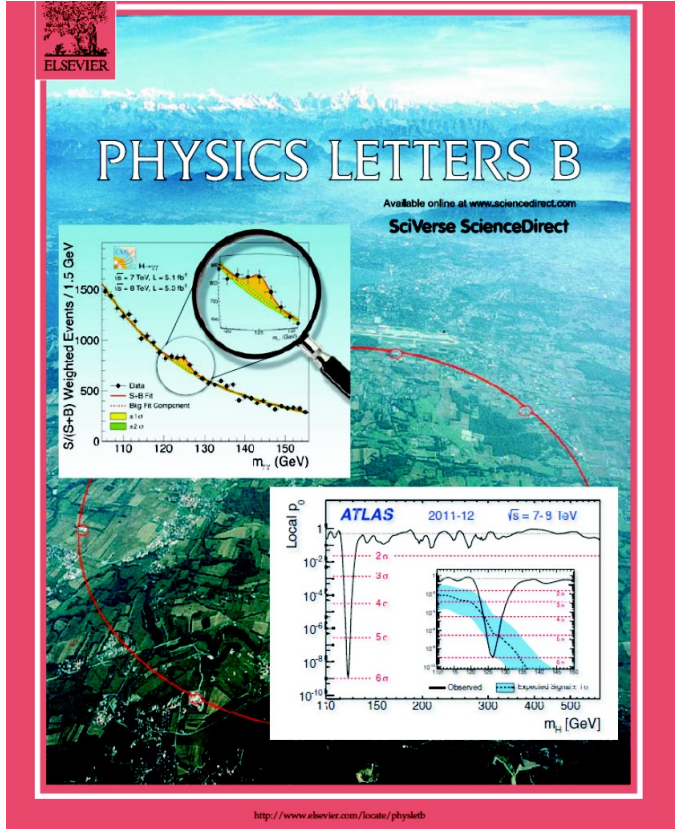
ATLAS Distributed Computing Jamboree  
10/12/2012



# Outline

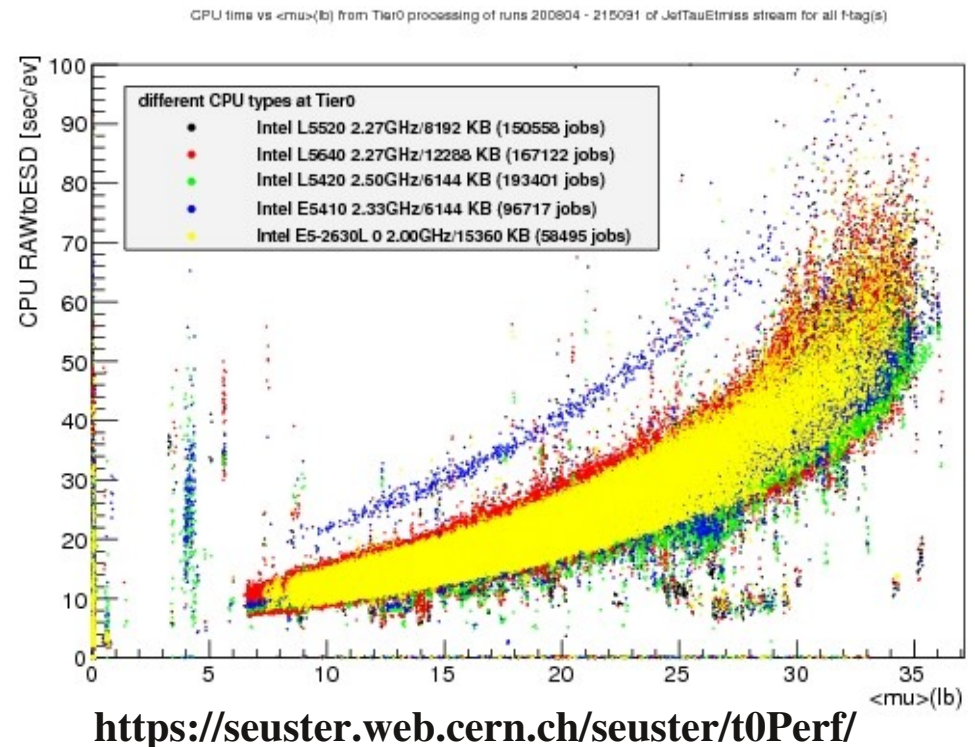
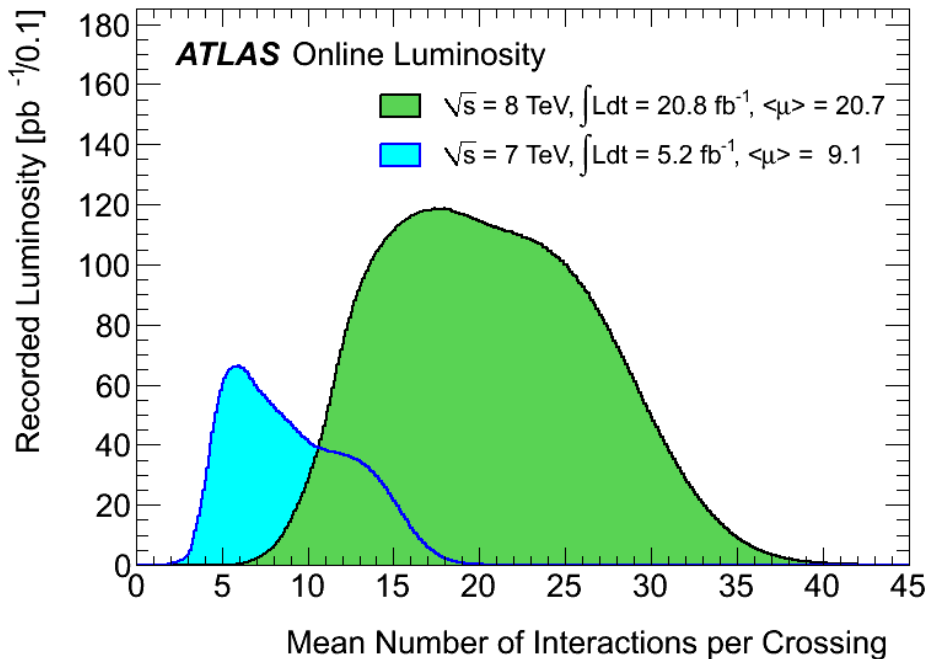
- Data taking
  - CPU usage
  - Data storage
  - Tier0 operations
- 2012 reprocessing
  - Motivations
  - CPU and disk usage on the GRID
- Reprocessing Plans for LS1
  - 2011 data reprocessing
  - FDR of the full chain
- Physics Validation for LS1
- ~~MC production for LS1~~ → *covered by Physics Coordination talk.*
- Challenges for Run2

