Data Analytics and CERN IT Hadoop Service

CERN openlab Technical Workshop
CERN, December 2016
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
- Complexity quickly goes up
  - Data ingestion tools and file systems
  - Storage and processing engines
  - ML tools that work at scale

Engineering Effort for Effective ML

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

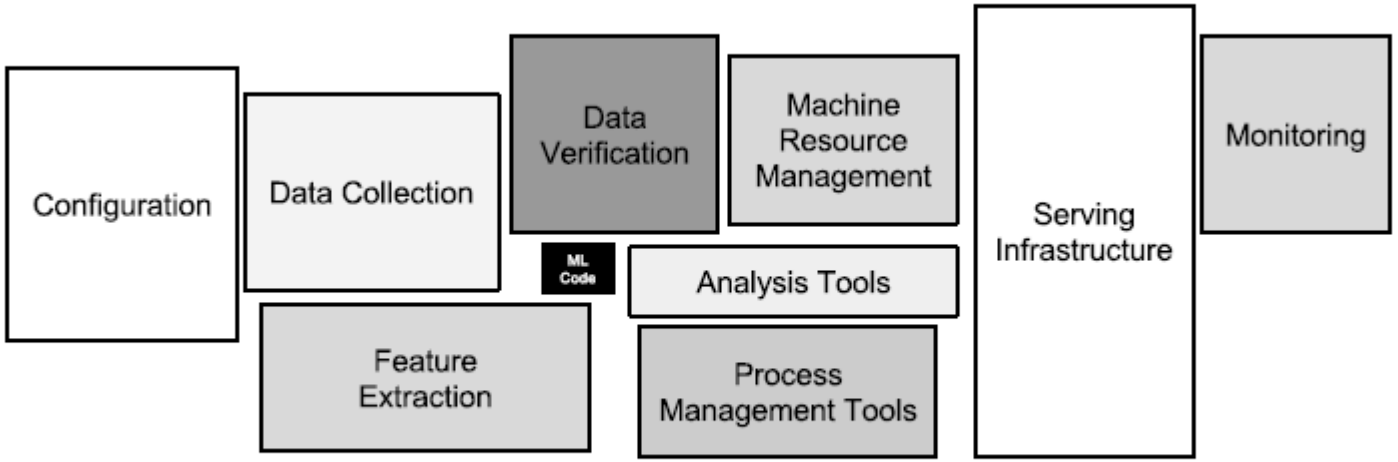


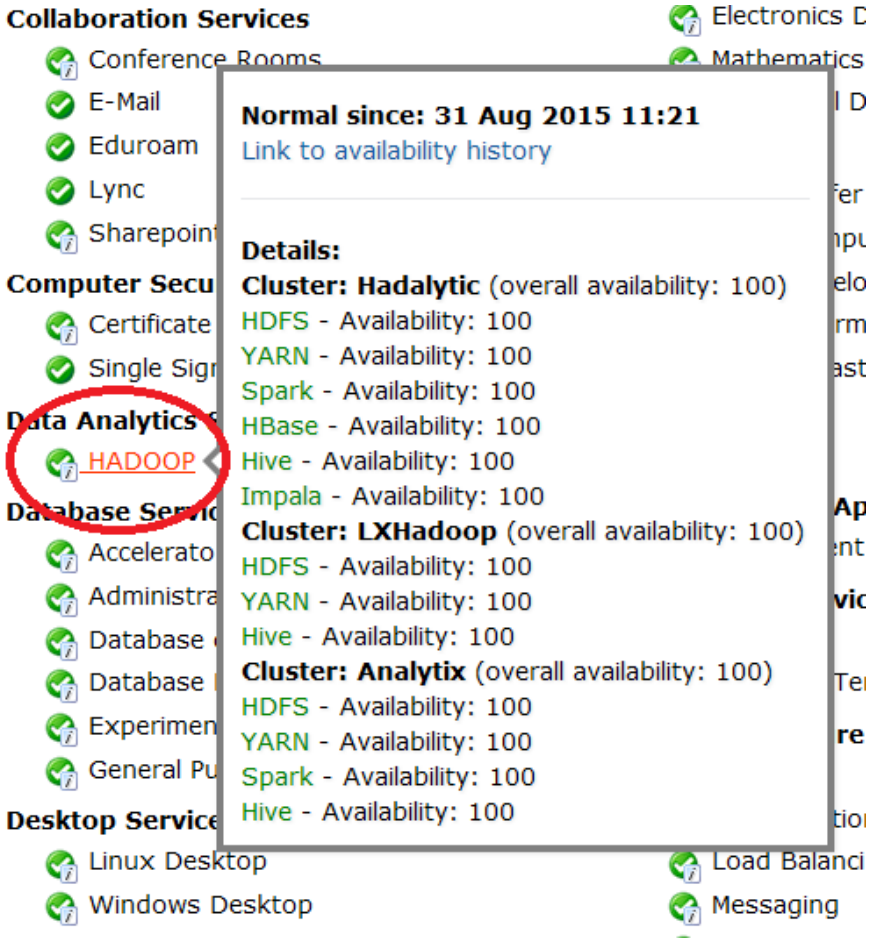
Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

