

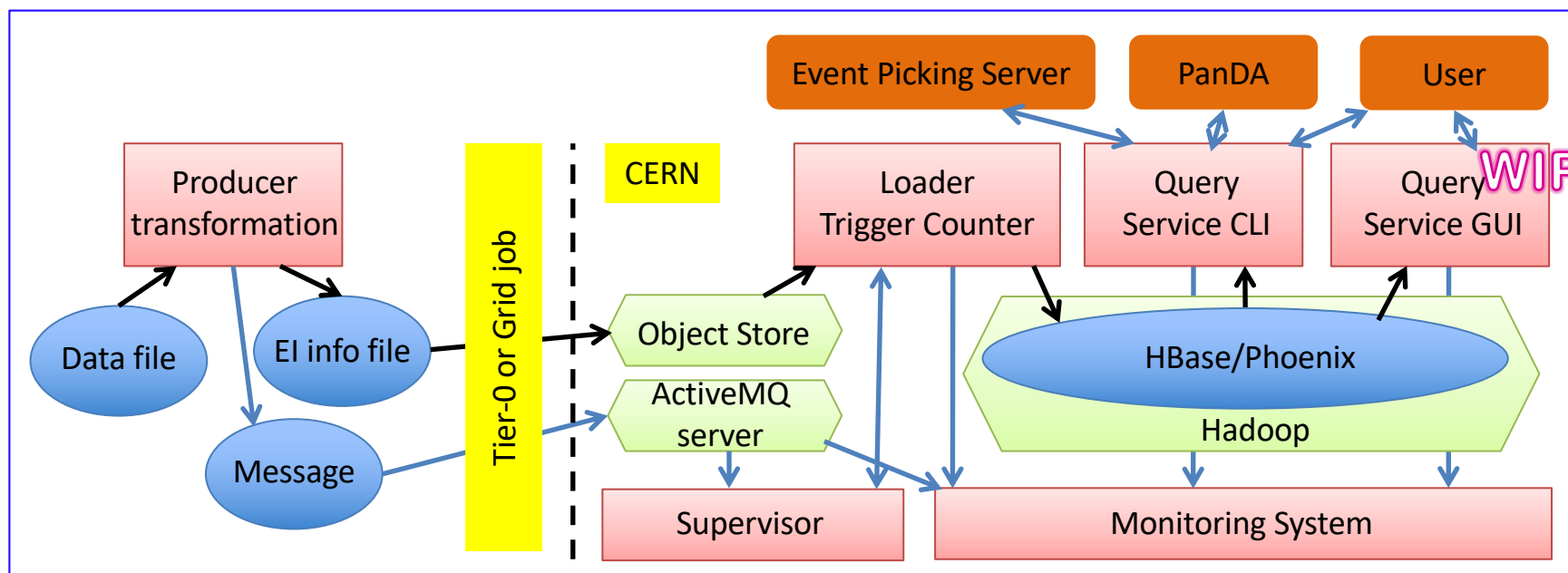
EventIndex

Progress and Evolution

Dario Barberis
University & INFN Genoa

EventIndex in Run 3

Functional schema and data flow

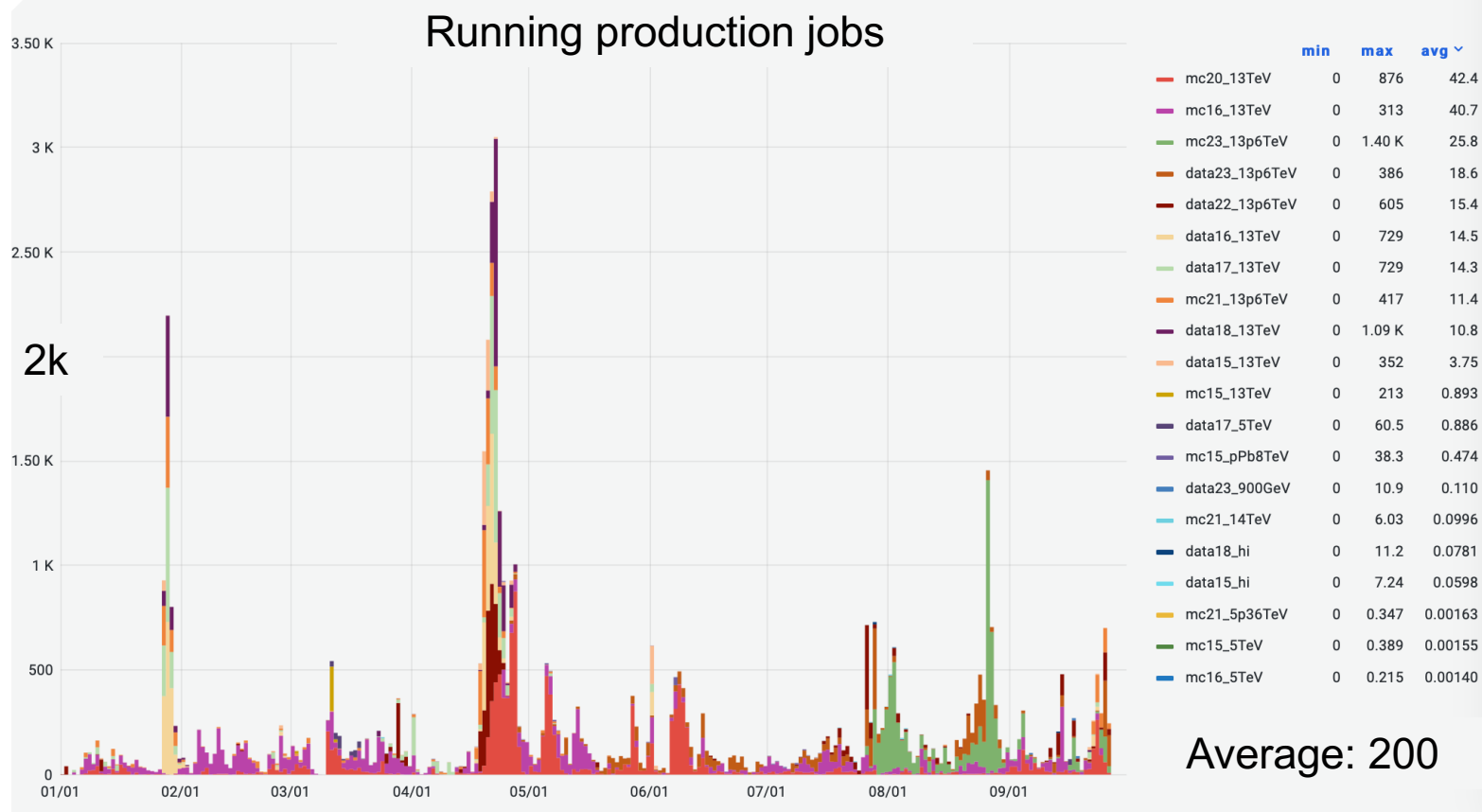


Most components updated or replaced during LS2

All major functionalities in stable operation

A few loose ends remaining

