Data Analytics and CERN IT Hadoop Service

Openlab workshop on machine learning and data analytics
CERN, April 2017
Luca Canali, IT-DB

Data Analytics at Scale – The Challenge

- When you cannot fit your workload in a desktop
  - Data analysis and ML algorithms over large data sets
  - Deploy on distributed systems
- Need specialized components for:
  - Data ingestion tools and file systems
  - Cluster storage and processing engines
  - ML tools that work at scale

Engineering Efforts to Enable Effective ML

- From “Hidden Technical Debt in Machine Learning Systems”, D. Sculley at al. (Google), paper at NIPS 2015

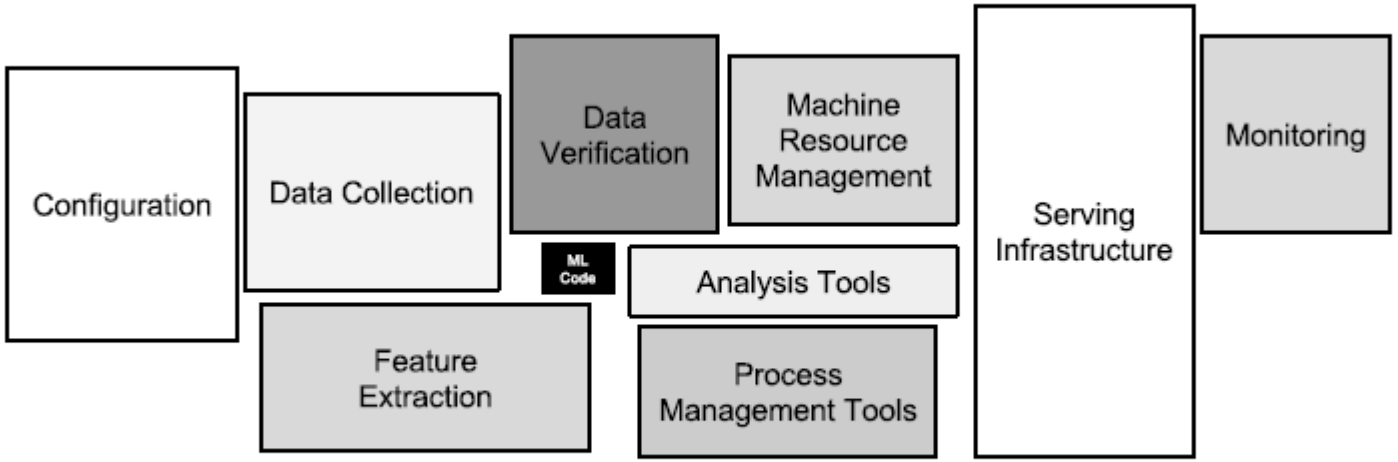
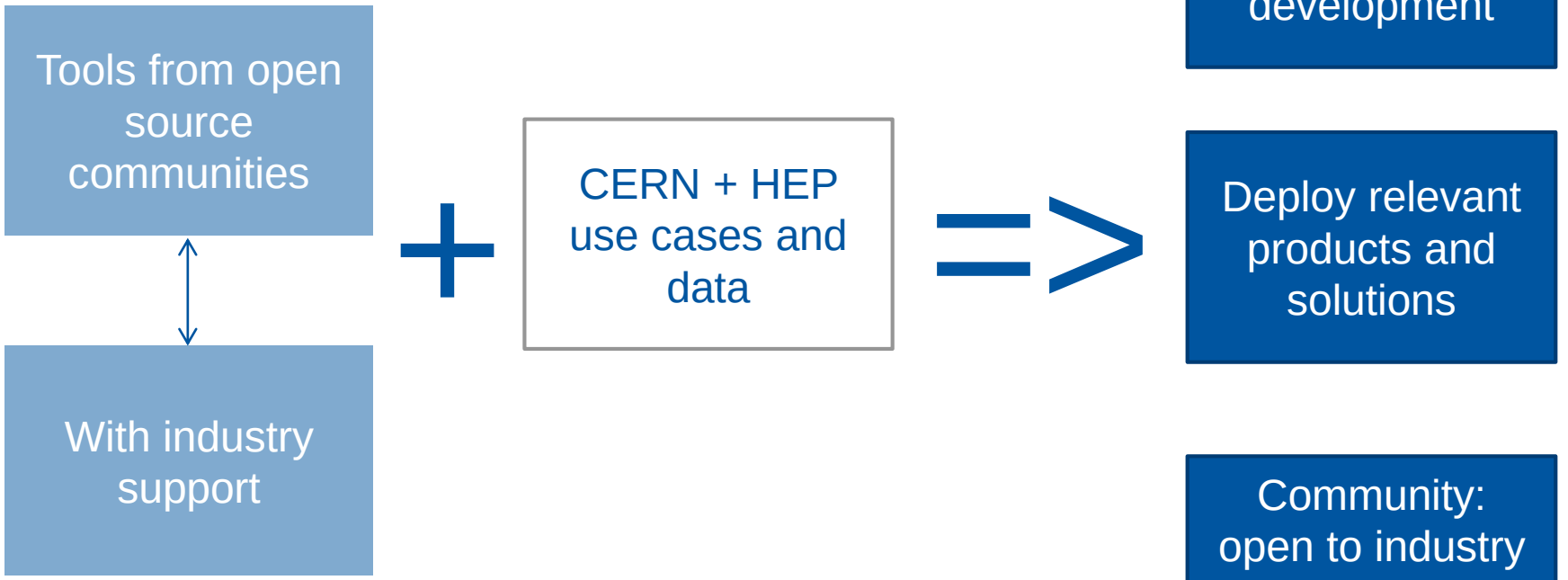