Managed Services for Data Engineering

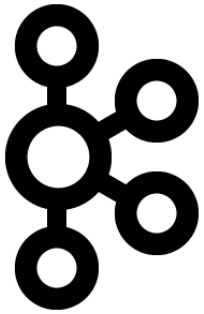
- Platform
  - Capacity planning and configuration
  - Define, configure and support components
- Running central services
  - Build a team with domain expertise
  - Share experience
  - Economy of scale

Hadoop Service at CERN IT

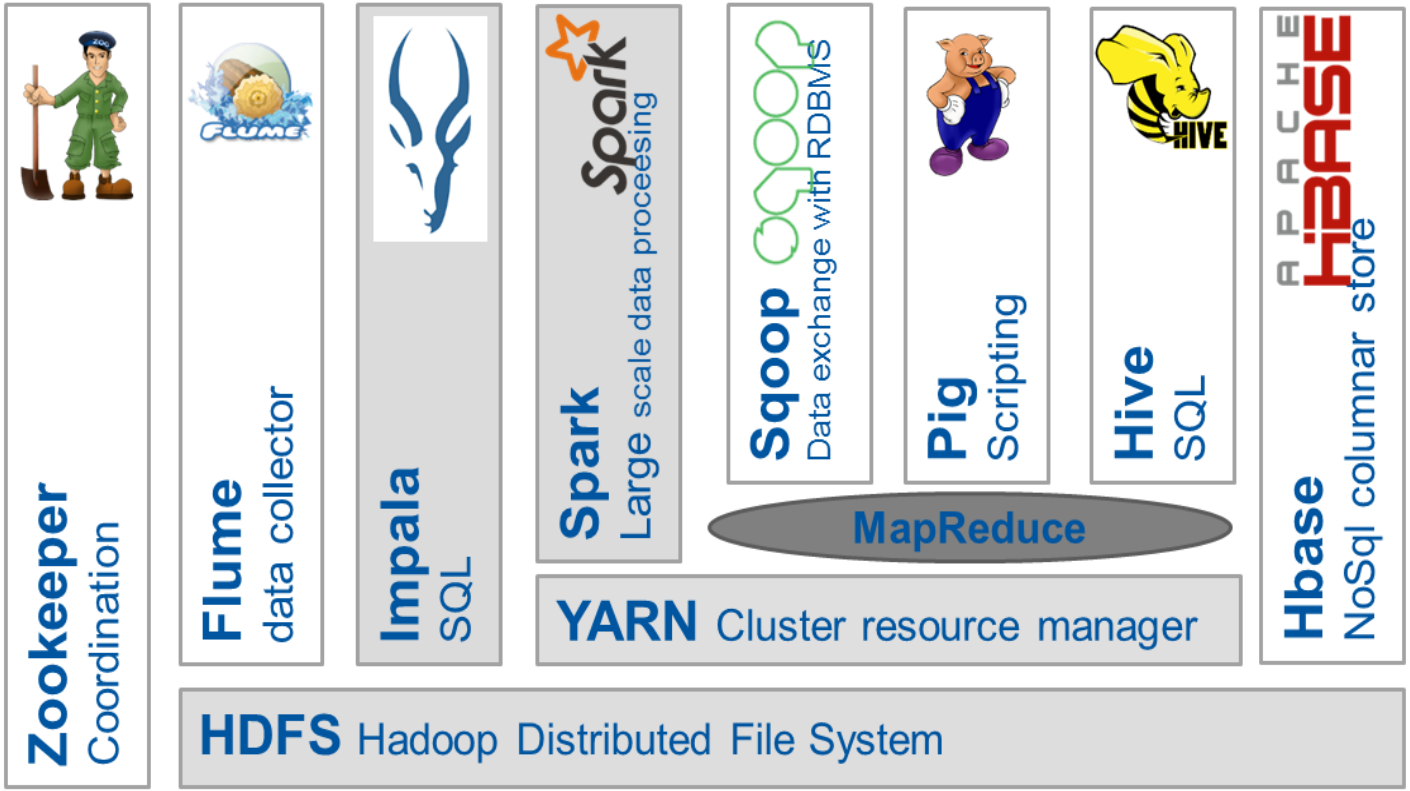
- Setup and run the infrastructure
- Provide consultancy
- Build user community
- Joint work
  - IT-DB and IT-ST