2023 Operations (1)



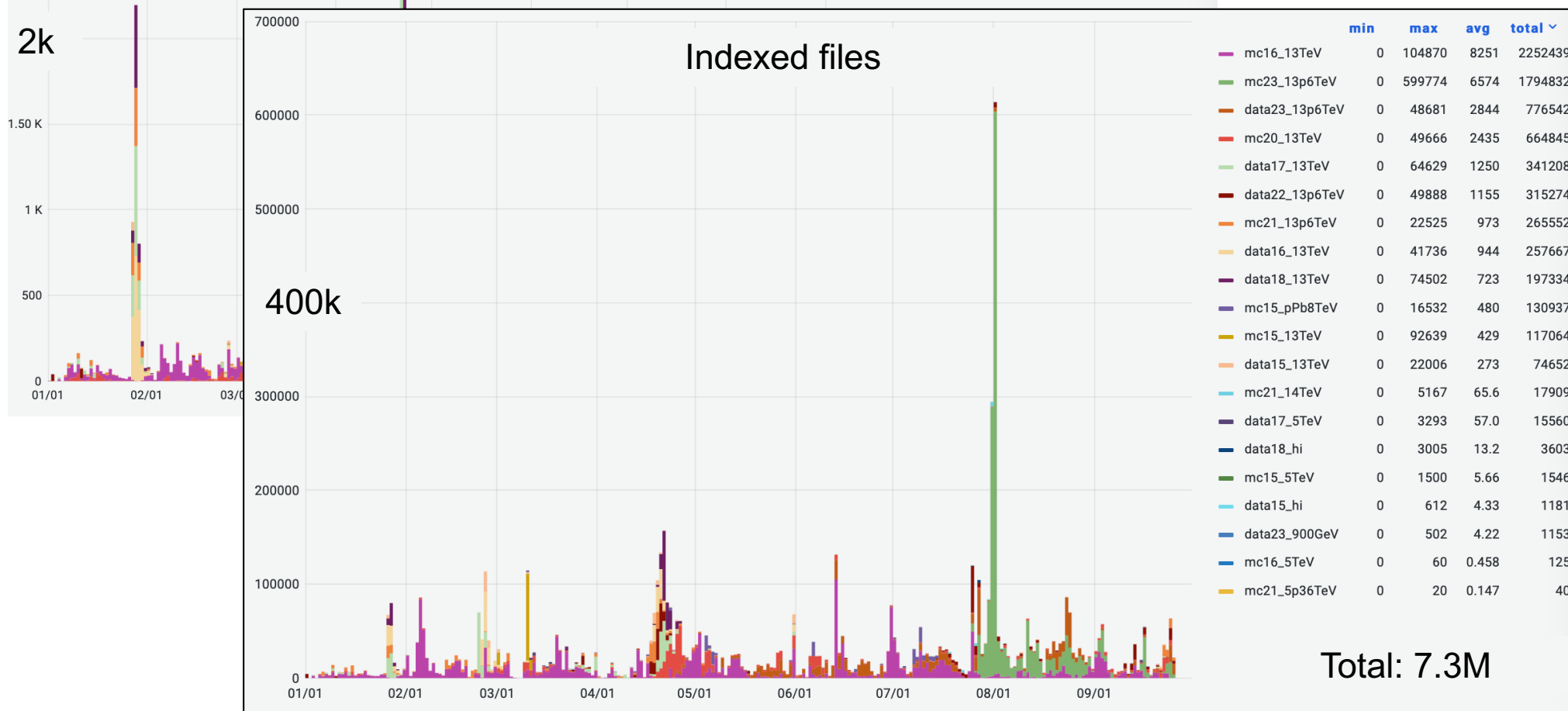
https://monit-grafana.cern.ch/d/000000696/job-accounting-historical-data?orgId=17&var-bin=1d&var-groupby=inputfileproject&var-processingtype=eventindex&var-processingtype=t0_eventindex&var-es_division_factor=24&var-pledges=CPU&from=1672531200000&to=169611

