

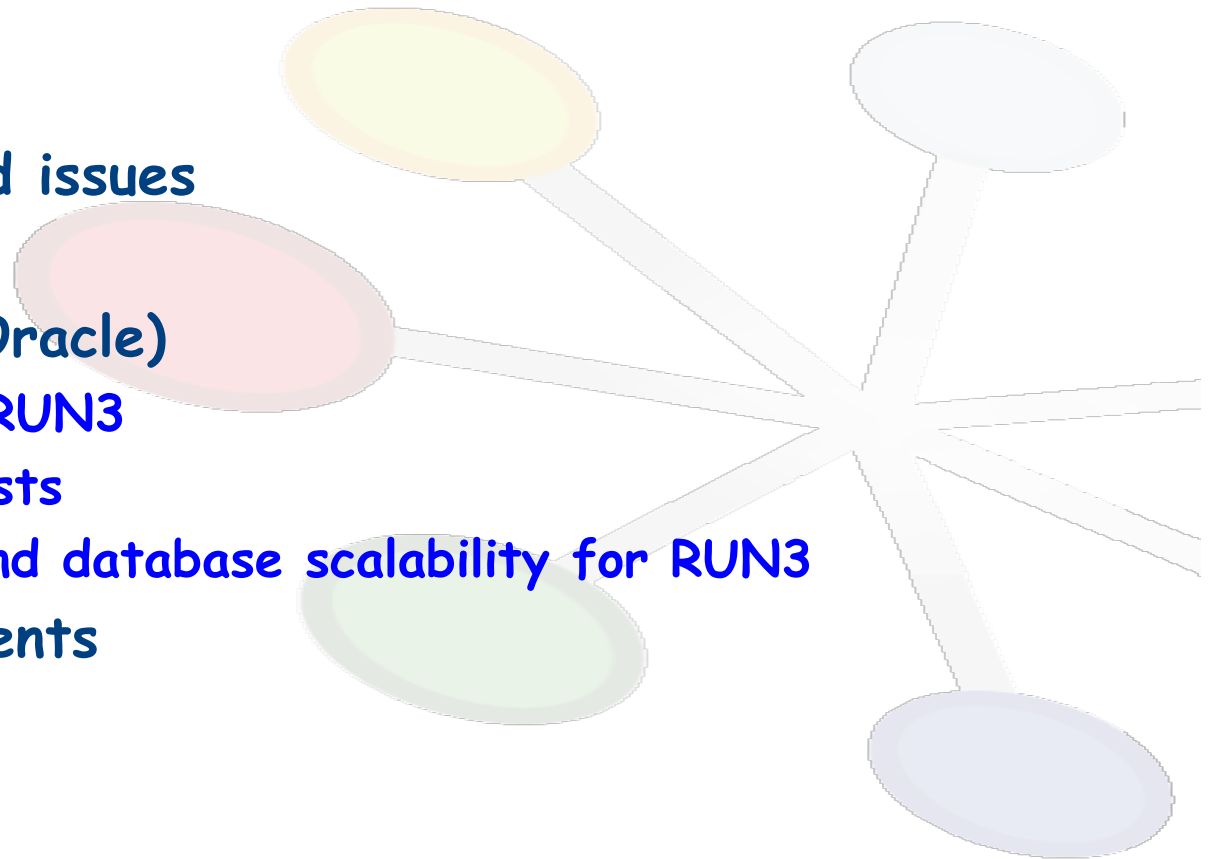


The LHCb experiment offline databases and LHCb requirements for RUN3

Luca Tomassetti and Zoltan Mathe
on behalf of LHCb experiment

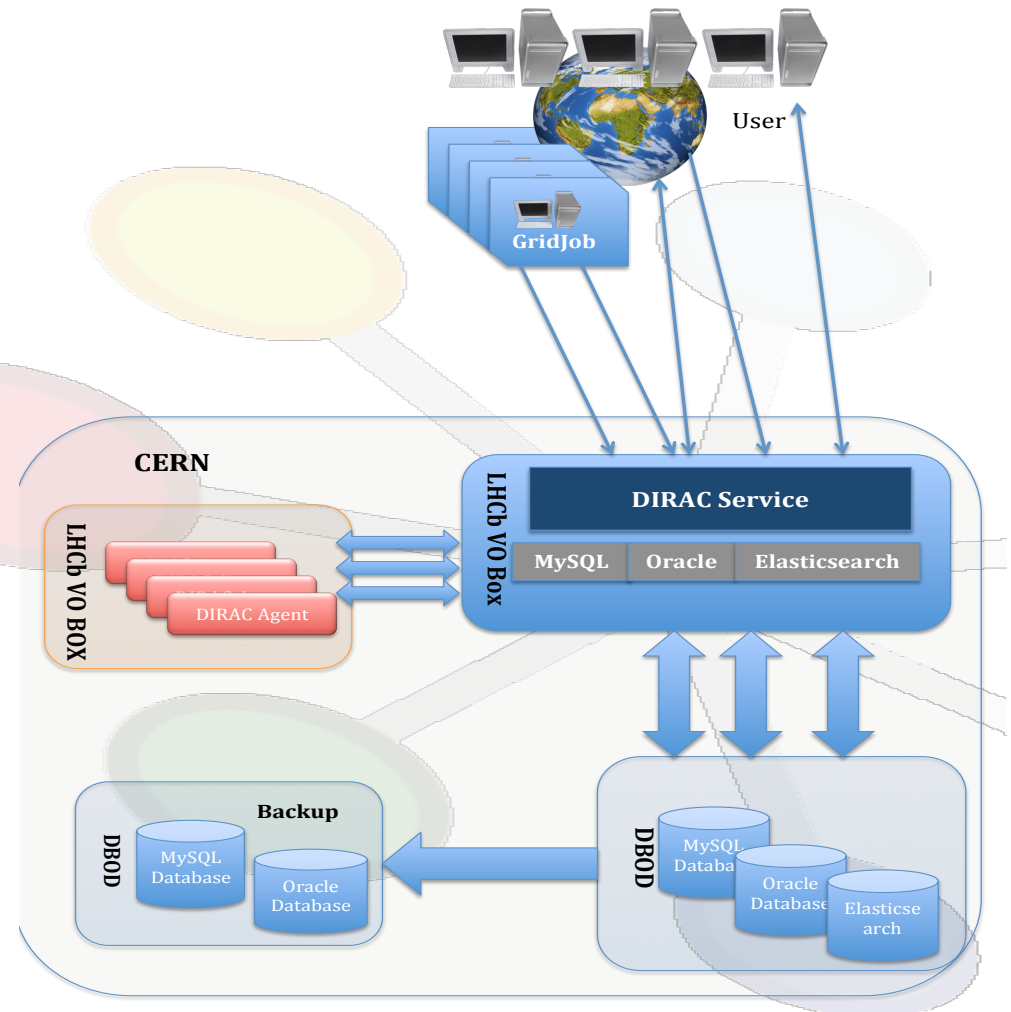


- LHCbDIRAC
- LHCb offline databases
 - MySQL
 - Oracle
 - NoSQL
- MySQL usage and issues
- MySQL in RUN3
- BookkeepingDB (Oracle)
 - Preparation to RUN3
 - Performance tests
 - Requirements and database scalability for RUN3
- NoSQL Requirements



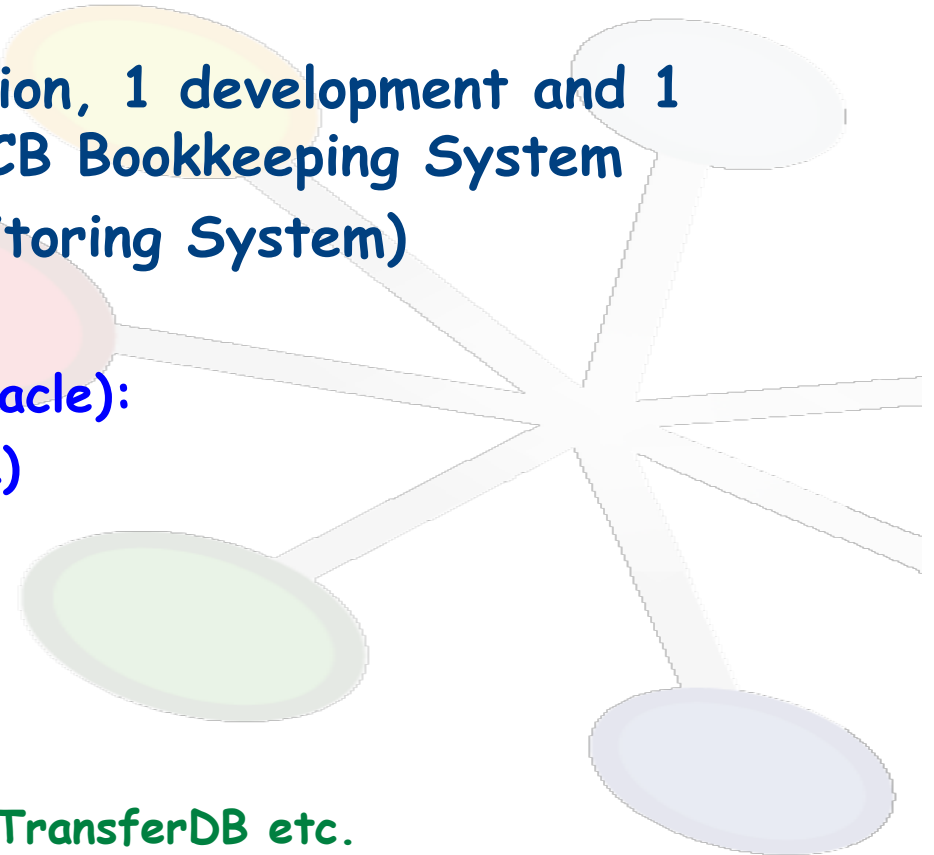


- DIRAC interware is a software framework, for distributed computing
- Very large communities: Belle II, ILC/CLIC, BES III, CTA, GridPP, France Grilles, EGI, etc.
- Main user is LHCb, which has an extension of DIRAC so called LHCbDIRAC (LHCb specific code, resources such as databases)





- LHCbDIRAC components are using Oracle, MySQL, Elasticsearch and SQLite (very limited) databases.
- 13 MySQL instances (production and certification), all are provided by IT (DBOD)
- 3 Oracle instances (1 production, 1 development and 1 certification) are used by LHCb Bookkeeping System
- 2 Elasticsearch clusters (Monitoring System)
- Various use cases:
 - Structured data (MySQL, Oracle):
 - Time Series (MySQL, NoSQL)
- MySQL DBs:
 - Accounting
 - ☆ AccountingDB
 - Data Management System
 - ☆ DIRAC File Catalog, FTSDb, TransferDB etc.
 - Workload Management System
 - ☆ JobDB, TaskQueueDB, PilotAgentDB, etc.

