Data Knowledge Base:
metadata integration system for HENP experiments

Marina Golosova[1], Vasilii Aulov[1], Mikhail Borodin[2], Maria Grigorieva[3], Anastasiia Kaida[4]

on behalf of ATLAS Collaboration

[1]National Research Center "Kurchatov Institute"
[2]University of Iowa
[3]Lomonosov Moscow State University
[4]National Research Tomsk Polytechnic University

Symposium on Nuclear Electronics and Computing (NEC'2019)

30.09.2019-04.10.2019, Montenegro, Budva

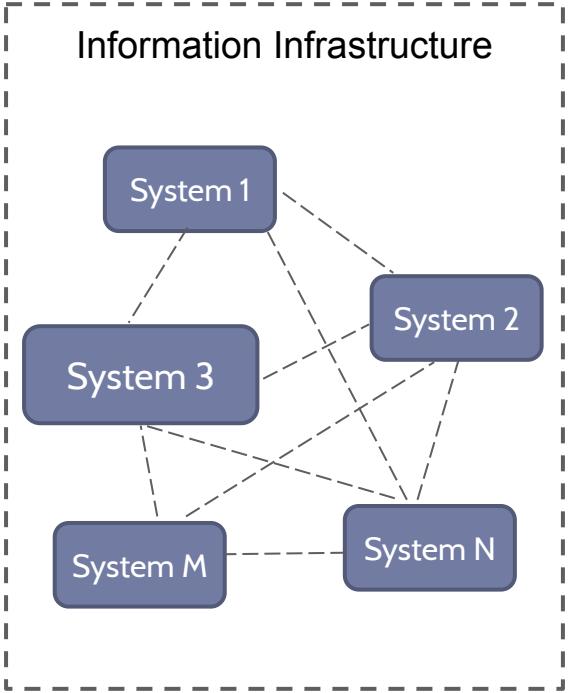
Outline

- HENP metadata
- Data Knowledge Base (DKB)
- DKB prototype: integration with the ATLAS Production System
- Future development: thoughts and plans

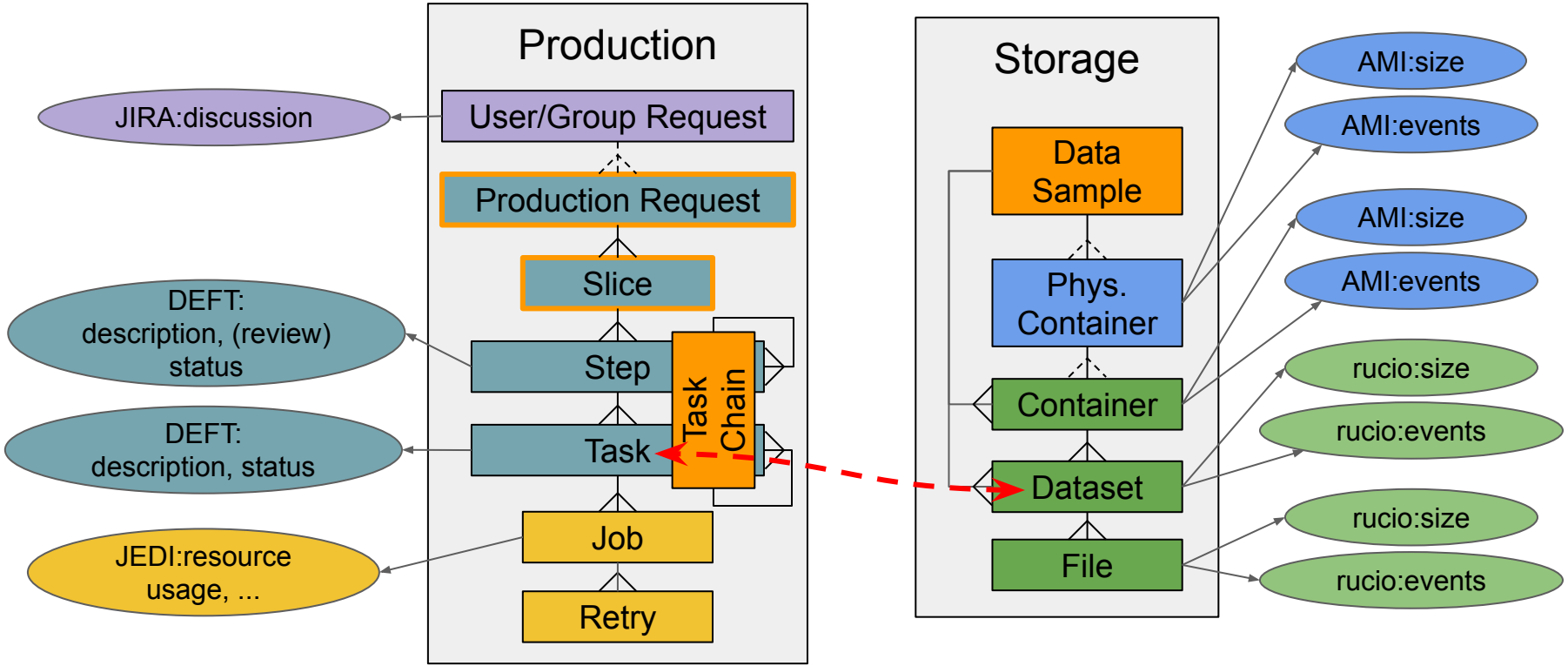
HENP metadata

- Experiment's data flow has **multiple aspects**: modelling, data taking, data processing, publications
- Information infrastructure reflects the diversity of the aspects and consists of **multiple subsystems** or even **standalone systems** taking care of a specific set of operations
- **Metadata** -- data about data, auxiliary data -- information about operations and objects, generated by those systems
- The more complicated systems are, the more **independent** they become -- however, they remain **semantically connected**
- Information about semantic links is not always stored together with other metadata, and sometimes is simply **known** by people
- Linking of separate pieces of information requires **manual** (or scripted) **work**, and analysis of joint metadata is performed **on the client side**

