Resource utilization of the ATLAS High Level Trigger during 2010 and 2011 LHC running

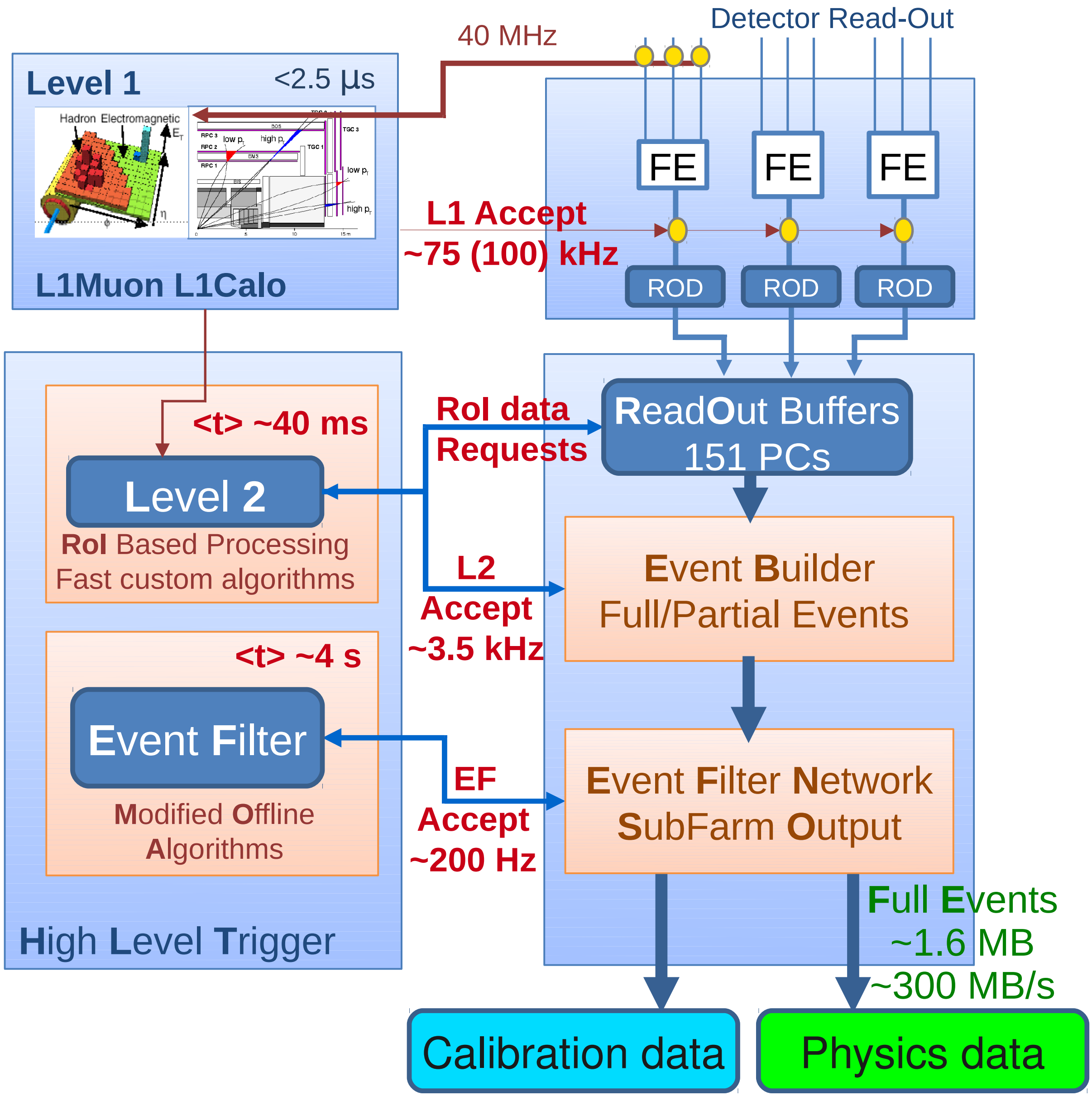
Rustem Ospanov for the ATLAS collaboration
University of Pennsylvania

Overview

- Trigger and DAQ design
- Trigger signatures and High Level Trigger (HLT)
- Detailed HLT monitoring data
- Resource utilization by HLT
- Trigger evolution with increasing LHC luminosity
- Conclusions

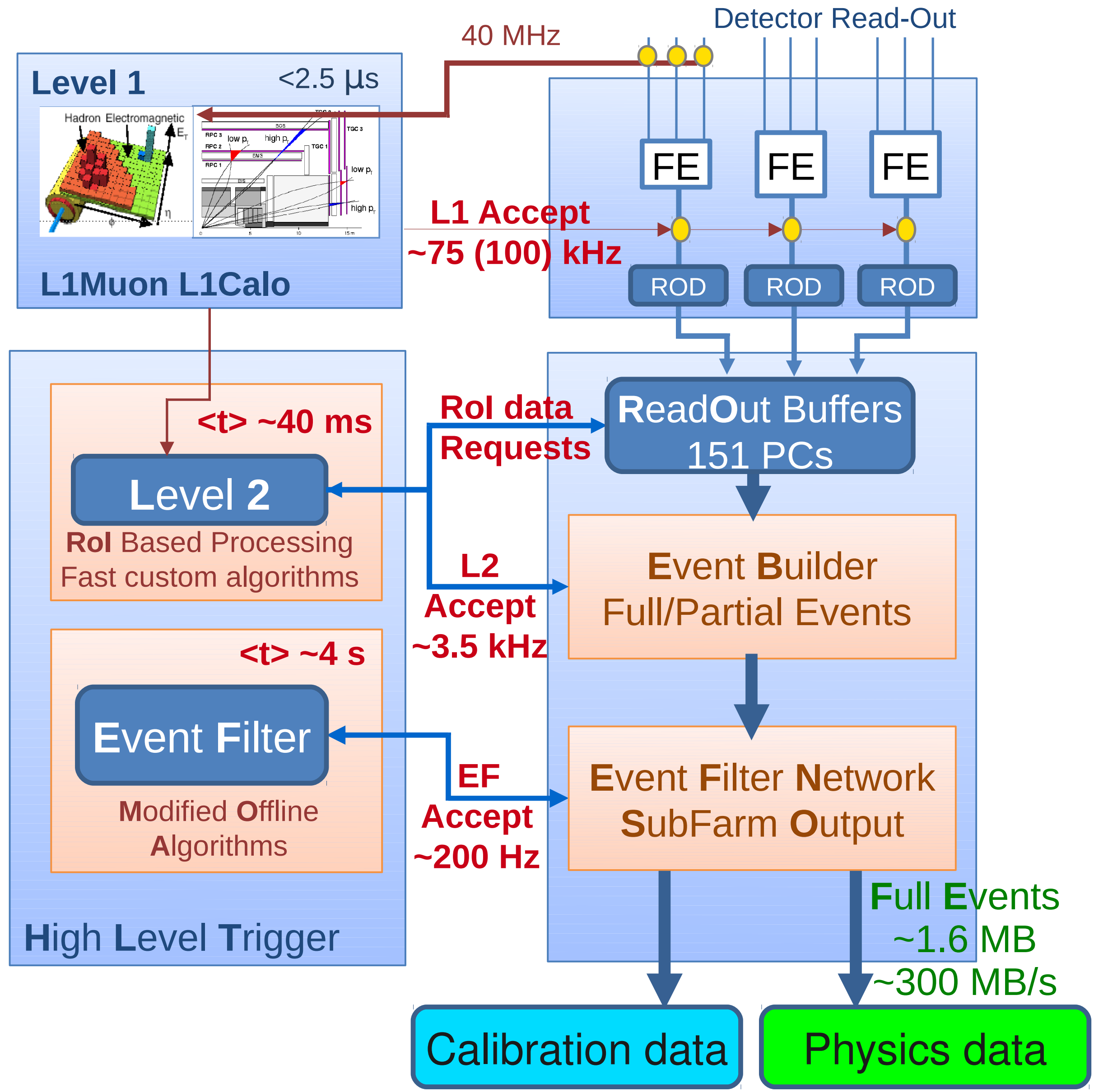
Trigger and DAQ Design

- Level 1 (L1)
  - Hardware based triggers at 40 MHz
  - Select Regions of Interest (RoI) with lepton, photon and jet candidates
  - Transverse missing energy (MET)
  - See talk by Martin Wessels
- Level 2 (L2)
  - Custom software algorithms
  - Read ~2% of detector data
  - Confirm L1 candidates within RoI at full detector granularity
- Event Filter (EF)
  - Modified offline algorithms
  - Read full detector data
  - Confirm L2 candidates within RoI
  - Search for new candidates
    - Full scan reconstruction for EF jets

