NEC'2019

XXVII International Symposium on Nuclear Electronics & Computing

Montenegro, Budva, Becici, 30 September - 4 October 2019

Monitoring and Accounting for the Distributed Computing System of the ATLAS Experiment

Dario Barberis[1], Alberto Aimar[2], Aleksandr Alekseev[3], Pedro Andrade[2], Thomas Beermann[4], Robert Gardner[5], Borja Garrido Bear[2], Tatiana Korchuganova[3], Luca Magnoni[2], Siarhei Padolski[6], Eric Schanet[7], Nikolay Tsvetkov[2], Ilija Vukotić[5], Torre Wenaus[6]

1) Genoa Univ./INFN, 2) CERN-IT, 3) Moscow State Univ. & Plekhanov Russian Univ. of Economics, 4) Wuppertal Univ., 5) Univ. of Chicago, 6) Brookhaven National Laboratory, 7) LMU Munich

1

Introduction

- ATLAS used during LHC Run 1 and Run 2 a monitoring and accounting infrastructure for the Distributed Computing (ADC) applications developed ~10 years ago by CERN-IT together with ATLAS members
  - These "old dashboards" started showing aging effects in the last few years:
    - Slowness of data retrieval due to the massive amount of data in Oracle databases
    - Lack of in-depth knowledge for maintenance as original developers left long ago
    - Lack of flexibility and impossibility to develop new views and/or data correlations across different data sources
  - This system worked well enough for general monitoring till the end of Run 2 last year but was evidently in need of a good refurbishing
- Since 2016 the CERN-IT MonIT group started developing a new infrastructure and environment for monitoring and accounting applications base on modern Open Source components
  - ATLAS started implementing "new" dashboards using this infrastructure, for data and workload accounting and global monitoring
- In the meantime the BigPandaMon application was developed for user and task oriented monitoring of the jobs submitted to the ATLAS Grid/Cloud/HPC resources through PanDA
  - This is now the workhorse of user-level job monitoring
- In the recent years many Analytics tools appeared on the market. They can be used for more detailed investigations and to correlate data from different sources
  - The Analytics cluster provided by the University of Chicago allows a more interactive use of monitoring data

2

Topics

- ### ADC dashboards in the MonIT infrastructure

- ### User level job/task monitoring with BigPandaMon

- ### Analytics cluster at UC and its applications

Entry page for ADC Monitoring:
https://atlas-adc-monitoring.web.cern.ch

Dario Bar

Topics

- ADC dashboards in the MonIT infrastructure

- User level job/task monitoring with BigPandaMon

- Analytics cluster at UC and its applications

The MonIT Infrastructure

- Monitoring Mission:
  - Provide Monitoring as a Service for CERN Data Centre (DC), IT Services and the WLCG collaboration
    - e.g. Dashboards, Alarms, Search, Archive
  - Collect, transport, store and process metrics and logs for applications and infrastructure
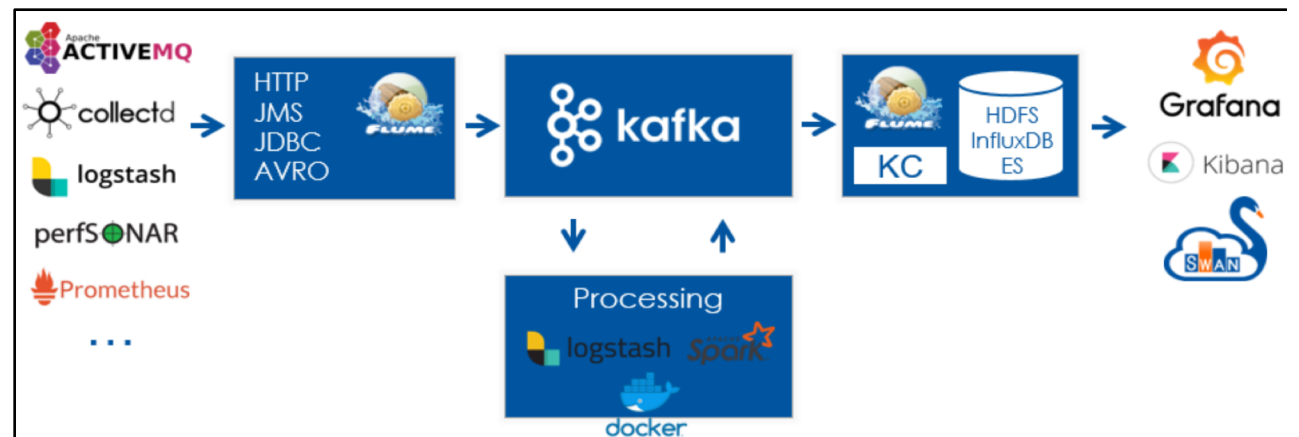- Challenges:
  - Rate & Volume
    - from ~ 40k machines
    - > 3 TB/day
    - ~ 100 kHz
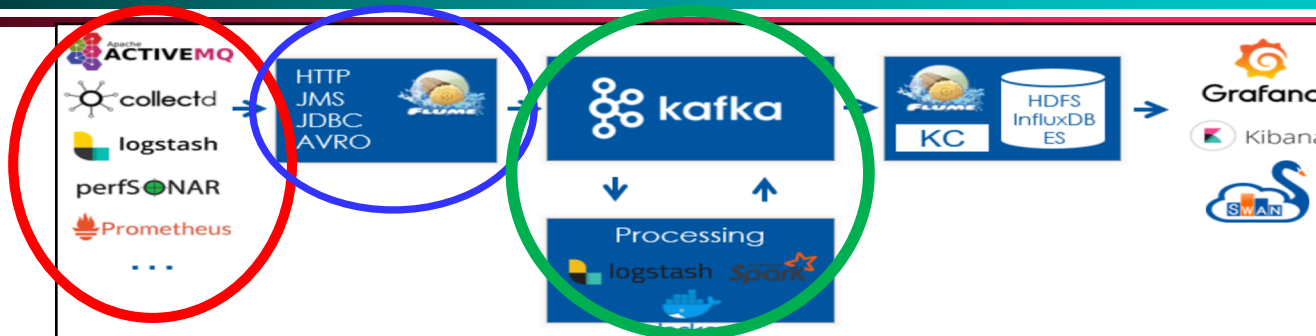 - Variety
    - > 150 producers
  - Reliability
    - spikes in rate and volume

The MonIT Infrastructure

Collectors / Collectd

- What is Collectd
  *"collectd is a **daemon** which **collects system and application performance metrics** periodically and provides mechanisms to store the values in a variety of ways"*
- Collectd for CERN DC and IT Services
  - Default HW/OS plugins on all DC machines
  - Additional plugins configured by IT Service Managers
    - Community-plugins or custom-made
  - Provide local alarms (~ Nagios Check) and actuators
  - Sampling intervals: 1 to 5 minutes