*HENP metadata are very **diverse**, have individual **specifics** for every project, and for long-living projects **volumes** of accumulated information make its **collective analysis** very resource-consuming.*

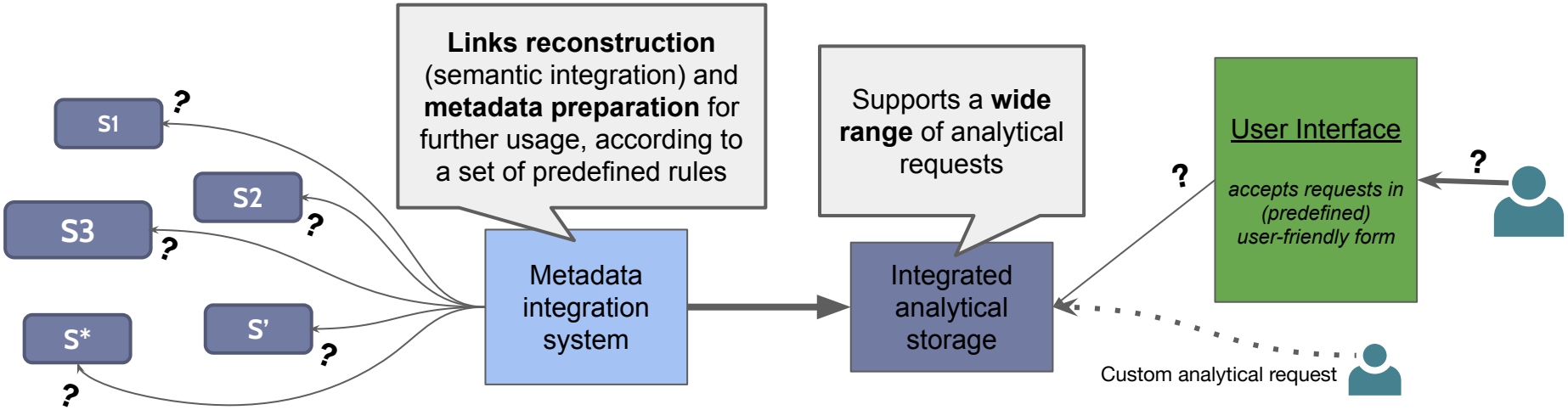


Hierarchy of data storage/processing concepts of the ATLAS experiment



Data Knowledge Base

- Semantics: simply **known** by people
- Linking: **manual** (or scripted) work **on the client side**
- High-level objects properties: the deeper into the hierarchy, the more **resource-consuming** is the aggregation



DKB integration with the ATLAS Production System

DKB instance is installed and configured at CERN

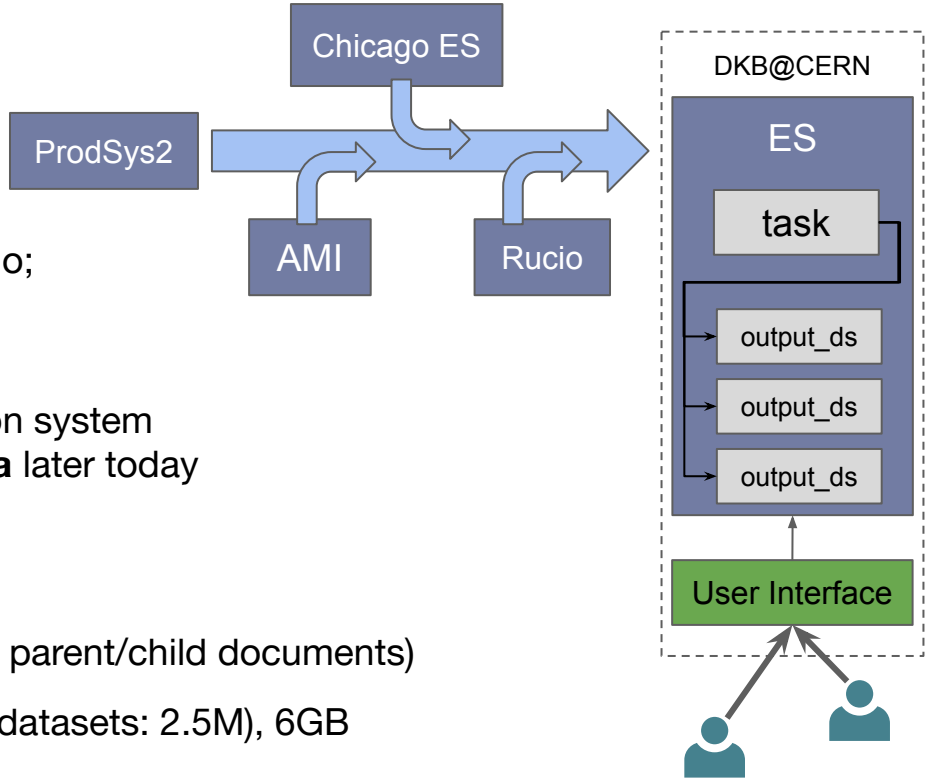
Metadata sources:
- ATLAS Production System 2;
- Rucio;
- ATLAS analytics platform in University of Chicago;
- ATLAS Metadata Interface (AMI).