# 2012: Annus Mirabilis for ATLAS and LHC?



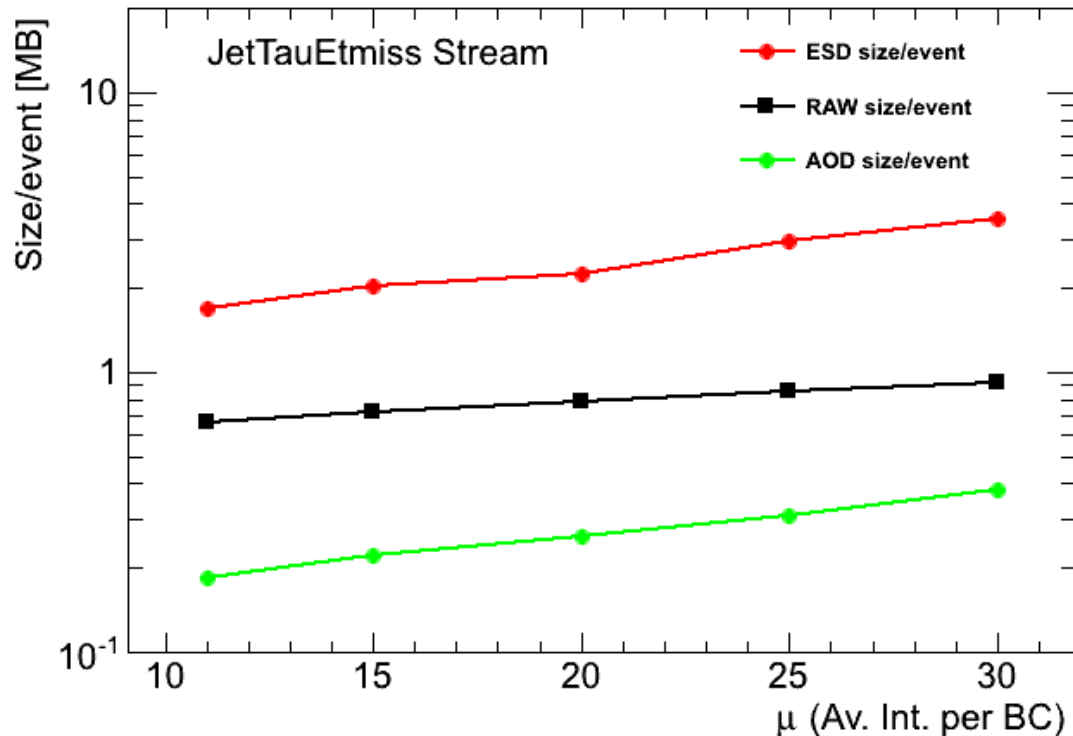
- 2012 was certainly an extraordinary year for LHC and ATLAS.
- Computing and ADC contributed to the success of this year making possible to have the results ready a few weeks after the data was stored on disk.
- LHC achievements in 2012 are remarkable:
  - Delivered integrated luminosity passed 23  $\text{fb}^{-1}$
  - Instantaneous luminosity was close to design one,  $>7.7 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$

# Pileup in 2012



- One of the main concerns in 2012 was dealing with high pileup environment ( ~factor 2 higher than in 2011), that have impact not only on the LHC physics, but on the full chain from the operation to the data distribution on the GRID.
- Significant optimization to reduce CPU time vs pileup was done in the 17.2.X release compared to 17.0.X (used for the 2011 run).
- RAWtoESD, that is the slowest part of the reconstruction, scales more than linearly with  $\mu$ , passing from ~20s for  $\mu=20$  to more than 40s for  $\mu=30$  (JetTauEtmisss stream).

# Pileup and Data storage



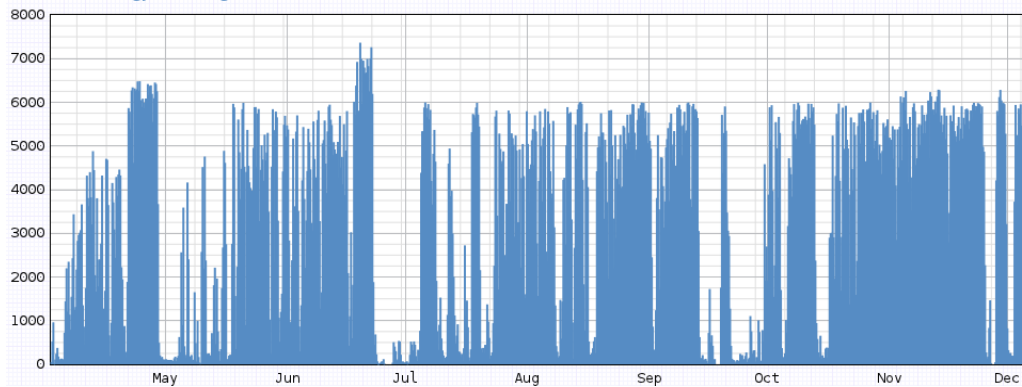
- Pileup not only affects the reconstruction time, but also the size of the output datasets.
- ESD and AOD data formats increase more rapidly than RAW wrt pileup.