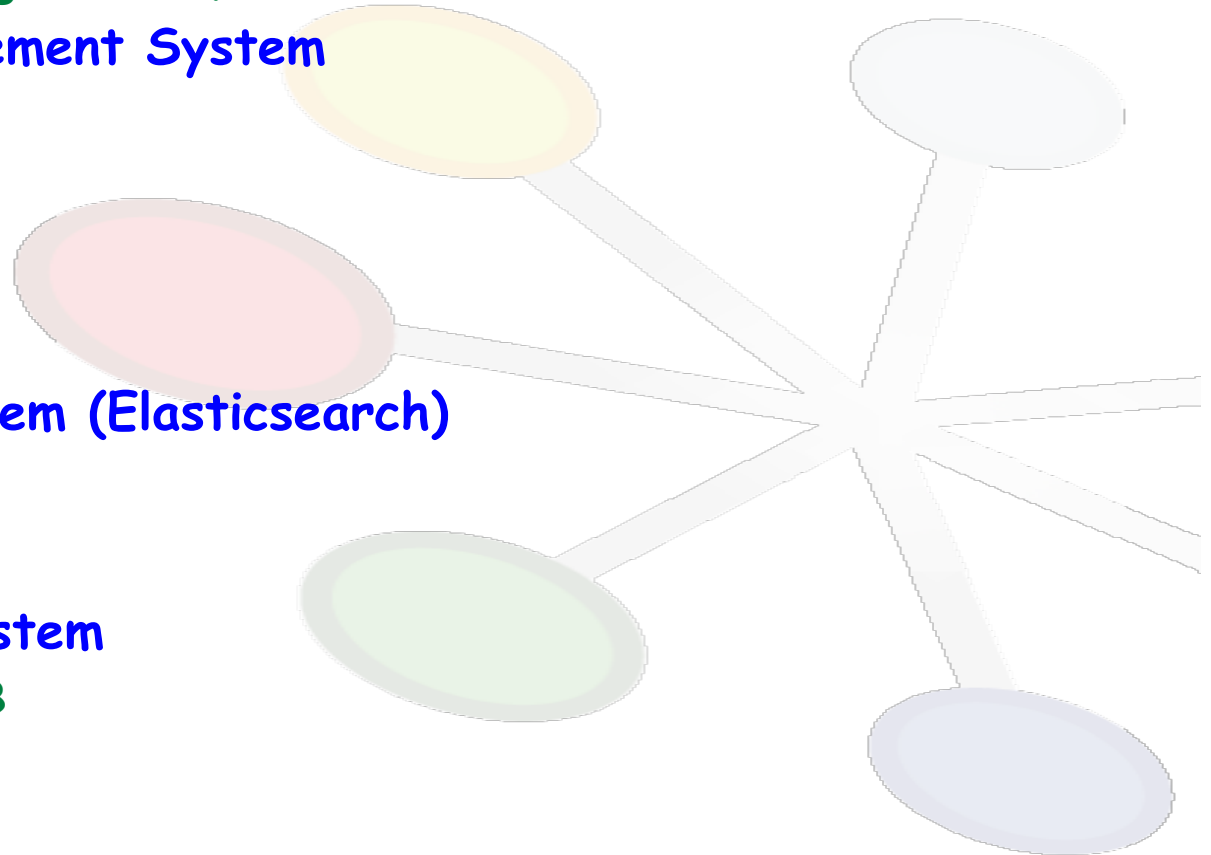




- Transformation System
 - ☆ TransformationDB
- Resource Status System
 - ☆ ResourceManagementDB, ResourceStatusDB
- Request Management System
 - ☆ RequestDb
- and many more