2023 Operations (2)

Running production jobs

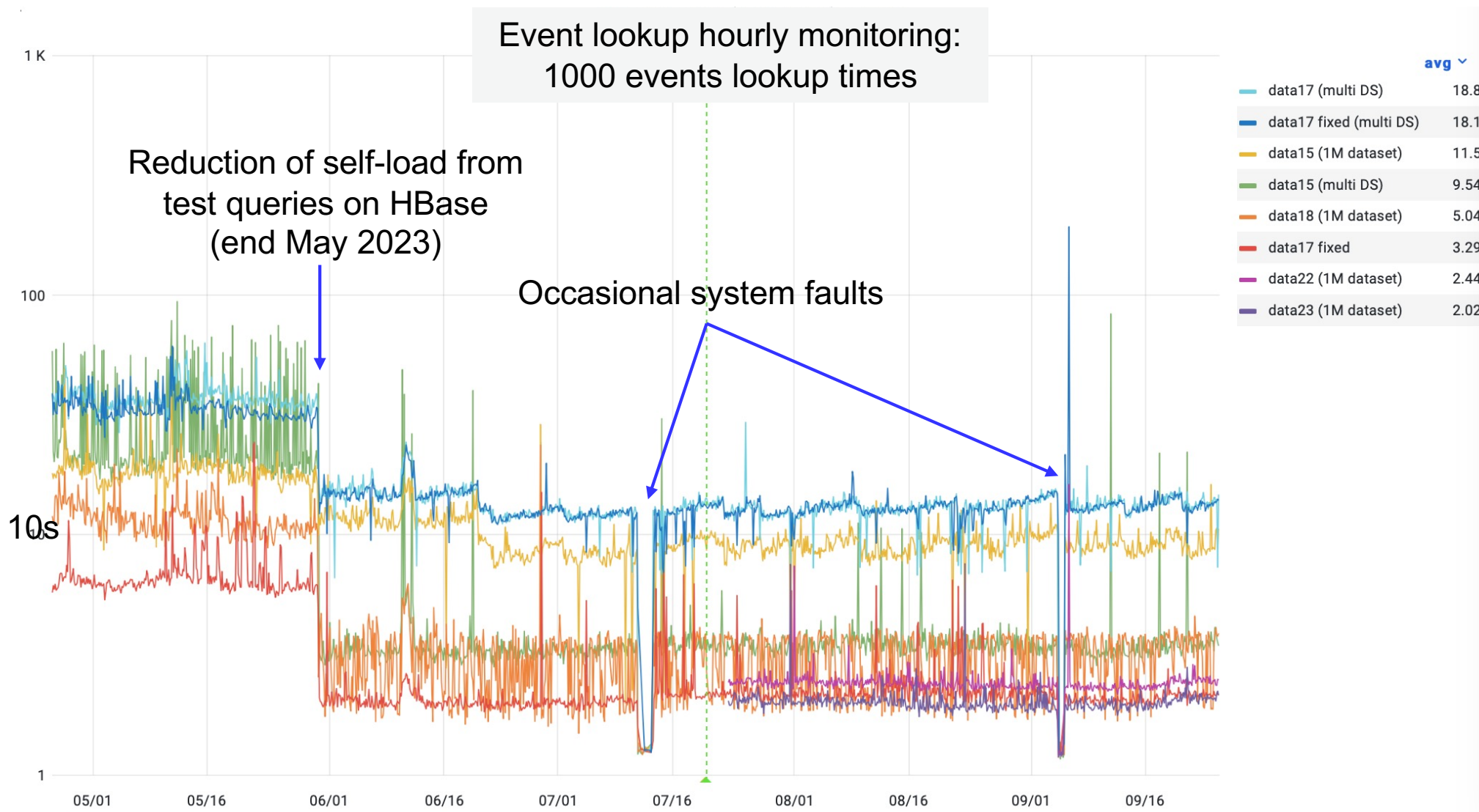


Indexed files



Total: 7.3M

2023 Operations (3)



2023 Operations (4)

Query Service access summary

Table: 2023-06-12 to 2023-10-01

Total 31140

OK 30924

Warn 204

Error 12

External 4358

OK 4223

Warn 126

Error 9

event-lookup 4050

event-list (counts) 160

event-list (events) 18

datasets-list 71

Work needed (1)

- Infrastructure:
 - Upgrade the operating system of all servers to AlmaLinux9
 - Best procedure to move (some) test servers first and then the corresponding production instances
 - ~~Or create a Kubernetes cluster for EventIndex services?~~
 - Timescale and exact sequence agreed with CSops team
 - First AlmaLinux9 server already available for testing
- Athena:
 - Move job configuration from jobOptions to Component Accumulator
 - JobOptions won't work any longer in future releases
 - Transition by the end of this year
 - A hackaton will be organised in December to help people
 - ART tests (discussed at [January workshop](#)) would have helped
 - But no personpower available
 - Affects Data Production and part of Event Picking in Athena

Work needed (2)

- Supervisor:
 - Move to Python 3
 - Priority for new hire at IFIC Valencia
 - Duplicated event counts
 - Recently found that the "Unique events" and "Duplicated events" fields in the dataset table are not filled properly
 - Trigger Counter
 - Still not fully implemented for Run 3