- Space used by 2012 data is (without considering replicas!) for the main physics streams (JetTauEtmis+Egamma+Muons):
  - RAW: ~ 2 PB
  - ESD: not saved (would of the order of the 6 PB)
  - AOD: ~600 TB
  - DESD: ~<600 TB
  - Calibration and Alignment activities: ~300 TB

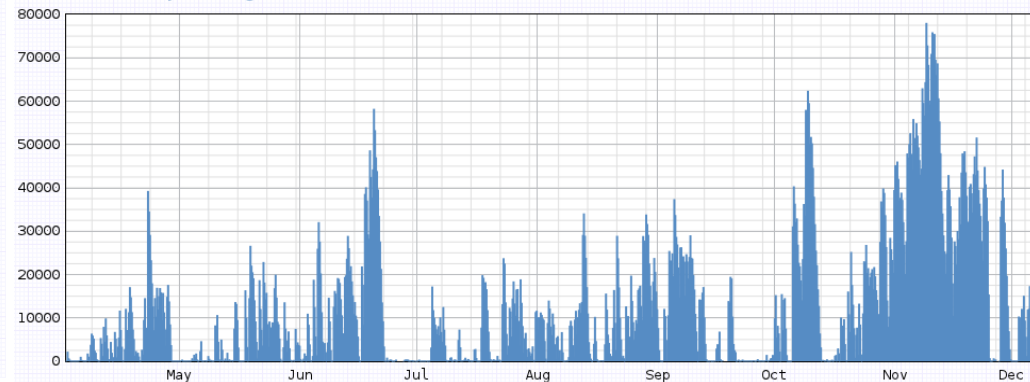
# Tier0 operations in 2012

- Tier0 operations went quite smooth in 2012, with few intense periods:
  - Tier0 operates with an **average** trigger rate from Point1 of **400 Hz** (calibration and delayed streams not included) and can run in parallel 6000 reconstruction jobs.
  - In current trigger/computing conditions, Tier0 is capable of reconstructing about  $1\text{fb}^{-1}$  of data per week.
  - More intense periods were the weeks close to ICHEP conference, in order to deliver the reconstructed data for the Higgs analysis in a timely matter and a large backlog period in November due to several LSF shortages.

recon runningjobs (avg), from: 15:32 01/04/2012, to: 14:32 05/12/2012



recon tobedonejobs (avg), from: 15:32 01/04/2012, to: 14:32 05/12/2012



# Data Replication on the GRID

- Data Replicated on the GRID to ensure safe storage and smooth access.
  - RAW: T1\_DATADISK + 1 copy on T1\_DATATAPE (2 different T1s).
    - RAW for 2012 were exceptionally been kept on disk (discovery case). We'll start deletion of the RAW when the reprocessing has been declared finished.
  - ESD: *“Life without ESD”* model. Only kept on disk for small streams (e.g. express). For physics streams deleted after 3 weeks on disk, no tape copy.
    - Time on disk has been adjusted a couple of times in 2012 passing from 6 weeks to 4 weeks in August and then from 4 to 3 weeks in November.
  - AOD: 2 T1 primary copies + 1 secondary copy
  - DESD: 1 T1 primary copy + 1 secondary copy
  - DRAW: 1 primary copy in CERN-PROD\_DATADISK
  - TAG and HIST: 1 primary copy in CERN-PROD\_DATADISK
- Express stream can be consider a small stream compared to the physics streams, having a flat rate of 10 Hz compared to 400 Hz average for physics.
- Space on disk is also of the order of 2.5-3% of the physics stream (designed to be 10% ).



# 2012 October Reprocessing

- Reprocessing of **period A-H data** ( $\sim 17 \text{ fb}^{-1}$  of data, 2 billion events).
- Purpose: Provide best 2012 data we can w/ rel 17 for 2013 winter conferences:
  - Improves alignments of ID and MS
  - Correct FCAL en. scale in C1-C8
  - Improves LAr DQ efficiency (gain 1-2% of statistics) and recover other isolated DQ issues
  - Apply all AODFix (apart from MET fix) in Reco.
  - Many other improvements. List here: <https://twiki.cern.ch/twiki/bin/viewauth/Atlas/Summer2012Reprocessing>
- Streams reprocessed: `express_express`, `physics_CosmicCalo`, `physics_CosmicMuons`, `physics_Background`, `physics_Egamma`, `physics_JetTauEtMiss`, `physics_MinBias`, `physics_ZeroBias`, `physics_Muons`, `physics_IDCosmic`, `physics_standby`, `debug_hltacc`
- Then the delayed streams (being prepared): `physics_jetCalibDelayed`, `physics_HadDelayed`, `physics_Bphysics`

Stream/Period	Period A	Period B	Period C	Period D	Period E	Period G	Period H	TOTAL
Egamma	43,266,159	177,569,377	51,382,952	113,173,567	86,690,531	43,857,431	53,569,900	569,509,917
Muons	43,661,420	163,597,791	80,935,368	112,743,072	81,667,914	41,182,058	49,412,952	573,200,575
JetTauEtmiss	64,345,937	189,247,789	53,773,871	121,460,633	111,745,415	57,078,247	66,807,066	664,458,958
Background	--	19,921,757	5,257,730	14,675,969	10,433,550	6,571,508	6,620,367	63,480,935
MinBias	6,000,436	16,156,158	40,824,856	10,456,014	7,276,542	3,652,824	22,526,657	106,893,487
TOTAL AOD	157,273,952	566,492,872	232,174,777	372,509,255	297,813,952	152,342,068	198,936,942	1,977,543,872



