

# Data processing at the CERN Tier-0/CAF: status and plans

**DataOps**, David Evans, Stephen Foulkes,  
**Stephen J. Gowdy**, Oliver Gutsche,  
Dirk Hufnagel, Peter Kreuzer, David Mason,  
Simon Metson, Michael Miller, Seangchan Ryu

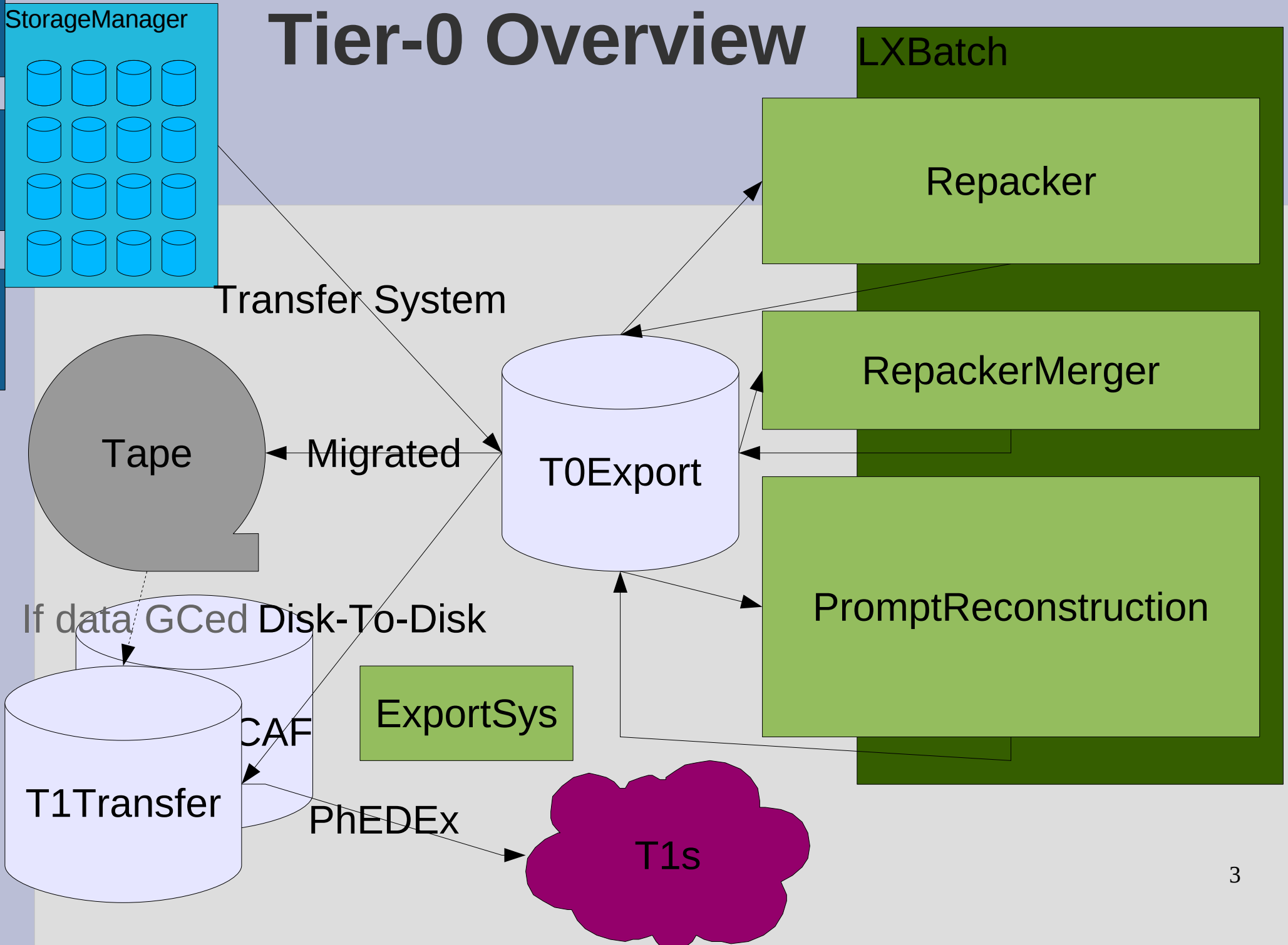
25<sup>th</sup> September 2008



# Overview

- Tier-0 Status
- *Middle bits*
- CAF Status
- Tier-0 Plans
- CAF Plans

# Tier-0 Overview



# Transfer System

- Copies data from Cessy to Tier-0
  - Storage Manager (SM) is told when data is copied and can delete it (Safe99)