Integration process:
- “Data streams processing in metadata integration system for HENP experiments”, talk of **Anastasiia Kaida** later today

Analytical storage: Elasticsearch

Metadata representation for analytical storage:
- objects of two types: task / output dataset (1:M, parent/child documents)

Current index volume: 4.9M documents (tasks: 2.4M, datasets: 2.5M), 6GB



DKB/ProdSys2: Web interface

ProdTask web-interface

- production steps resource usage statistics *(~10 sec)*
  https://prodtask-dev.cern.ch/dkb/#/steps_stat/?hashtag=MC16e_CP
  https://prodtask-dev.cern.ch/dkb/#/deriv_outputs_stat/?request_id=18030
- derivation efficiency statistics *(~12 sec)*
  https://prodtask-dev.cern.ch/dkb/#/deriv_ratio/?amitag=p3553&project=data18_13TeV
- task keyword search
  (task or request ID, campaign, …)
  (112 results, ~5.5 sec)
  https://prodtask-dev.cern.ch/dkb/#/task_keywords/18030,DAOD_EXOT4

Hashtags: MC16e_CP

Step Name	Evgen 99.36%	Evgen Merge 100.00%	Simul 99.79%	Merge 99.97%	Reco 100.00%
Tasks	403	403	403	65	403
Input_events	~2,259,931,000	~2,245,390,410	2,245,240,710	~470,998,050	2,240,389,166
Total_Events	2,245,390,610	2,245,355,560	2,240,545,216	470,868,346	2,240,328,166
Running/Pending	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%
Input_bytes	0 B (0 tasks)	3.07 TB (4 tasks)	60.98 TB (403 tasks)	100.48 TB (20 tasks)	1297.08 TB (403 tasl
Output_bytes	68.03 TB (57 tasks)	60.98 TB (403 tasks)	1094.80 TB (358 tasks)	302.76 TB (65 tasks)	35.13 TB (1 tasks)
Total HS06	4.27E13	2.87E10	8.97E13	5.39E10	2.44E13
Total Failed HS06	4.81E12	2.59E			
Duration	9.61 days	6.27			

Keywords: "18030" AND "DAOD_EXOT4"

Tasks found: 112 Tasks displayed: 112

Manage tasks

Filter

Default

data18_13TeV.00357713.physics_Main.deriv.f958_m2015_p3583 **34669139**

data18_13TeV.00358175.physics_Main.deriv.f961_m2015_p3583 **32050813**

TASK		EXPERIMEN	
taskID	15133829 BigPanda	AMI tag	p358
Request	18030	Campaign	data1
status	done	Subcampaign	
Description	open-ended production trains on data18_13TeV physicsMain with 21.2.34.0 starting with run 351359	Project	data1
User	atlas-dpd-production	Energy	1300
timestamp	19-08-2018 01:55:37	Physics Group	PHYS
start - end time	16-08-2018 18:42:04 18-08-2018 20:25:05	Physics Category	["Und
HS06 per event	44	Hashtags	
Outputs	["DAOD_EXOT13","DAOD_SUSY5","DAOD_SUSY7","DAOD_EXOT4","DAOD_HIGG4D2"]		
EVENTS Requested / Processed			

transPath	Reco_tf.py
transUses	AthDerivation-21.2.34.0
run number	358175

project: data18_13TeV

ami tag: p3553

Output	Ratio	Events ratio	Tasks
DAOD_BPHY1	0.3880%	6.2404%	5
DAOD_BPHY4	0.4969%	4.3871%	51
DAOD_BPHY5	0.2424%	0.6005%	8
DAOD_BPHY7	0.3241%	2.0861%	5
DAOD_EGAM1	1.2525%	4.5007%	57
DAOD_EGAM2	0.2130%	1.2257%	53
DAOD_EGAM3	0.1222%	0.4385%	56
DAOD_EGAM4	0.2017%	0.6928%	56