Connectors / Apache Flume

We rely on several Flume features:
- Transactional nature
- Pull/push modes
- Protocol-based agents (sources and sinks)
  - AVRO, JDBC, JMS, HDFS, Elasticserch, HTTP, Kafka
- Interceptor / Morphlines for event transformation and validation

Collectors / Flume & Logstash

- Flume for basic log forwarding
  - Running on all DC nodes for syslog
- Logstash for advanced use cases
  - Many read/write plugins
  - Allows log parsing and tokenization
- Other lightweight options getting popular
  - Filebeat, Fluentd

Transport / Apache Kafka

- Kafka as rock-solid core of our pipeline
  - decouples producers / consumers
  - enables stream processing
  - resilient (72 hours data retention)

The MonIT Infrastructure

Storage / InfluxDB

- Time Series DB for time series data (metrics)
- Data kept for ~5 years (auto down-sampling)
- >30 instances (from 8 to 128GB mem each)
- Data cardinality matters
- Aggregated write: ~65k points per second
- Gross read (single instance): ~100k pps

Storage / HDFS

- HDFS for long term archive
- Compressed JSON (or Parquet)
- Data kept ~ forever (or by GDPR agreement)
- Works well with batch and data-intensive analytics workflows
  - e.g. Hadoop, Spark

Storage / Elasticsearch

- Distributed search and indexing engine
- Powerful query language to browse and explore
  - Apache Lucene based
- 3 clusters (syslog, service logs and metrics)
  - ~100 TB (total storage size)
- Data kept for 1 month

Visualization

- Grafana for dashboards
  - Users can create their own
- Kibana for data exploration
  - Data discovery and logs
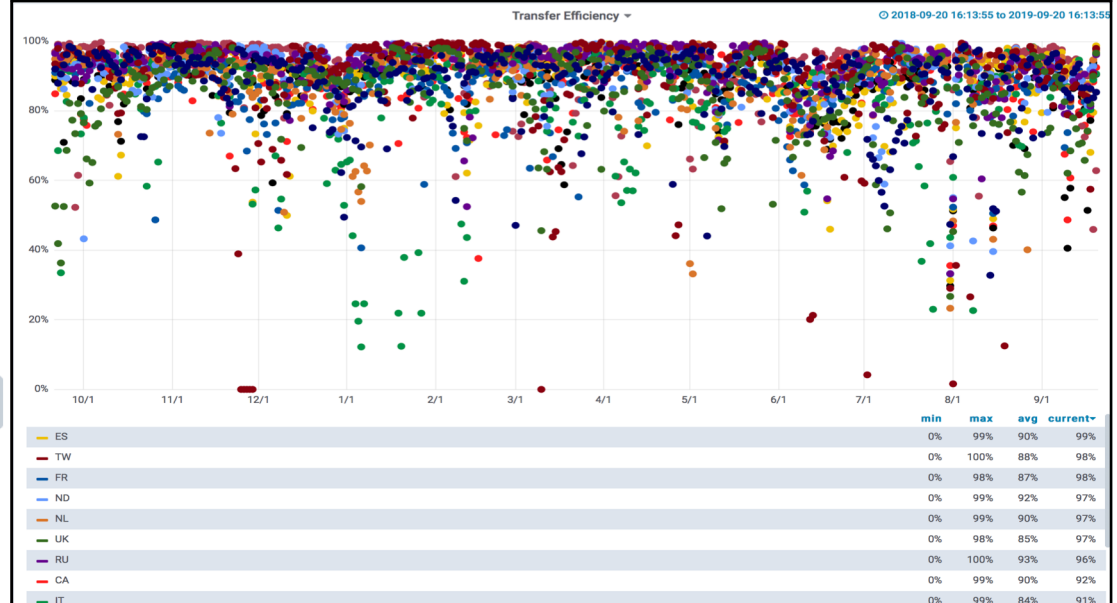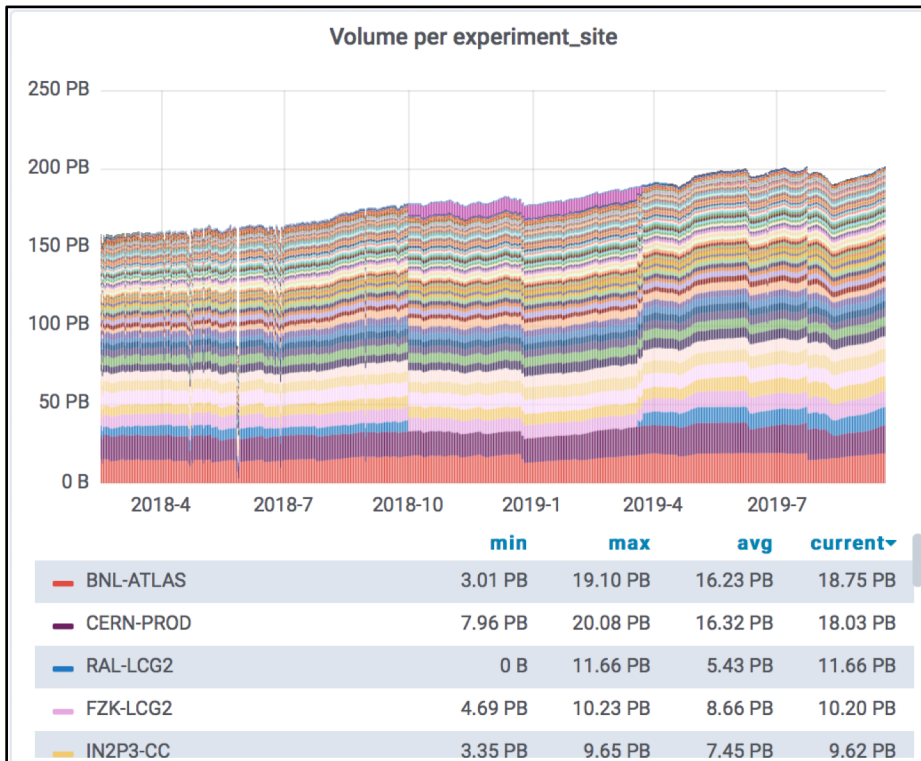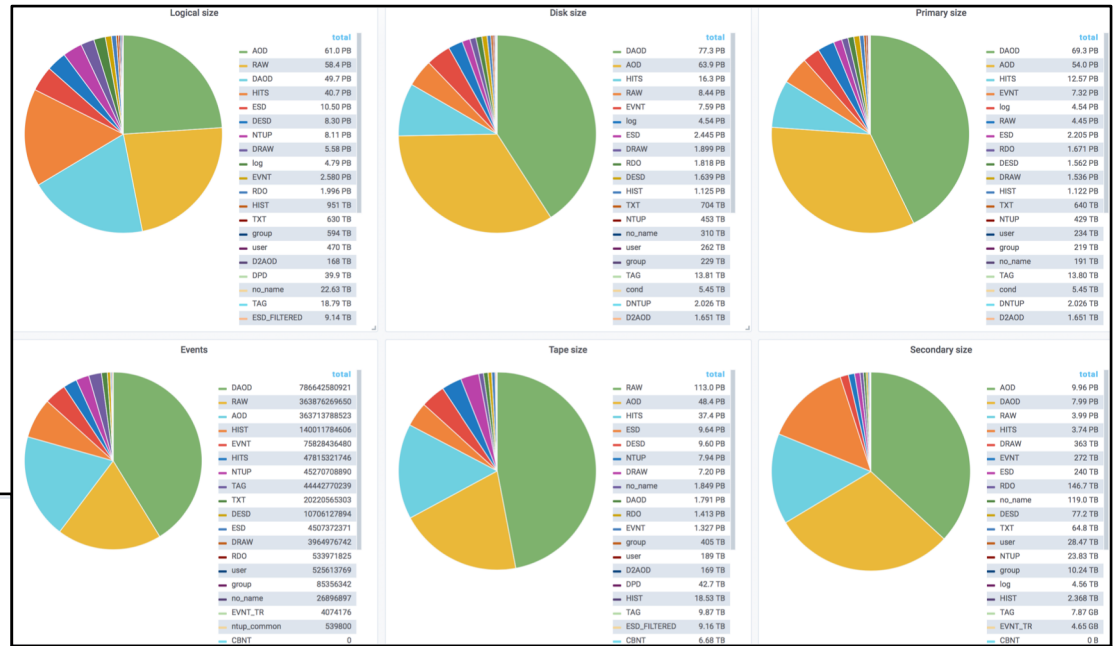- SWAN for analytics (notebooks)
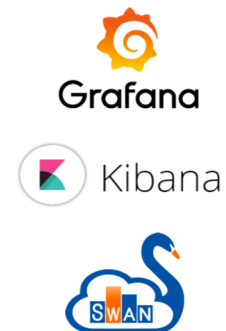
ATLAS Dashboards