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

The High Level Picture

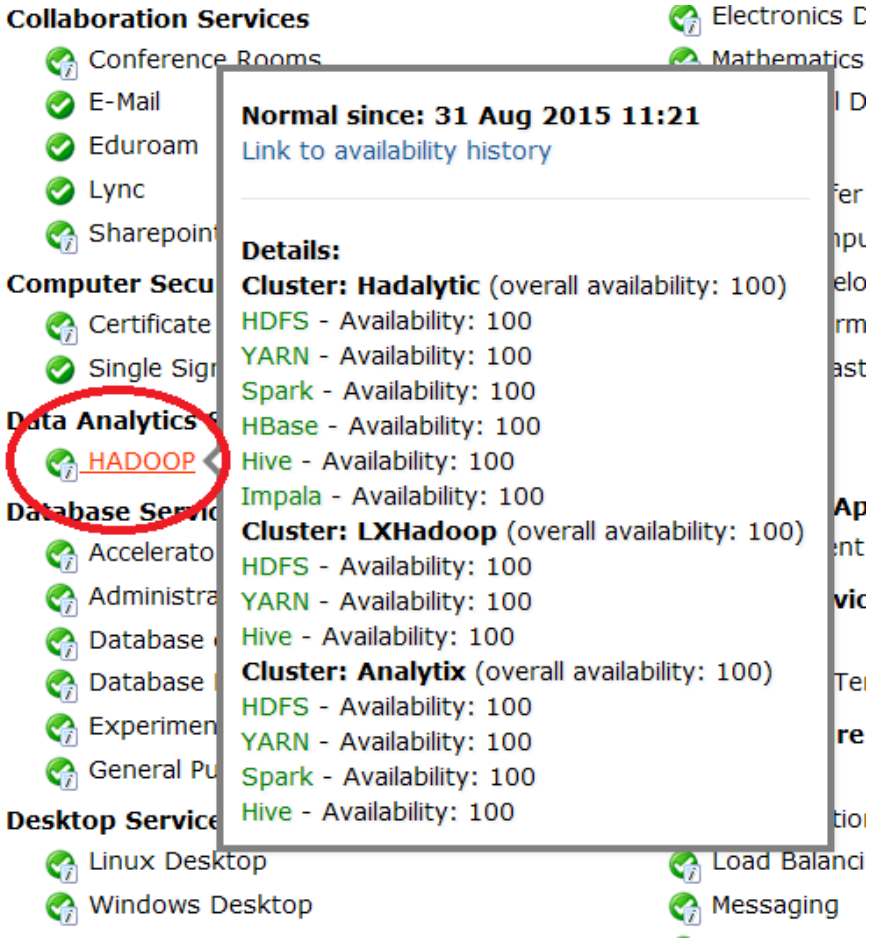


Managed Services for Data Engineering

- Platform
  - Capacity planning and configuration
  - Define, configure and support components
- Running central services
  - Build a team with domain expertise
  - Share experience