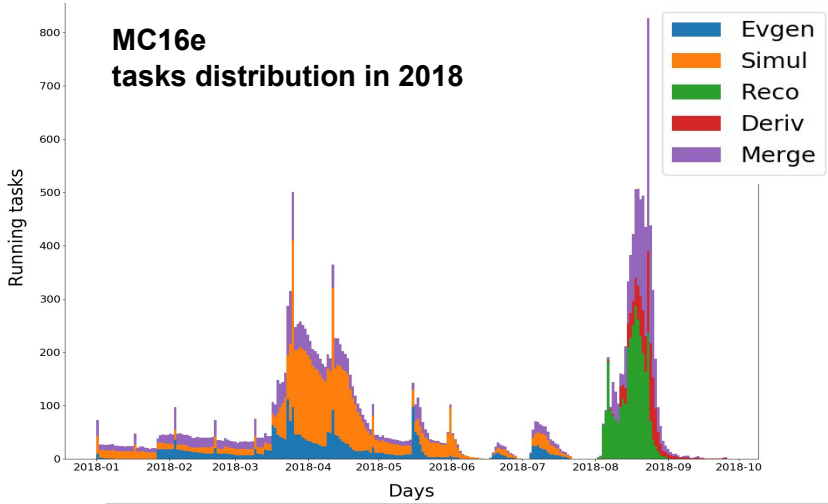
Marina Golosova

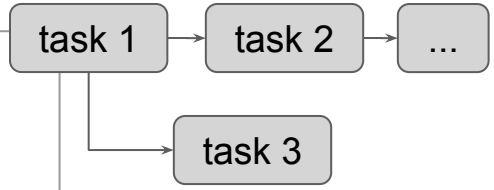
DKB/ProdSys2: REST API

REST API

- tasks distribution vs production step for given campaign *(here: MC16e, Monte-Carlo simulation campaign in 2019)*
  http://aiatlas172.cern.ch:5080/task/hist?htags=mc16e_cp&start=2018-01-01&end=2019-01-01&rtype=img
  http://aiatlas172.cern.ch:5080/task/hist?htags=mc16e_cp&start=2018-01-01&end=2019-01-01

- task chain reconstruction:
  http://aiatlas172.cern.ch:5080/task/chain?tid=16655409&pretty






```
{"status": "OK",
 "data": {taskID_1: [taskID_2,
                     taskID_3,
                     ...],
          taskID_2: [...],
          taskID_3: [], ...}}
```

```
{"status": "OK",
 "data": {
  "data": {"x": [[...],[...],...],
           "y": [[...],[...],...]},
  "legend": ["Reco", "Deriv", ...]
}
```

Response time performance

In **1993** Jakob Nielsen in his *Usability Engineering* stated that **10 sec** is the **upper limit** for keeping user's attention. But life's getting faster, and now people tend to add adjustment comments to this limit: with **2-5 sec** the system considered "responsive", while with **5-10 sec** it is closer to "slow" (e.g. *UI Response Times* note by Steve Henty (2015)).

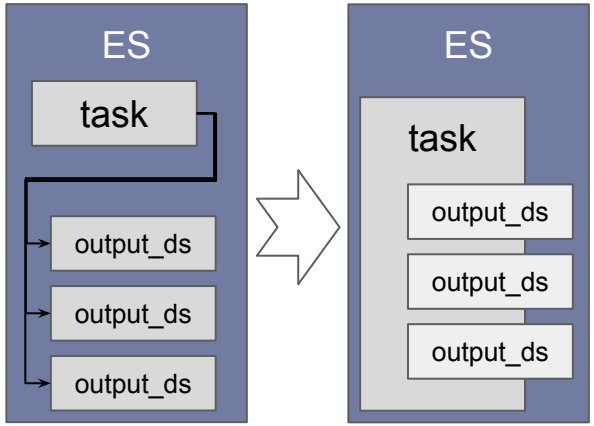
Currently we have ~10 sec; can we have get closer to the "responsive"?

Hypothesis:
single-object scheme with nested fields (instead of currently used two-objects parent/child scheme) may improve performance:

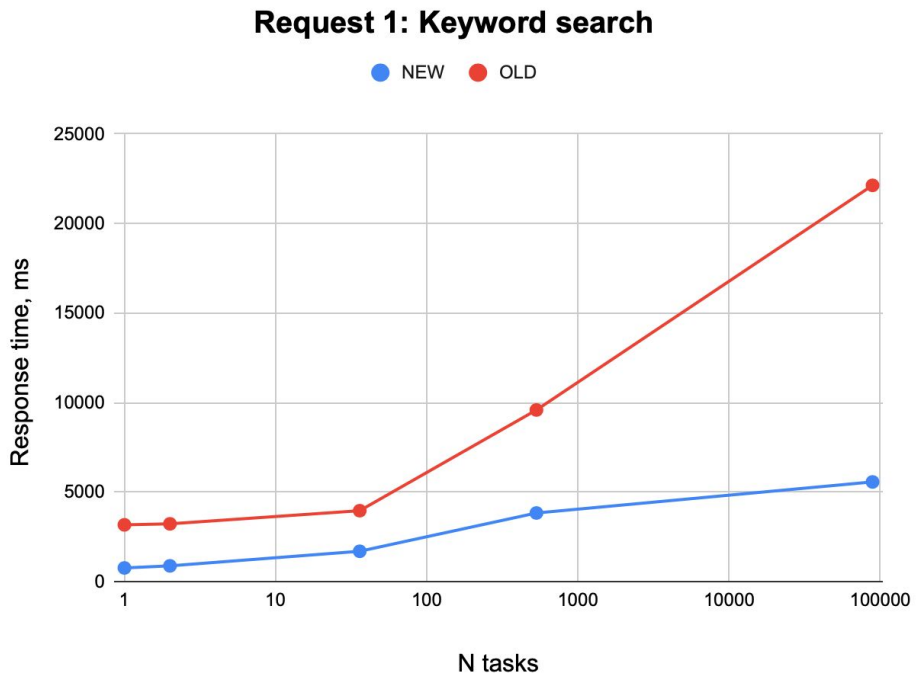
- queries will be more simple *(e.g. single query instead of two or multiple ones)*;
- better search performance in general.