SM Page 1

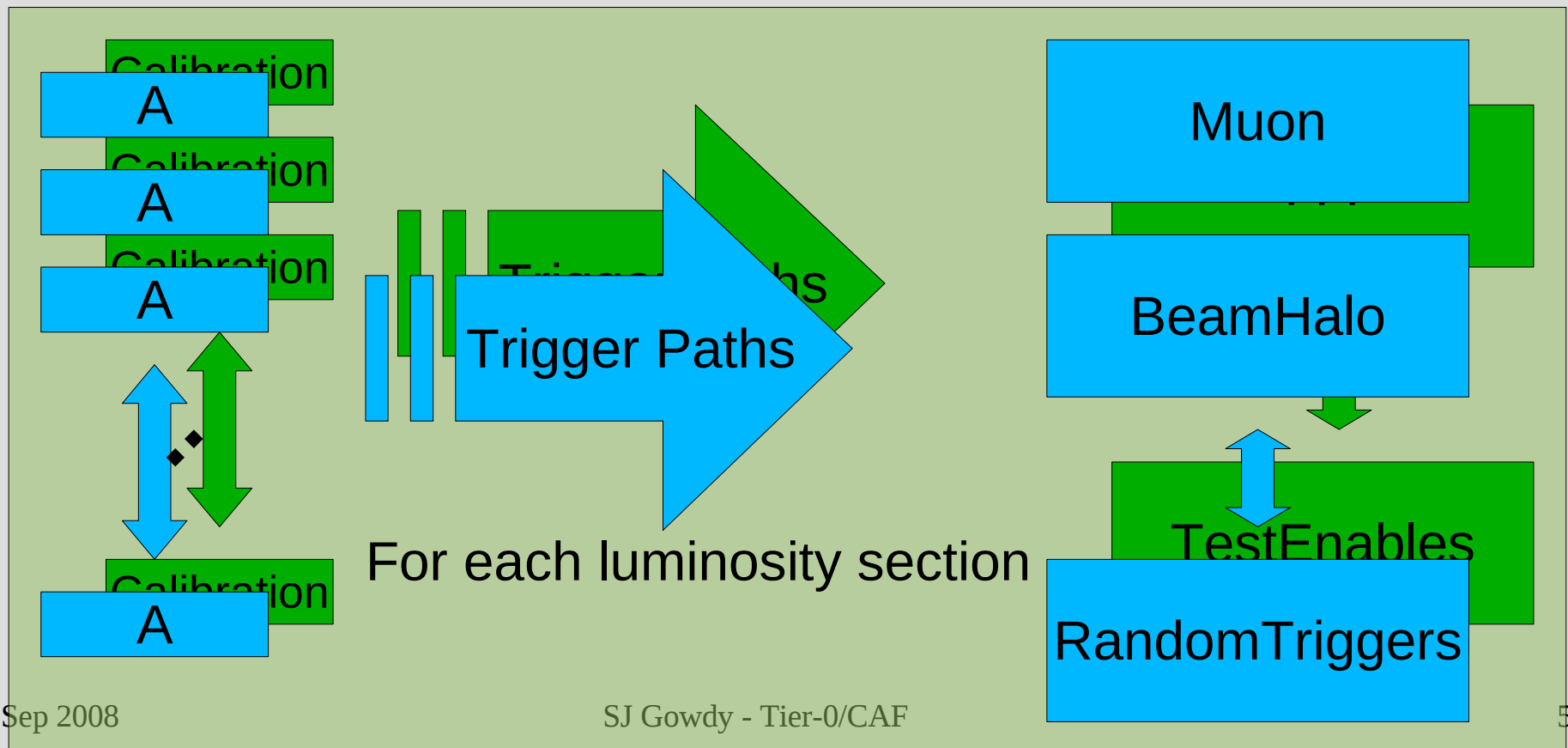
Run Number	Start Time	Setuplabel	Size [GB]	Events	Files	Disk [MB/s]	Tier0 [MB/s]	Open	Closed	Safe-0	Safe-99	Deleted
63463	2008/09/20 15:53	GlobalBeamCommissioning08	285.27 GB	5792808	744	33.98	34.34	0	744	744	744	744
63462	2008/09/20 15:00	GlobalBeamCommissioning08	142.79 GB	1745285	220	58.79	59.71	0	220	220	220	220
63459	2008/09/20 13:38	GlobalBeamCommissioning08	140.88 GB	2856134	368	33.79	35.49	0	368	368	368	368
63451	2008/09/20 12:38	GlobalBeamCommissioning08	7.52 GB	1379978	208	3.3	30.45	0	208	208	208	208
63446	2008/09/20 10:59	GlobalBeamCommissioning08	295.61 GB	2703343	468	85.39	37.1	0	468	468	468	468
63441	2008/09/20 08:36	GlobalBeamCommissioning08	600.23 GB	5507511	935	85.16	58.2	0	935	935	935	935
63440	2008/09/20 03:39	GlobalBeamCommissioning08	822.31 GB	7760400	1212	61.35	52.02	6	1206	1204	1204	1204
63436	2008/09/20 02:47	GlobalBeamCommissioning08	51.6 GB	694290	112	42.82	46.72	0	112	112	112	112
63419	2008/09/20 01:38	GlobalBeamCommissioning08	0 GB	6	6	1.43	.14	0	6	6	6	6
63393	2008/09/19 23:48	GlobalBeamCommissioning08	98.74 GB	182833	129	345.08	26.33	4	125	123	123	123