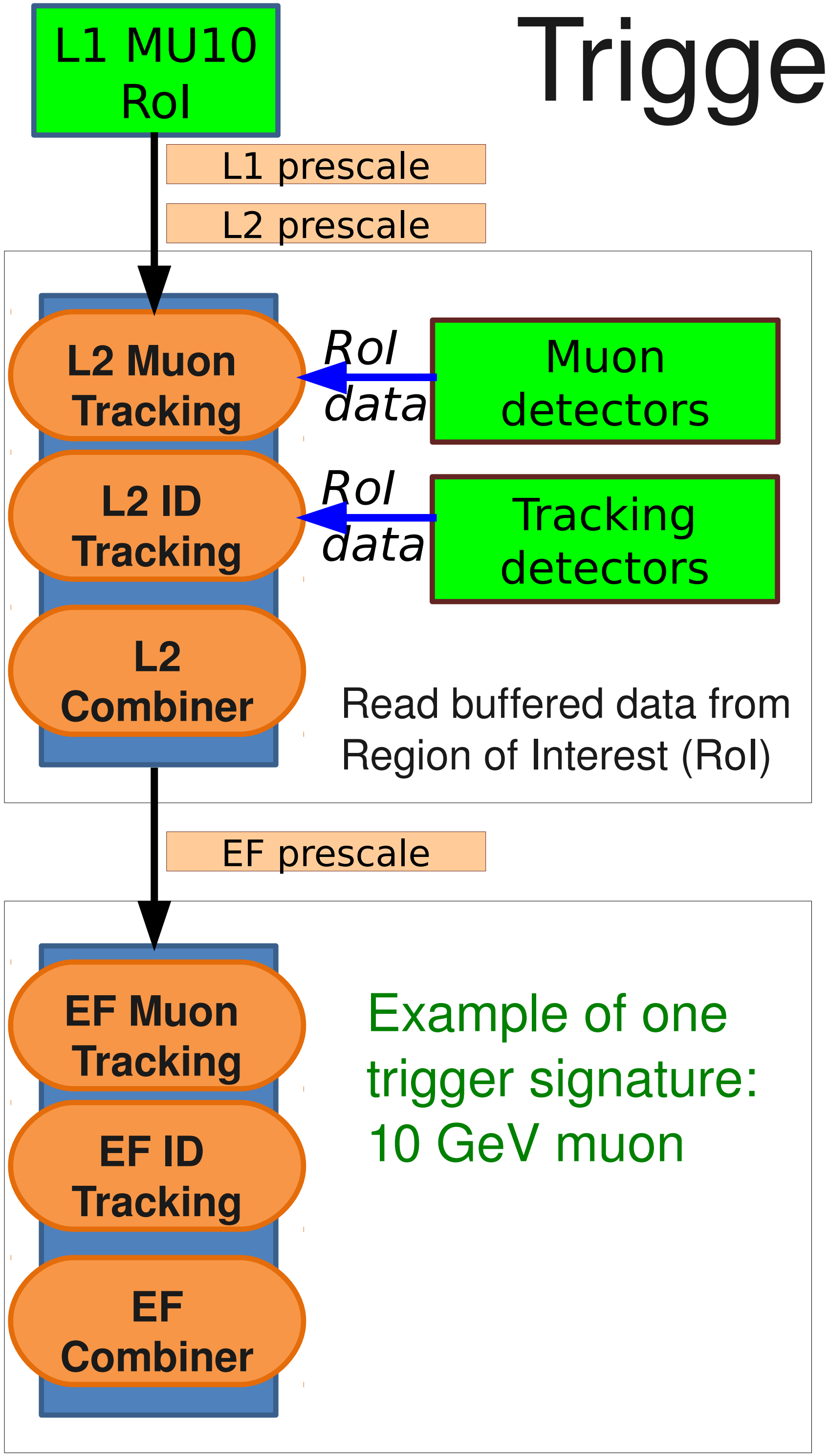


Trigger and DAQ Design

- TDAQ Design
  - Shown are design limits @1e34
  - Typical rates so far:
    - L1 accept up to 50 kHz
    - L2 accept up to 5 kHz
    - EF accept 300-400 Hz
  - See talks by Sergio Ballestrero and Srini Rajagopalan
- Limits:
  - Event rate at L1 and L2
  - CPU load for L2 and EF
  - L2 data request rate
- Exceeding these limits results in deadtime and loss of efficiency
  - Output rate to permanent storage is limited by offline resources
- Anticipate luminosity increases
  - Predict rates and L2 data requests
  - Optimize trigger selections and resource usage



Trigger signatures

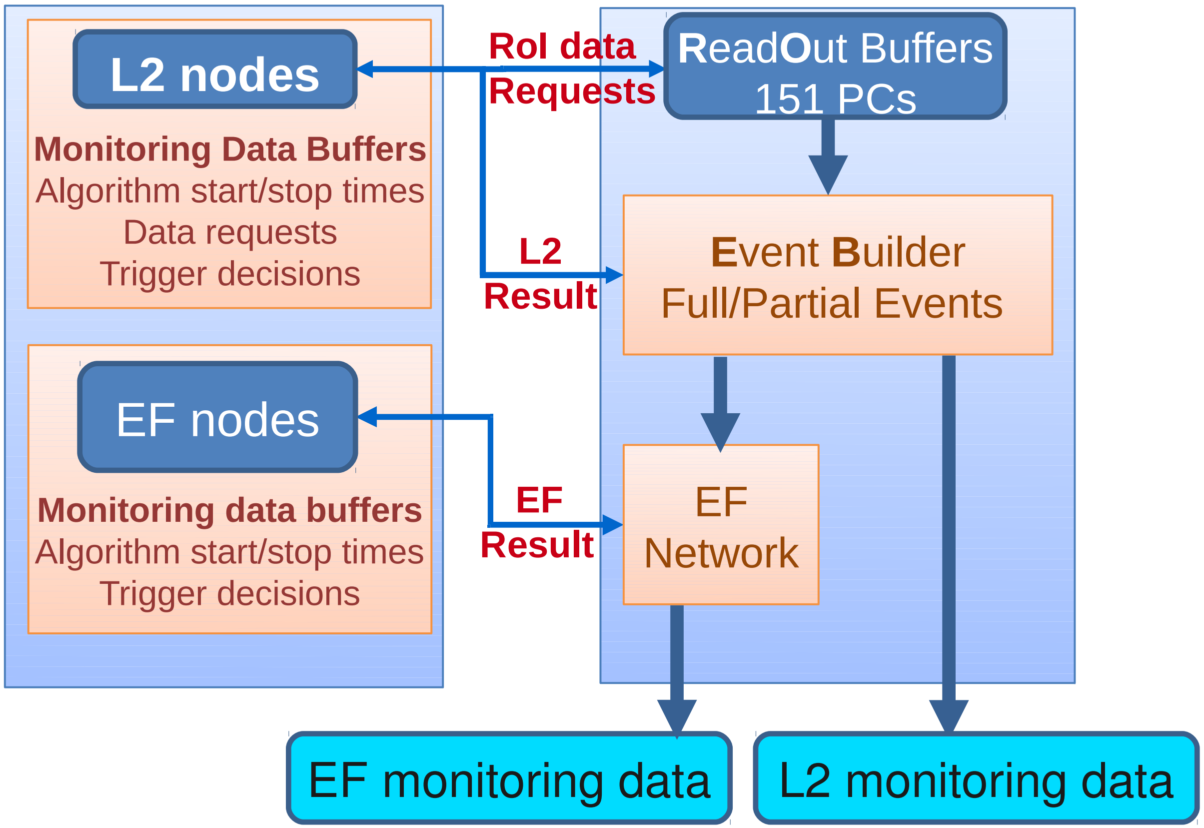


- ATLAS trigger menu contains ~300 physics trigger signatures
  - Leptons, photons, jets and transverse missing energy plus combinations
  - Signatures contain multiple algorithms
  - Algorithms are shared with results cached and reused
- Trigger signatures incur resource costs
  - Rates at three trigger levels
  - CPU utilization at L2 and EF
  - RoI data request rate at L2
- Steering framework for High Level Triggers
  - Seed trigger signatures
  - Execute algorithms
  - Collect results and monitoring data
- Prescaling mechanism
  - Executes algorithms for a fraction of input events
  - Controls trigger rates