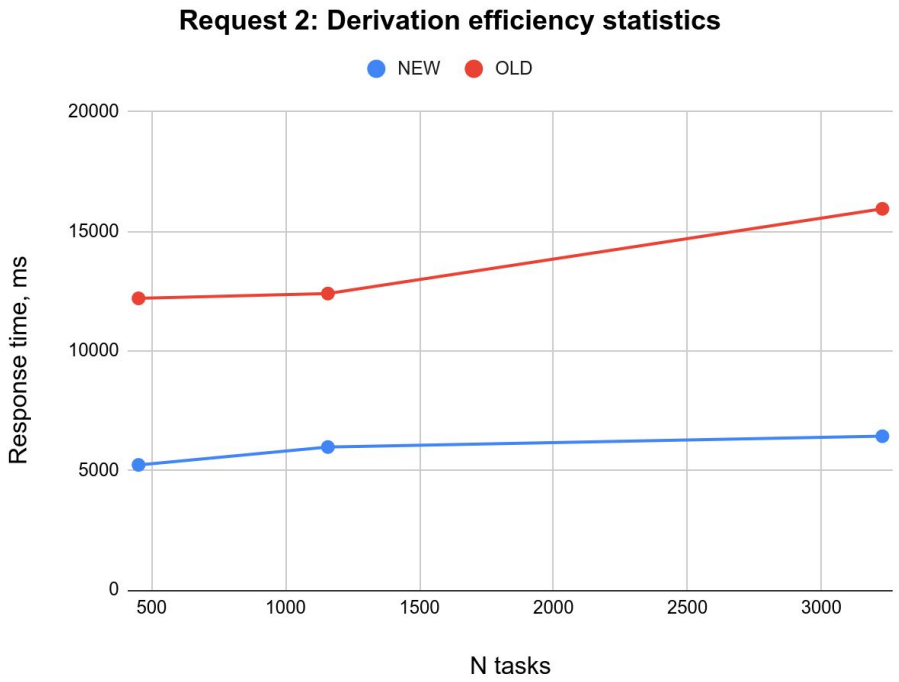
Testing:

- 1 node, 4GB heap space;
- index volume: 4M records (~2M tasks, ~2M datasets) (5GB);
- no caching (both ES and disk cache cleaned after every request);
- 4 real-life requests adapted to the new scheme.