- NoSQL:
 - Monitoring System (Elasticsearch)
 - ☆ MonitoringDB

- Oracle:
 - Bookkeeping System
 - ☆ BookkeepingDB

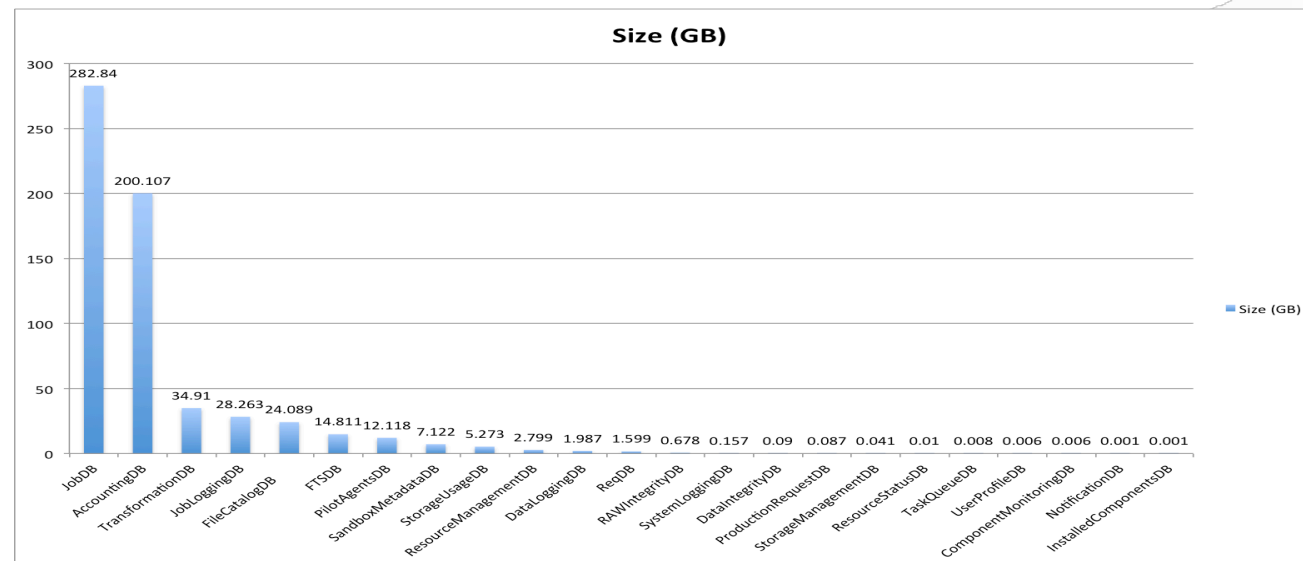




MySQL usage and issues

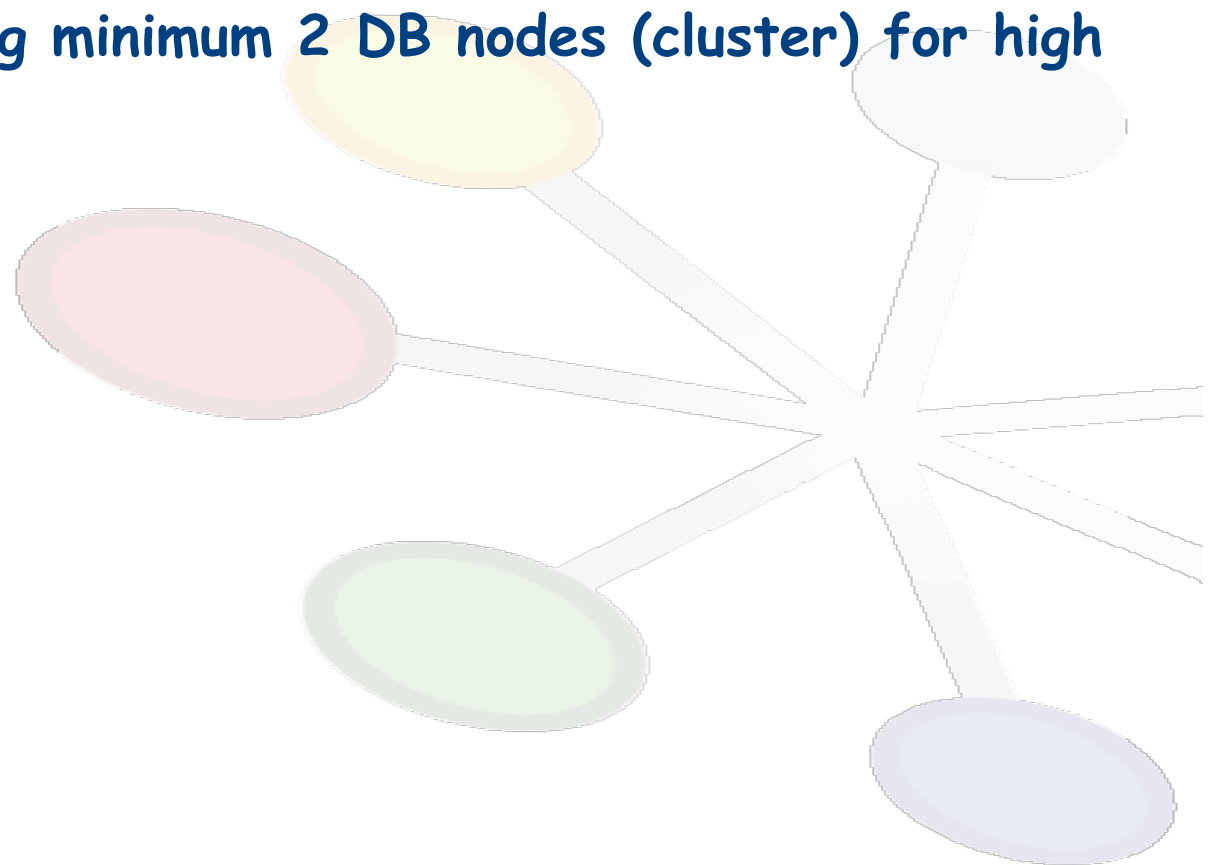
LHCb Offline DBs

- The average load on the database servers around a few thousand Hz (For DFC 14k, WMS 2,6k)
- MySQL was not stable at the beginning, required interventions
 - DB back-ends changes
 - Updates
 - One instance was holding two big databases
- DBOD team has been always very responsive
- Now, it is stable





- Continue using MySQL
- All MySQL DBs are very critical
 - Long Downtime is a big issue
- prefers for having minimum 2 DB nodes (cluster) for high availability