- Now also passes some information on online release (**sometimes overridden**) and HLT key
  - Previously from cmsmon, for a week was manual

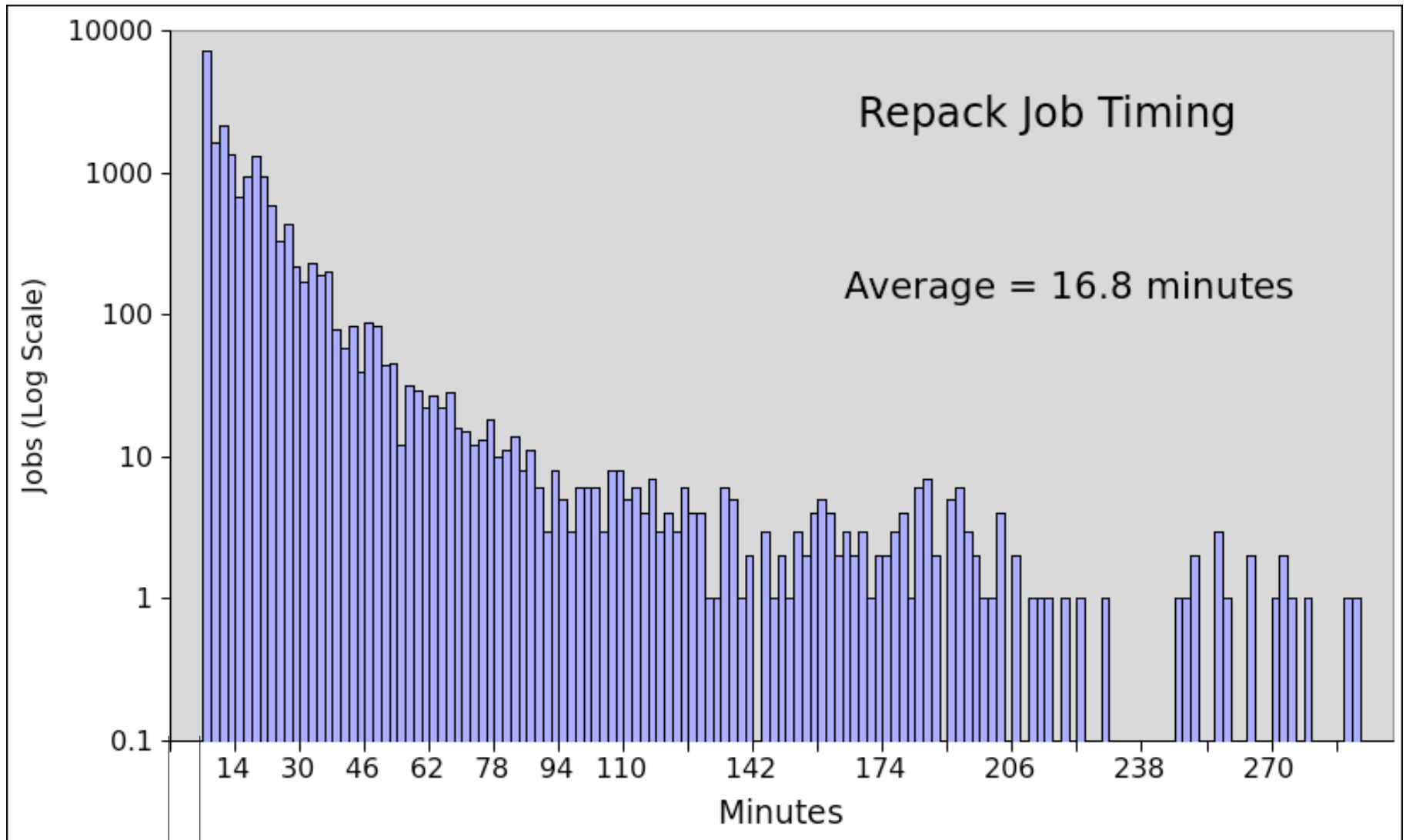
Framework  
Validation  
Required

# Repacking

- “Temporary” streamer files have Primary Datasets (PD) extracted