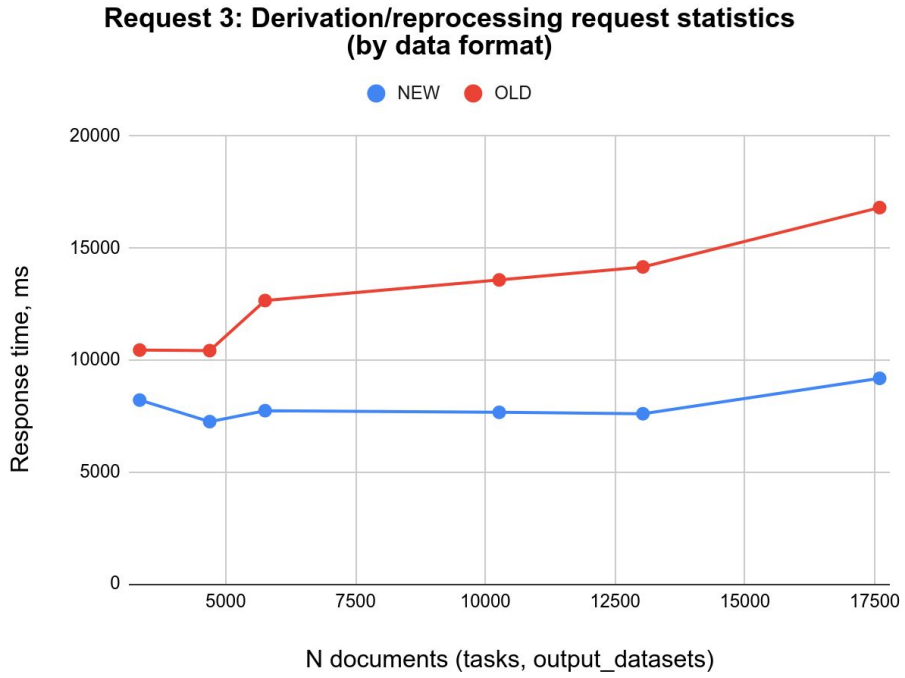


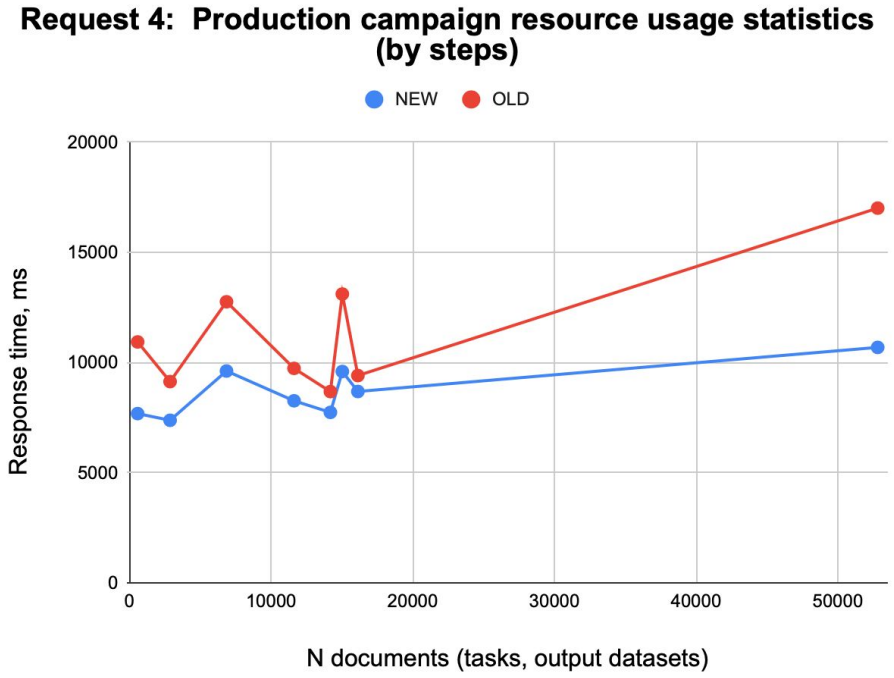
Test results (1)





Test results (2)





Summary

- Metadata integration makes analytics tasks easier (even if analytical storage scheme is not optimal)
- In the addressed use-cases response time is tolerable, but still does not fit the **interactive** operations requirement (should be less than **5 sec**)
- Performance tests demonstrated that ES scheme with nested fields will serve current needs better than the originally designed one, so in the future DKB instance integrated with ATLAS Production System 2 will switch to the new scheme

Acknowledgements

This work is supported by Russian Science Foundation
under contracts №18-37-20003

Special thanks to:

Alexei Klimentov
Brookhaven National Laboratory

Alexei Saveliev
National Research Tomsk Polytechnic University

Dmitry Golubkov
National Research Center "Kurchatov Institute"

Backup

Glossary

ATLAS Production System 2

top level workflow management system

ProdTask

web-interface for production processes management (https://prodtask-dev.cern.ch/)

Rucio

Distributed Data Management system (https://rucio.cern.ch)

AMI

ATLAS Metadata Interface (https://ami.in2p3.fr)

Derivation and reprocessing statistics

https://prodtask-dev.cern.ch/dkb/#/deriv_outputs_stat/?request_id=18030

ATLAS PanDA | Requests | Tasks | Jobs | Train | Datasets | Meta | PanDA Config

DKB

18030

Step Name	DAOD_EXOT4 p3583 100.00%	DAOD_SUSY7 p3583 100.00%	DAOD_EXOT22 p3583 100.00%	DAOD_STDM5 p3583 100.00%	DAOD_HIGG8D1 p3583 100.00%	DAOD_HIGG2D5 p3583 100.00%	DAOD_JETM10 p3583 100.00%	DAOD_JETM11 p3583 100.00%	DAOD_EGAM2 p3583 100.00%	DAOD_EXOT17 p3583 100.00%	DAOD_HIG p3583 100.0
Tasks	112	210	210	210	210	210	210	210	159	209	210
Input_events	2,247,080,254	5,075,183,617	5,075,183,617	5,075,183,617	5,075,183,617	5,075,183,617	5,075,183,617	5,075,183,617	3,722,825,086	5,033,071,758	5,075,183,61
Total_Events	1,843,863,199	4,147,624,965	5,883,944,254	347,187,348	9,259,102,795	94,664,361	71,396,859	5,901,736,048	269,904,038	707,743,747	1,900,867,07
Running/Pending	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%
Input_bytes	490.73 TB (112 tasks)	1198.49 TB (210 tasks)	1198.49 TB (210 tasks)	1198.49 TB (210 tasks)	1198.49 TB (210 tasks)	1198.49 TB (210 tasks)	1198.49 TB (210 tasks)	1198.49 TB (210 tasks)	867.36 TB (159 tasks)	1187.22 TB (209 tasks)	1198.49 TB (tasks)
Output_bytes	50.23 TB (477 tasks)	110.47 TB (824 tasks)	151.89 TB (1675 tasks)	14.77 TB (1156 tasks)	444.86 TB (1787 tasks)	1.24 TB (630 tasks)	4.18 TB (628 tasks)	152.21 TB (1676 tasks)	11.73 TB (952 tasks)	35.92 TB (1672 tasks)	89.01 TB (1 tasks)
Total HS06	5.25E11	1.01E12	1.05E12	3.06E11	2.13E12	3.63E11	2.60E11	1.06E12	2.26E11	3.59E11	7.82E11
Total Failed HS06	3.88E10	7.98E10	1.16E11	1.24E10	2.19E11	1.92E10	1.84E10	1.16E11	9.68E9	3.68E10	4.93E10
Duration	2.26 days	3.33 days	3.58 days	2.58 days	3.99 days	2.41 days	2.52 days	3.57 days	2.68 days	2.71 days	3.21 days

Keywords (Google-like) search

- Find tasks and related datasets by any information known: task or request ID, run number, AMI tags, data format, geometry version, project, campaign, ...
- Get all the information available in a summary table on one page.