Overview of Available Components (Dec 2016)



Kafka
Streaming/In
gestion



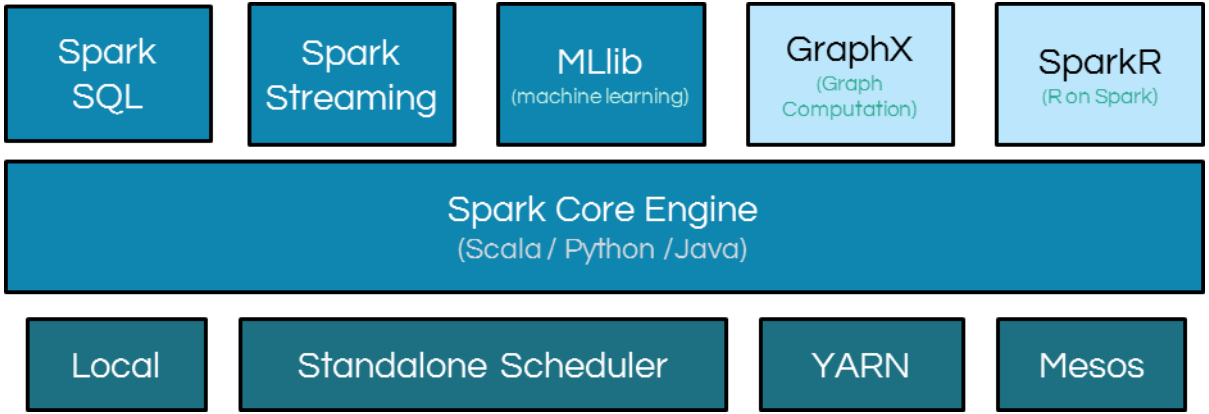
Hadoop clusters at CERN IT

- 3 production clusters (+ 1 for QA) as of December 2016

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

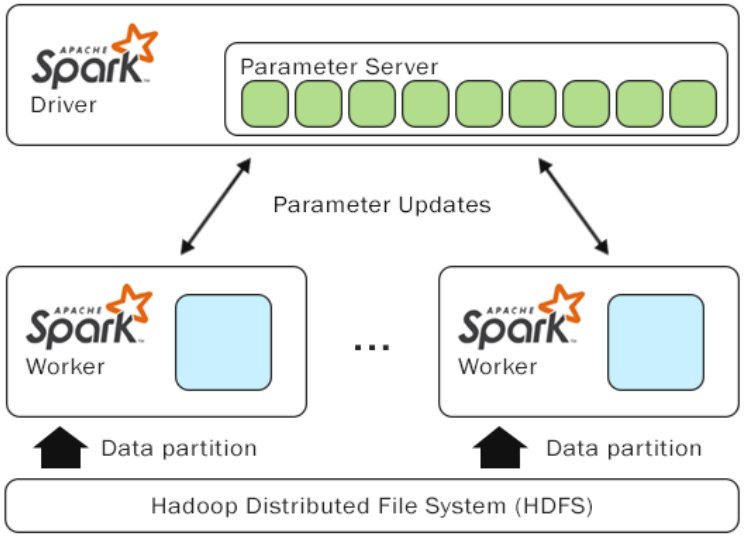
Apache Spark

- Spark evolution from map reduce ideas
- Powerful engine, in particular for data science and streaming
  - Aims to be a “unified engine for big data processing”