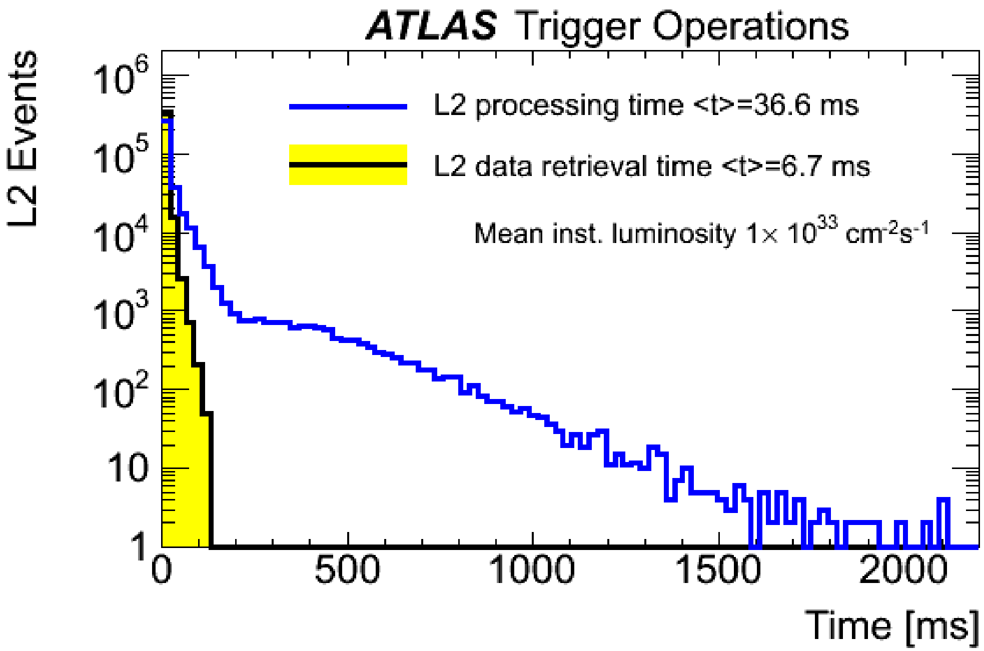
Detailed HLT monitoring data

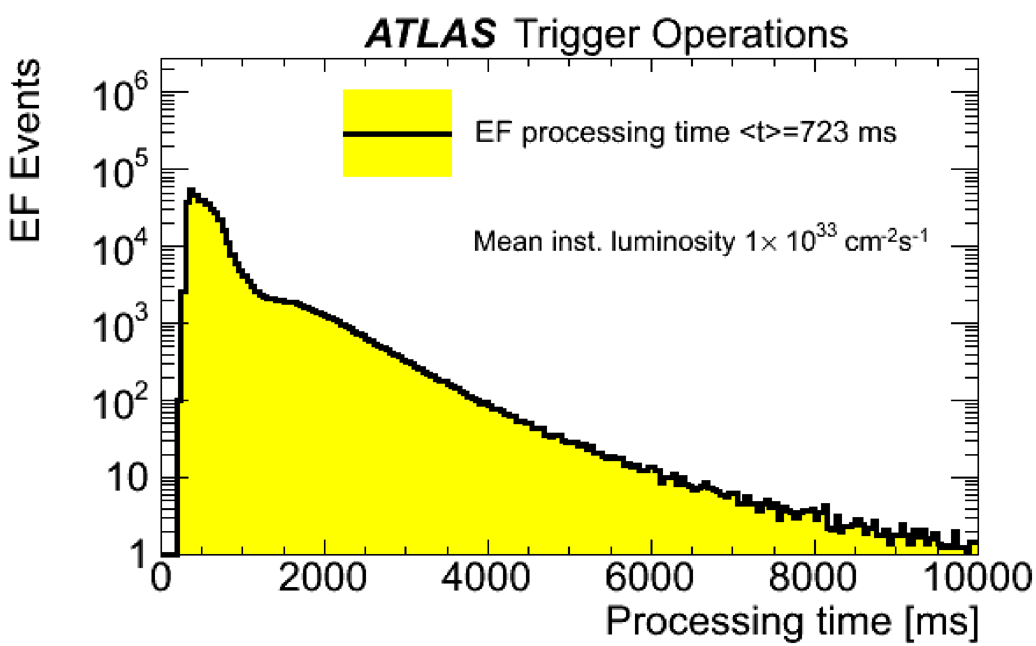
- Cost monitoring data
  - Each event
    - Event id
    - L1 and HLT decisions
  - Every 10th event
    - RoI positions
    - Start/stop times for execution of HLT algorithms
    - L2 data request parameters
- Monitoring data are recorded via DataFlow system using otherwise rejected events
- Partial Event Building
  - Remove detector data payload
  - Only record L2 and EF results with monitoring data

- HLT steering framework is instrumented to record detailed cost monitoring data
  - Step by step records of HLT execution