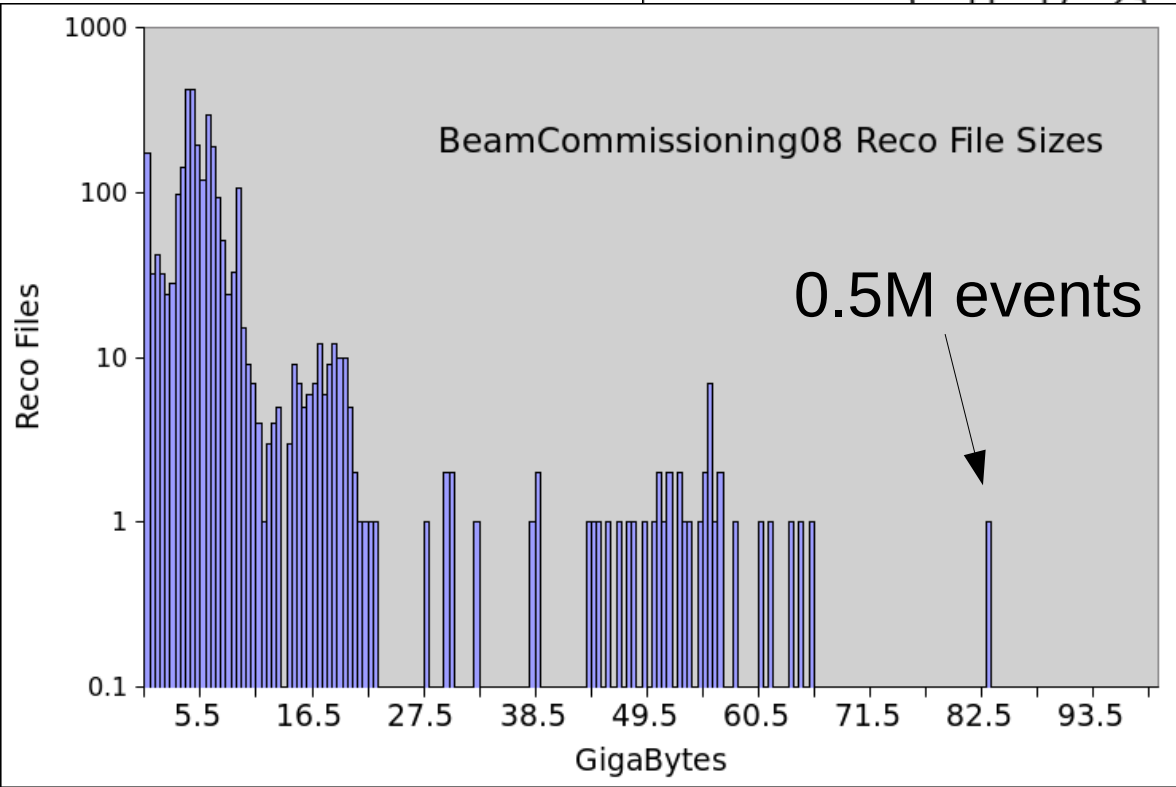
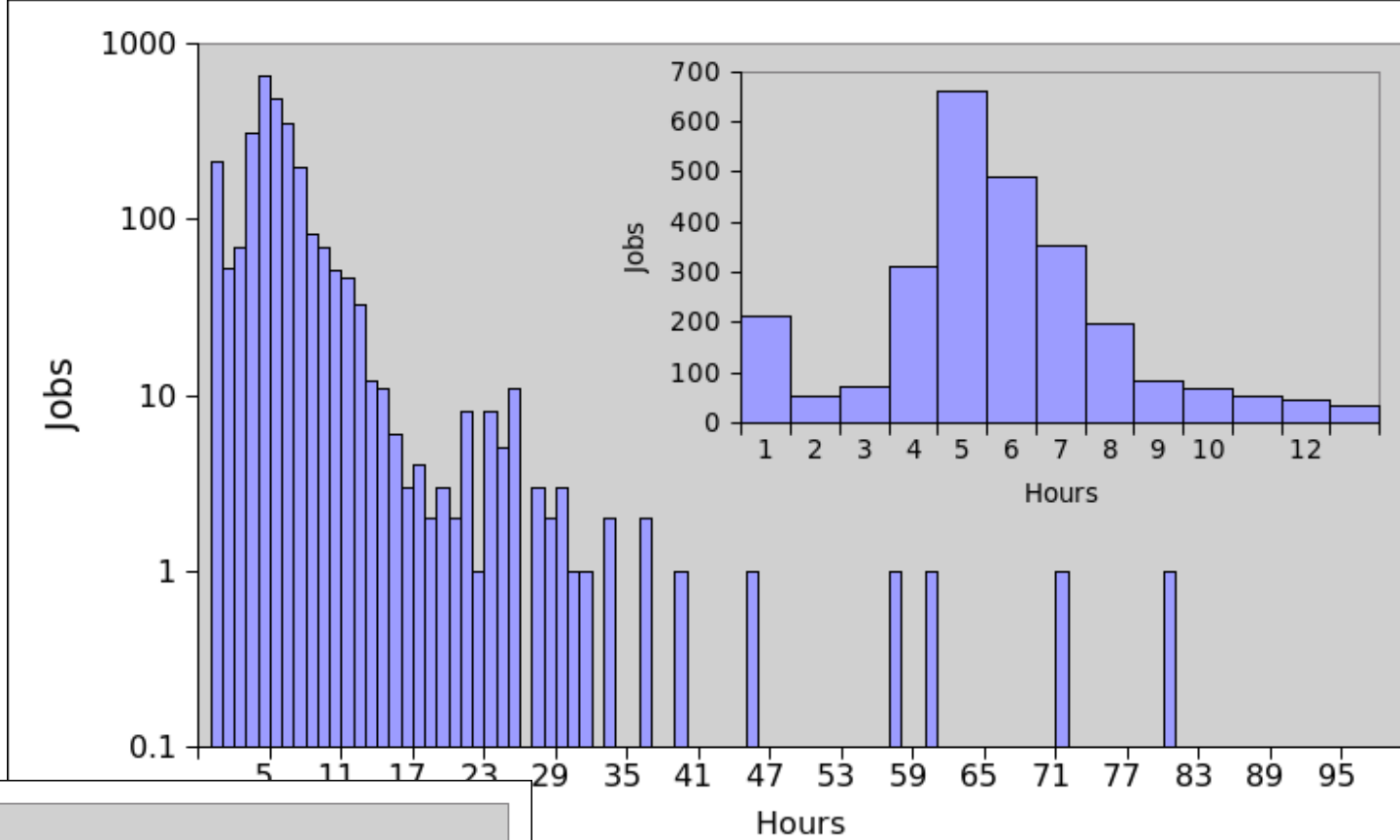
Currently no check on complete luminosity section arriving at Tier-0