- "MonIT Dashboards" for Data and Job monitoring and accounting (and a number of other applications too)
    - Collect data from the Rucio and PanDA databases in Oracle and other sources
    - Transfer to BigData infrastructure using Kafka – possibility of data enrichment and correlations with other sources of information
        - (e.g. AGIS for the relation between PanDA queues, sites, federations, countries, pledges etc.)
    - Storage in HDFS with aggregation in InfluxDB
    - Display with Grafana
        - Still far from perfect as display tool for us but usable
        - Used by ADC shifters and site admins (monitoring) and computing managers (accounting)
- Three groups of dashboards:
    - Production
    - Development (pre-production)
    - Playground (free for all)
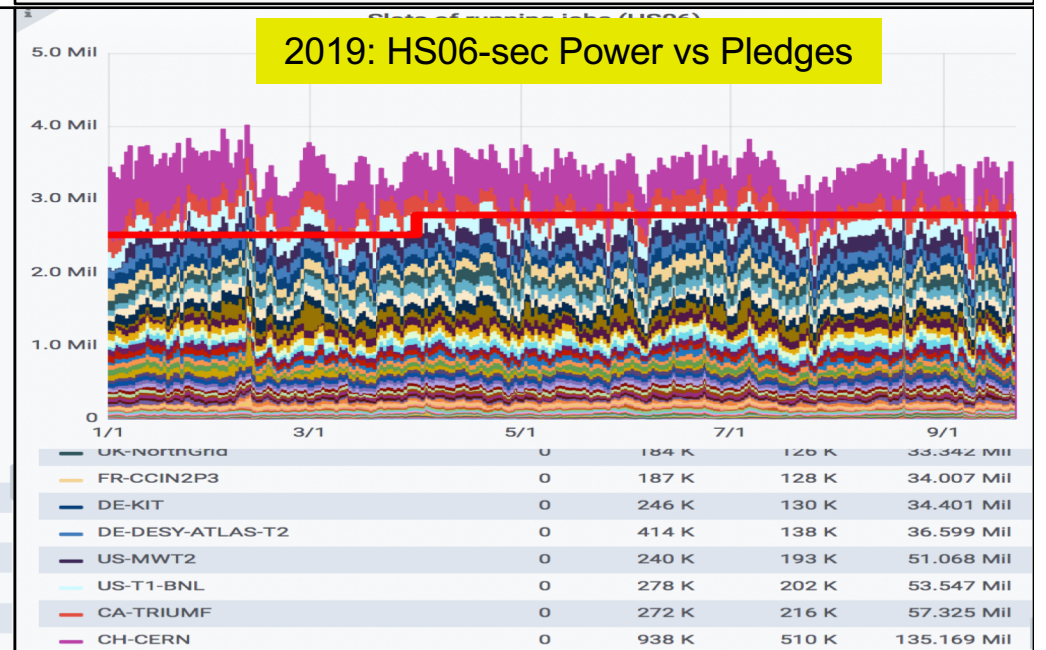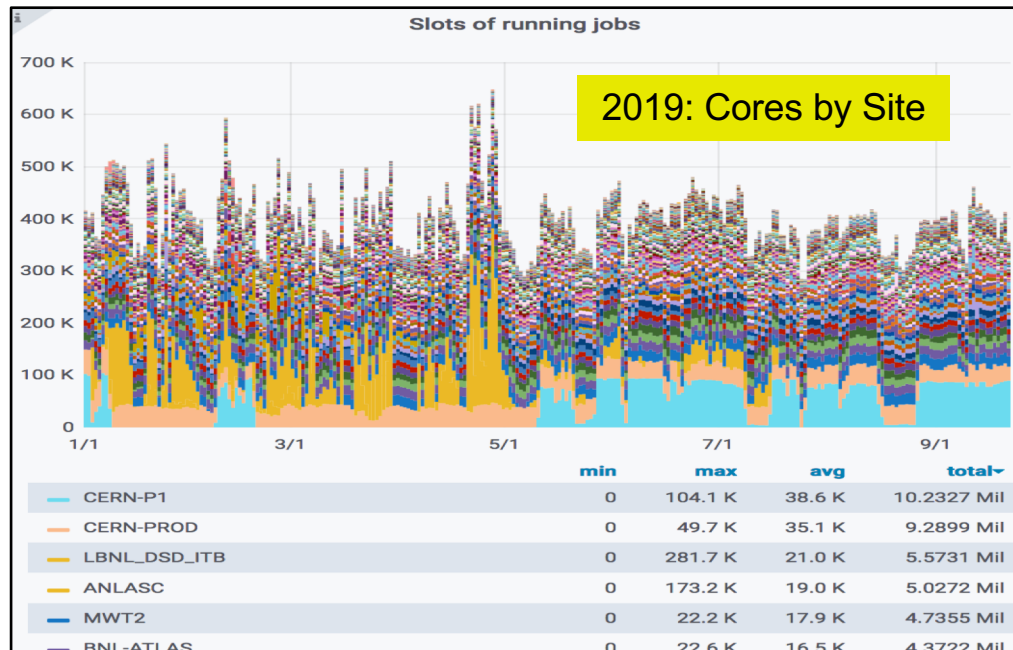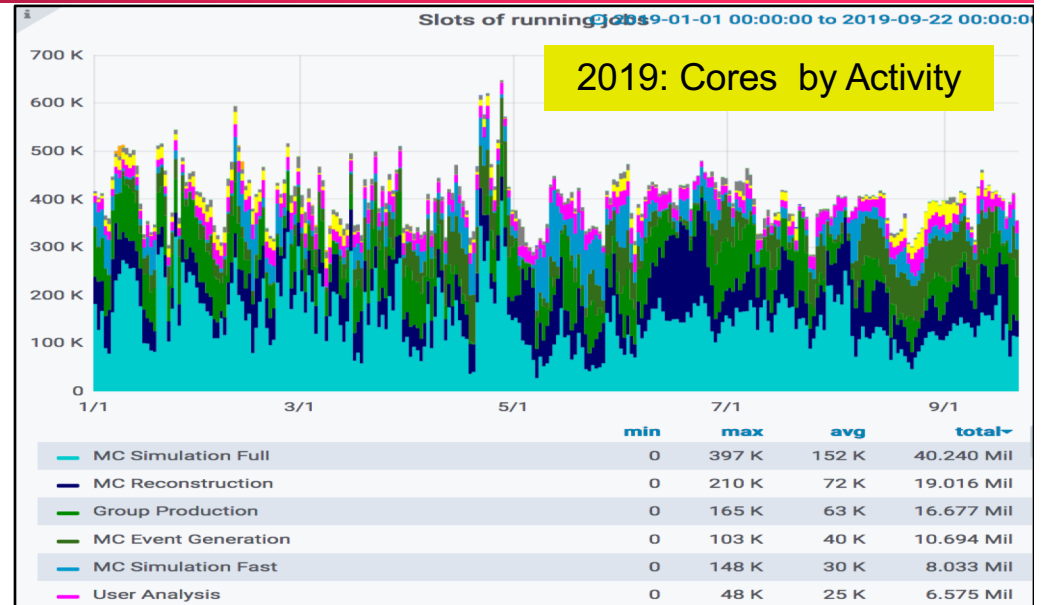
Data Management (DDM) Dashboards