# 2012 Reprocessing

A. Lucotte

Data accumulated up to MD3 + period H

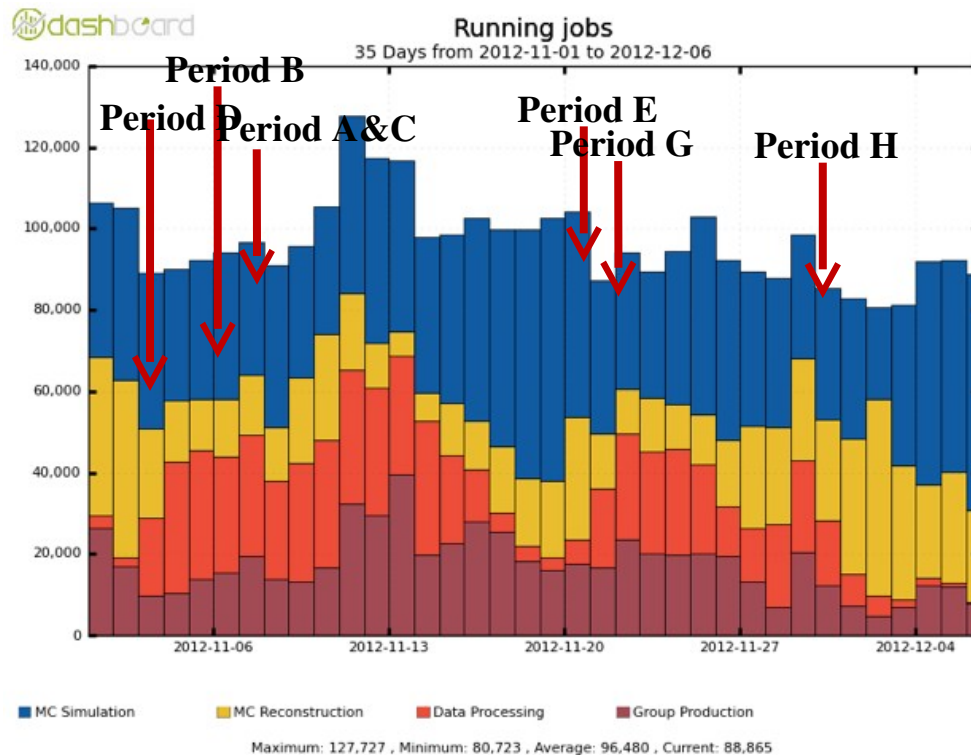
2 billion events → twice the size of 2011 reprocessing

CPU Timing : ~25s/evt (pile-up)

→ Double that of 2011, due to higher pileup

Reprocessing was dispatched to all Tier1 to cope with the CPU requirements

- AOD Size: 300kB/evt resulting in ~0.6 PB
- ESD size too large : *life w/o ESDs*. Only kept for 2 periods (B11+D3)



MC simulation

MC reconstruction

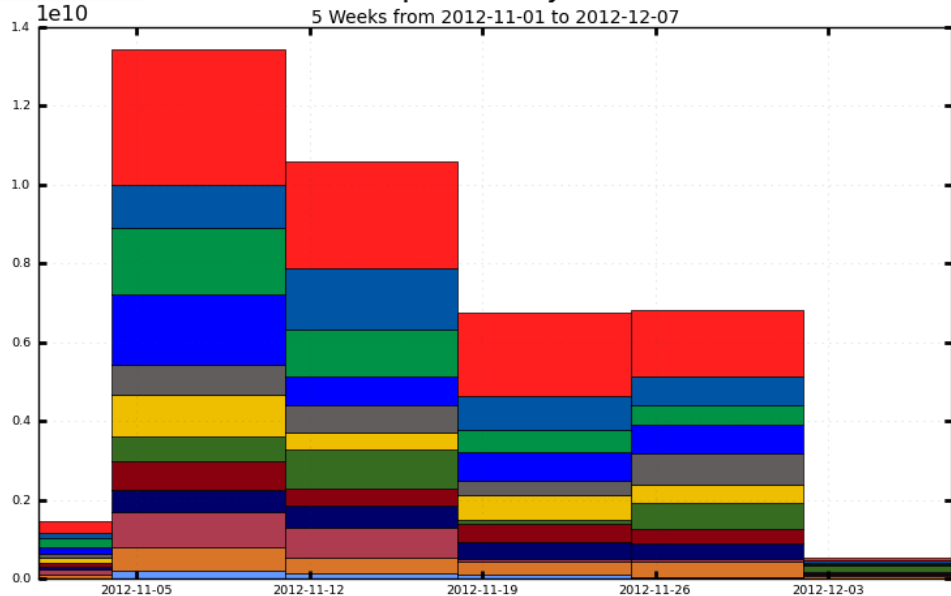
Reprocessing

Group production

# 2012 Reprocessing (2)



CPU consumption Good Jobs in seconds

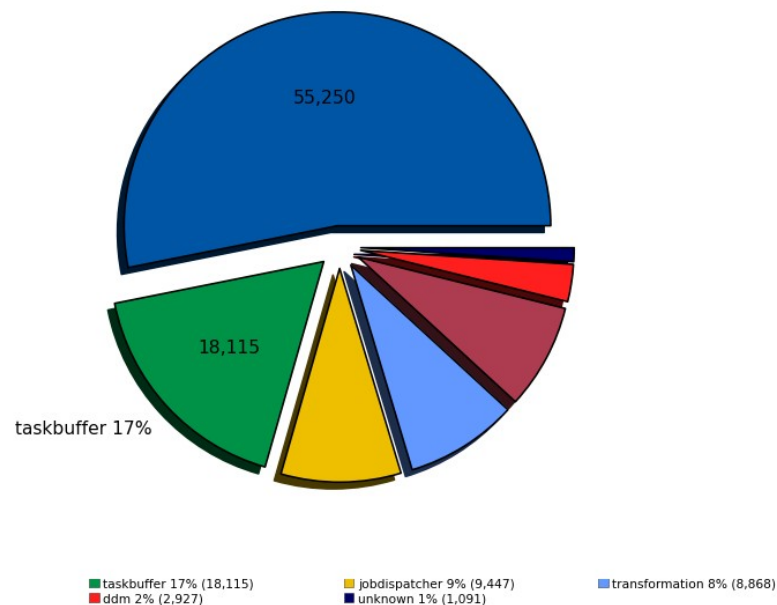


- BNL-ATLAS
- FZK-LCG2
- NIKHEF-ELPROD
- IN2P3-CC
- NDGF-T1
- INFN-T1
- TRIUMF-LCG2
- SARA-MATRIX
- TAIWAN-LCG2
- PAL-LCG2
- RIC
- CERN-PROD