  - Economy of scale

Hadoop Service at CERN IT

- Setup and run the infrastructure
- Provide consultancy
- Support user community

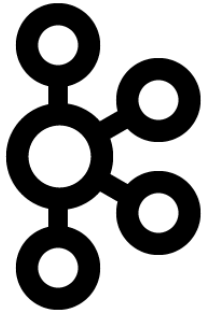


Hadoop clusters at CERN IT

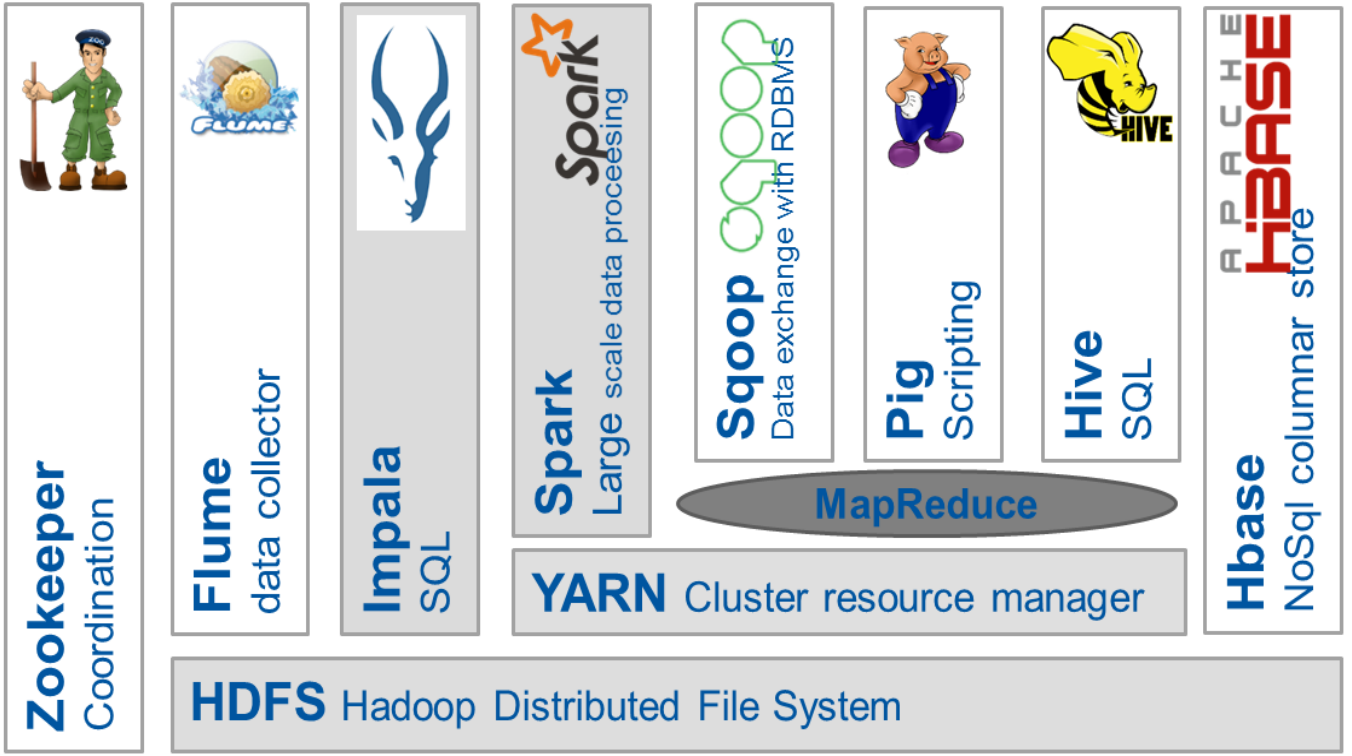
- 3 production clusters (+ 1 for QA)
- A new system for BE NXCALs (accelerator logging) platform
  - Later in 2017