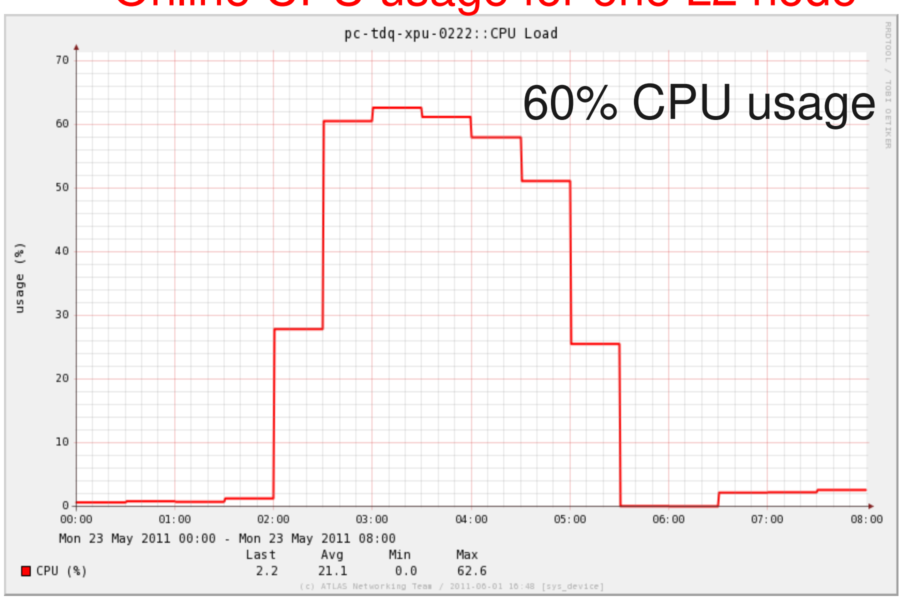
HLT CPU resource monitoring





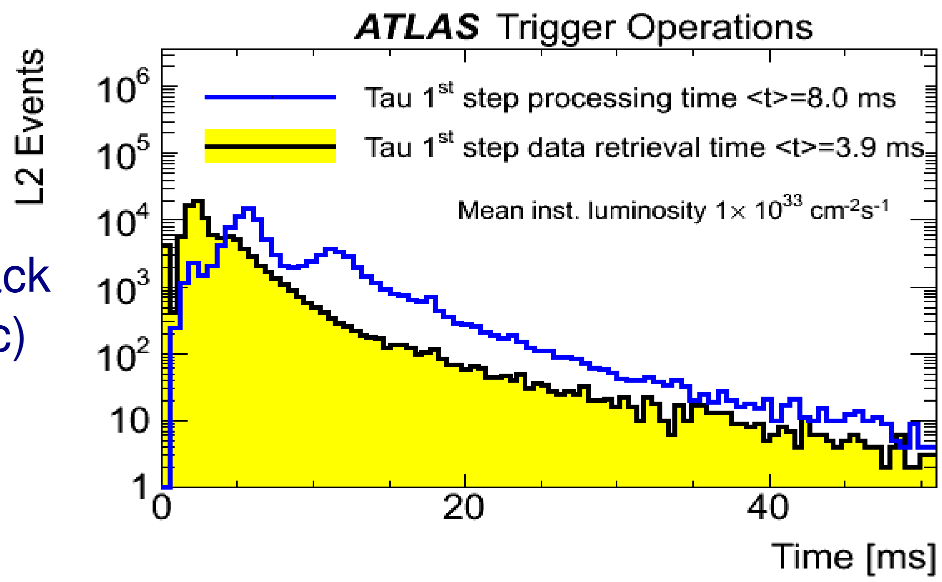
- L2 and EF execution times are within TDAQ design limits
- Real time online monitoring of entire HLT infrastructure:
  - CPU loads and memory usage
  - Network bandwidth
- Detailed offline monitoring of individual trigger signatures:
  - Breakdown total CPU usage by algorithms and signatures

Online CPU usage for one L2 node



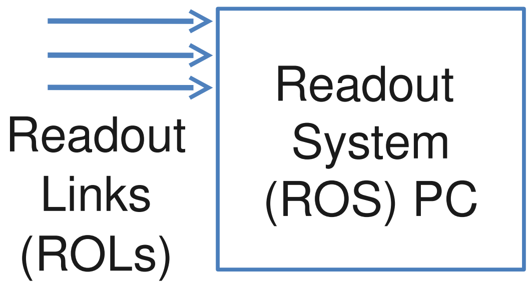
Resource utilization by algorithm

- Rank algorithm by resource utilization
  - CPU usage per event and total
  - Data request time, size and frequency
- CPU usage for L2 trigger groups @1e33:
  - ~60% for triggers requiring inner detector track pattern recognition (di-muons, beamspot, etc)
  - ~40% all others: muons, electrons, photons, taus, jets, combinations, etc
  - Nontrivial overlaps because of caching
  - Resource priority depend on luminosity



Algorithm	Events	Calls	Calls > 500 ms	Event rate (Hz)	Call rate (Hz)	Alg time/call (ms)	Alg time/event (ms)	Total alg time (s)	Total alg time (%)	ROS request rate (Hz)	ROS retrieval rate (Hz)
TrigSiTrack_BeamSpot	49	49	9	940	940	323.49	323.49	15.9	16.12	18804	18804
TrigIDSCAN_FullScan	36	36	2	691	691	282.02	282.02	10.2	10.32	25404	25270
TrigSiTrack_Tau_robust	202	302	0	3876	5795	27.07	40.48	8.2	8.31	53591	42040
TrigIDSCAN_Muon	511	756	0	9805	14506	9.54	14.12	7.2	7.34	100869	55893
T2CaloTau_Tau	760	1119	0	14583	21471	6.06	8.92	6.8	6.89	877235	223093
TrigIDSCAN_Bphysics	157	247	0	3012	4739	23.43	36.87	5.8	5.89	64624	32408
T2CaloEgamma_eGamma	1139	2143	0	21855	41119	2.26	4.26	4.8	4.93	735458	169311
TrigSiTrack_Jet	135	316	0	2590	6063	14.29	33.44	4.5	4.59	37646	21586
TrigSiTrack_FullScan	14	14	2	269	269	292.92	292.92	4.1	4.17	10189	2955
TrigIDSCAN_BeamSpot	49	49	0	940	940	73.38	73.38	3.6	3.66	18804	0
muFast_Muon	1119	1488	0	21471	28551	2.38	3.17	3.5	3.61	73642	13527
TrigIDSCAN_eGamma	183	452	0	3511	8673	7.58	18.72	3.4	3.48	39680	33712
muFast_900GeV	545	909	0	10457	17441	3.52	5.87	3.2	3.25	46395	28800
muFast_Mcal	606	611	0	11628	11724	4.53	4.57	2.8	2.81	28877	25692
T2CaloJet_Jet	283	587	0	5430	11263	4.35	9.02	2.6	2.60	236294	66657
TrigSiTrack_Muon	176	288	0	3377	5526	6.73	11.01	1.9	1.97	37377	13911