→ ← ~7 minutes minimum time due to overheads

# PromptReconstruction

- Certain datasets are configured to get reconstructed
  - BeamHalo, ZeroBias, TriggerTest, Cosmics, BarrelMuon, EndcapsMuon
- Configuration is currently from CVS
  - Different configurations for magnet on and off
    - **Not possible to re-reconstruct with two configs**
  - Changing release not completely synchronised with config change
  - Not well controlled, but (**should** be) stored in DBS
- Till last weekend attempt to reco lumisection
  - Logging single PD at 10kHz == long or **failed** job



Deployed max events  
per RAW file  
50k  $\approx$  500Hz  
Splits up lumi sections so  
**don't trust** luminosity in  
these cases

**Need sensible Primary  
Dataset definitions**

# Export System

- Registers files with PhEDEx & Tier-0 DBS
  - Can be transferred straight away
- Added to global DBS after block closes
- Closes block for each run
  - Or when max number of files reached (unusual)
- Few troubles initially as it didn't register the Storage Element in DBS
  - Meant data had to be on tape before CAF folk could use it
- **Currently doesn't store job config in DBS**

T0 status @ Wed Sep 24 15:40:47 2008

Runs [count=323, query=35ms, analysis=471 ms]												
All 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33												
Run ID	Start	End	Process	Version	Reco version	Global tag	Status	Streamers	Repacked	Merged	Reconstructed	
63463	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	744	469 (100%)	138 (100%)	43 (100%)	
63462	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	220	131 (100%)	76 (100%)	2 (7%)	
63459	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	368	215 (100%)	68 (100%)	22 (100%)	
63451	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	CloseOutPromptReco	208	100 (100%)	7 (100%)	0 (0%)	
63446	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	468	195 (100%)	128 (100%)	35 (92%)	
63441	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	935	368 (100%)	260 (100%)	71 (91%)	
63440	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	1204	575 (100%)	440 (100%)	123 (84%)	
63436	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	112	51 (100%)	22 (100%)	7 (100%)	
63419	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	6	3 (100%)	3 (100%)	1 (100%)	
63393	🕒	🕒	FU	CMSSW_2_1_7	CMSSW_2_1_9	CRUZET4_V5P::All	Complete	123	28 (100%)	17 (100%)	5 (100%)	

Repack stats [10ms]	
New jobs	0
Used jobs	0
Successful jobs	19351
Failed jobs	52
Abandoned jobs	0

Merge stats [7ms]	
New jobs	0
Used jobs	0
Successful jobs	4124
Failed jobs	0
Abandoned jobs	0

Reco stats [7ms]	
New jobs	4
Used jobs	0
Successful jobs	2723
Failed jobs	90
Abandoned jobs	7

Repack running jobs [count=0, query=10ms]				
Job ID	Run ID	Running time	Streamers	Stream

Merge running jobs [count=0, query=6ms]				
Job ID	Run ID	Running time	Repacked	Dataset

Reco running jobs [count=4, query=7ms]				
Job ID	Run ID	Reco time	Merged	Dataset
2723	63451	4d 1h 20m 25s	1	BeamHalo
2372	62904	4d 23h 58m 6s	1	Cosmics
2371	62904	4d 23h 58m 6s	1	Cosmics
2370	62904	4d 23h 58m 6s	1	Cosmics

# Monitoring Tier-0: T0Mon

- Looks into T0AST Database for information
  - Tier 0 Activity State Tracker
- Next version allows complete history views
  - Jobs or files

Repack job 19047 history [[Back to main page](#)]

Streamer files [count=4, query=4ms]							
Id	Lumi Id	Filesize	Events	Name	E	S	Lfn
62726	14	41.84 MB	7077	A	0	0	
62670	14	41.75 MB	7158	A	0	0	
62576	14	40.25 MB	6881	A	0	0	
62568	14	41.62 MB	7103	A	0	0	

Repacker jobs [count=1, query=7ms]						
Job ID	Run ID	Definition time	Streamers	Stream	Status	
19047	63451	2008-09-20 13:52:38	4	A	Success	

