

The ATLAS Event Index:

The Architecture of the Core Engine

Dario Barberis⁽¹⁾, Zbigniew Baranowski⁽²⁾, Andrea Favareto⁽¹⁾, Julius Hřivnáč⁽³⁾, Fedor Prokoshin⁽⁴⁾, Grigori Rybkin⁽³⁾, Rainer Toebbicke⁽²⁾, Ruijun Yuan⁽³⁾ on behalf of the **ATLAS Collaboration** (1) Universite INFN Genova, (2) CERN, (3) LAL Orsay, (4) Federico Santa Maria Technical University Valparaíso

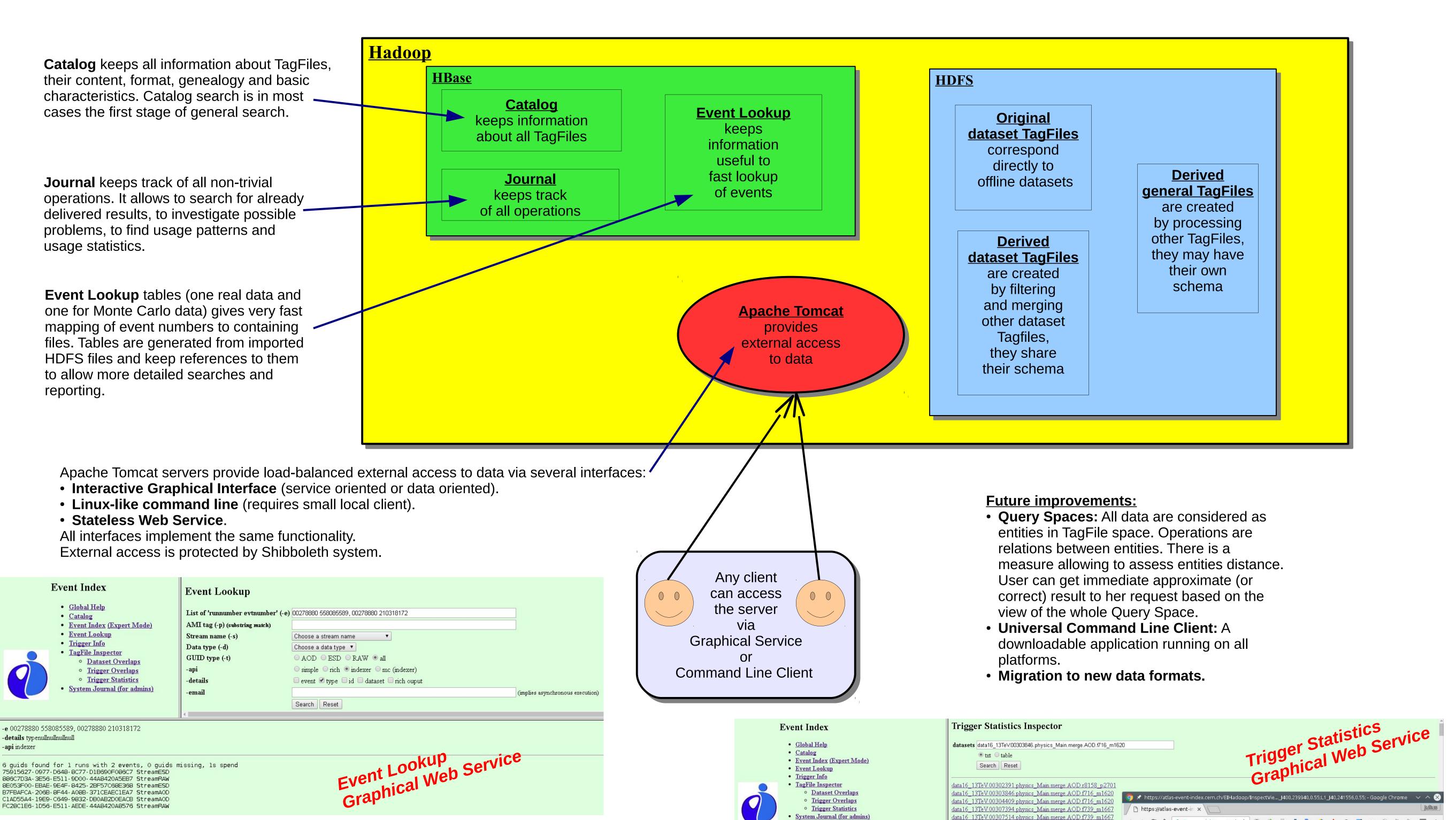


The **ATLAS Event Index** is a system which provides experts and users with event-level metadata services (finds information about specific events from the Event Index data rather than from event dataset files). Event Index stores data GUIDs as well as a variety of other event-wise information. All data are organized in **TagFiles**, which are implemented as Hadoop MapFiles. A set of TagFiles is a TagSet, which can be accessed in the same way as a TagFile. All TagFiles are registered in a Catalog. Each operation on the Event Index database is implemented as a transformation of TagFiles, its result is again a standard TagFile with common interface. TagFiles can be processed in the most general way by applying any valid (Java) code on a (set of) TagFiles. All operations are registered in a Journal. Some frequently-used operations have special interfaces.

Special interfaces:

- Import of event metadata from ATLAS offline data.
- Searching with very flexible specification of queries and reports.
- Event Lookup for finding the location of events.
- Generation of **histograms** of any variable.
- Generation of **trigger statistics** histogram.
- Generation of dataset overlaps table.
- Generation of **trigger overlaps** table.

The system is integrated in the ATLAS offline environment and is available from all its distributions. It runs on the CERN Hadoop cluster. Currently, it handles over 80 billion events (data and Monte Carlo).



