



ATLAS EventIndex Data Collection Supervisor and Web Interface

A. Fernández Casani, <u>C. García Montoro</u>, J. Sánchez

On behalf of the ATLAS Collaboration



Instituto de Física Corpuscular (IFIC) CSIC- Universitat de València

CGW'16 Kraków, Poland – 24-26 October 2016



Outline

- LHC and ATLAS
- Events in ATLAS
 - How ATLAS production works
- What is ATLAS EventIndex
 - Current Messaging
 - New Data Flow
- ATLAS EventIndex Data Collection Supervisor
- ATLAS EventIndex Web Interface
- Conclusions and Future Work



LHC and ATLAS







Large Hadron Collider

(LHC) is a particle accelerator and collider located at CERN in the border of Switzerland and France.

ATLAS is one the 4 detectors at LHC, devoted to test the predictions of the Standard Model, and to physics beyond the Standard Model.

"ATLAS is an experiment at CERN designed to explore the secrets of the universe."



Events in ATLAS







Tier0

CERN (Geneva) + Wigner (Budapest)

- First copy of raw data.
- First pass reconstruction.
- **Distribution** of raw and reconstruction data to Tier1.

Tier1

13 centres (11 for ATLAS)

- Safe-keeping of a proportional share of raw and reconstructed data.
- Large-scale
 reprocessing and safe-keeping the output.
- **Distribution** to T2 and safe-keeping of a share of T2produced simulated data.

Tier2

Around 160 sites

- Computing power for specific analysis tasks.
- Handle analysis requirements and proportional share of simulated event production and reconstruction.





Tier0

- Processes from raw data to first pass reconstruction.
- First copy of data.

PanDA

- Production ANd Distributed Analysis system.
- A manager of tasks and jobs for analysis and production.

Rucio

- Distributed Data Management.
- A catalogue of files, datasets and containers.





- Just in 2015, ATLAS has produced:
 - 12 billion real events in one million files.
 - 5 billion Monte Carlo simulation events in 8 million files.
- EventIndex aims to build a complete catalogue of all ATLAS physics events, real and Monte Carlo simulation, for all processing stages.
- To allow:
 - **Event picking**: Give the reference to specific events depending on constraints:
 - Order of hundreds of concurrent users requesting from 1 event to 30k.
 - Provenance: Chain of processes.
 - Production consistency checks:
 - Duplicate event checking and overlap detection.
 - Trigger checks and event skimming:
 - Count or give an event list based on trigger selection.
 - Trigger Overlap: number of events in real data that satisfies two different triggers.



Current ATLAS EventIndex Messaging





CGW 2016 Kraków – 24-26 Oct.

Distributed Data Collection for the ATLAS EventIndex





- Current EventIndex is working reliably, but...
 - We have to cope with future necessities, i.e. more data to be indexed.
 - It has a significant messaging overhead.
 - Requires frequent human intervention of developers and experts.
- A new data flow architecture prototype is ongoing:
 - To reduce messaging complexity.
 - To ease operational procedures.





Data production is shared between producers and consumers by means of CERN's Object Store facility based on Ceph, using its Amazon S3 compatible interface





- Python + SQLAlchemy:
 - Not tied to a particular DBMS.
 - Shared models with web interface.
 - Thread-local scoped sessions with the database
- Asynchronous multithreaded application with queues:
- Threads retrieve info from external resources asynchronously
 - Main supervisor loop.
 - Message loop.
 - Active Task loop.
 - Rucio loop.
 - Validation loop.

- Queues store small transient entities, mainly ids, on their own tables
 - Active tasks queue.
 - Rucio queue.
 - Validation queue.



EventIndex Data Collection Supervisor







- Shows the data flow and processing state.
- Provides information about the supervisor and the indexing tasks.
 - Able to alert when some problems are detected.
- Written in Python with Flask framework and SQLAlchemy.
 - Shared models with supervisor.
- Uses Apache's GridSite module to authenticate through certificates.
- JSON is used to provide data to clients.
 - Will ease interactions with external applications.



EventIndex Web Interface: Components





Filters										+
Show 10 • entries						Search wit	hin results			
Dataset	Taskid 💡	Creation 🕴	Task State	EventIndex State	Jobs					
Dalaset					Total	Done	Run.	Proc.	Produced	Consumed
notknownye	11	1970-01-01 00:00		EI_REGISTERED	0	0	0	0	0	0
data16_13TeV.00012354.physics_Main.merge.AOD.f999_m999	10	2016-09-23 19:31	FINISHED	EI_VALID	15	15	0	0	0	0
data16_13TeV.00012353.physics_Main.merge.AOD.f999_m999	9	2016-09-23 19:31	RUNNING	EI_RUNNING	20	15	0	0	0	0
data16_13TeV.00012352.physics_Late.merge.AOD.f999_m999	8	2016-09-23 18:59	FINISHED	EI_VALID	35	35	0	0	0	0
data16_13TeV.00012351.physics_Late.merge.AOD.f999_m999	7	2016-09-23 18:58	FINISHED	EI_VALID	20	20	0	0	0	0
data16_13TeV.00012350.physics_Main.merge.AOD.f999_m999	6	2016-09-23 18:49	FINISHED	EI_VALID	35	35	0	0	0	0
data16_13TeV.00012349.physics_Main.merge.AOD.f999_m999	5	2016-09-23 18:48	FINISHED	EI_VALID	20	20	0	0	0	0
data16_13TeV.00012348.physics_zeroBias.merge.AOD.f999_m999	4	2016-09-23 17:53	FINISHED	EI_VALID	20	20	0	0	0	0
data16_13TeV.00012347.physics_Main.merge.AOD.f999_m999	3	2016-09-22 20:13	FINISHED	EI_VALID	20	20	0	0	0	0
data16_13TeV.00012346.physics_Late.merge.AOD.f999_m999	2	2016-09-22 20:13	FINISHED	EI_VALID	20	20	0	0	0	0
Showing 1 to 10 of 11 entries								Previous	1 2 Next	
Tacks in EventIndex Oueves										

Tier0 Tasks in EventIndex Queues

		Active Tacks Queue	Validation Quana			
Total	Waiting	Recoverable	Unrecoverable	Active lasks Queue	Valuation Queue	
1	1	0	0	2	0	

CGW 2016 Kraków – 24-26 Oct. Distributed Data Collection for the ATLAS EventIndex





- ATLAS EventIndex handles billions of events in millions of files per year.
- Though current EventIndex is working reliably, we are making it future-proof.
 - Object Store will reduce messaging considerably.
 - New supervisor and web interface will reduce operational efforts.
- The web interface already implements some REST endpoints, but further development is required.









From http://atlasexperiment.org/fact_sheets.html

- 3,000 scientists, 38 countries, 180 universities and labs.
- An event is a proton-proton collision.
 - Around 1 billion collisions per second.
 - Trigger and Event Filter select about 1000 *interesting* events/sec.
 - Raw data: The data after trigger 1GB/s, 1MB/event, 10PB/year.





- Around 200 producer tasks simultaneously running on Tier0.
- Around 20 events/message.
- Typically around 300 messages/file.
- In some unusual case, up to 20,000 events/file.
- 5 ActiveMQ Brokers.
- Consumers = 2 x Brokers = 10.