Cluster Name	Configuration	Primary Usage
lxhadoop	22 nodes (cores – 560,Mem – 880GB,Storage – 1.30 PB)	Experiment activities
analytix	56 nodes (cores – 780,Mem – 1.31TB,Storage – 2.22 PB)	General Purpose
hadalytic	14 nodes (cores – 224,Mem – 768GB,Storage – 2.15 PB)	SQL-oriented engines and datawarehouse workloads

Overview of Available Components (Apr 2017)

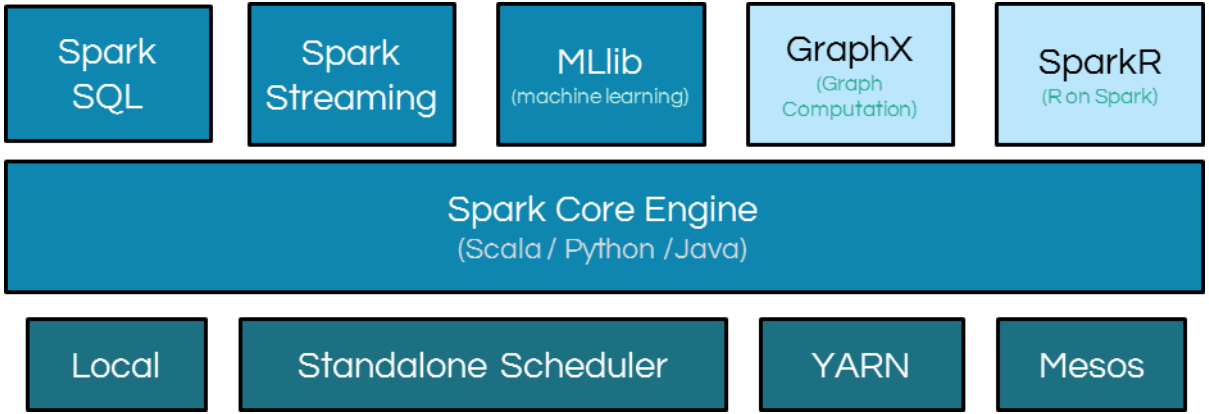


Kafka:
streaming
and ingestion



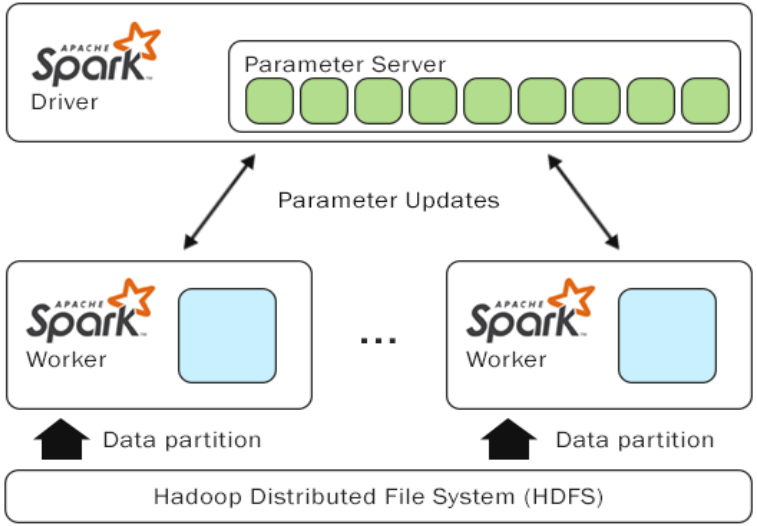
Apache Spark

- Powerful engine, in particular for data science and streaming
  - Aims to be a “unified engine for big data processing”
- At the center of many “Big Data” solutions