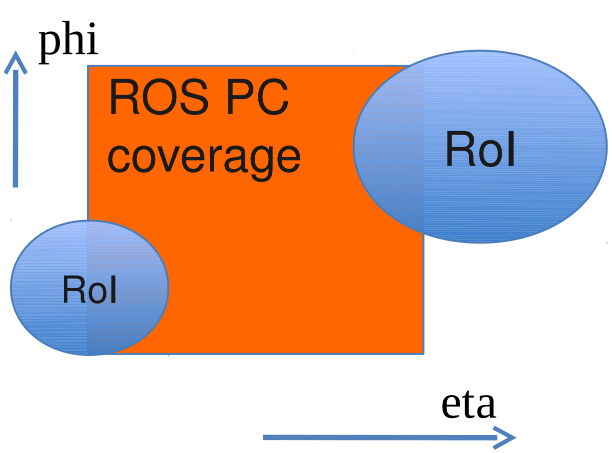
L2 data requests

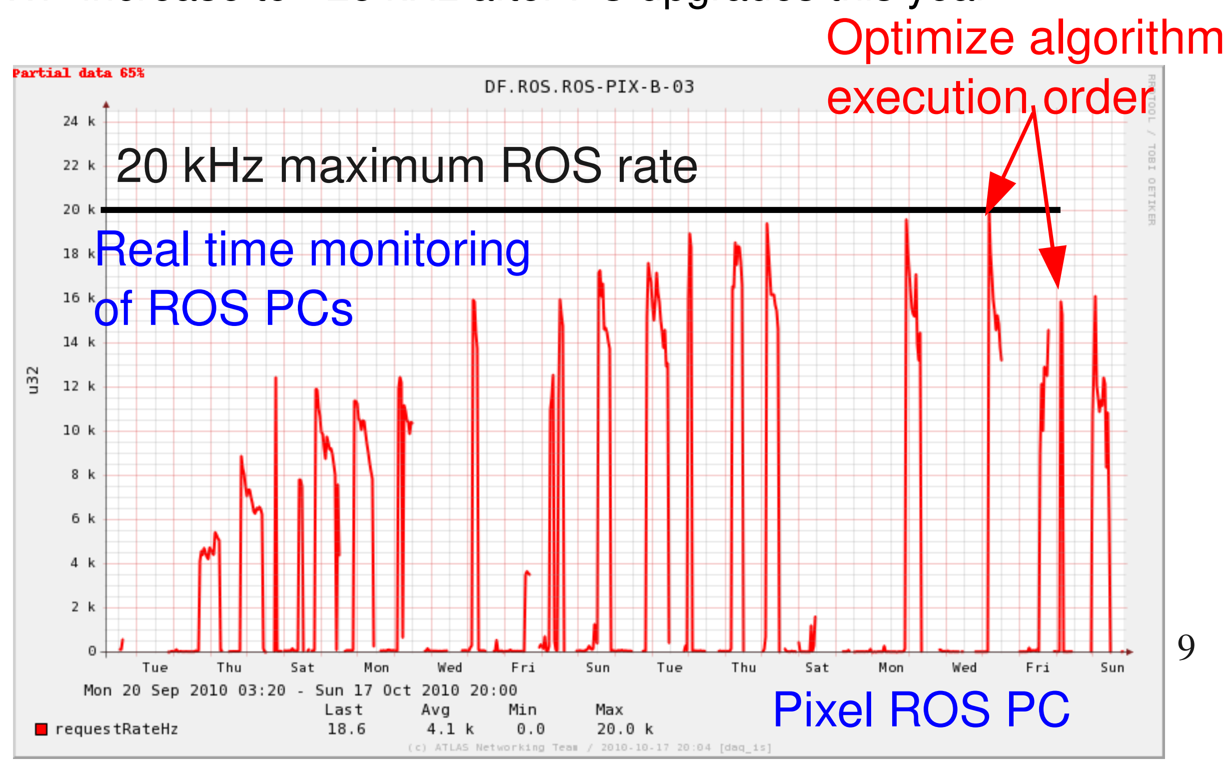


<u>Total in ATLAS:</u>
~1600 ROLs
151 ROS PC

- Detector data are buffered by Readout System (ROS)
  - 151 PCs store event data until event rejected or accepted by L2
  - Event Builder reads data from all ROSs for accepted events
- L2 algorithms request data in Region of Interest
  - Can read all or subset of ROS data
  - Fixed time delay constant per request
  - ROS processing time proportional to data size
  - Data caching (and pre-fetching)
  - Current maximum rate per ROS ~20 kHz (depends on detector)
  - Will increase to ~28 kHz after PC upgrades this year

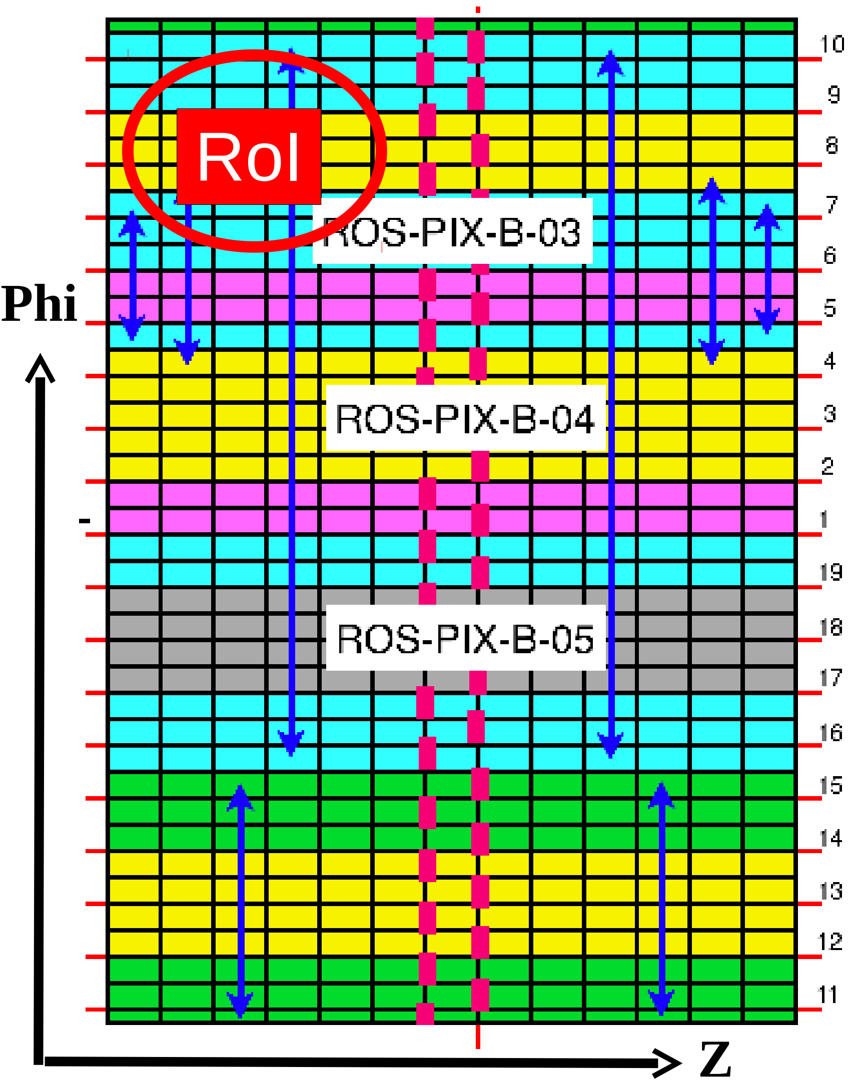
ROS coverage and RoI size determine data request rate