- Traces and events reported constantly by Rucio to the message brokers and further processed in the MonIT infrastructure
- Rucio data are dumped periodically from Oracle to HDFS for site accounting views
  - Then used from there to fill InfluxDB
- Many many views available
  - Historical views and snapshots
  - Data storage volumes
  - Transfer rates and efficiencies

Volume per experiment_site

	min	max	avg	current▼
BNL-ATLAS	3.01 PB	19.10 PB	16.23 PB	18.75 PB
CERN-PROD	7.96 PB	20.08 PB	16.32 PB	18.03 PB
RAL-LCG2	0 B	11.66 PB	5.43 PB	11.66 PB
FZK-LCG2	4.69 PB	10.23 PB	8.66 PB	10.20 PB
IN2P3-CC	3.35 PB	9.65 PB	7.45 PB	9.62 PB

Transfer Efficiency

|	min	max	avg	current▼
ES	0%	99%	90%	99%
TW	0%	100%	88%	98%
FR	0%	98%	87%	98%
ND	0%	99%	92%	97%
NL	0%	99%	90%	97%
UK	0%	98%	85%	97%
RU	0%	100%	93%	96%
CA	0%	99%	90%	92%
IT	0%	99%	84%	91%

Job Dashboard

- Data collected from the PanDA DB every 10 minutes and grouped in 1-hour bins
- Transient job stats:
  - Pending, running, finalising jobs
- Permanent job stats:
  - Submitted and completed jobs
- All job parameters on submission and completion
  - Plus site info from AGIS
  - All available for selections and display
- Data imported from start of Run 1

2019: Cores by Activity

	min	max	avg	total▾
MC Simulation Full	0	397 K	152 K	40.240 Mil
MC Reconstruction	0	210 K	72 K	19.016 Mil
Group Production	0	165 K	63 K	16.677 Mil
MC Event Generation	0	103 K	40 K	10.694 Mil
MC Simulation Fast	0	148 K	30 K	8.033 Mil
User Analysis	0	48 K	25 K	6.575 Mil

2019: Cores by Site

	min	max	avg	total▾
CERN-P1	0	104.1 K	38.6 K	10.2327 Mil
CERN-PROD	0	49.7 K	35.1 K	9.2899 Mil
LBNL_DSD_ITB	0	281.7 K	21.0 K	5.5731 Mil
ANLASC	0	173.2 K	19.0 K	5.0272 Mil
MWT2	0	22.2 K	17.9 K	4.7355 Mil
BNL-ATLAS	0	22.6 K	16.5 K	4.3722 Mil

2019: HS06-sec Power vs Pledges

	min	max	avg	total
UK-NorthGrid	0	184 K	126 K	33.342 Mil
FR-CCIN2P3	0	187 K	128 K	34.007 Mil
DE-KIT	0	246 K	130 K	34.401 Mil
DE-DESY-ATLAS-T2	0	414 K	138 K	36.599 Mil
US-MWT2	0	240 K	193 K	51.068 Mil
US-T1-BNL	0	278 K	202 K	53.547 Mil
CA-TRIUMF	0	272 K	216 K	57.325 Mil
CH-CERN	0	938 K	510 K	135.169 Mil