Machine Learning with Spark

- Spark has tools for machine learning at scale
  - Spark library MLlib
- Distributed deep learning
  - Working on use cases with CMS and ATLAS
  - We have developed an integration of Keras with Spark
- Possible tests and future investigations:
  - Frameworks and tools for distributed deep learning with Spark available on open source:
    - TensorFrame, BigDL, TensorFlowonSpark, DL4j, ..
  - Also of interest HW solutions: for example FPGAs, GPUs etc



https://github.com/cerndb/dist-keras

Main developer: Joeri Hermans (IT-DB)

Jupyter Notebooks

- Jupyter notebooks for data analysis
  - System developed at CERN (EP-SFT) based on CERN IT cloud
  - SWAN: Service for Web-based Analysis
  - ROOT and other libraries available
- Integration with Hadoop and Spark service
  - Distributed processing for ROOT analysis
  - Access to EOS and HDFS storage
  - Opens the possibility to do physics analysis on Spark using Jupyter notebooks as interface

Highlight of Use Cases

- Accelerator logging
- Industrial controls
- Analytics on monitoring data
- Physics analysis
  - Development of Big Data solutions for physics

Performance and Testing at Scale

- Challenges with scaling up the workloads
  - Example from the CMS data reduction challenge: 1 PB and 1000 cores
    - Production for this use case is expected 10x of that.
    - New territory to explore
- HW for tests
  - CERN clusters + external resources, example: testing on Intel Lab equipment (16 nodes) in February 2017

Architecture and Components Evolution

- Challenge: Mixing write/update workloads with analytics

- Testing Apache Kudu
  - Currently good candidate as infrastructure component for the next generation controls system
    - We are working on this with BE-ICS and openlab with Intel
  - Kudu is storage engine
    - Alternative to HDFS, however works with many Hadoop components, so evolution not revolution
    - Kudu provides features that currently one has to implement using a combination of HBase and Parquet for example

Architecture and Components Evolution

- Architecture decisions on data locality
- Currently we deploy Spark on Yarn and HDFS

- Investigating: Spark clusters without directly attached storage?
  - Using EOS and/or HDFS accessed remotely?
  - EOS integration currently being developed for Spark
  - Spark clusters “on demand” rather than Yarn clusters?
  - Possibly on containers

Hadoop and Spark on OpenStack

- Tests of deploying Hadoop/Spark on OpenStack are promising
- Appears a good solution to deploy clusters where local storage locality is not needed
  - Example: possible candidates for Spark clusters for physics data processing reading from EOS (or from remote HDFS)
- Also run tests of Hadoop clusters with local storage
  - Using ad-hoc and “experimental configuration” in particular for the storage mapping, thanks to the collaboration with OpenStack team at CERN
  - Promising results, we plan to further explore

Training and learning efforts

- Upcoming classroom course
  - "Apache Spark for Developers“, June 12-15, in the training catalogue https://cern.ch/course/?173SPK02
- Intro material, delivered by IT-DB
  - "Introduction and overview to Hadoop ecosystem and Spark“, April 2017. Slides and recordings at: https://indico.cern.ch/event/590439/
  - 2016 tutorials: https://indico.cern.ch/event/546000/
- We plan more training sessions: tutorials and possibly hands-on
  - For announcements, subscribe to the egroup it-analytics-wg
  - See also presentations at the Hadoop Users Forum: https://indico.cern.ch/category/5894/

Conclusions

- CERN Hadoop and Spark service
  - We deploy and support open-source and industry “Big Data” solutions for CERN user cases
    - Data engineering
    - Solutions for data analytics and machine learning
- Ongoing openlab projects and collaborations
  - with CMS Big Data, CERN BE-ICS and Intel
- Contact us:
  - http://information-technology.web.cern.ch/services/Hadoop-Service

Acknowledgements

The following have contributed to the work reported in this presentation

- Members of IT-DB and IT-ST
  - Supporting Hadoop core and components
- CMS Big Data
- Thanks to Intel for their collaboration in openlab and tests on their lab
- Thanks to IT-CM for the tests on OpenStack