Machine Learning and Spark

- Spark addresses use cases for machine learning at scale
- Distributed deep learning
  - Working on use cases with CMS and ATLAS
  - Custom development: library to integrate Keras + Spark
  - Testing also other solutions



Some Important Challenges

- Infrastructure
- Components
- Evolution
- Running Services
- Knowledge sharing and technology adoption

Infrastructure and configuration

- What HW to use, how to configure it?
  - We are starting small (~ 1000 cores, 3TB RAM, 6 PB disk)
  - Commodity HW (standard components at CERN datacenter)
- Cloud
  - Pilot in the pipeline using private cloud at CERN
    - collaboration with OpenStack team
  - Future: test public cloud?

Components, Engines and File Formats

- State of the art evolves quickly
  - Challenge: reviewing application and architecture choices on short cycles (~2 years?)
  - Need to minimize technical debt
- Examples:
  - Yesterday: it was all about map reduce
  - Today: Spark, Impala
- Service evolution and testing in progress
  - Example: Kudu promising to add fast insert and updates and key-based search

Deep Learning

- Quickly evolving field
  - New tools and platforms
  - Role of HW also very important (GPUs, FPGAs, etc)
  - Many challenges to understand with distributed learning
- Hadoop service at CERN working on
  - Integration with Spark
- Comment: an area to further explore

Running the service

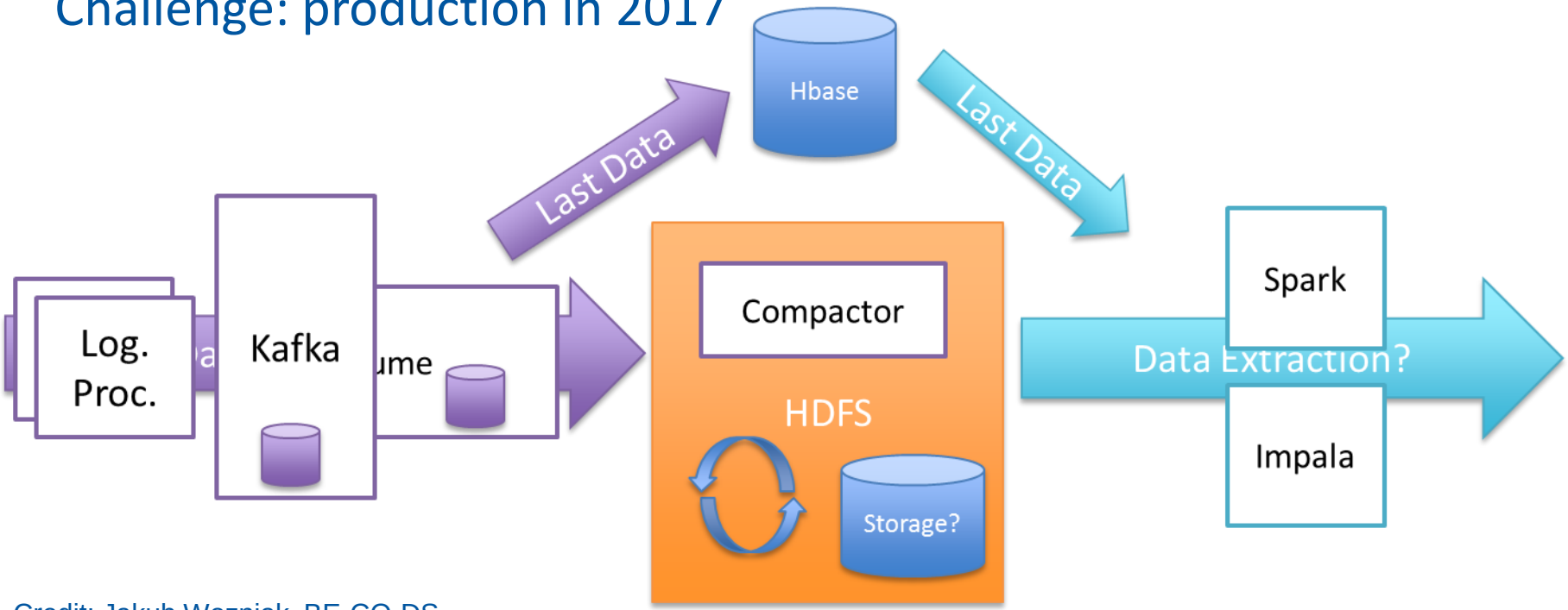
- Configuration:
  - Currently using CDH (Cloudera) distribution + puppet
  - Plans to test use of Cloudera Manager (at least for monitoring, possibly for installation)
  - Backup and recovery: work in progress
  - More work needed on monitoring and security, building from current experience
  - Further work and understanding on workload management and performance