Maximum: 13,447,988,512 , Minimum: 0.00 , Average: 5,653,034,091 , Current: 532,995,419



Panda Failures by Category (Sum: 103,973)



Total CPU consumption was  $\sim 45 \cdot 10^9$  s (CPU time). 1/4 of it in BNL.  
CPU efficiency was high ( $>90\%$ ).

CPU usage was high in the beginning (periods D+B running), then it decreased quickly.

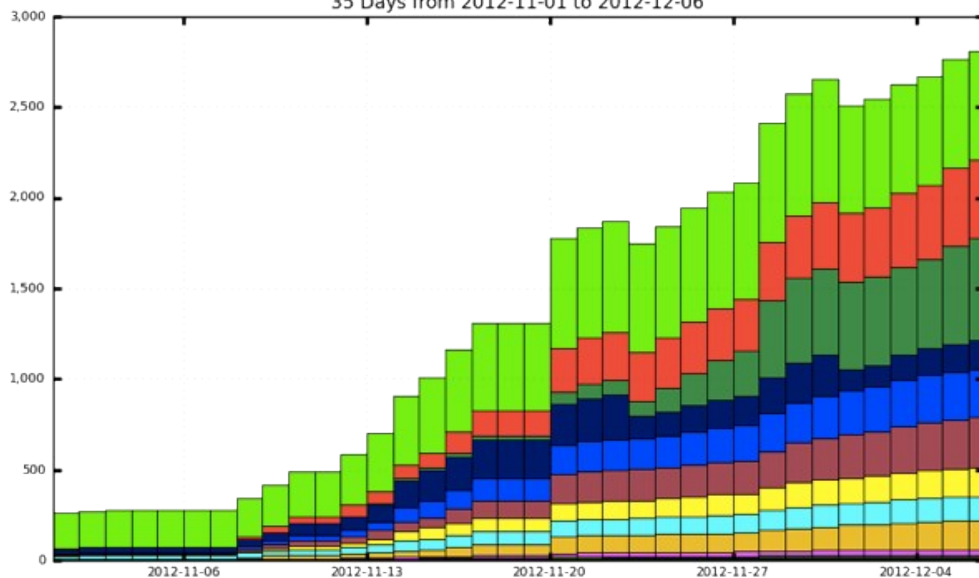
Jobs failure rate was of the order of 6%.

# 2012 Reprocessing (3)

A. Lucotte



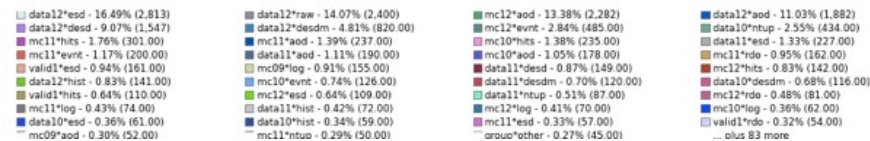
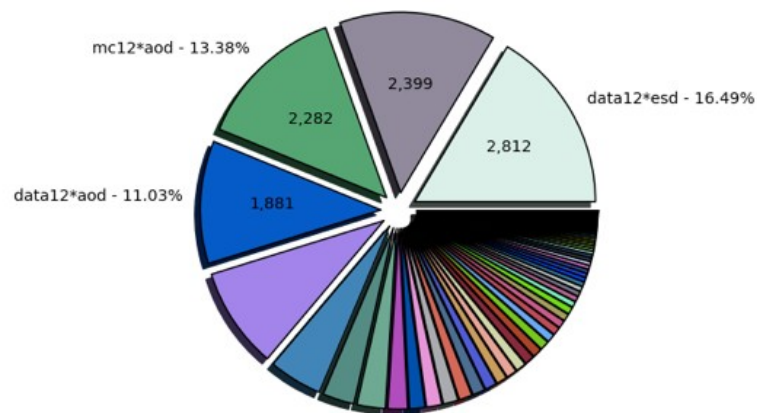
Number of Physical Bytes (in TBs)  
35 Days from 2012-11-01 to 2012-12-06



Maximum: 2,811 , Minimum: 261.61 , Average: 1,407 , Current: 2,811



Number of Physical Bytes (in TBs) (Sum: 17,107)  
data12\*raw - 14.07%



Space used for the reprocessing quickly increased in less than 1 month by 2 PB, creating issues to few Tier1 sites.

Situation improved reducing the lifetime of the ESD from 4 weeks to 3, removing the 2011 HI ESD (on disk for reprocessing) and by removing the transient unmerged files.

# *Processing in 2013*

- p+Pb run scheduled at the beginning of 2013: ~1 month of data taking, with eventually few days of pp collisions at 2.76 TeV.
- This is the last data taking period before LS1.
  - Pilot runs in September 2012 allowed us to test reconstruction and computing requirements for this kind of data (runs 210184, 210185, 210186).
  - From the computing point of view, these events are more similar to pp MinBias events than HI events:
    - Reco CPU time: ~ 12s/event
    - RAW: 1.1 MB/event
    - ESD: ~1 MB/event
    - NTUP: 0.5 MB/event
  - Hard probes stream was not enabled for the pilot run, but we don't expect the overall picture to change significantly.