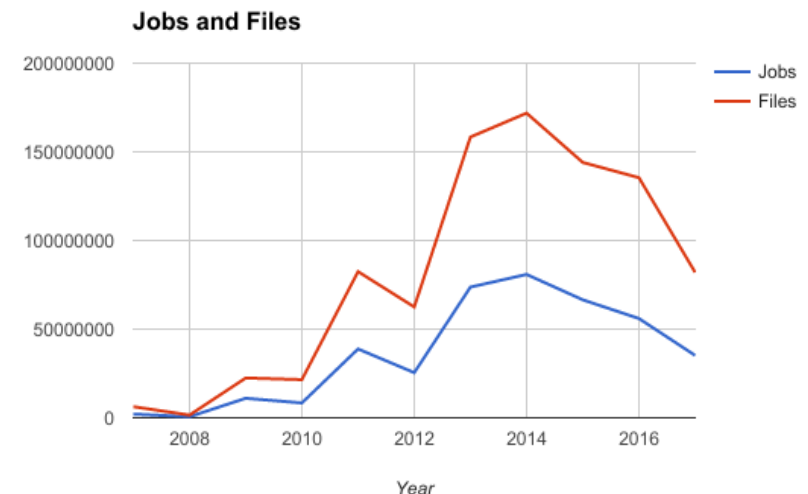




LHCb Bookkeeping System

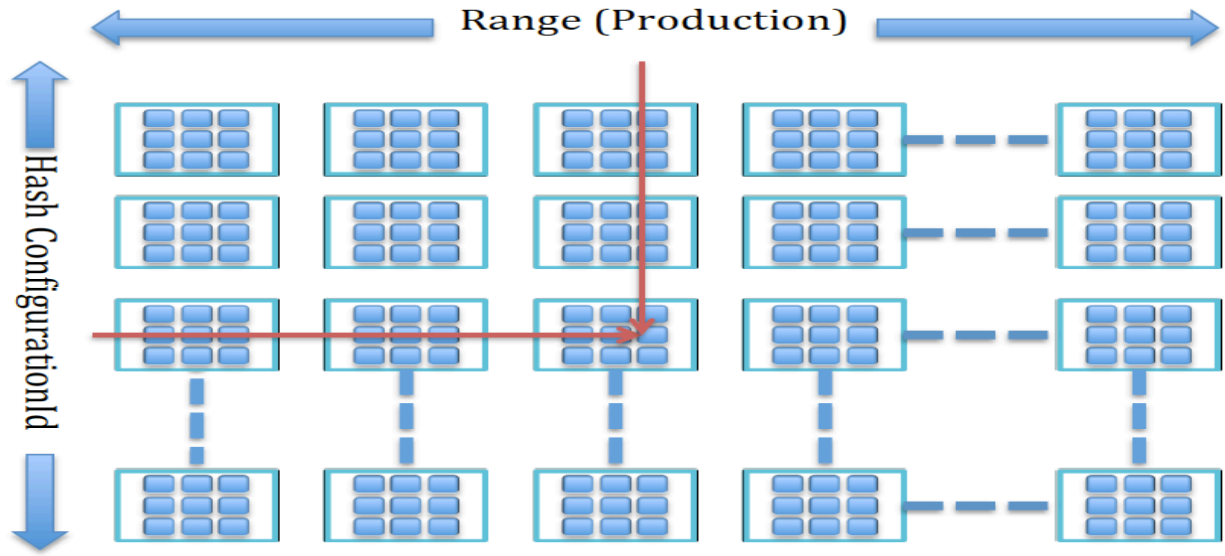
LHCb Offline DBs

- LHCb Bookkeeping System is a metadata management system which stores the conditions relative to jobs and files as metadata associated to them, as well as their provenance information in an organized way
- It is heavily used by users and production system
 - Last 24 hours the system is used by 149 unique users
 - It has to cope with ~15 queries per second
- Based on Oracle, the db consists of 17 tables and 3 materialized views
 - 3 biggest tables
 - ☆ Jobs 397 million rows
 - ☆ Files 892 million rows
 - ☆ Inputfiles 480 million rows
 - Size 1.6 TB





- New partition schema (in production since 21 of February)
 - Range-Hash composite partitioning on *Production* and *Configuration*
- New hardware 1th of March
 - Increased the Oracle compatibility mode to 12.1.0.2
 - ★ We are using Oracle Hint for a certain queries (bug reported to Oracle)
- Enable in-memory store (ongoing)

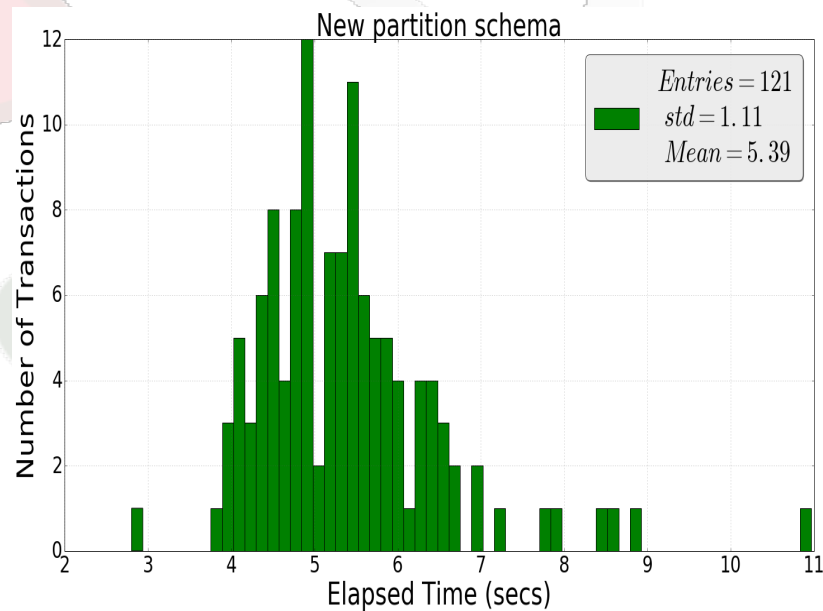
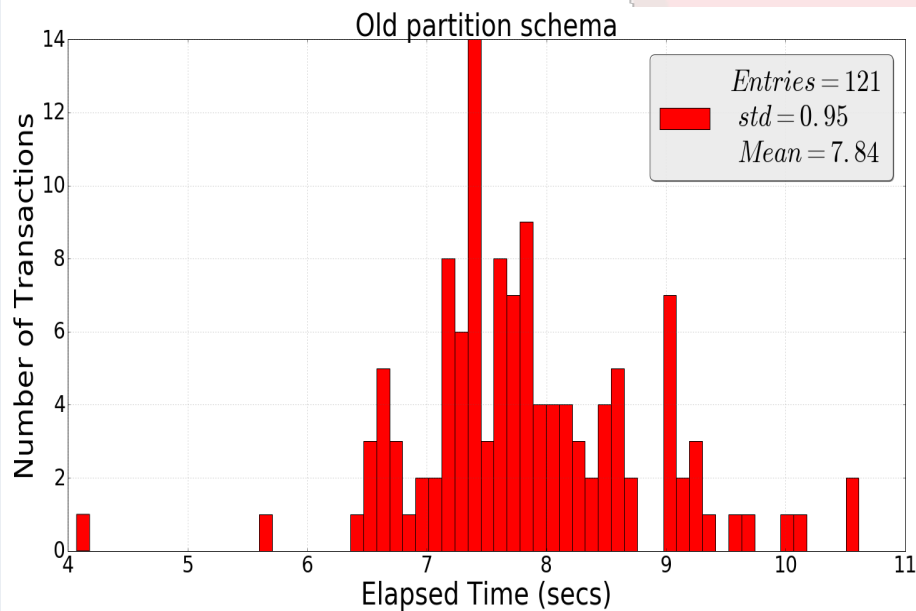




Performance test(1)