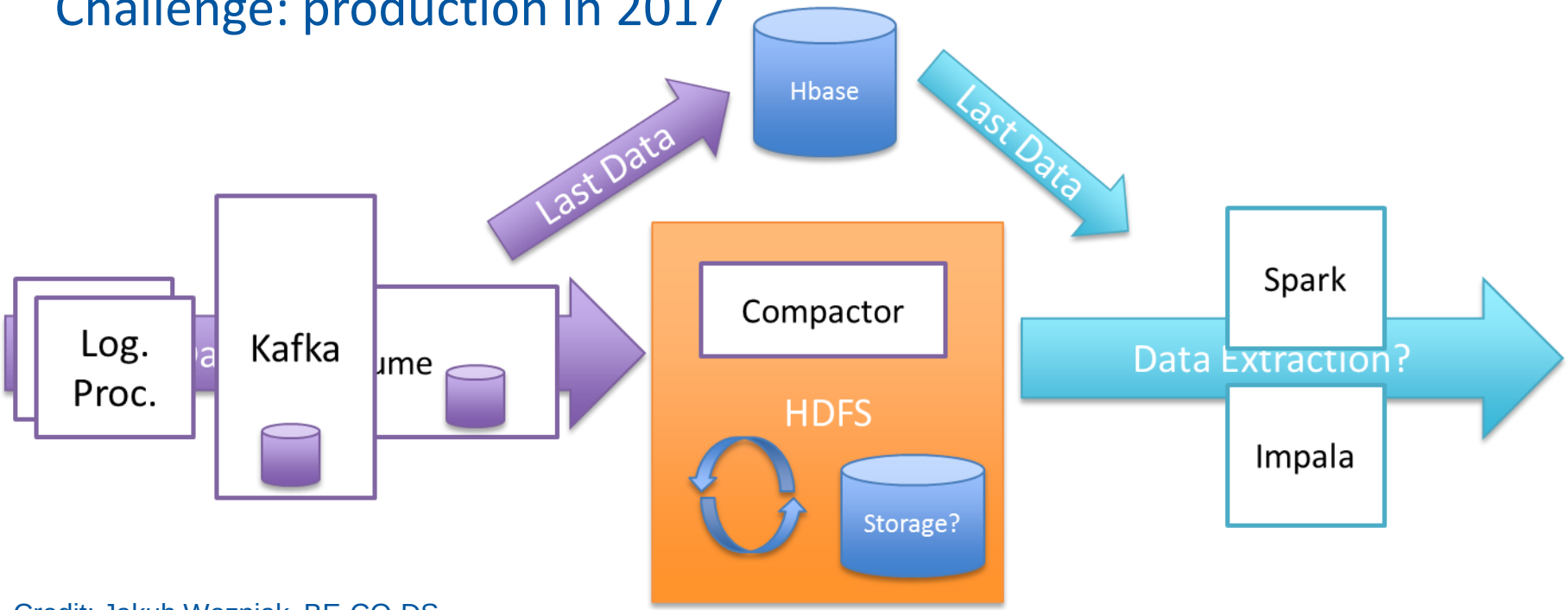
Backup Slides

Next Gen. Archiver for Accelerators Logs

Pilot architecture tested by CERN Accelerator Logging Services

Critical system for running LHC - 500 TB today, growing 200 TB/year

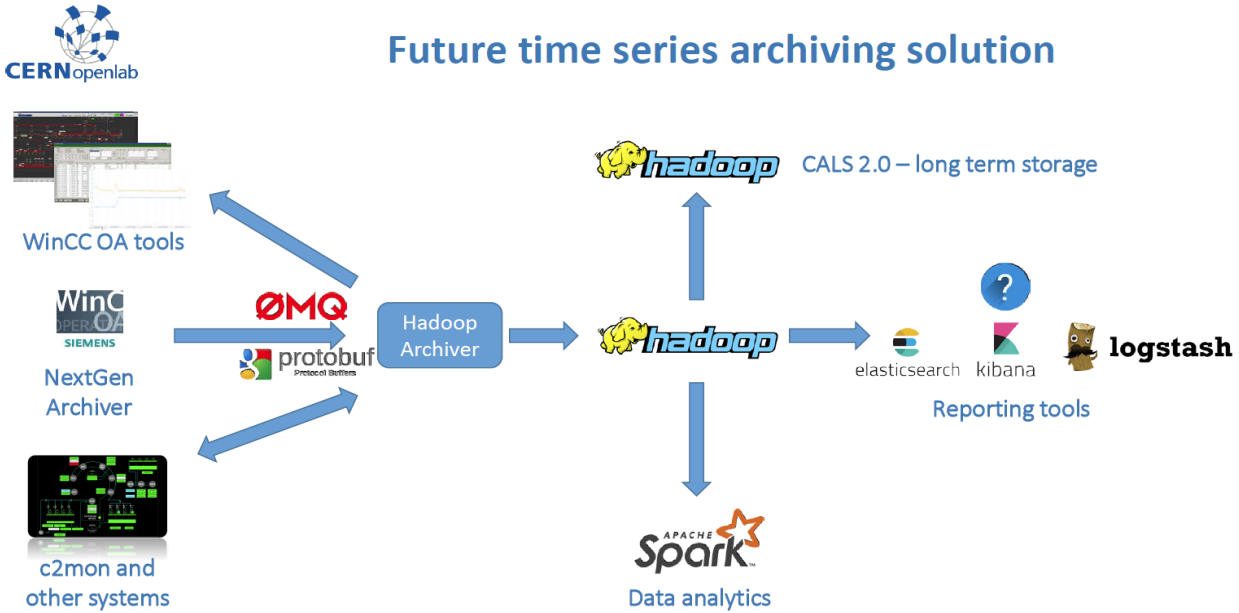
Challenge: production in 2017



Credit: Jakub Wozniak, BE-CO-DS

Industrial Controls Systems

- Development of next generation archiver
- Currently investigating possible architectures (openlab project)
  - Including potential use of Apache Kudu