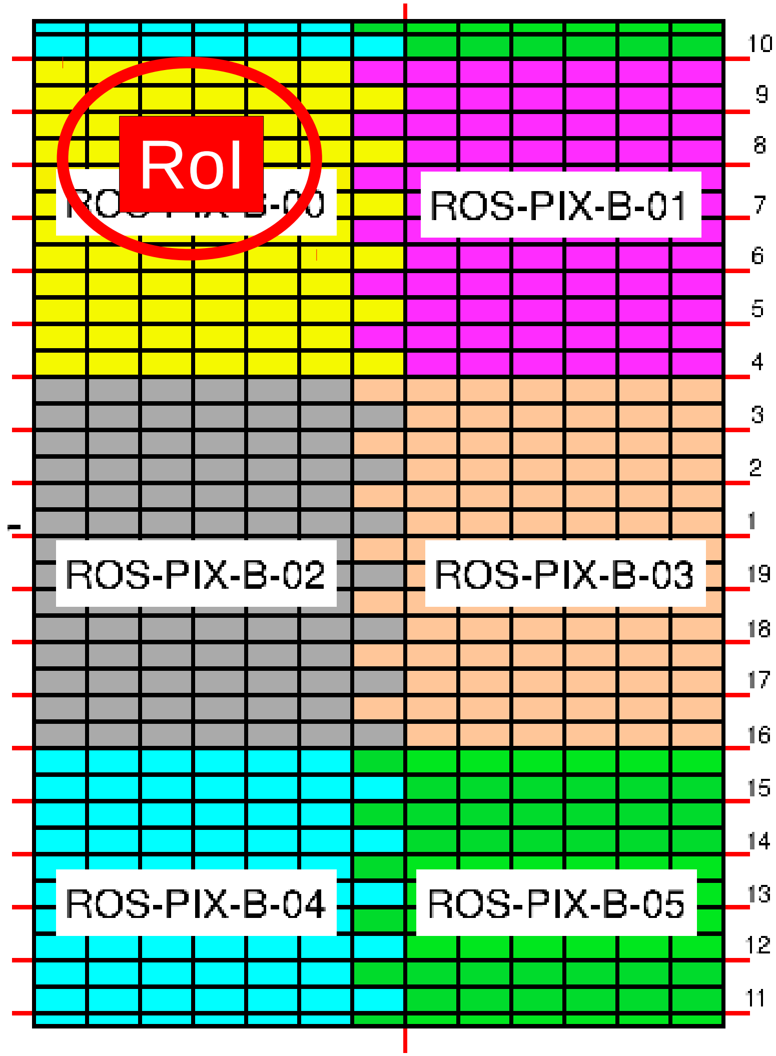


L2 data requests from Pixel Detector

Pixel Layer 1 readout map in 2010

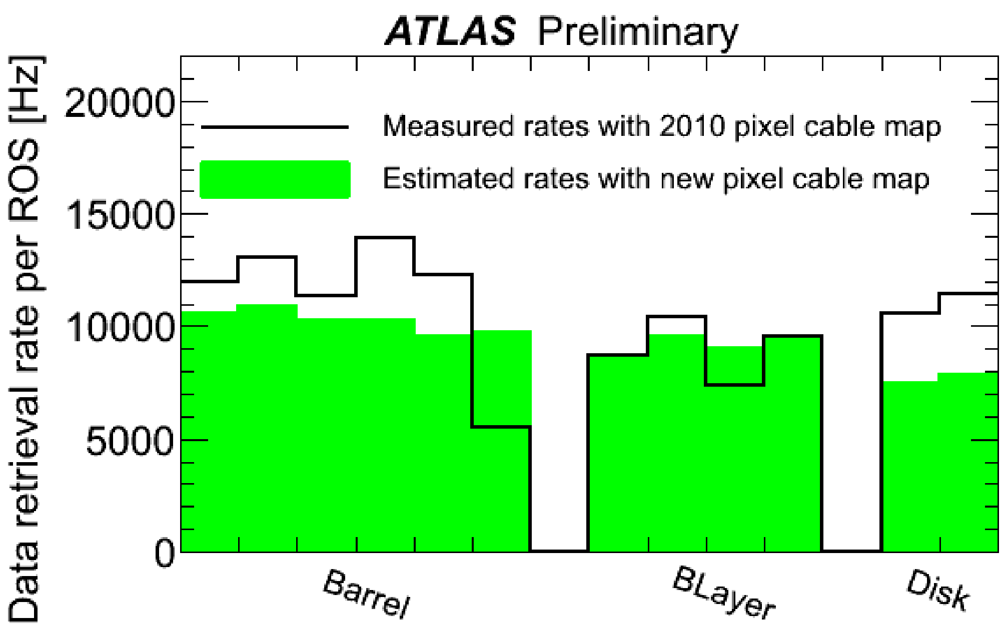


Pixel Layer 1 readout map in 2011

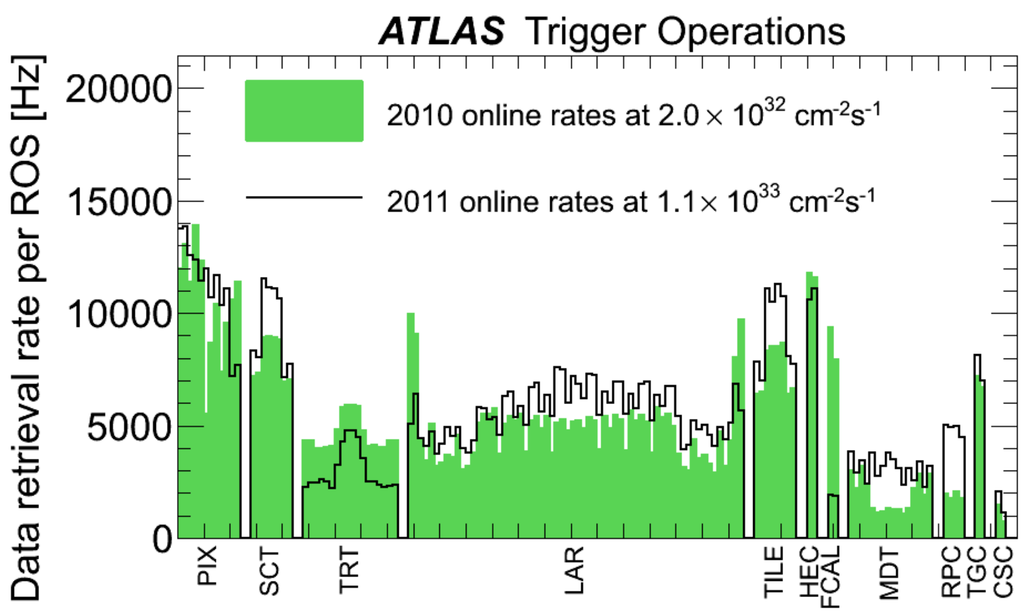


- In 2010 observed large differences in L2 data access rates for different regions of the pixel
- The pixel detector was recabled during the 2010/2011 Winter shutdown

L2 data requests from Pixel Detector

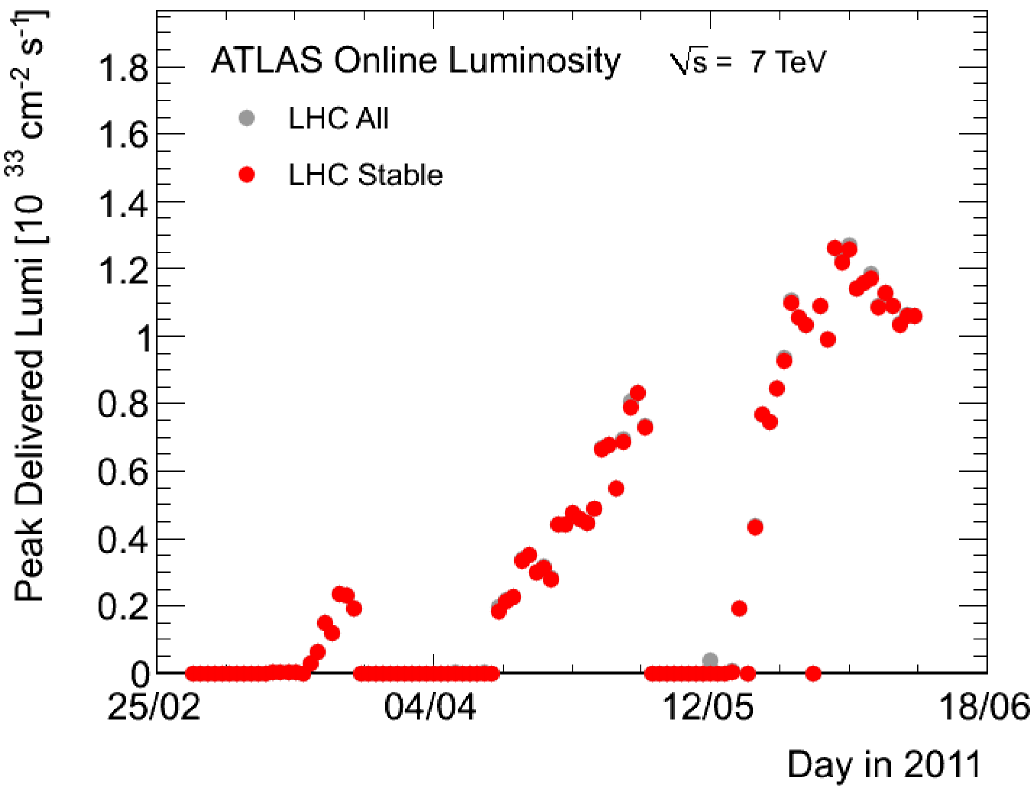


- Recompute offline expected L2 data requests rates with different detector layouts
- Optimize pixel readout map for more uniform data access rates per pixel ROS PC