L1 TAU60

L1 MU20

L1 Л00

L1 EM15VH 3EM7

L1_EM13VH_3J20

L1 2EM15VH

L1_J120

L1_TAU40

L1 2MU6

L1_J75_3J20

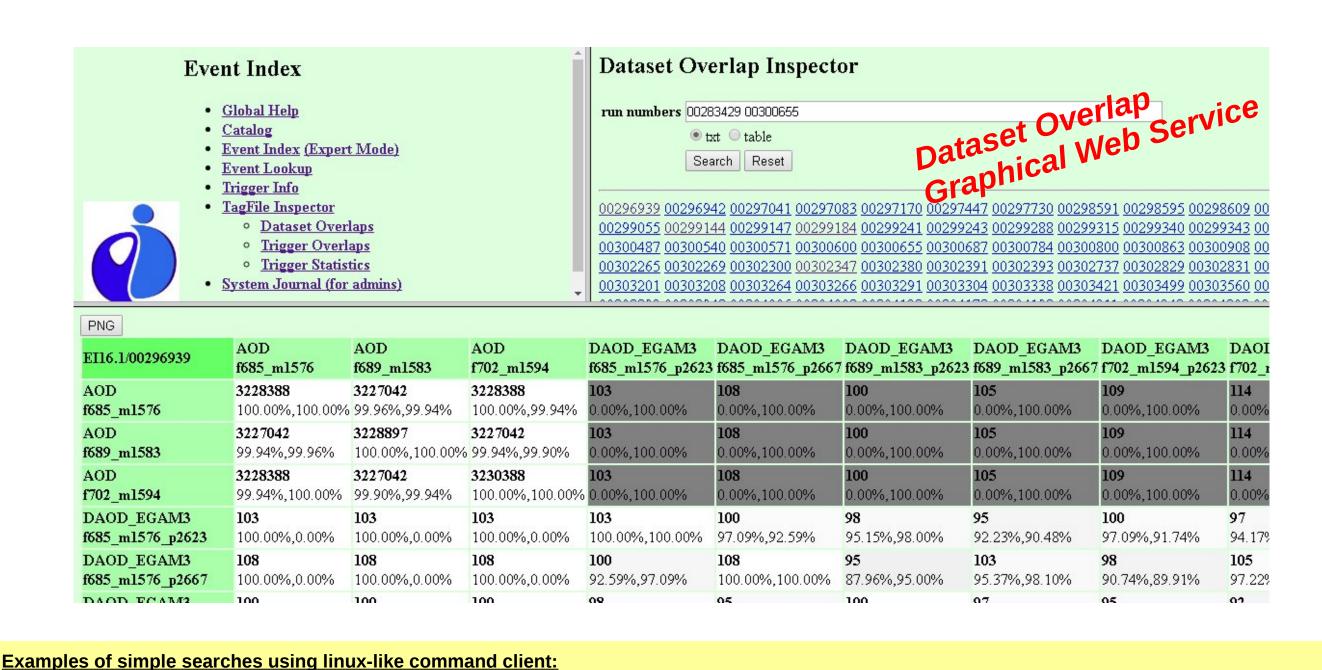
L1 MU6 2MU4

L1_EM15HI_TAU40_2TAU15

L1_TAU20IM_2TAU12IM_J25_2J20_3J12

