CERN openlab

Data Analytics

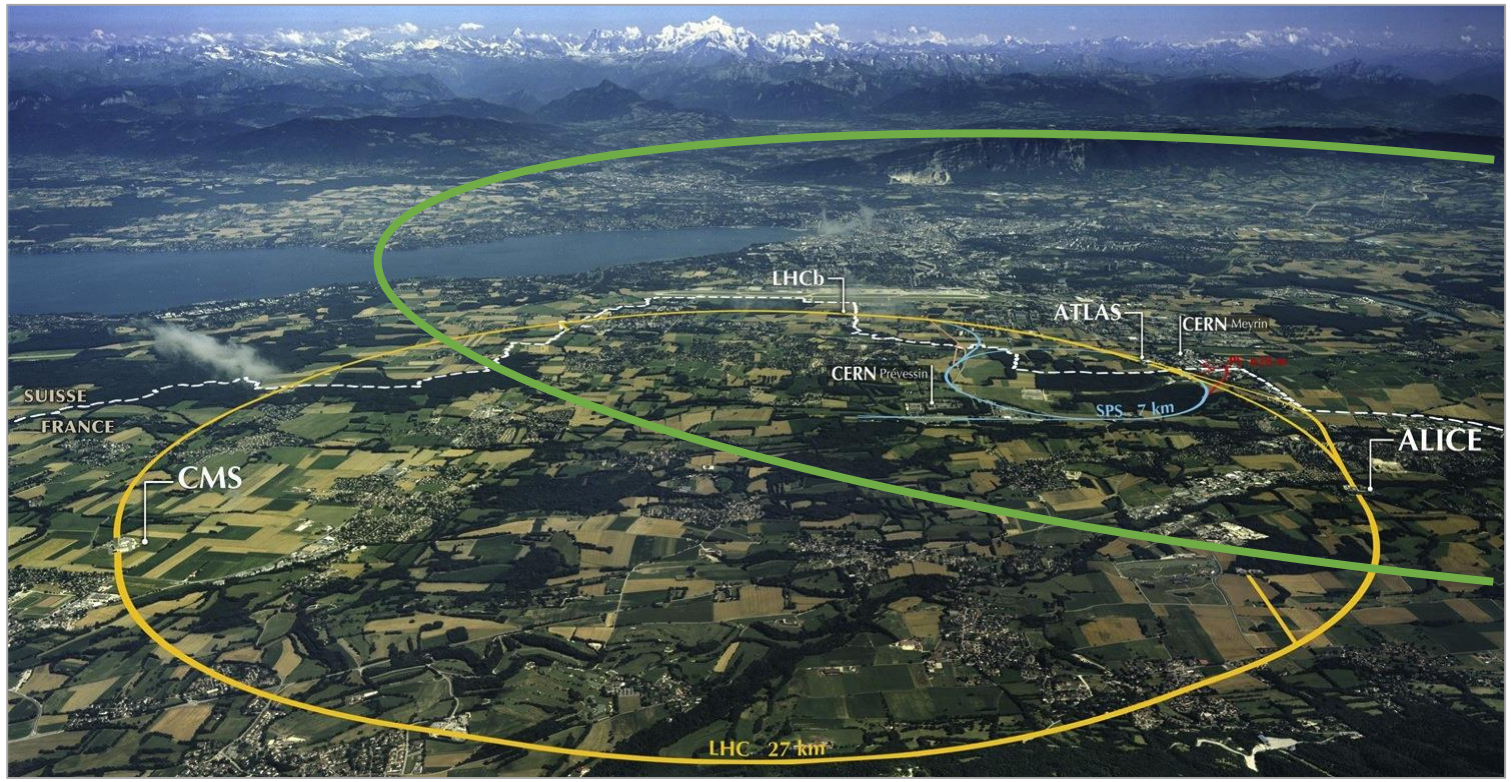
CERN openlab Open Day

Manuel Martin Marquez

Self-Service Data Analytics

Future Circular Collider Study

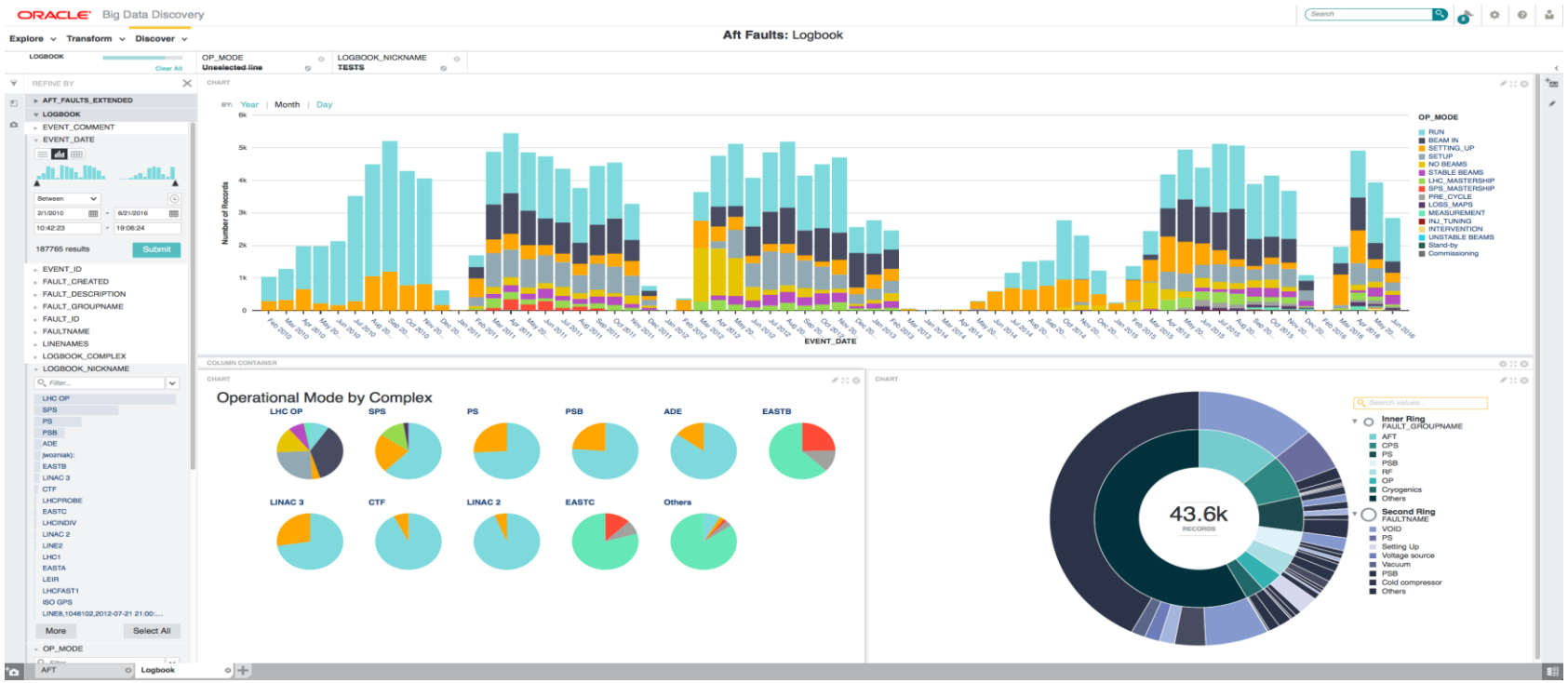
- Collaboration with more than 70 institutes from all over the world
- Need for a central data repository, which integrates several data sources
  - Control Configuration
  - Fault tracking data
  - Control Iot Systems
  - Operation Logbook
- Leverage analytics capabilities for highly heterogonous objectives
- Different expertise and backgrounds involved



Self-Service Data Analytics

Future Circular Collider – **Oracle Big Data Discovery**

- Agile data governance
- User driven ETL
- Hide technology complexity
- Self-Service discovery tools