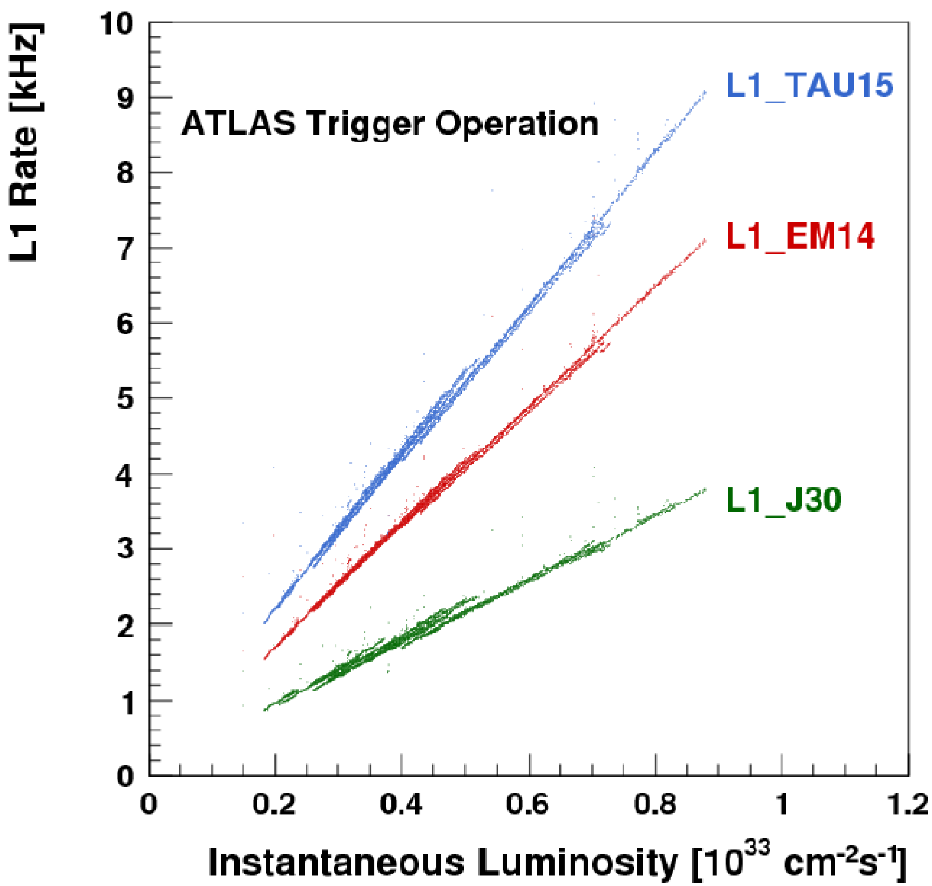


- L2 data requests depend on luminosity and deployed menu of trigger signatures
- The trigger menu evolves with increasing luminosity to fit within allowed limits
    - Make tighter trigger selections
    - Raise thresholds

Trigger dependence on luminosity



- Rapid increases in daily LHC luminosity
- Trigger rates scale linearly for single leptons and central jets
- Multiple interactions per event (pileup) create non-linear effects for missing energy and forward/multi-jet triggers

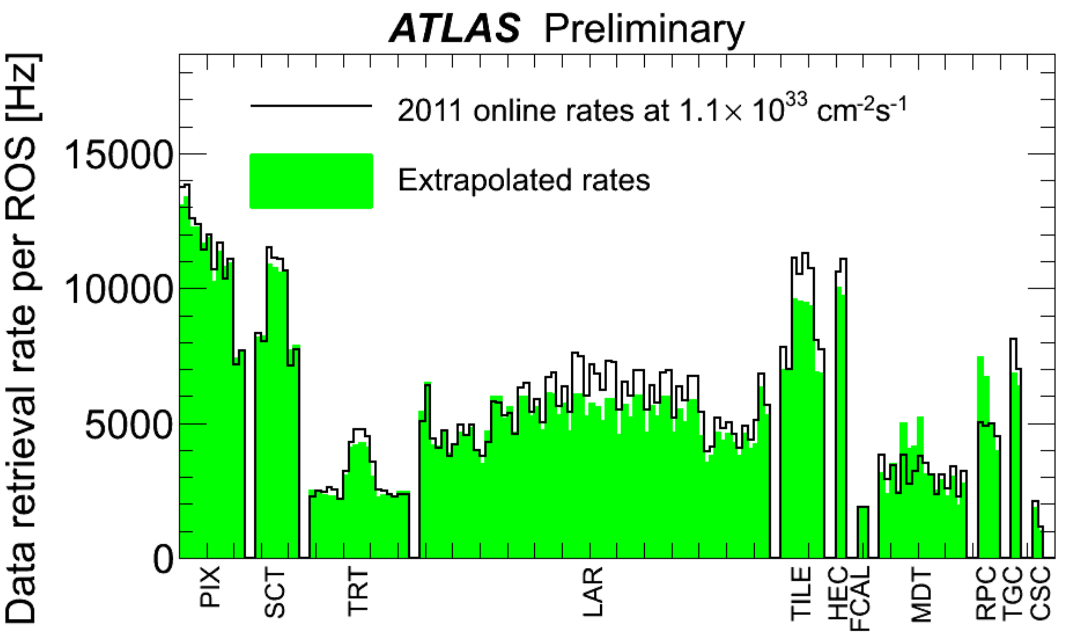


- Predict increases in trigger rates and L2 data request rates
- Make tighter trigger selection with higher instantaneous luminosity
- Tune selections to stay within TDAQ and offline resource limits

Trigger resources at higher luminosity

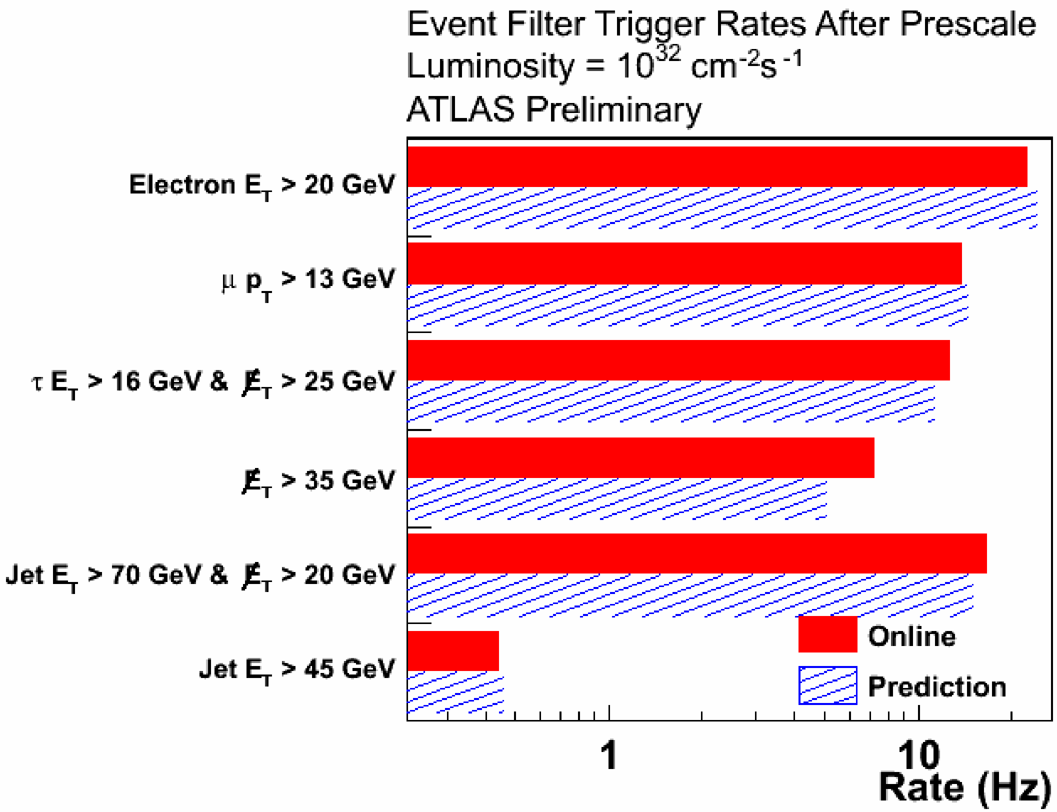
- "Enhanced bias" data:
  - LHC collisions unbiased by HLT
  - ~2 million of L1 triggers: lepton, photon, jet and MET
  - A small number of random triggers
  - Re-record data after L1 updates or changes in LHC conditions

- Uses of "enhanced bias" data:
  - Compute trigger rates and L2 data requests at any luminosity point
  - Compute group trigger rates taking into account overlaps
  - Validate new HLT triggers