LHCb Offline DBs

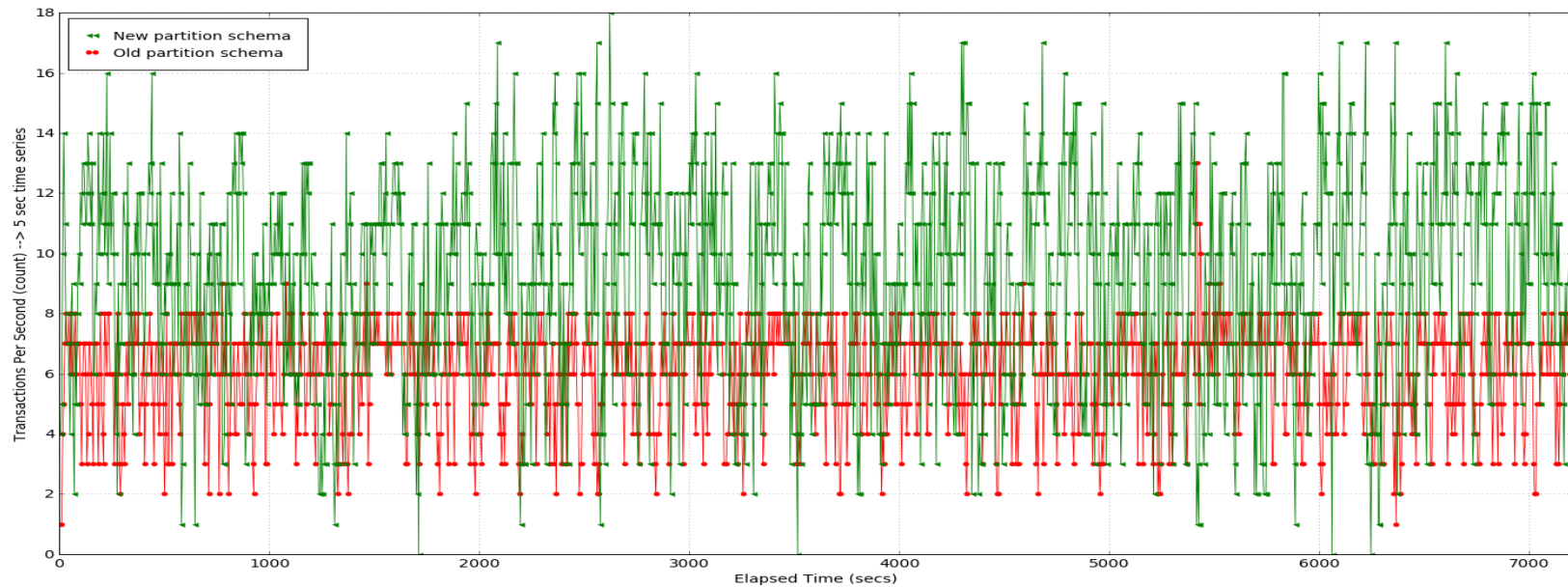
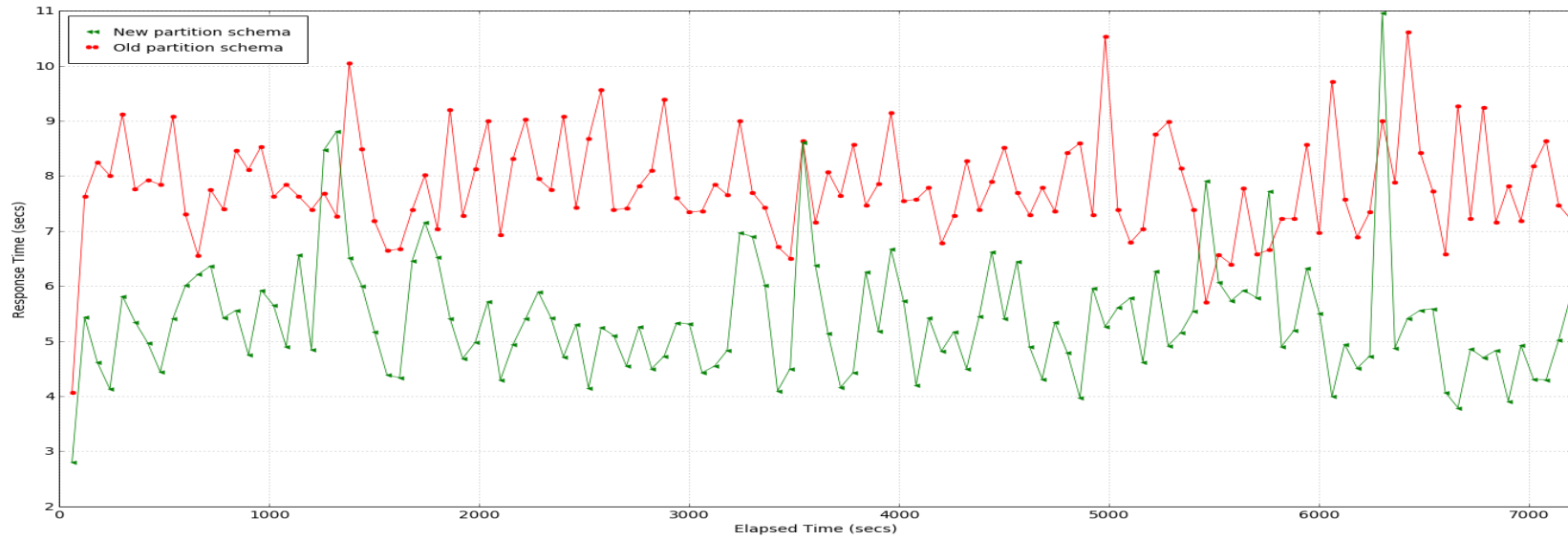
- New partition schema and 12.1.0.2 compatibility parameter
- High workload generated by 50 virtual users (python threads)
- 240 distinct datasets
- We run the tests during 7200 second and measured the response time and throughput





Performance test(2)