Pilot Implementation – NXCALS

Pilot architecture tested by CERN Accelerator Logging Services

Critical system for running LHC

Challenge: production in 2017



Credit: Jakub Wozniak, BE-CO-DS

Performance and Testing at Scale

- Challenges with ramping up the scale
  - 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
- Action: access to HW for tests
  - CERN clusters + external resources from partners?

Technology Adoption

- Build community at CERN
  - Hadoop users forum, analytics working group, openlab workshops, seminars, …
  - Build a HEP community around ML
    - How to position ML with “current tools for HEP” analysis?
- Offer and provide training
  - Examples: tutorials by IT-DB, also discussing with tech training at CERN on courses for the catalog,
- Link with other scientific communities and industry
  - Knowledge sharing

Conclusions

- CERN Hadoop and Spark service
  - Established and evolving
  - Bring “big data” solutions from open source into CERN use cases
  - Several production implementations more in pipeline
- Brings value for analytics and large datasets
  - Machine learning at scale
- IT Hadoop service provides consultancy, platforms and tools

Acknowledgements

The following have contributed to the work reported in this presentation

- Members of IT-DB-SAS section
  - Supporting Hadoop components FE
- Rainer Toebbicke, Dirk Duellmann, Luca Menichetti from IT-ST
  - Supporting Hadoop Core FE