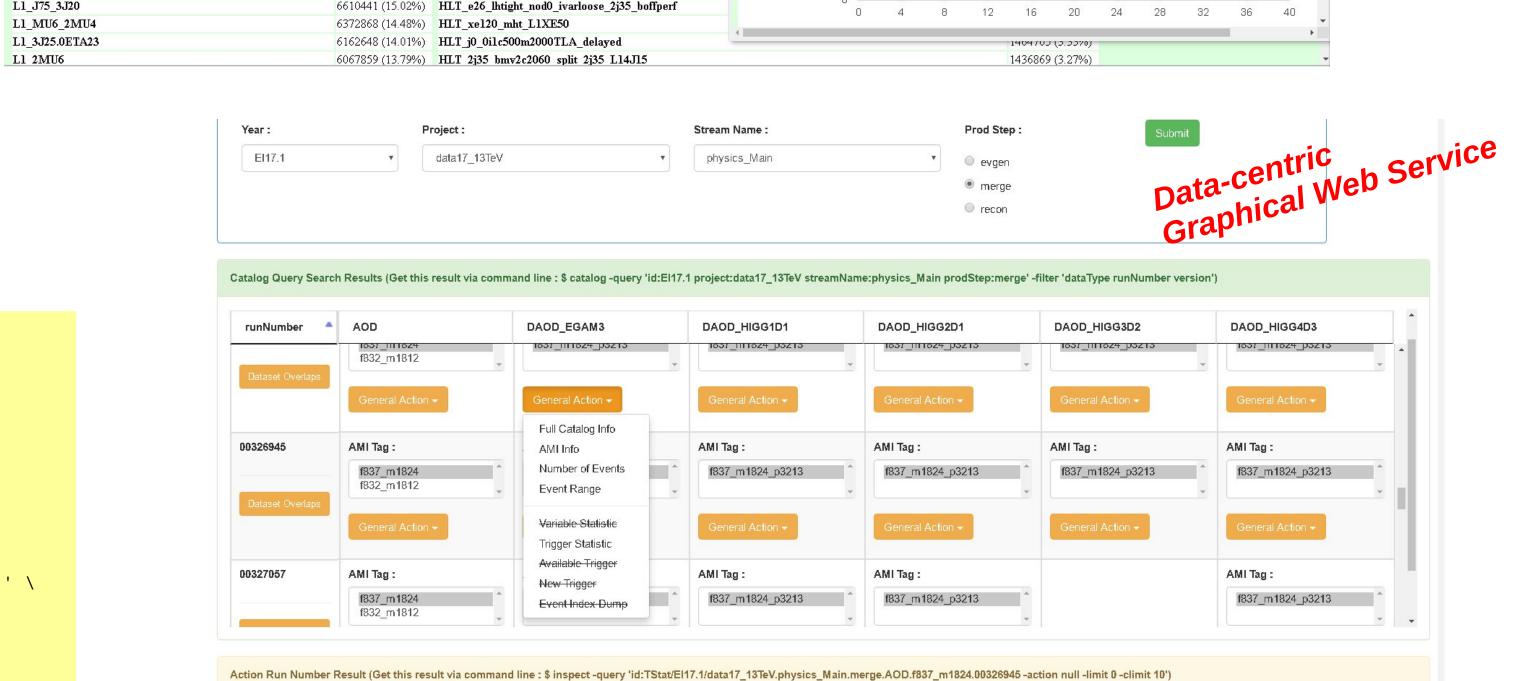
L1 EM15HI 2TAU12IM J25 3J12

L1_EM15HI_TAU20IM_2TAU15_J25_2J20_3J15



to get all 2016 datasets imported within last 24 hours (between 0 and 1440 minutes from now)

catalog -query id:EI16.1 -period '0-1440'