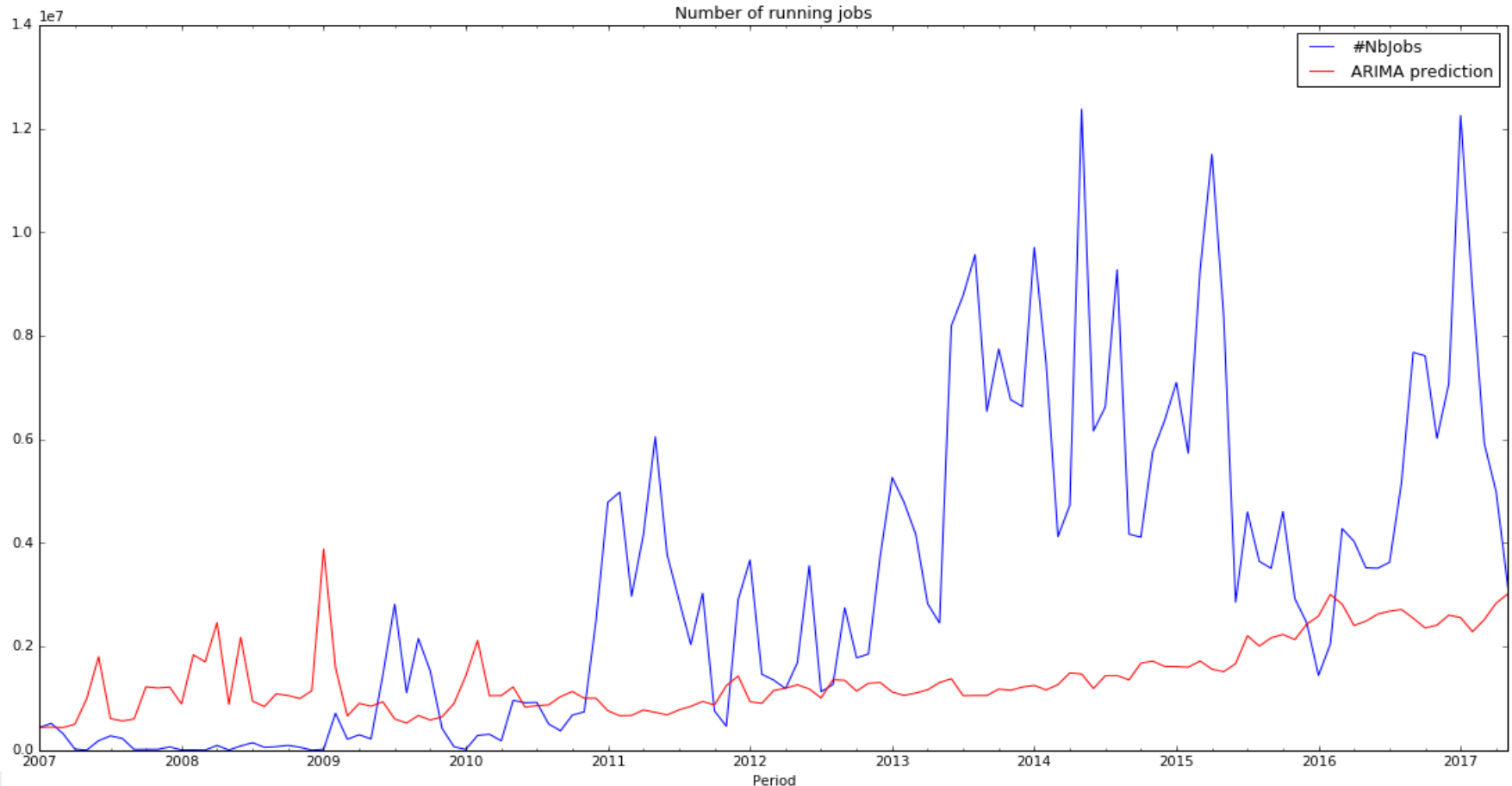
LHCb Offline DBs





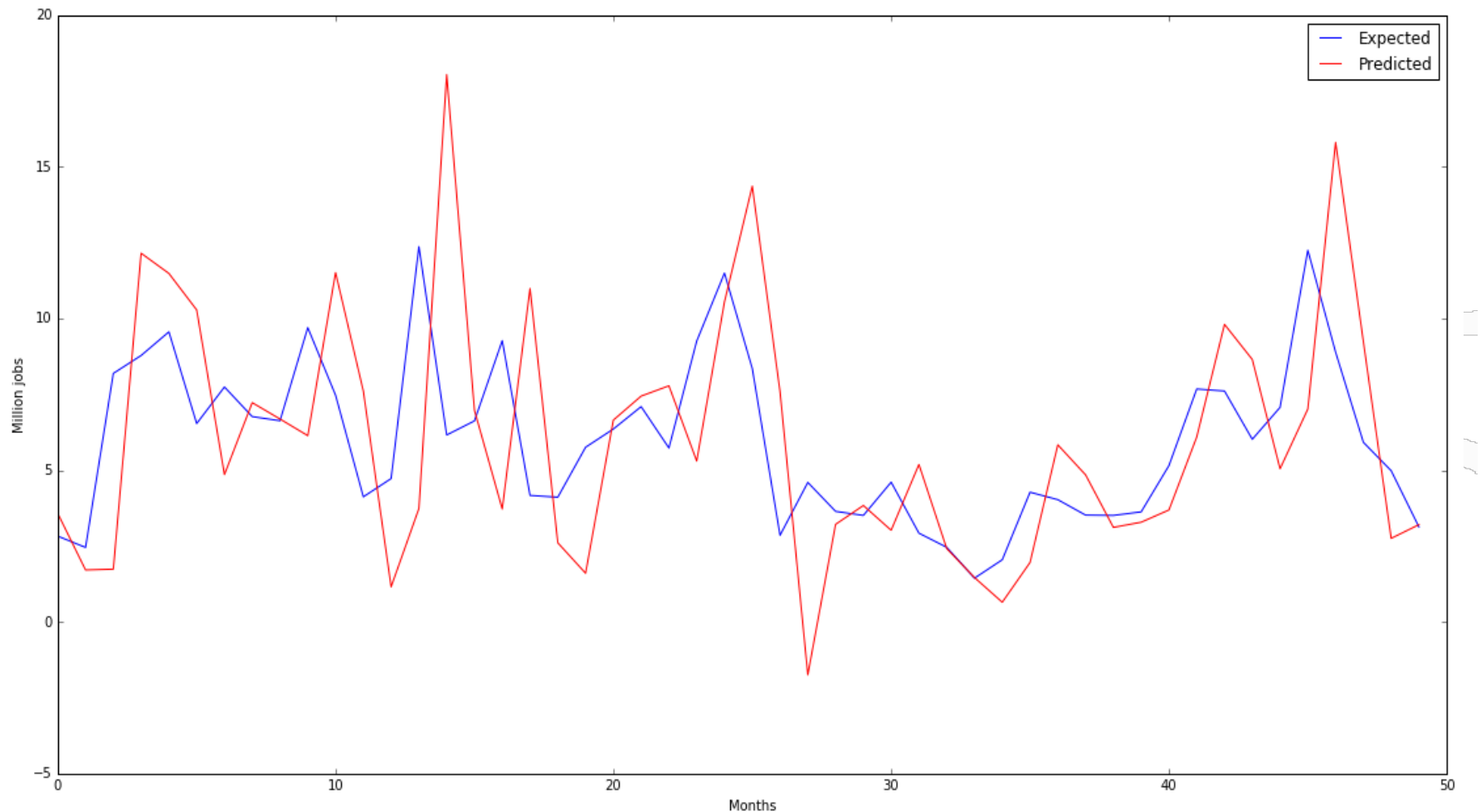
Requirements and Database scalability for RUN3