Repacked files [count=3, query=75ms]						
Id	Filesize	Events	Status	Export status	Dataset	Lfn
27722	111.52 MB	26478	Merged	NotExportable	Muon	
27721	111.52 MB	26478	Merged	NotExportable	BeamHalo	
27720	4.79 MB	1741	Merged	NotExportable	RandomTriggers	

Merger jobs [count=3, query=15ms]						
Job ID	Run ID	Definition time	Repacked	Dataset	Status	
3894	63451	2008-09-20 14:11:55	5	BeamHalo	Success	
3893	63451	2008-09-20 14:11:55	22	RandomTriggers	Success	
3892	63451	2008-09-20 14:11:55	5	Muon	Success	

Merge files [count=3, query=195ms]							
Id	Filesize	Events	Status	Export status	Dataset	Lfn	
27764	558.10 MB	134212	ReconstructionQueued	InFlight	BeamHalo		
27762	558.10 MB	134212	ReconstructionDisabled	InFlight	Muon		
27761	100.97 MB	37220	ReconstructionDisabled	InFlight	RandomTriggers		

Reco jobs [count=1, query=7ms]						
Job ID	Run ID	Definition time	Merged	Dataset	Status	
2723	63451	2008-09-20 14:20:22	1	BeamHalo	New	

Reco files [count=0, query=6ms]						
Id	Filesize	Events	Status	Export status	Dataset	Lfn

# Middle bits

- Currently run by cron driven ProdAgent
- AICaReco generation
  - run independently of Tier-0, watches Global DBS
    - **Adds latency**, only complete runs in Global DBS
  - Two sets of jobs for each of Cosmics and BeamHalo Primary Datasets (and for each Processing Era) producing five outputs
    - TkAICosmics0T, TkAICosmics, MuAlStandAloneCosmics, MuAlBeamHaloOverlaps, MuAlGlobalCosmics
- DQM Harvesting
  - Run on Reco and AICaReco output

# CAF Status

- 1.4MSpecINT2k (696 batch slots)

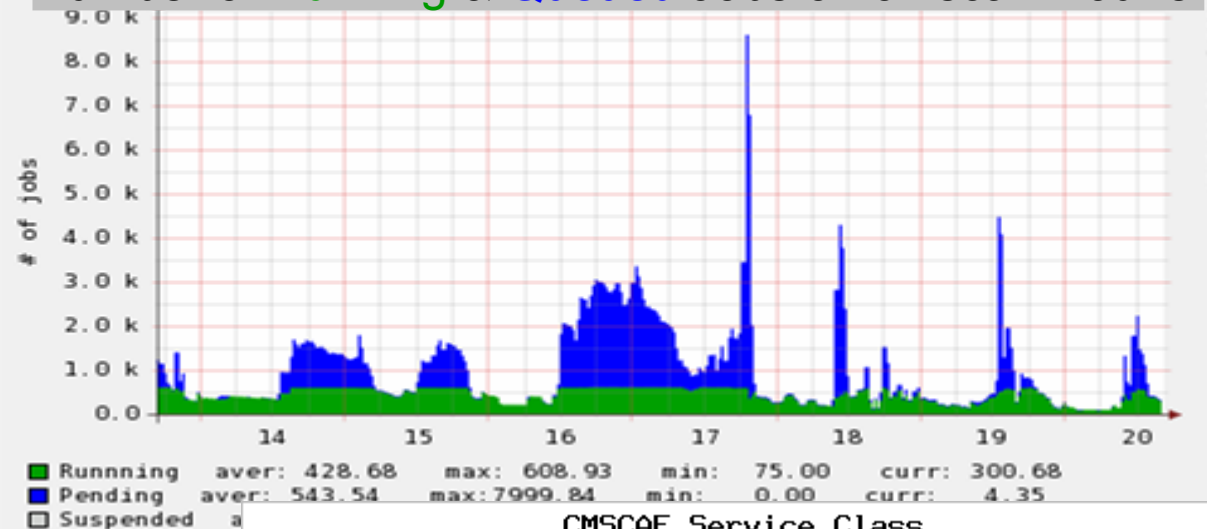
- Fully occupied a few times in June
- Low CPU Efficiency of jobs (mostly I/O?)
- 40:40:20 fairshare between

Commissioning,  
AlCa & Physics

- 854TB (400TB free)