# *Plans for reprocessing during LS1*

- 2012 Reprocessing campaign with rel 17.2.X over before Christmas.
- **Currently there are no firm plans for reprocessing campaigns during LS1.**
  
- We are doing preparatory studies for a reprocessing of 2011 in the first part of 2013, which is not decided yet:
  - Goal is to deliver a consistent data+ MC dataset of both 2011 and 2012 for the final results of ATLAS Run1 data (e.g. Higgs results) reconstructed with release 17.2.X. This campaign would include:
    - the reprocessing of 2011 data, that is less data and faster to reconstruct than 2011 (ETA<1 month).
    - the reprocess (DIGI+RECO) of MC11c samples, that is the most demanding part of the 2011 reprocessing.  
~3 months were needed in the original processing and we assume that it will not going to change significantly.

# *Plans for reprocessing during LS1 (2)*

- PhysCoord is evaluating the Pros&Cons of the 2011 reprocessing:
  - It has to be ready for final ATLAS papers of 2011+2012 data (summer 2012?).
  - Some analysis will benefit from better 2012 reconstruction+object ID.
  - Systematics uncertainties of the 2011+2012 combination could be reduced.
  - Enough extra time after the end of the reprocessing should be granted to re-derive the PC and analysis corrections and scale factors.
  - The 2011 data reprocessing will compete for resources during several months with other production campaigns, such as MC production
  - Trigger simulation in 17.2.X is not exactly the same used in rel16 for MC11c.
  - Validation schedule of the reprocessing is rather long and it cannot be easily compressed:
    - Software validation : sample A comparing to mc11c (with new / old pileup sample) and comparing one big data run.
    - Express stream reprocessing and DQ signoff to validate the data over all the chain.



# *Plans for reprocessing during LS1 (3)*

- Plans for other reprocessing besides 2011 data are purely speculative
- ATLAS is undergoing a full review to improve the analysis model during LS1.
  - The work of the analysis model panel already started.
  - The review is closely tied to all other aspects of ATLAS and their plans for LS1 :
    - Trigger, DP, reconstruction, computing, ADC, data format, streaming, replication, etc.
- We should foresee at least 1 Full Dress Rehearsal during LS1 to test the full chain from Tier0 to the plots for the press.

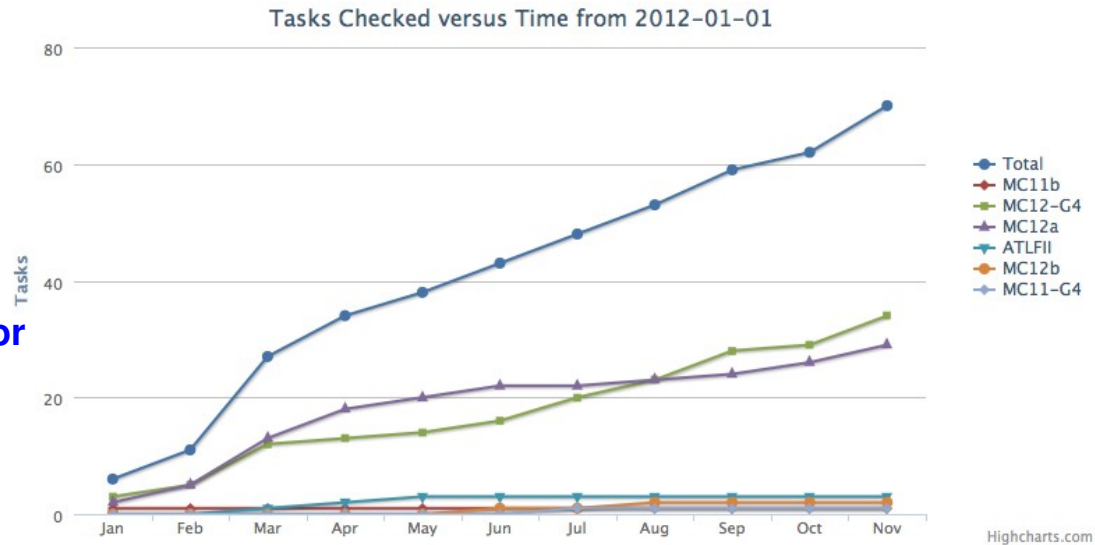
# Physics Validation MC Production

J. E. Garcia

Cloud	Datasets	Total Files in datasets	Total CpFiles in datasets	Completed	Transfer	Subscribed
TO	501	123135	121663	409	91	1
CA	6	4139	4139	6	0	0
DE	31	8284	8284	31	0	0
ES	12	2603	2603	12	0	0
FR	501	123254	122463	439	61	1
IT	7	1907	1907	7	0	0
NG	24	17477	17477	24	0	0
NL	0	0	0	0	0	0
TW	2	779	779	2	0	0
UK	39	5590	5590	39	0	0
US	501	123371	122137	427	73	1

Datasets	Total Files in datasets	Last Subscription	LFC Checked	Last Transfer
25535	123371	Dec 05 13:05:43	Dec 05 09:20:50	Dec 05 09:20:50