Other Dashboards

- Several lighter dashboards have been implemented for service monitoring

- Also for short-term site oriented job monitoring
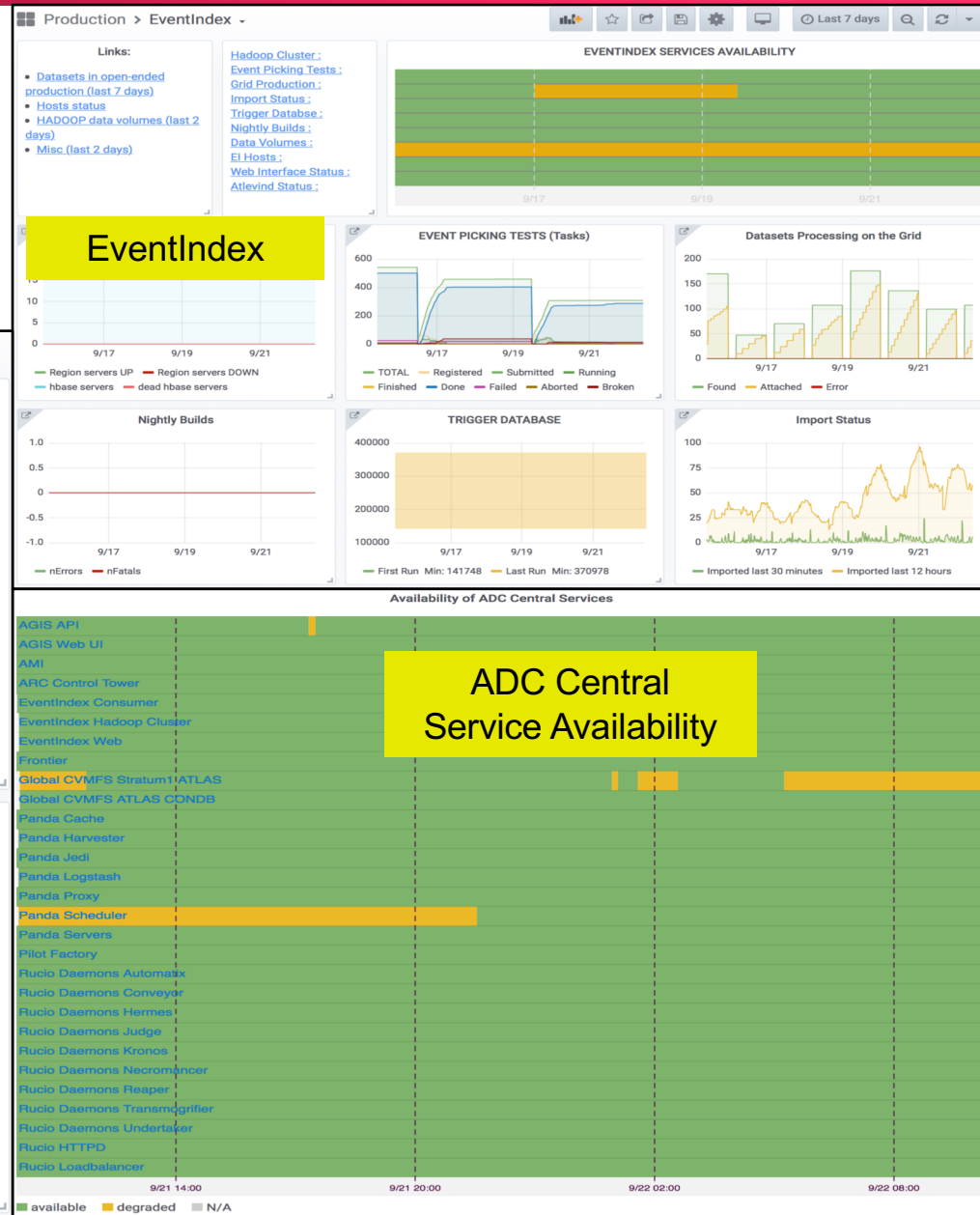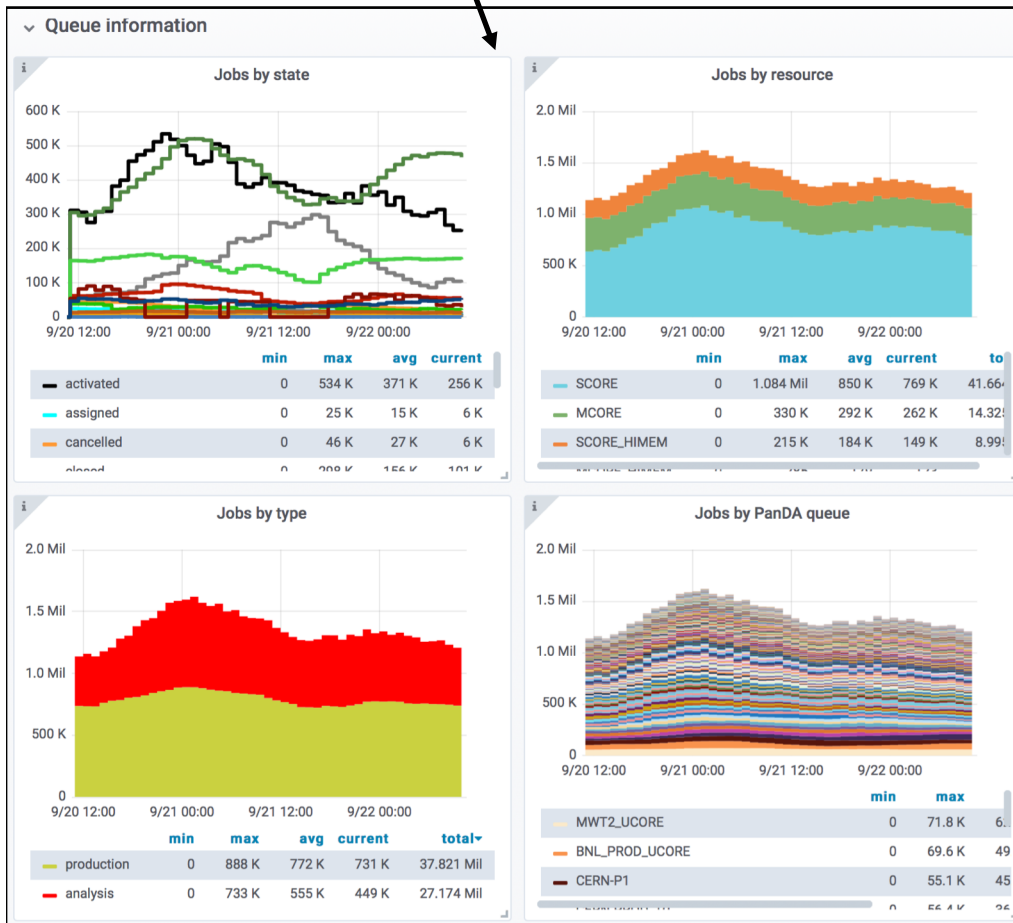
Live Page

- Derived collection of statistics and plots
  - Jobs, data storage, transfers, services
  - Inputs from different dashboards
  - Also directly from services and Elogs
- Contains links to information sources for deeper investigations
- Used by shifters and managers for daily monitoring
  - Refreshed hourly

Topics

- ADC dashboards in the MonIT infrastructure

- User level job/task monitoring with BigPandaMon

- Analytics cluster at UC and its applications

BigPandaMon

- A window into the PanDA system
- \> 100 different views
- From production dashboards to logs
- Covers scope in range $1...10^{11}$ events