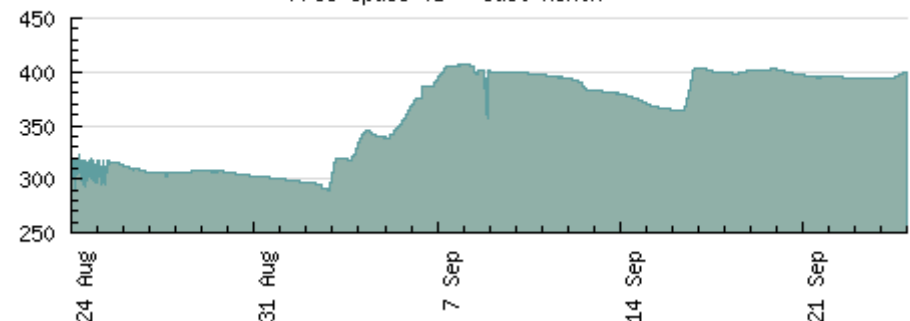
- Heavy clean up recently
- Agree 50:10:40 split

Number of **Running** & **Queued** Jobs on cmscaf in June



CMSCAF Service Class

Free Space TB - last month



# Tier-0 Plans

- Some system needs to decide when a lumi-section is complete to specify # files
  - Storage Manager or Tier-0 (latter would add time)
- Multiple jobs to process a single RAW file
  - Reduce RECO latency at cost of extra merge
  - Should be deployed end of October
- New effort for monitoring system
  - Collecting requirements
- Need configuration system
  - Primary Datasets require different reco configs
  - AlCaReco requires different configs
- AlcaReco production & PromptCalib system

# CAF Plans

- Monitoring for major stakeholders on use
  - Enable educated decisions for allowing data to be added and if their user's are using the expected level of resources Simon's talk
- Move user data to new 50TB CASTOR pool
  - Only allow PhEDEx to write to large existing pool my talk
- CRAB Server to be deployed
  - Testing multiple user support (currently all jobs are run under user “spiga”)
  - Will eventually provide automatic job submission as new data arrives on CAF Daniele's talk

# Summary

- Many parts of Tier-0 system deployed in the last several months
- System designed around Computing Model
  - Has been stretched by current and proposed usage (eg one Primary Dataset for all data, Open Trigger data taking, who defines complete luminosity sections)
- Using double RAW size on tape
  - Must validate cmsRun repack process
- Need to evaluate plans given new data taking expectations

# BACKUP

# Dataset Names

- Based on a few elements
  - Primary Dataset (PD)
  - Processing Era (PE): defined ad-hoc currently
  - Processing Info (PI): information on processing
  - Processing Version (PV): usually starts as -v1
  - Data Tier (DT): RAW, RECO, ALCARECO, (AOD)

**/<PD>/<ProcEra><PI><ProcVer>/<DataTier>**

- Some recent datasets

/BeamHalo/BeamCommissioning08-v1/RAW

/BeamHalo/BeamCommissioning08-PromptReco-v1/RECO

/BeamHalo/BeamCommissioning08\_CRUZET4\_V5P\_StreamALCARECOTkAICosmics\_v3/ALCARECO

# Logical File Names

- Built out of the same aspects
    - <SPLIT> is way to split files up among directories, two methods
      - some include run number broken across three path elements, NNN/NNN/NNN
        - Run 62442 would be 000/062/442
      - Others just a counter of job number etc
        - With AlcaReco it is four digit number like 0013
    - UUID is automatically generated
      - eg 88034D41-7780-DD11-A3F6-000423D6A6F4
- /store/data/<PE>/<PD>/<DT>/<PI><PV>/<SPLIT>/<UUID>.root*

- Recent examples

*/store/data/BeamCommissioning08/BeamHalo/RAW/v1/000/062/399/FAC7D166-1183-DD11-8B4C-000423D98B6C.root*  
*/store/data/BeamCommissioning08/BeamHalo/RECO/v1/000/062/236/88034D41-7780-DD11-A3F6-000423D6A6F4.root*  
*/store/data/BeamCommissioning08/BeamHalo/ALCARECO/CRUZET4\_V5P\_StreamALCARECOTkAlCosmics\_v3/0011/5CB91478-E985-DD11-8C0A-001617E30D12.root*