 - Integration of Data Production scripts (and conversion to Python 3)
- COMA interface:
 - Needed for production completion checks and web interface to event lookup

Work needed (3)

- Event Picking Server:
 - Robust error recovery
 - Formation and correct naming of final dataset
 - Periodic clean-up of EOS space used for output
- Monitoring:
 - Streamlining of Grafana dashboards
 - Removal of obsolete data sources and plots
 - Dashboard re-organisation
 - Completion of data sources to include (all vs time)
 - Number of datasets (from Data Collection or HBase table)
 - Number of events (from Data Collection or HBase table)
 - Size in HBase (from HBase cluster)
 - Client access statistics (from Data Query server)

2023 publications

- Run 2 paper (published):
 - The ATLAS EventIndex: A BigData Catalogue for All ATLAS Experiment Events
[*Comput.Softw.Big Sci.* 7 \(2023\) 1, 2](#)
- CHEP2023 proceedings (public but not published yet):
 - Deployment and Operation of the ATLAS EventIndex for LHC Run 3
<https://cds.cern.ch/record/2870346>
 - HBase/Phoenix-based Data Collection and Storage for the ATLAS EventIndex
<https://cds.cern.ch/record/2870347>
 - Query Service for the new ATLAS EventIndex system
<https://cds.cern.ch/record/2868540>
 - Testing framework and monitoring system for the ATLAS EventIndex
<https://cds.cern.ch/record/2868449>
- Grid2023 proceedings (public but not published yet):
 - The ATLAS Event Picking Service and its evolution
<https://cds.cern.ch/record/2871735>