Dario Barberis: ATLAS ADC Monitoring

Architecture & current usage
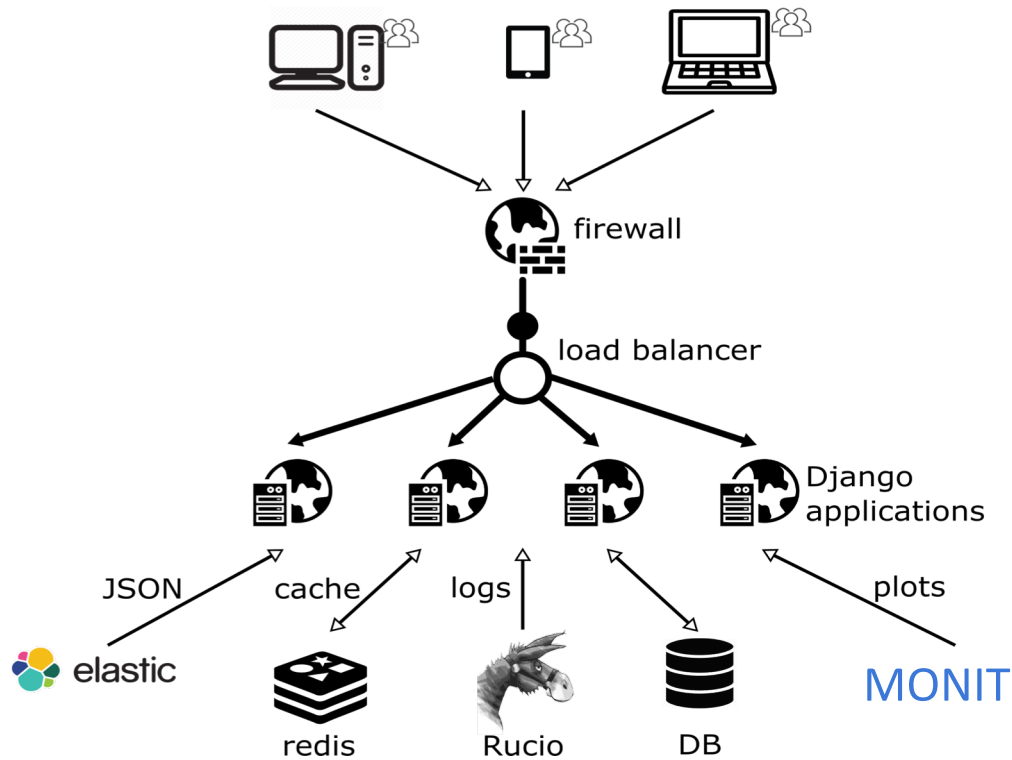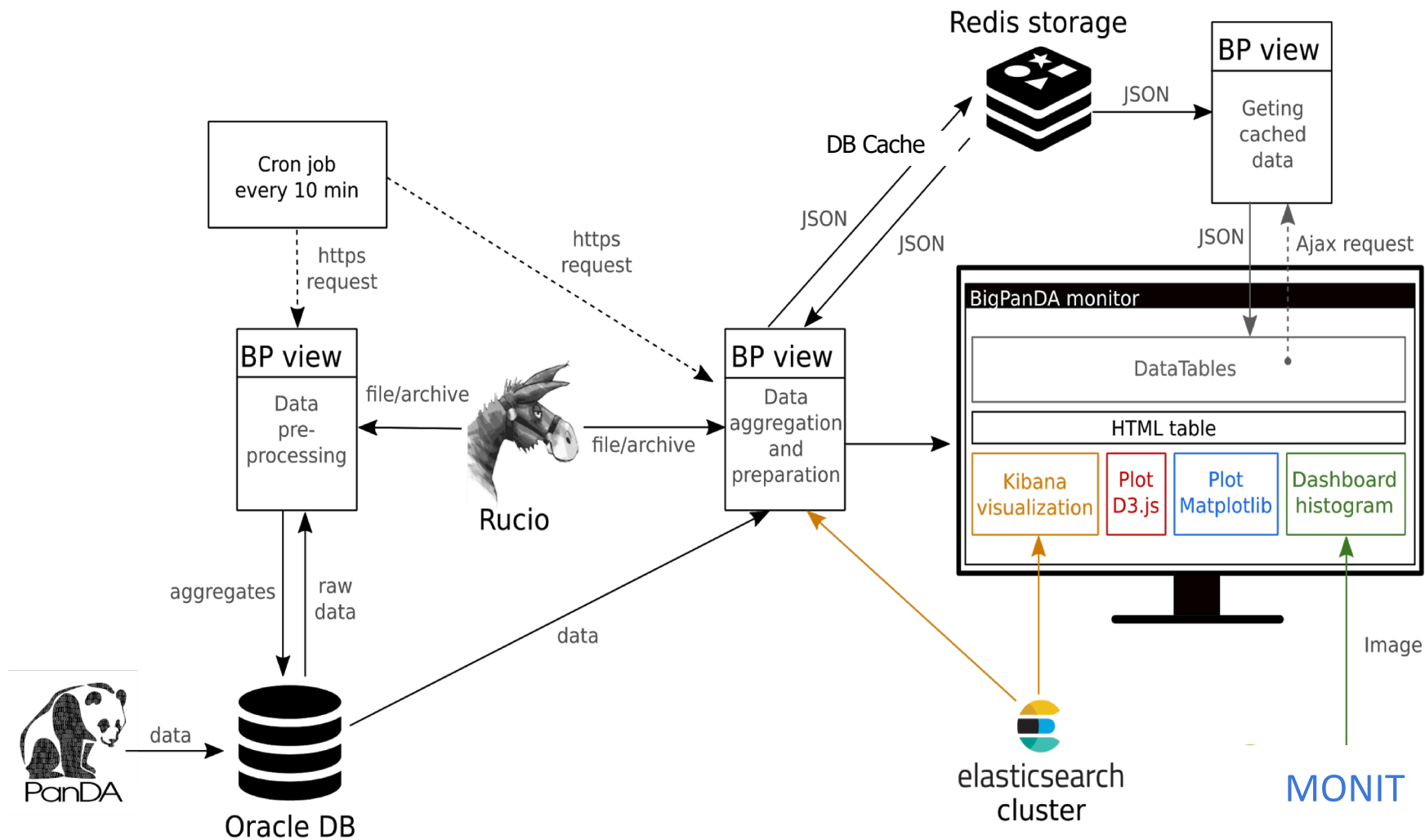
17000 json requests a day

- 6.5 (+3%) user queries a day
- From 1 to 626 pages a day per user
- 1110 monthly active users
- 342 daily users