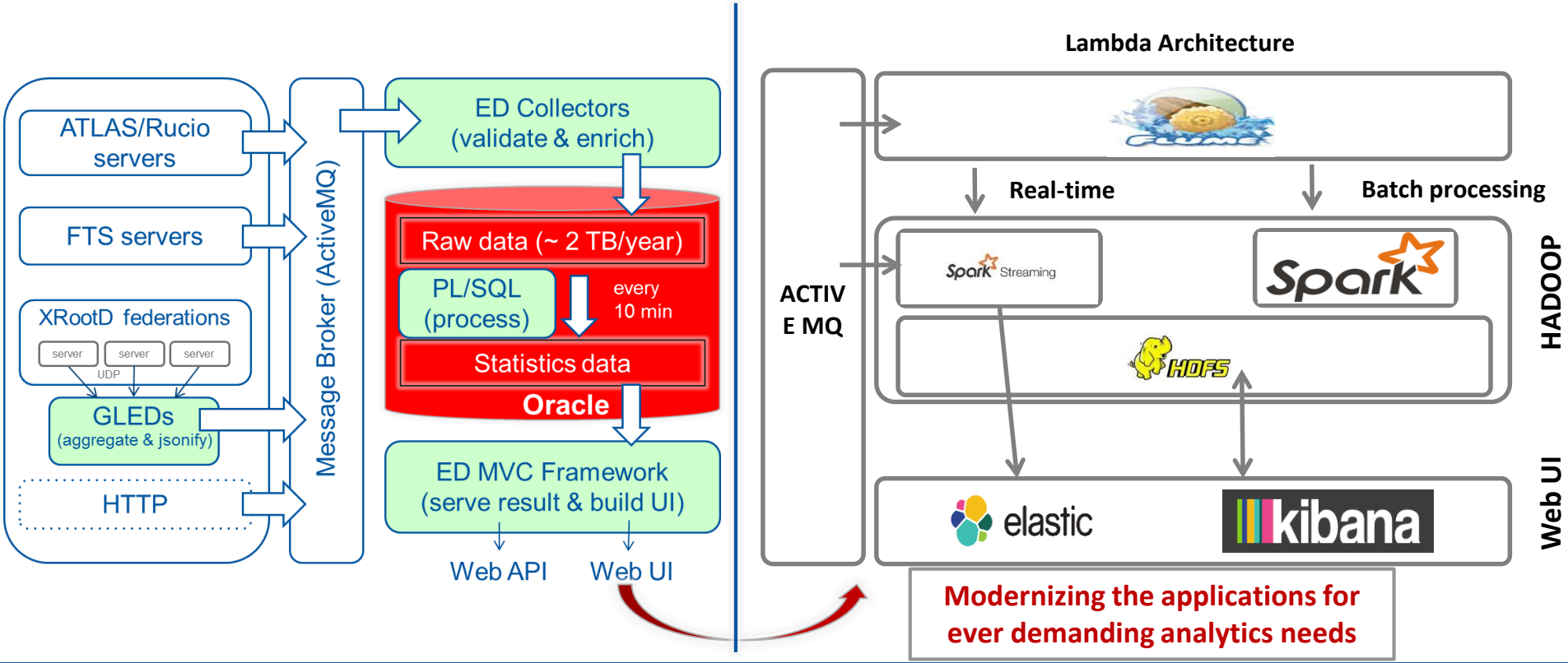


Credits: CERN BE Controls team

Analytics platform for controls and logging

- Use distributed computing platforms for storing analyzing controls and logging data
  - Scale of the problem 100s of TBs
- Build an analytics platform
  - Technology: focus on Apache Spark
  - Empower users to analyze data beyond what is possible today
  - Opens use cases for ML on controls data

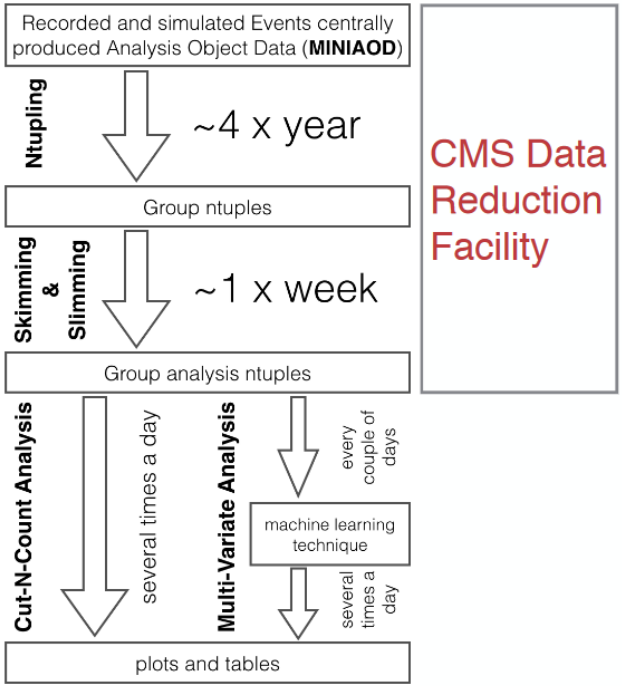
Production Implementation – WLCG Monitoring



Credit: Luca Magnoni, IT-CM

CMS Big Data Project and Openlab

Proposal: CMS Data Reduction Facility



- Demonstration facility optimized to read through petabyte sized storage volumes
  - Produce sample of reduced data based on potentially complicated user queries
  - Time scale of hours and not weeks
- If successful, this type of facility could be a big shift in how effort and time is used in physics analysis
  - Same infrastructure and techniques should be applicable to many sciences

Physics Analysis and “Big Data” ecosystem

- Challenges and goals:
  - Use tools from industry and open source
    - Current status: Physics uses HEP-specific tools
    - Scale of the problem 100s of PB – towards hexascale
  - Develop interfaces and tools
    - Already developed first prototype to read ROOT files into Apache Spark
  - Challenge: testing at scale