Discussion / Feedback

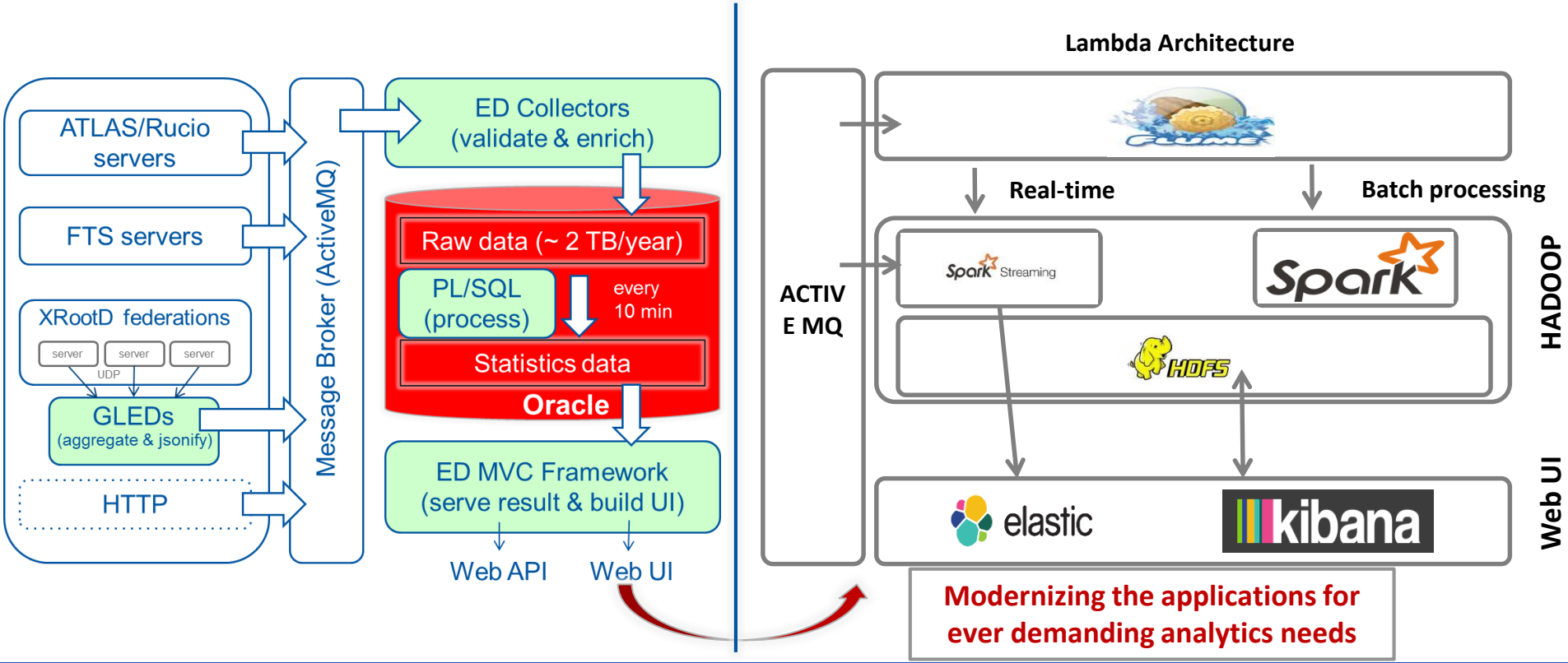
Backup Slides

- Backup Slides with additional details on ongoing projects

Impala - SQL on Hadoop

- Distributed SQL query engine for data stored in Hadoop
- Based on MPP paradigm (no MapReduce, Spark)
- Designed for high performance
  - Written in C++
  - Runtime code generation using LLVM
  - Direct data access

Production Implementation – WLCG Monitoring



Credit: Luca Magnoni, IT-CM

Projects

- Atlas Event Index (Production Service)
  - HBASE for fast lookup events; 40 TB/year
- LHC Postmortem Analysis
  - Real-time Postmortem Analytics of LHC monitoring data – Kafka + Spark
- Analysis of industrial controls data
- Future Circular Collider: Reliability and Availability analysis
  - Integrating heterogeneous data sources
  - Correlation between different domains

Connecting Hadoop and Oracle

Advantages

- ✓ No changes to the application
- ✓ Data sources are transparent to the users
- ✓ Opens up the possibility for new analytical queries

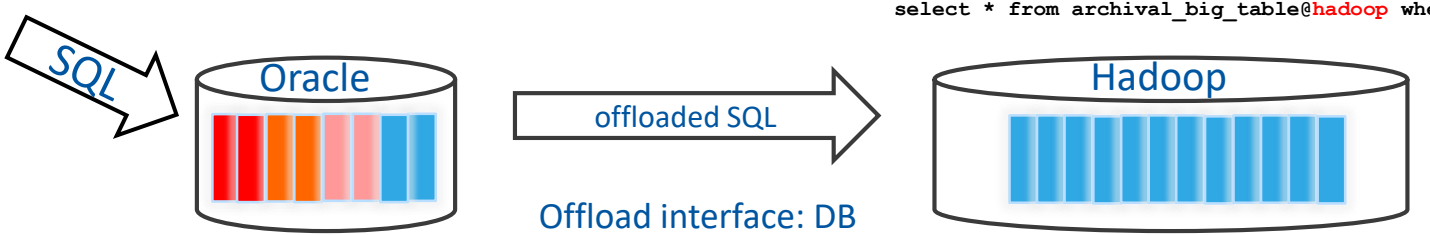
- Offload data from Oracle to Hadoop
  - recent data in Oracle; archive data in Hadoop

limited storage
throughput

scalable storage

- Offload queries to Hadoop

```
create view big_table as
    select * from online_big_table where date > '2016-05-05'
    union all
    select * from archival_big_table@hadoop where date <= '2016-05-05'
```