Is a primary tool ATLAS wide for shifters, experts and ADC in general

Dario Barberis: ATLAS ADC Monitoring

BigPandaMon Data-flow diagram

User & task page

- User page: all tasks
- Task page:
  - All jobs
  - Exec CPU & memory plots
  - Task exec time
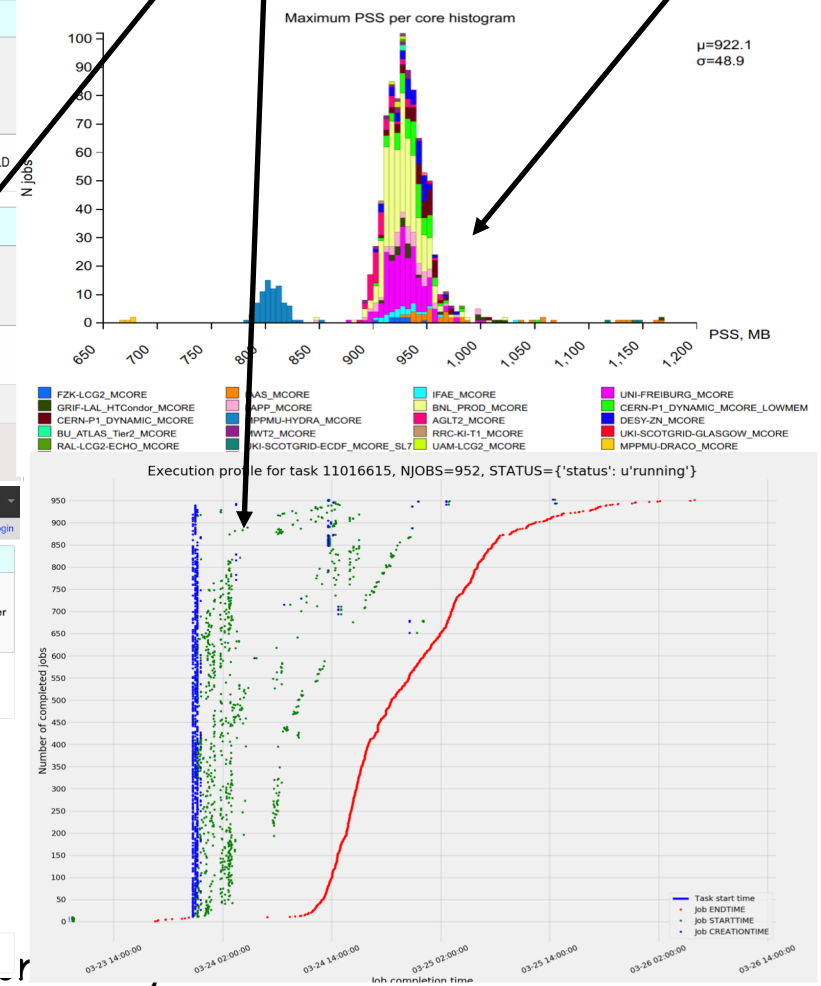
Jobs Page

specified parameters

total number of jobs matched to specified parameters

ATLAS PanDA   Dash ▾   Tasks ▾   Jobs ▾   Errors ▾   Users ▾   Sites ▾   Incidents ▾   Search   Admin   Prodsys ▾   Services ▾   VO ▾   Help ▾

PanDA jobs, last 12 hours.   **Params:**   **limit=10000**   Warning: limit 20000 per job table   Task has 0 jobs in table Jobsarchived, limit is set to 2000   **Total jobs found ≈ 1470000**   aipanda105 08:51:18, Reload Login

30013 jobs in this selection

Job **modification times** in this listing range from **2017-09-25 20:50:36** to **2017-09-26 08:50:36**.

Job **current priorities** in this listing range from **1000000** to **-1000000**. See priorityrange in the job attribute summary to see how priorities are distributed.

Job attribute summary	Sort by count, alpha	
atlasrelease (41)	Atlas-17.2.2 (7) Atlas-17.2.8 (428) Atlas-17.6.0 (10) Atlas-19.2.3 (176) Atlas-19.2.4 (142) Atlas-19.2.5 (5533) Atlas-2.6.4 (8) Atlas-20.1.5 (20) Atlas-20.1.8 (2) Atlas-20.20.10 (9) Atlas-20.20.6 (5) Atlas-20.20.7 (1) Atlas-20.20.8 (486) Atlas-20.7.5 (600) Atlas-20.7.6 (84) Atlas-20.7.7 (393) Atlas-20.7.8 (120) Atlas-21.0 (18) Atlas-21.0.13 (2) Atlas-21.0.14 (95) Atlas-21.0.15 (9358) Atlas-21.0.19 (708) ... more	
attemptnr (59)	1 (16519) 2 (10147) 3 (1207) 4 (348) 5 (267) 6 (83) 7 (34) 8 (46) 9 (21) 10 (29) 11 (6) 12 (22) 13 (76) 14 (90) 15 (5) 16 (2) 17 (25) 19 (1) 20 (1) 21 (2) 22 (25) 23 (2) 24 (1) 26 (2) 30 (1) ... more	
cloud (13)	CA (1204) CERN (1146) DE (1556) ES (360) FR (1406) IT (322) ND (663) NL (942) RU (467) TW (153) UK (2477) US (3002) WORLD (16291)	
computingsite (381)	AGLT2_LMEM (45) AGLT2_MCORE (56) AGLT2_SL6 (139) ALCF_Theta (826) ANALY_AGLT2_SL6 (134) ANALY_AGLT2_TEST_SL6-condor (1) ANALY_AGLT2_TIER3_TEST (5) ANALY_ARNES (18) ANALY_ARNES_DIRECT (78) ANALY_AUSTRALIA (15) ANALY_AUSTRALIA_TEST (2) ANALY_BEIJING (6) ANALY_BHAM_SL6 (5) ANALY_BNL_CLOUD (2) ANALY_BNL_LONG (207) ANALY_BNL_LONG_CONTR (16) ANALY_BNL_MCORE (3) ANALY_BNL_SHORT (180) ANALY_BNL_SHORT_CONTR (38) ANALY_BNL_Test_2_CE-1 (4)	

Overall error summary			
Category:code	**Attempt list**	**Nerrors**	**Sample error description**
ddm:200	jobs	3	Could not get GUID/LFN/MD5/FSIZE/SURL from pilot XML
exe:-1	jobs	67	LRMS error: (257) Job timeout

Job list   Only the most recent 100 jobs are shown. Remove the limit and sort by PandaID, time since last state change, ascending mod time, descending mod time, priority, attemptnr, ascending duration, descending duration

PanDA ID Attempt#	Owner Group	Request Task ID	Transformation	Status	Created	Time to start d:h:m:s	Duration d:h:m:s	Mod	Cloud Site	Priority	Job info
	sgonzale AP_SOFT	11341 11081388	POOLtoEI_tf.py	activated	2017-04-01 06:07:53	0:0:52:59		2017-04-01 07:00:43	WORLD SARA-MATRIX online no active blacklisting rules defined	880	
3315291794 Attempt 1											

Job name: data_evind.00311365.physics_Main.eventIndex.r9264_p3083_i14.3315291794 #1

Datasets: In: data16_13TeV:data16_13TeV.00311365.physics_Main.merge.AOD.r9264_p3083_tid11038617_00

     Out: data_evind.00311365.physics_Main.eventIndex.log.r9264_p3083_i14_tid11081388_00

Topics

- ADC dashboards in the MonIT infrastructure

- User level job/task monitoring with BigPandaMon

- **Analytics cluster at UC and its applications**

Analytics Cluster at UC

- It provides an interactive environment to:

  - Develop additional or alternative dashboards
  - Investigate correlations between several data sources

- Complementary to the monitoring and accounting infrastructure at CERN

Analytics Infrastructure

CERN

- DB - Oracle, Hadoop, Ingress, Elastic
- Processing - Sqoop, pig, Spark, SWAN, Kubernetes cluster

UC

- Elasticsearch
- K8s cluster, ML platform

Data sources:

- Oracle (panda, jedi, rucio) `CERN` `UC`
- MySQL (BOINC) `CERN` `UC`
- AMQ (FTS) `CERN` `UC`
- RMQ (PerfSonar) `UC`
- Pilot (benchmarks) `UC`
- User codes (xAOD usage) `CERN` `UC`
- Frontiers (full logs) `UC`
- HC jobs (cost matrix) `UC`
- SLATE `UC`
- XCache `UC`

2

Ilija Vukotic, S&C Week, CERN October 1st, 2019

Data collection infrastructure - CERN

- **All services dockerized**
- **All running in a single kubernetes cluster at CERN. Very stable.**
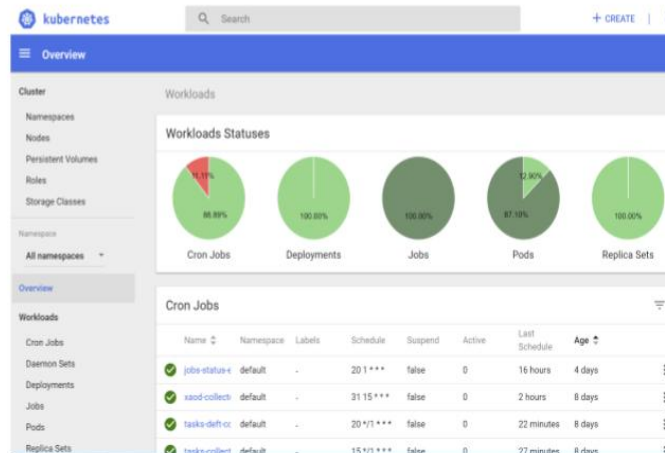
Ilija Vukotic, S&C Week, CERN October 1st, 2019

Data collection infrastructure - UChicago

- **All services dockerized. All running at UChicago kubernetes cluster.**
- **Running:**
  - **logstash collectors**
  - **rollup jobs (DDM)**
  - **Alarm & Alert service**
  - **REST interfaces (benchmarking, cost matrix, xcache backend)**

Ilija Vukotic, S&C Week, CERN October 1st, 2019

Analytics Cluster at UC

- It provides an interactive environment to:

  - Develop additional or alternative dashboards

  - Investigate correlations between several data sources

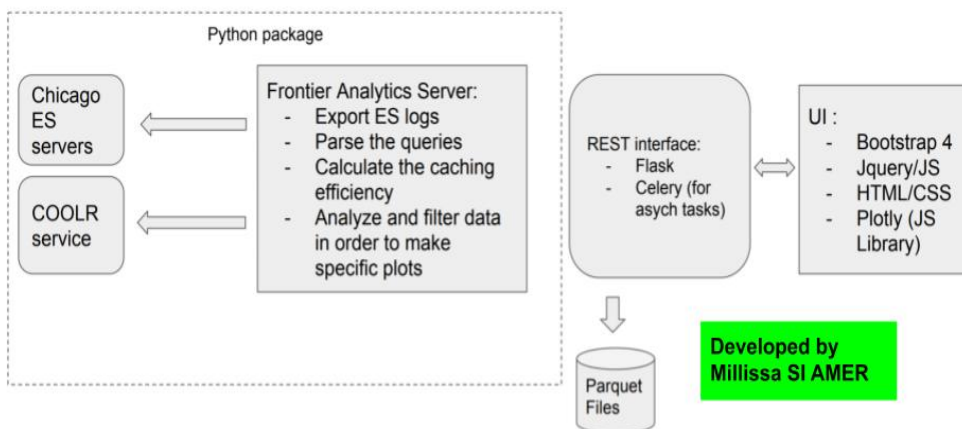- Complementary to the monitoring and accounting infrastructure at CERN

Analytics platforms

While all the data can be analysed using kibana, or JupyterLab notebooks, some projects would benefit from a custom analytics platforms:

- Extra processing and result caching
- Dedicated web site
- Custom searches
- Custom visualizations

We have a way to run these platforms in UC River2 K8s cluster. Currently we support two - Frontier and Perfsonar analytics platforms. Expect more to come.

Frontier analytics tools

Python package

Chicago ES servers

COOLR service

Frontier Analytics Server:
- Export ES logs
- Parse the queries
- Calculate the caching efficiency
- Analyze and filter data in order to make specific plots

REST interface:
- Flask
- Celery (for asych tasks)

UI :
- Bootstrap 4
- Jquery/JS
- HTML/CSS
- Plotly (JS Library)

Parquet Files

Developed by Millissa SI AMER

Ilija Vukotic, S&C Week, CERN October 1st, 2019

Perfsonar analytics platform

- Made by Евгений Третьяков Сергеевич (MEPhi)
- Django + ES + JavaScript libraries: three.js, d3-force-3d, jquery
- Future (short term)
  - Full production deployment
  - Pagination
  - Adaptive interface Full HD+ (Now only FullHD)
  - Advanced search
- Future (long term)
  - Integration of derived data
  - Other Perfsonar indices

Ilija Vukotic, S&C Week, CERN October 1st, 2019
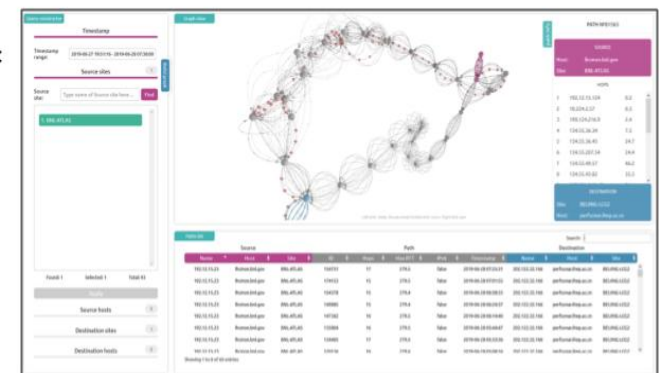
Outlook

- ATLAS Distributed Computing has a coherent set of monitoring and accounting dashboards and interactive tools

- Technologies evolve all the time – we follow them
    - Trying to use Open Source solutions as much as possible
    - Even if at times some home-made parts are inevitable
        - See the low number of display options in Grafana

- The future is in more interactive environments providing the possibility to correlate information from many different sources
    - This is real BigData in action
        - Not just tens of billions of statistically equivalent event records!

- Any way we are quite well set for the start of LHC Run 3

Thank you!

Thank you for listening

Thanks to all authors of this work