## Snapshot of validation jobs in panda monitor

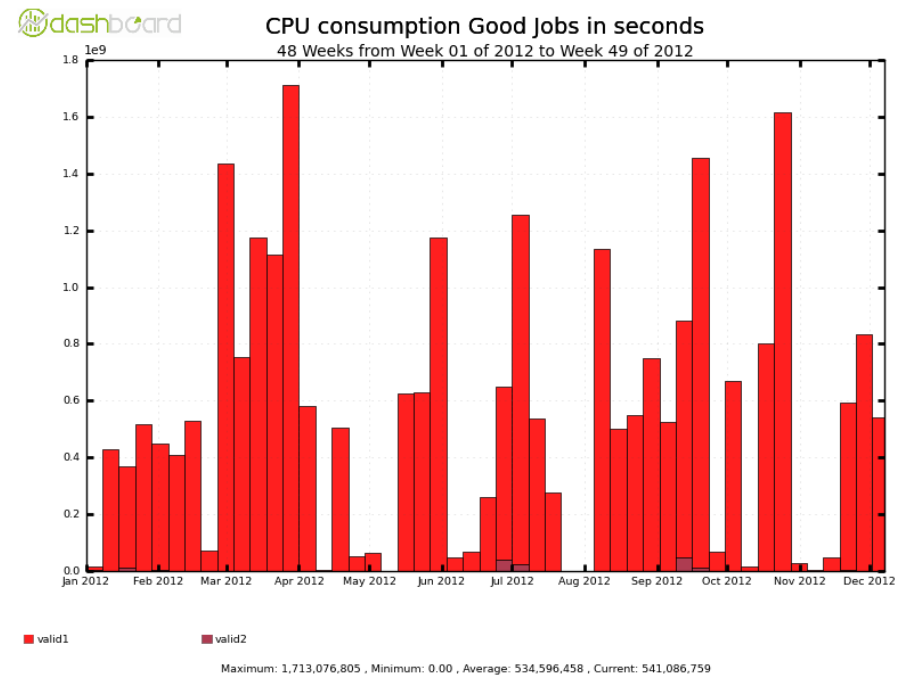
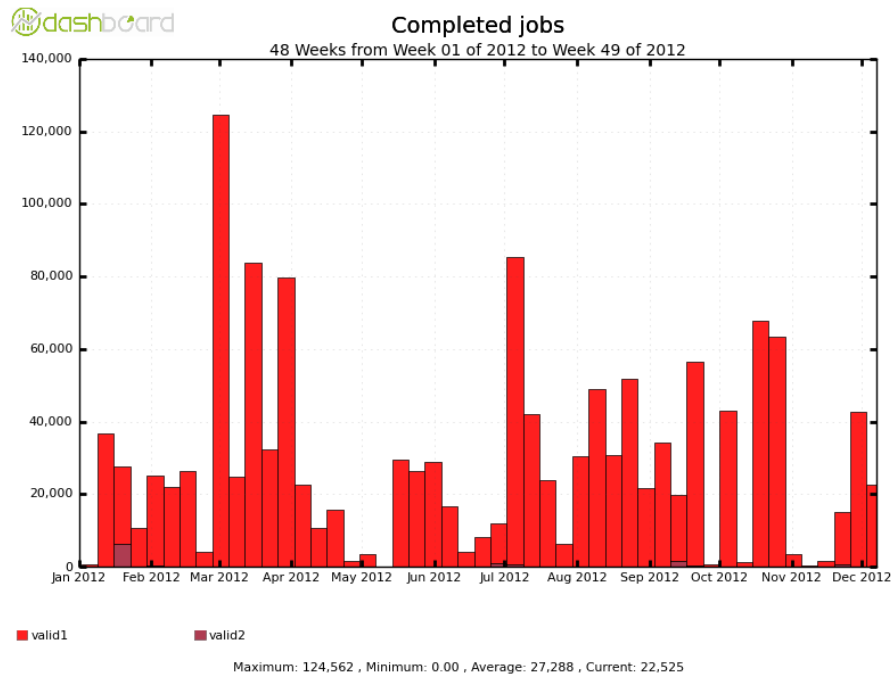


Validation jobs increase for main deliveries:  
MCXXa, MCXXb,... and data reconstruction release

- Around 70 validations tasks during 2012.
- Typically each of them means a Sample A, which corresponds to 5.5 TB for digi+reco and 1.3 TB for simulation.

# Physics Validation CPU in 2012

J. E. Garcia



## CPU Consumption through 2012

Validation jobs are not constant in time.

Validation schedule has peaks and holes.

2012 is probably a good representative of of what may happen next year.

# Challenges for Run2

- Several challenges are awaiting us for the restart of operations after LS1.
- Pileup conditions in 2015 will depend mainly on the bunch spacing.
  - For 50 ns, it would be in the range of 40-80 for  $1-2+ \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ .
  - For 25 ns, it would be in the range of 25-50 for  $1-2+ \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ .
- We have to be prepared to the worst case scenarios for 2015.
  
- The Trigger rate should be close to read-out saturation, as usual. Possibly at the level of 1 kHz output passed to a 4<sup>th</sup> level of trigger at Tier0, where events will be filtered further (to be decided).
  
- It is necessary to make software and computing faster, more scalable and reduce the memory footprint.
  
- ATLAS Streaming and data replication could also be revisited during LS1.

# Conclusions

- WLCG was crucial to deliver physics rapidly during the first years of LHC data.
- 2012 was an exciting year: high pileup, high trigger rates and stringent deadlines put the computing and DP under strain.
- Reprocessing of 2012 data happened in parallel of data taking.
  - Main issue was data space available at Tier1 and sharing of resources with MC production.
- Currently there are no firm plans for reprocessing and physics validation campaigns during LS1:
  - assume requests will be  $\leq 2012$ .
- LS1 is the only opportunity in many years for disruptive changes in the ATLAS work-flow model.
- Restart of data-taking will bring new challenges for DP and computing with possibly much higher pileup and trigger rates.

# *Support Material*



## **Inner Detector**

ID alignment

B-layer conditions problems for part of period A

TRT fix for the run in C8 which had the wrong dead straw list uploaded

## **LAr Calorimeter**

**FCAL HV calibration corrections** [*Improve 1.4/fb*]

Problems with HV power supply system of FCAL channels for which HV readings were unreliable. Affects HV scale determination, resulting in ~10% energy shift in 1/4  $\phi$ -region;

- Need correct HV reading which required new measurement of faulty modules
- Use corrected HV readings to recompute scale corrections
  - ... applied at reconstruction levels (when building raw channels)

**Cell-level OFC timing corrections**

Cell-level timing corrections to be applied to the OFC-reconstructed time to improve LAr timing.