Search keywords:
18030,DAOD_EXOT4

Search

Analysis

https://prodtask-dev.cern.ch/dkb/#/task_keywords/18030,DAOD_EXOT4

Keywords: "18030" AND "DAOD_EXOT4"

Tasks found: 112 Tasks displayed: 112

Manage tasks

Filter

Default

14865286 data18_13TeV.00357713.physics_Main.deriv.f958_m2015_p3583 **34669139**

15133829 data18_13TeV.00358175.physics_Main.deriv.f961_m2015_p3583 **32050813**

TASK		EXPERIMENT		CONFIGURATION	
taskID	15133829 BigPanda	AMI tag	p3583	Step Name	Reco
Request	18030	Campaign	data18_13TeV	ticket_id	ATLPSTASKS-1557948
status	done	Subcampaign		Architecture	x86_64-slc6-gcc62-op
Description	open-ended production trains on data18_13TeV physicsMain with 21.2.34.0 starting with run 351359	Project	data18_13TeV	Core Number	8
User	atlas-dpd-production	Energy	13000	ATLAS Geometry	
timestamp	19-08-2018 01:55:37	Physics Group	PHYS	Conditions Tags	
start - end time	16-08-2018 18:42:04 18-08-2018 20:25:05	Physics Category	["Uncategorized"]	trigger_config	
HS06 per event	44	Hashtags		transPath	Reco_tf.py
Outputs	["DAOD_EXOT13","DAOD_SUSY5","DAOD_SUSY7","DAOD_EXOT4","DAOD_HIGG4D2"]			transUses	AthDerivation-21.2.34.0
EVENTS Requested / Processed 32050813 / 32050813				run number	358175
Datasets					
Input Datasets	data18_13TeV.00358175.physics_Main.merge.AOD.f961_m2015				9759088099934
Output Datasets					**Size(bytes)**
data18_13TeV.00358175.physics_Main.deriv.DAOD_EXOT13.f961_m2015_p3583_tid15133829_00					221350172557
cross section:		filt eff:		events:	10914429

Marina Golosova

Derivation statistics

ATLAS PanDA | Requests | Tasks | Jobs | Train | Datasets | Meta | PanDA Config

DKB

project: data18_13TeV

ami tag: p3553

Output	Ratio	Events ratio	Tasks
DAOD_BPHY1	0.3880%	6.2404%	5
DAOD_BPHY4	0.4969%	4.3871%	51
DAOD_BPHY5	0.2424%	0.6005%	8
DAOD_BPHY7	0.3241%	2.0861%	5
DAOD_EGAM1	1.2525%	4.5007%	57
DAOD_EGAM2	0.2130%	1.2257%	53
DAOD_EGAM3	0.1222%	0.4385%	56
DAOD_EGAM4	0.2017%	0.6928%	56
DAOD_EGAM5	1.3750%	8.9922%	28
DAOD_EGAM7	0.4618%	2.2212%	53
DAOD_EGAM9	0.5317%	2.5934%	22
DAOD_EXOT0	0.4847%	2.8099%	55
DAOD_EXOT10	0.0324%	0.1167%	5
DAOD_EXOT12	0.2466%	2.5883%	54

Output data format

Output/input dataset size (bytes) ratio

Ratio of the number of events in output/input datasets

Number of tasks

Derivation statistics in the aggregated form is not available **'out of the box'**, by a single click, anywhere else.

https://prodtask-dev.cern.ch/dkb/#/deriv_ratio/?amitag=p3553&project=data18_13TeV

Production steps statistics

https://prodtask-dev.cern.ch/dkb/#/steps_stat/?hashtag=MC16e_CP

ATLAS PanDA | Requests | Tasks | Jobs | Train | Datasets | Meta | PanDA Config

DKB

Hashtags: MC16e_CP

Step Name	Evgen 99.36%	Evgen Merge 100.00%	Simul 99.79%	Merge 99.97%	Reco 100.00%	Rec Merge 99.97%	Deriv 100.00%	Deriv Merge 99.13%
Tasks	403	403	403	65	403	403	403	403
Input_events	~2,259,931,000	~2,245,390,410	2,245,240,710	~470,998,050	2,240,389,166	~2,240,559,000	2,239,858,166	~2,259,500,000
Total_Events	2,245,390,610	2,245,355,560	2,240,545,216	470,868,346	2,240,328,166	2,239,858,166	2,239,858,166	2,239,858,166
Running/Pending	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%	0%/0%
Input_bytes	0 B (0 tasks)	3.07 TB (4 tasks)	60.98 TB (403 tasks)	100.48 TB (20 tasks)	1297.08 TB (403 tasks)	35.13 TB (1 tasks)	1128.97 TB (403 tasks)	0 B (0 tasks)
Output_bytes	68.03 TB (57 tasks)	60.98 TB (403 tasks)	1094.80 TB (358 tasks)	302.76 TB (65 tasks)	35.13 TB (1 tasks)	1128.97 TB (403 tasks)	0 B (0 tasks)	16.56 MB (403 tasks)
Total HS06	4.27E13	2.87E10	8.97E13	5.39E10	2.44E13	1.81E11	9.60E9	6.17E7
Total Failed HS06	4.81E12	2.59E9	9.89E12	3.07E9	8.13E11	2.03E10	8.12E8	2.55E6
Duration	9.61 days	6.27 days	17.84 days	19.83 days	6.82 days	4.29 days	1.99 days	0.26 days