C ↑ 🔒 Secure | https://atla ☆ 🚳 🧶 🌷 🥾 🤚 🐞 🖴 🚆 🔠 🕢 🖔 🗅 🔲

🚻 Apps 🖿 G 🖿 TR 🖿 Search 🖿 Compt 🖿 Conf 🖿 Talk 🖿 Cafe 🖿 Action 💮 🧪 🖿 Other bookmarks

L1-Triggers

%: 1.16

L1_2MU4 (509442

to get all 2016 datasets with concrete set of characteristics catalog -query 'id:EI16.1.data16 13TeV.physics Main.merge version:f701 m1598 dataType:DESDM TILEMU' -filter id # to get all 2016 datasets with multiplicated events ('multi' value is not null) catalog -query 'id:EI16.1 multi:~null' -filter 'id multi' # to get a Dataset Overlap table for a run inspect -query id:D0verlap/EI16.1/00311321 -limit 0 # to search for an event ei -query id:EI15.1.data15 13TeV.physics Main.merge.AOD.f594 m1435.00267069 -key 00267069-00000008169 # to search for any field (full Map/Reduce job), restrict output to some fields ei -query id:EI15.1.data15_13TeV.physics_Main.merge.AOD.f594_m1435.00267069 -mr 'hasGuid("710982CD-72FD-0D45-8865-5417419103B4")' \ -filter RunNumber EventNumber # to do some result counting ei -query id:EI15.1.data15_13TeV.physics_Main.merge.AOD.f594_m1435.00267069 -mr 'true' -count "EventWeight" # to get GUIDs for a set of runNumber-eventNumbers specified as an argument el -e '00278880 558085589, 00278880 255538473' # the same request, give more informations about found events el -e '00278880 558085589, 00278880 255538473' -details 'type event dataset'

-query id:EI17.1.data17_13TeV.physics_Main.merge.AOD.f837_m1824.00326945 status:good

data16_13TeV.00307539.physics_Main.merge.AOD.f739_m1667

17209353 (39.11%) HLT_j0_perf_ds1_L1Л100

10971036 (24.93%) HLT_mu26_ivarmedium

9087995 (20.65%) HLT mu28 ivarmedium

7658118 (17.40%) HLT_xell0_mht_L1XE50

6943766 (15.78%) HLT_xel10_mht_L1XE55

16400481 (37.27%) HLT_e26_lhtight_nod0_ivarloose

10156918 (23.08%) HLT_e26_lhtight_smooth_ivarloose

8811984 (20.03%) HLT e28 lhtight nod0 ivarloose

8195903 (18.63%) HLT e28 lhtight smooth ivarloose

7058129 (16.04%) **HLT_xell0_mht_wEFMu_L1XE50**

6884852 (15.65%) HLT_xel10_mht_wEFMu_L1XE55

7729228 (17.57%) HLT_e28_lhtight_nod0_ringer_ivarloose

7006299 (15.92%) HLT_2mu4_bJpsimumu_L1BPH-2M8-2MU4

6678604 (15.18%) HLT_e26_lhtight_nod0_ivarloose_2j35_bperf

10073776 (22.89%) HLT_e28_lhtight_nod0_ivarloose_L1EM22VHI

data16_13TeV.00307394.physics_Main.merge.AOD.f739

Examples of more complex searches using linux-like command line client: # to check a statistical distribution of LumiBlockN variable of a dataset ei -query 'dataset:data16 13TeV.00307454.physics L1Topo.merge.AOD.x464 m1667' -mr 'true' -aux 'net.hep.atlas.Database.EIHadoop.Accessor.Aux.VarStat;-v LumiBlockN' # this request can be combined with a search, for example on a fired trigger ei -query 'dataset:data15 13TeV.00279515.physics Main.merge.AOD.r7562 p2521' \ -mr 'trigFired("HLT tau35 medium1 tracktwo tau25 medium1 tracktwo L1TAU20IM 2TAU12IM") && !trigFired("HLT tau35 tight1 tracktwo tau25 tight1 tracktwo L1TAU20IM 2TAU12IM") \ -aux 'net.hep.atlas.Database.EIHadoop.Accessor.Aux.VarStat;-v BunchId'