



Event Index Core *(Hadoop & HBase)*

- Import
- Trigger Overlap
- Other News

[Julius Hrivnac](#)
[Fedor Prokoshin](#)
[Grigorij Rybkin](#)
[Rainer Toeblcke](#)
[Ruijun Yuan](#)



Import

- Very smooth and fast after migration to ObjectStore Consumer
 - Also because legacy importing procedure re-written
- The only (temporary) problem:
 - We haven't taken into account that new Consumer files are not sorted, which breaks 'ei -key' and 'el -api simple/rich'
 - Import has been fixed and already imported files have been sorted (by Grigorij)
- Import of 37 datasets have failed during cluster upgrade
 - Automatic re-import successful
- What about conf-directory OSDatasetsConf_2 ?
 - It contains a lot of PARTIAL conf files
- Import executed in several stages (change in one stage triggers re-execution of following stages)
 - Actual import from Consumer, sorting, indexing, adding trigger info, registering in Catalog - as datasets arrive, usually within minutes
 - Uploading to HBase for EL - 4x a day (may be slower depending on number of data)
 - Generating derived tables (DOverlaps, TOverlaps, TStats) - nightly
 - Old TagFiles processed with 100 per day, all will be done till end 2017
 - Testing (consistency, multi-events, EL,...) - nightly (EL postponed one day)
- All import steps
 - Recorded in Journal
 - Registered in Catalog (if relevant)
 - Mailed to atlevind

IMPORTED EVENTS

=====

EI 2009	34	444	939	
EI 2010	1	196	985	327
EI 2011	1	844	680	412
EI 2012	3	313	189	845
EI 2013		331	475	303
EI 2014		436	763	032
EI 2015	38	527	094	863 (was 37 875 871 406)
EI 2016	63	203	930	561 (was 57 524 958 034)
EI 2017	17	004	632	956 (was 5 948 137 819)
MC 2012		300	000	
MC 2015	28	763	092	958
MC 2016	3	825	805	416 (was 1 571 035 156)
ALL	> 158 billion events			



Trigger Overlap

- Generated from subset of events because full generation would take too long
 - For all newly imported datasets (and some old datasets)
 - Parallelism (M/R,...) doesn't help much because task can't be easily divided into independent subtasks (big common data)
 - Full statistics generation for subset of triggers is fast and can be done on request (via CLI or WS)
- Visualisation may be slow (as a lot of data should be parsed), esp. *the second arrow*
- Trigger Statistics is contained in Trigger Overlap (but Trigger Statistics table use all data)
- Working on it....

Event Index

- Global Help
- FAQ, Use Cases, Home
- Data-Centric View (prototype)
- Catalog
- Event Index (Expert Mode)
- Event Lookup
- Trigger Info
- TagFile Inspector
 - Dataset Overlaps
 - Trigger Overlaps
 - Trigger Statistics
- System Journal (for admins)
- External Services
 - AMI

data17_5TeV.00340718.physics_ZeroBias.merge.AOD.896_m1902

data17_5TeV.00340814.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340844.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340849.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340850.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340910.physics_ZeroBias.merge.AOD.898_m1907

data17_5TeV.00340918.physics_ZeroBias.merge.AOD.898_m1907

data17_5TeV.00340925.physics_ZeroBias.merge.AOD.898_m1907

data17_5TeV.00340973.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00341027.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00341123.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00341184.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00340308.physics_Background.merge.AOD.894_m1902

data17_5TeV.00340308.physics_CosmicCalo.merge.AOD.894_m1902

data17_hi.00338037.physics_Background.merge.AOD.882_m1885

data17_hi.00338037.physics_CosmicCalo.merge.AOD.882_m1885

data17_hi.00338037.physics_Main.merge.AOD.882_m1885

PNG

data17_hi.00338037.physics_Main.merge.AOD.882_m1885

evts=549637,t3=550098.0

HLT_Lamnoiseburst_rerun	550098 (100.08%)	✓
L1_2EM10VH	154818 (28.17%)	✓
L1_2EM15	152073 (27.67%)	✓
L1_2EM15VH	151524 (27.57%)	✓
L1_2EM15VHI	151524 (27.57%)	✓
L1_2EM20VH	151524 (27.57%)	✓
L1_2EM3	537471 (97.79%)	✓
L1_2EM7	194346 (35.36%)	✓
L1_2EM8VH_MU10	1647 (0.30%)	✓
L1_2J15_XE55	152073 (27.67%)	✓
L1_2J50_XE40	184464 (33.56%)	✓
L1_2MU4 B	549 (0.10%)	✓
L1_2MU4	2196 (0.40%)	✓
L1_2MU4_J20_XE30_DPHI-J20+2XE30	1647 (0.30%)	✓
L1_2MU4_J20_XE40_DPHI-J20+2XE30	1647 (0.30%)	✓