-- Gain of 1-2%.

**LAr noise mask and event veto** [*recover 55/pb*]

Apply proper noisy cell masking and noise burst vetos for a couple of runs in A4 and A8

## **Muons**

### **Calibrations**

MDT calibration update for B1-B5 (correct a wrong flag in DB)

Calibration streams were off for some runs in period D

### **Alignment**

- Barrel+Endcap alignment now includes EC alignment from I-lines (CSC internal alignment)
- Using new Muon Spectrometer layout, ie ATLAS-GEO-20-00-02

## **Tile Calorimeter**

Fix for the hot cell affecting B1+B2

Timing for one run in B4.

## **AODFix**

AODFix currently applied to the MC production :

- new  $e/\gamma$  isolation, reurun track isolation for  $\mu$ , update of  $\tau$ ID,  $\tau$ -  $\mu/e$  veto,  $\tau$ -energy scale, MetRef bugFix

Applied those changes at the RECO level for data to save significant time in D3PD making

- Validated by comparing AODFix vs reconstruction

## Bulk reprocessing

Phase II of the reprocessing :

- using the Beam spot & Conditions derived in ES
- using latest Conditions (Lar and Tile updates)

Streams reprocessed:

- all main physics streams, *then* JetCalibDelayed streams

Reprocessing order :

- Period D, B, (highest stat) then A, C, E and G
- **No ESD kept for main physics streams (keep only B11+D3)**

## Bulk reprocessing

*Physics streams:*

- physics\_CosmicMuons
- physics\_CosmicCalo
- physics\_Egamma
- physics\_Muons
- physics\_JetTauEtMiss
- physics\_Background
- physics\_MinBias
- physics\_IDCosmic
- physics\_ZeroBias
- express\_express
- debug\_hltacc
- physics\_standby

+ *Delayed streams:*

- physics\_jetCalibDelayed
- physics\_HadDelayed
- physics\_Bphysics

## Bulk reprocessing outputs : “Life w/o ESD”

### 1) *Reconstruction step (r-tag)*

- AOD, HIST, TAG , DRAW
- DESD (stream-dependent):
  - [MinBias] DESD\_MBIAS, DESD\_SGLMU, DESDM\_MET
  - [EGamma] DESD\_EGAMMA, DESD\_PHOJET, DESD\_SGLEL, \_RPVLL
  - [Muons] DESD\_SGLMU, DESD\_RPVLL
  - [JetTauEtMiss] DESD\_CALJET, DESDM\_TRACK, DESDM\_MET, \_RPVLL

### 2) *Merging of output formats (p-tag)*

- AOD, HIST, TAG and DESDs will be merged

### 3) *Production of DAOD and DESD (p-tag)*

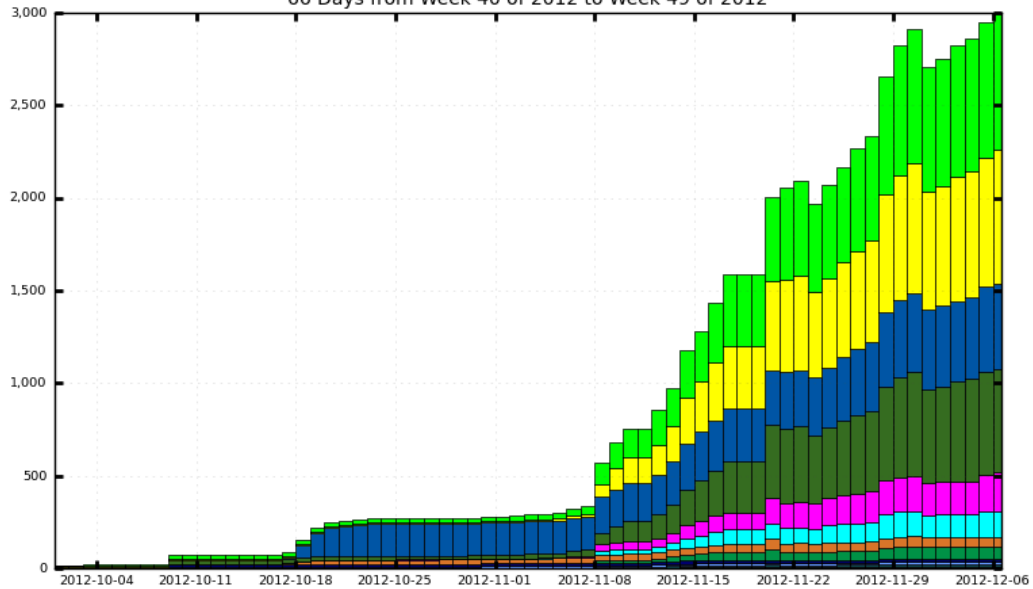
- DAOD\_ZLL, DESD\_ZLL from DRAW

## DRAW reprocessing

*Physics streams:*

- Muons and Egamma streams
- Required by RecoPerf

### Number of Physical Bytes (in TBs) 66 Days from Week 40 of 2012 to Week 49 of 2012



- physics\_muons
- physics\_jettauetmiss
- express\_express
- physics\_egamma
- physics\_minbias
- physics\_standby
- physics\_cosmiccalo
- physics\_background
- physics\_haddelayed
- physics\_bphysics
- physics\_zerobias
- physics\_cosmicmuons
- physics\_idcosmic
- physics\_jetcalbdelayed
- debugrec\_hltacc
- debugrec\_hltrej
- debugrec\_crashed
- physics\_zerobiasoverlay

Maximum: 2,997 , Minimum: 18.23 , Average: 928.73 , Current: 2,997

### WallClock consumption in seconds