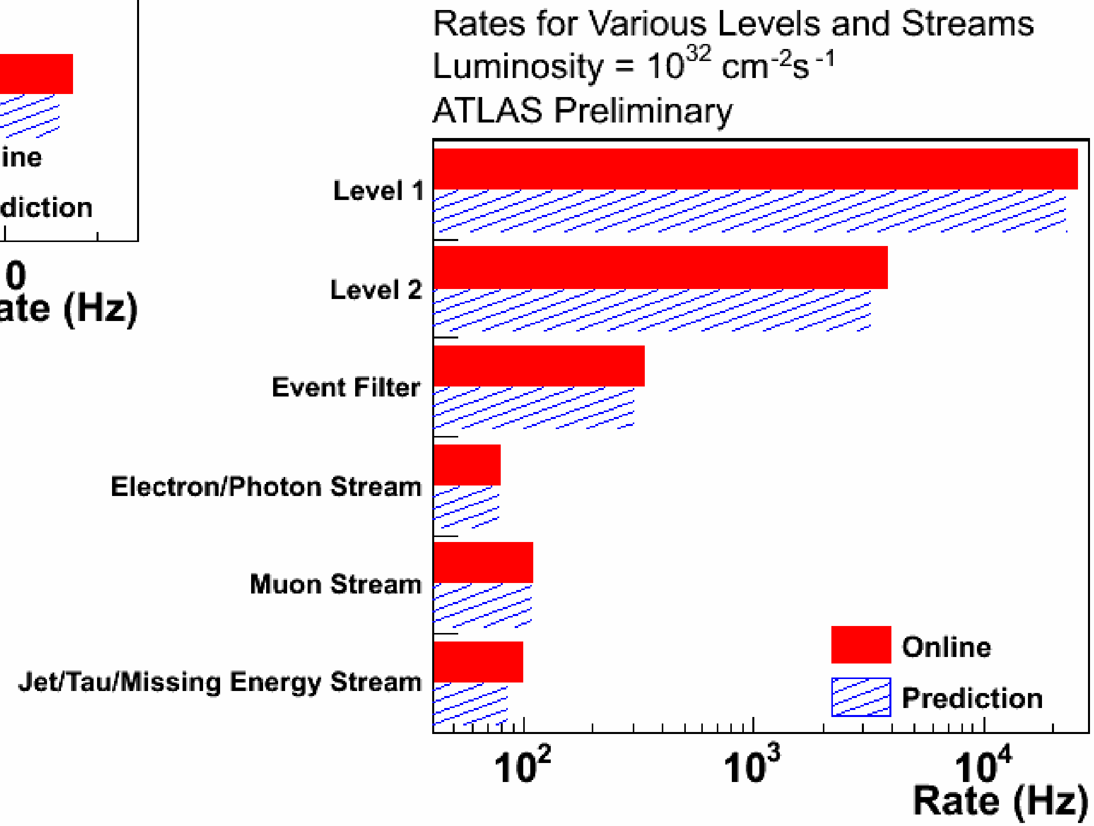


- L2 data request rates have complex dependency on trigger menu composition and algorithm order because of data caching
- Run HLT triggers with online prescales and algorithm ordering to predict L2 data request rates per ROS PC

Trigger rates at higher luminosity



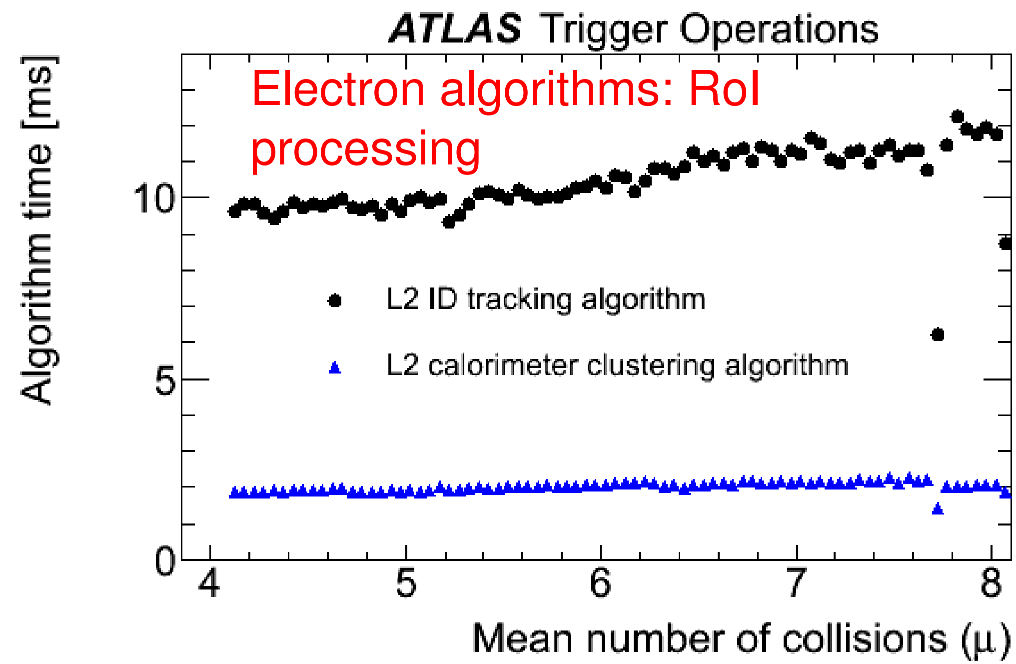
- Compute trigger rates (with errors) for individual triggers for each increase in LHC luminosity
- Can now correct trigger rates for different pileup conditions

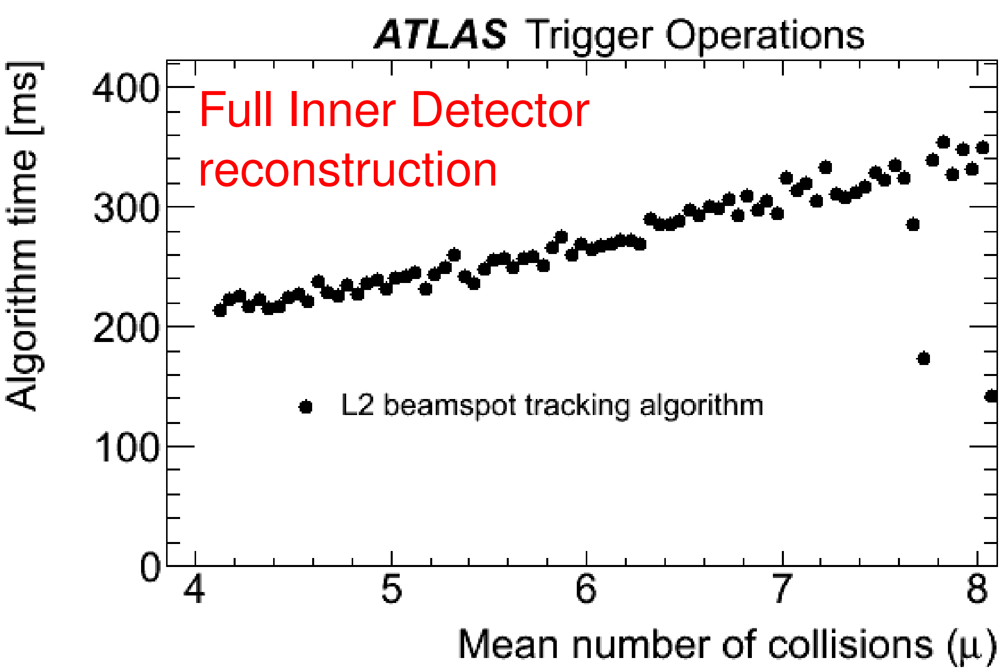


- 3 physics data streams:
  - Muons, Electrons/Photons and Jets/Taus/MET
  - Predict evolution of stream rates and tune triggers to stay within allowed limits

Resource costs at higher luminosity

- Mean number of collisions
  - Up to ~3 in 2010
  - Up to ~8 in 2011
  - Soon expect up to 20
- Primary physics triggers have weak scaling with a number of collisions because of Region of Interest based processing





- Use MC ttbar events with up to 30 collisions per crossing for higher pileup studies
- MC scaling for algorithm times matches data where two overlap
- Reasonable scaling for CPU loads for up to 30 collisions per event

Conclusions

- Excellent performance by ATLAS TDAQ
  - Thanks to robust design and careful preparation
- Many monitoring tools for system diagnostics
- Detailed monitoring tools used to:
  - Anticipate luminosity increases
  - Predict trigger rates and L2 data requests
  - Menu optimization with higher luminosity
  - Diagnose problems
  - Test improvements