Event Index

- Global Help
- FAQ, Use Cases, Home
- Data-Centric View (prototype)
- Catalog
- Event Index (Expert Mode)
- Event Lookup
- Trigger Info
- TagFile Inspector
 - Dataset Overlaps
 - Trigger Overlaps
 - Trigger Statistics
- System Journal (for admins)
- External Services
 - AMI

data17_5TeV.00340718.physics_ZeroBias.merge.AOD.896_m1902

data17_5TeV.00340814.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340844.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340849.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340850.physics_ZeroBias.merge.AOD.897_m1902

data17_5TeV.00340910.physics_ZeroBias.merge.AOD.898_m1907

data17_5TeV.00340918.physics_ZeroBias.merge.AOD.898_m1907

data17_5TeV.00340925.physics_ZeroBias.merge.AOD.898_m1907

data17_5TeV.00340973.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00341027.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00341123.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00341184.physics_ZeroBias.merge.AOD.899_m1912

data17_5TeV.00340308.physics_Background.merge.AOD.894_m1902

data17_5TeV.00340308.physics_CosmicCalo.merge.AOD.894_m1902

data17_hi.00338037.physics_Background.merge.AOD.882_m1885

data17_hi.00338037.physics_CosmicCalo.merge.AOD.882_m1885

data17_hi.00338037.physics_Main.merge.AOD.882_m1885

PNG

data17_hi.00338037.physics_Main.merge.AOD.882_m1885

evts=549637,t3=550098.0

HLT_Lamnoiseburst_rerun	550098 (100.08%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM10VH	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)	
L1_2EM15	152073 (27.67%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM15VH	151524 (27.57%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM15VHI	151524 (27.57%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM20VH	151524 (27.57%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM3	537471 (97.79%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM7	194346 (35.36%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2EM8VH_MU10	1647 (0.30%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2J15_XE55	152073 (27.67%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2J50_XE40	184464 (33.56%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2MU4 B	549 (0.10%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2MU4	2196 (0.40%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2MU4_J20_XE30_DPHI-J20+2XE30	1647 (0.30%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)
L1_2MU4_J20_XE40_DPHI-J20+2XE30	1647 (0.30%)	154818 (28.17%)	537471 (97.79%)	152073 (27.67%)	184464 (33.56%)

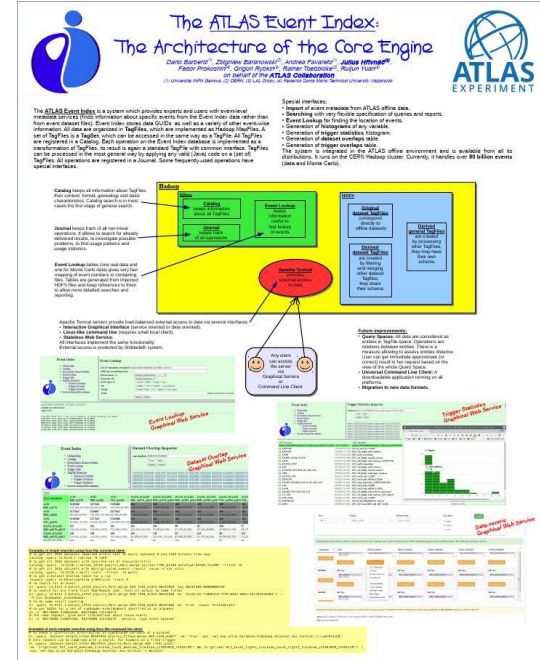
Get this result via [command line](#)

```
$ inspect -query id:TOverlap/EI17.1/data17_hi.physics_Main.merge.AOD.f882_m1885.00338037 -action null -limit 0 -climit 10
```



Other News

- All TagFiles are now using Record gzip compression
 - Gives factor 2x-3x, completely transparent to clients
 - More aggressive compression (5x-10x) would break some clients
 - Before only newly imported TagFiles were compressed, derived TagFiles were uncompressed
- New nightly tests
 - Mostly for EL: Just the same tests as Panda, but without Panda
- Many small fixes and improvements (Justin, Grigorij, Rainer, Fedor)
- Poster in ACAT/Seattle ([ATL-COM-SOFT-2017-073](#))





Info

Web Service: <https://atlas-event-index.cern.ch/ElHadoop>

Documentation & Distribution: <https://atlas-event-index.cern.ch/doc>

Frequently Asked Questions: <https://atlas-event-index.cern.ch/doc/faq>

Sources: svn+ssh://svn.cern.ch/repos/atlasoff/Database/TAGHadoop/TagConvertor

AFS: /afs/cern.ch/project/jps/repos/atlas-eies/TagConvertor

CVMFS: \$ lsetup eiclient

EOS: /eos/atlas/atlascerngroupdisk/proj-evind/Results