- Collaboration and findings sharing
- Integrate seamlessly with advance analytics

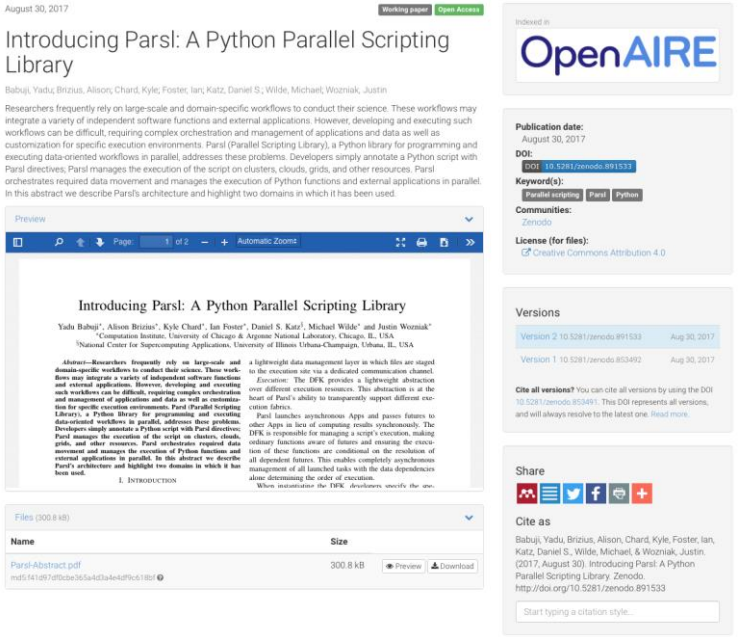


Self-Service Data Analytics

Graph Data Bases – **Oracle Parallel Graph Analytics**

- Assess the Data Analytics advantages of Graph Databases
- Find patterns and indirect relations not easy to obtain with traditional database systems
- Evaluate the potential applications of the technology within CERN

- Proof-of-Concepts:
  - Zenodo
  - Accelerator Faults analysis



Self-Service Data Analytics

Some Future Plans – **Cloud Analytics**

- Oracle Data Visualization and Discovery Cloud solutions
- Streaming Data – Kafka, Oracle Stream Explorer and Golden Gate
- Push analytics computation to the cloud
- Evaluate services integration