REST API

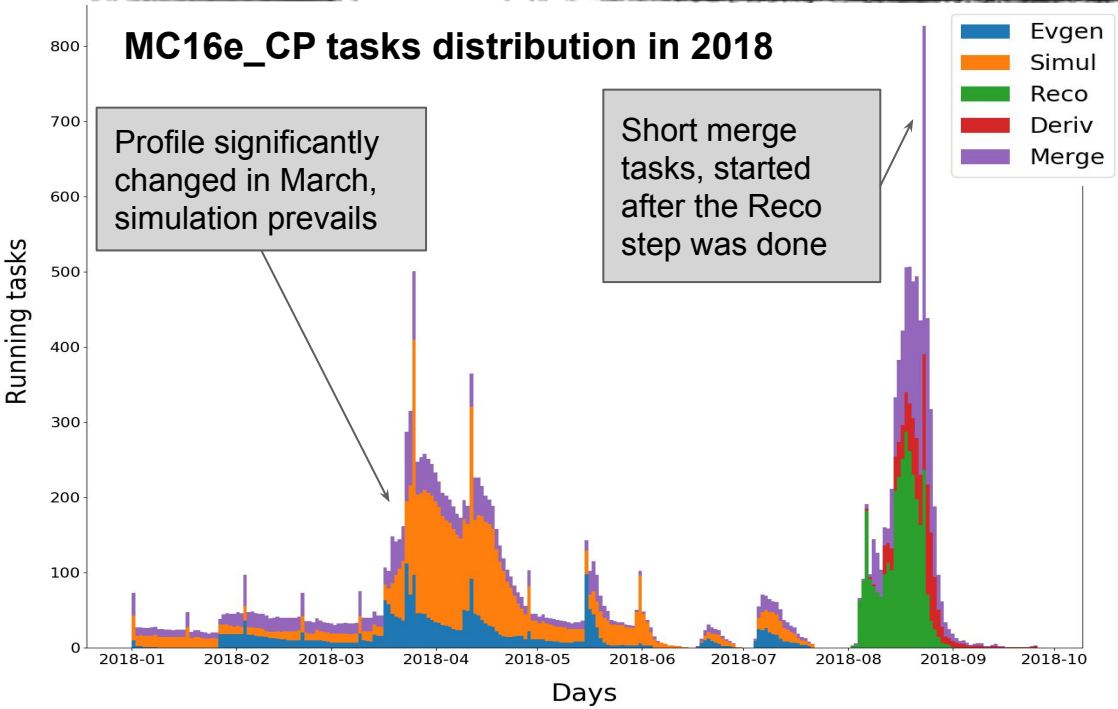
Status: development
Basic data format: JSON
Methods available:

- /server_info
- /info
- /task/hist (img, json): *production step tasks distribution over time*
  - ?htag=HASHTAG1&htag=HASHTAG2&... (“OR”)
  - ?htag=HTAG1,HTAG2,... (“OR”)
  - ?start=YYYY-MM-DD[THH:MM:SS]
  - ?end=YYYY-MM-DD[THH:MM:SS]
  - ?bins=N (allows to get less/more detailed histogram)
  - ?detailed (allows to see individual merge steps)
- /task/chain: *task chain reconstruction data (BigPanDA version)*
  - ?tid=TASK_ID

Available from CERN:
http://aiatlas172.cern.ch:5080

More methods and proper documentation are coming!

REST API: task steps statistics



http://aiatlas172.cern.ch:5080/task/hist?htags=mc16e_cp&start=2018-01-01&end=2019-01-01&rtype=img

JSON:

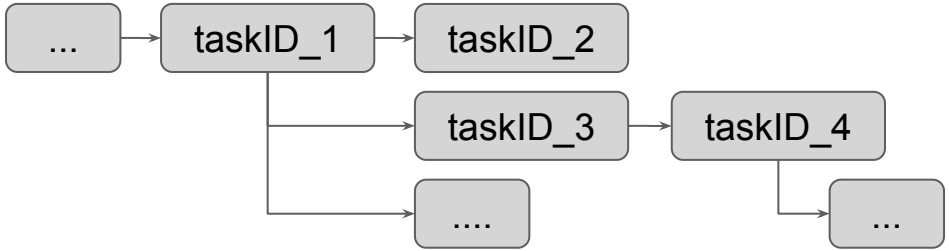
```
{
  "status": "OK",
  "data": {
    "legend": ["Evgen", "Simul", …],
    "data": {
      "x": [
        [evgenX_1, evgenX_2, …],
        [simulX_1, simulX_2, …],
        …
      ],
      "y": [
        [evgenY_1, evgenY_2, …],
        [simulY_1, simulY_2, …],
        …
      ]
    }
  }
}
```

REST API: task chain reconstruction

Task chain reconstruction:
http://aiatlas172.cern.ch:5080/task/chain?tid=16655409&pretty

```
{
  "status": "OK",
  "data": {
    taskID_1: [taskID_2,
               taskID_3,
               …],
    taskID_2: [],
    taskID_3: [taskID_4],
    …
  }
}
```

More information can be added to the structure or retrieved for each task individually.



https://bigpanda.cern.ch/ganttTaskChain/?jeditaskid=7999893