Offload interface: DB LINK, External table

SQL engines: Impala, Hive

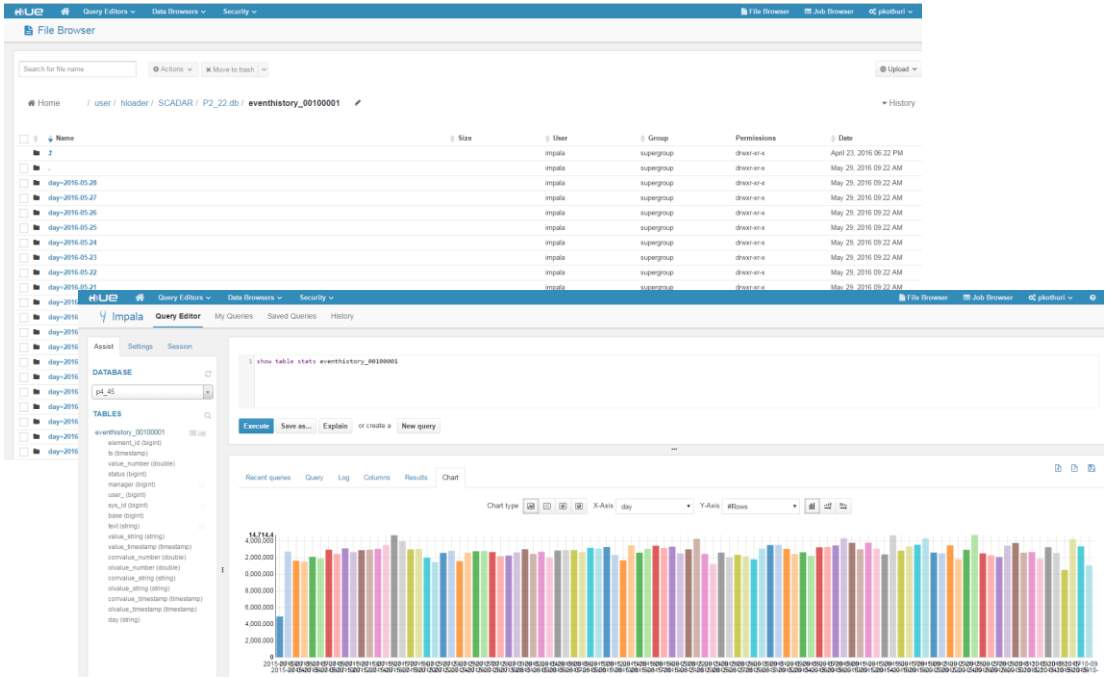
Jupyter Notebooks

- 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

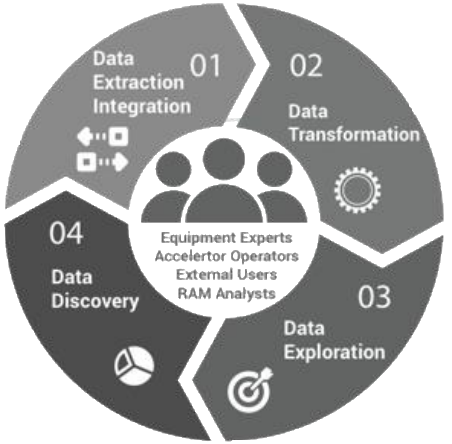
Hadoop User Experience - HUE

- Hue is a web interface for analyzing data with Apache Hadoop
  - View your data using HDFS filebrowser
  - Enhance and Analyze using Query editors for Impala, HIVE
  - Analyze & visualize using Spark notebooks (beta)
  - Requested by the user community

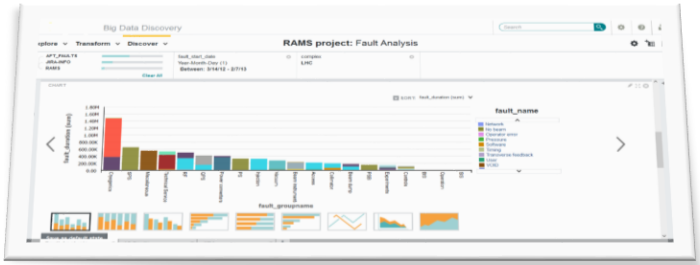


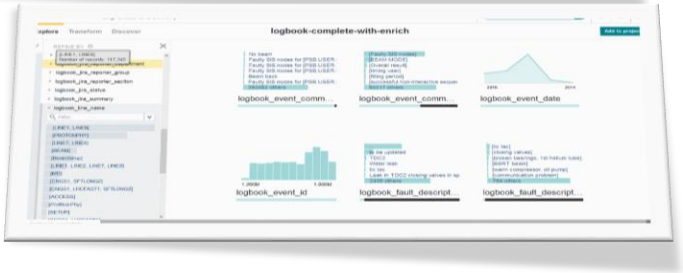
- Available on Hadoop clusters
  - https://hue-hadalytic.cern.ch
  - https://hue-analytix.cern.ch (soon)

Oracle Big Data Discovery



- Features
  - Data Exploration & Discovery
  - Data Transformation with Spark in Hadoop
    - Apply built-in transformations or write your own scripts
    - Data Enrichment: Text analytics, geolocation, etc.
  - Collaborative environment
- CERN SSO integrated
- Already available for Hadoop test cluster
- Some demos
  - https://www.youtube.com/watch?v=Jyw9NtUZ_ks



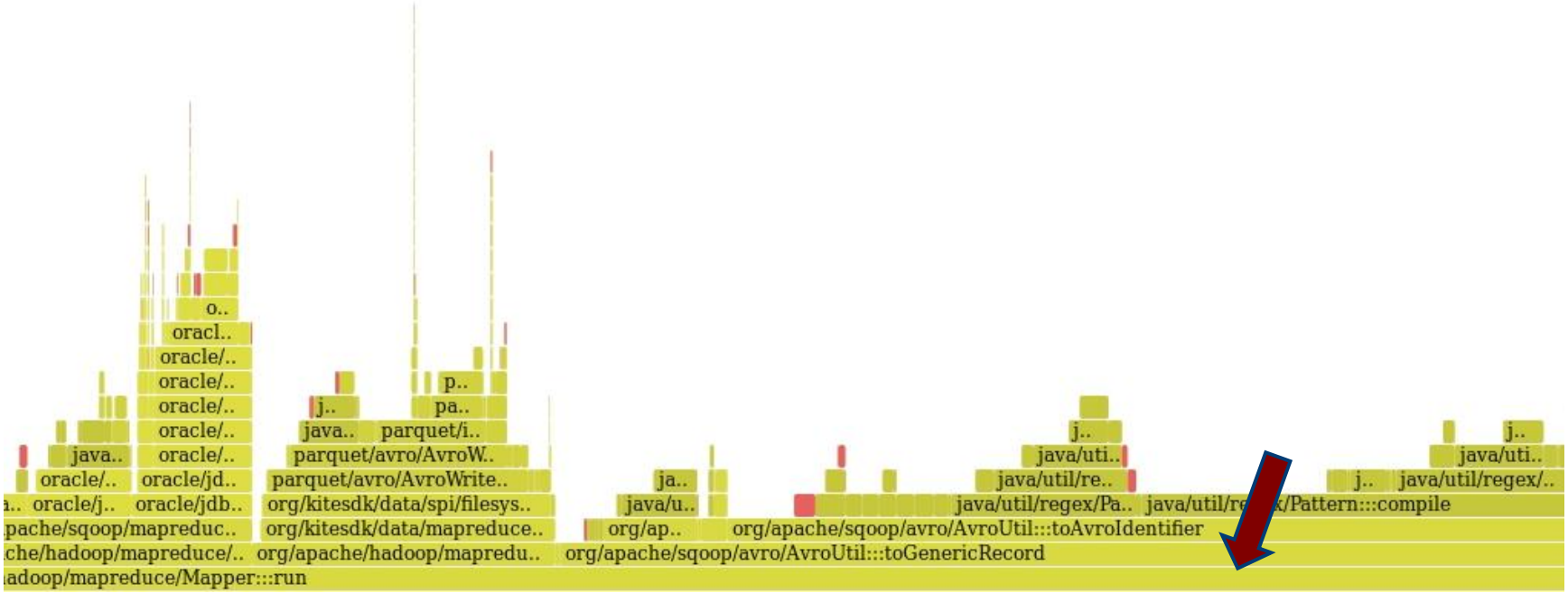


Hadoop performance troubleshooting

- hprofile
  - Tool developed by IT Hadoop service to troubleshoot application performance on Hadoop
  - Ability to identify part of the code the application is spending most time on and visualize this in a Human readable manner using flamegraphs
- Usage and more information
  - https://github.com/cerndb/Hadoop-Profiler
  - Blog - http://db-blog.web.cern.ch/blog/joeri-hermans/2016-04-hadoop-performance-troubleshooting-stack-tracing-introduction

Hadoop performance troubleshooting

- This profiler helped to identify the performance bottlenecks in sqoop when importing data in parquet format



Service Evolution

- **Kudu** – New Hadoop storage for faster analytics
  - Complements HDFS and HBASE
  - Fills the gap in capabilities of HDFS (optimized for analytics on extremely large datasets) and HBASE (optimized for fast ingestion and queries over small datasets)

- **Backups** for Hadoop

- Evaluation and possible deployment of **Alluxio** – in-memory distributed filesystem