- We will continue using Oracle in RUN3
- DB size:
 - Forecast of the data size growth
 - History of the database from 2007
 - Defined ARIMA model





- 4 years forecast using ARIMA
- We expect ~300 million jobs during 4 year (assuming that we will have the same amount of CPU during this period)





- Elasticsearch for WMS Monitoring
 - 108 GB, 1200 indexes, 1280 million documents

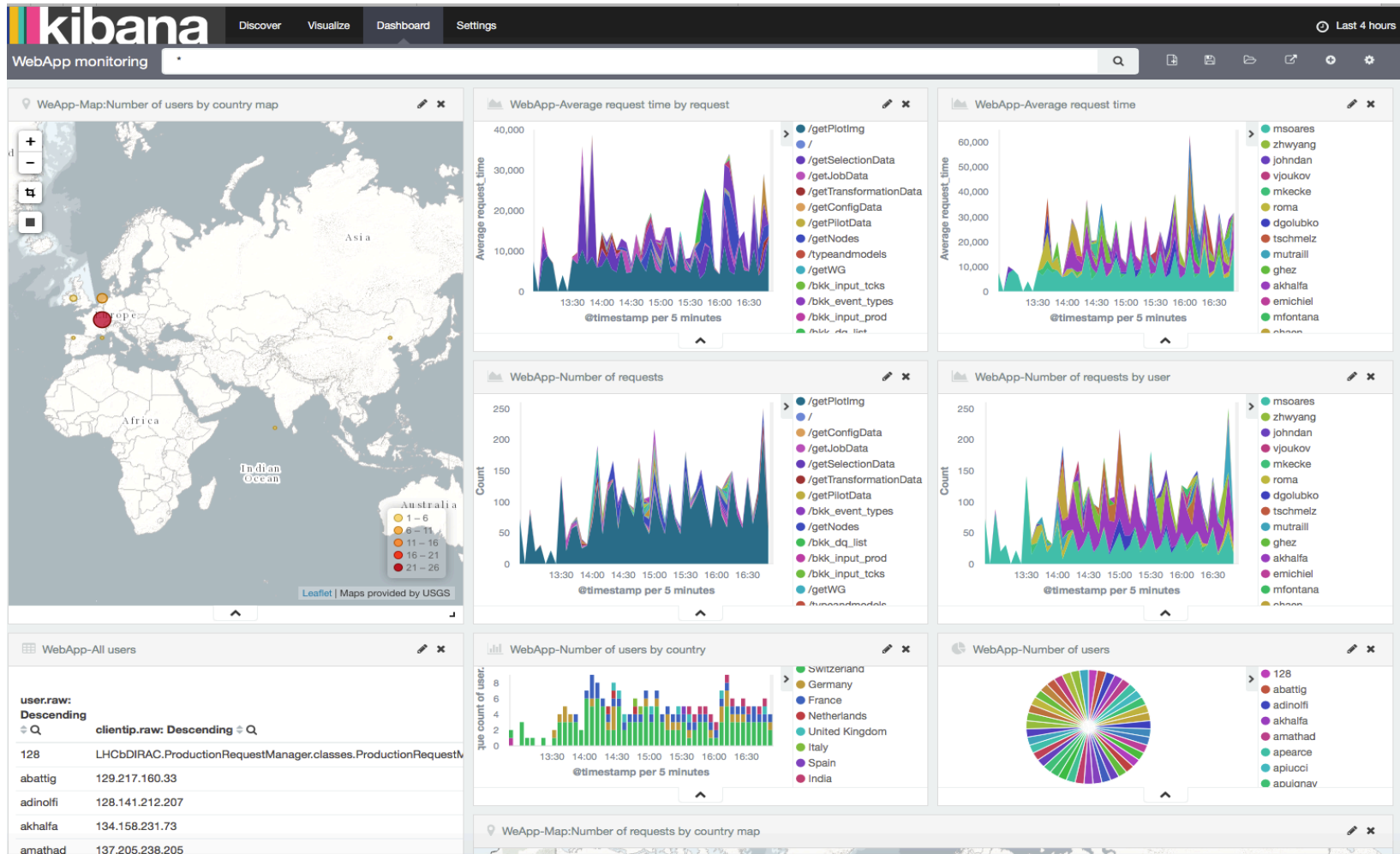
LHCb Offline DBs





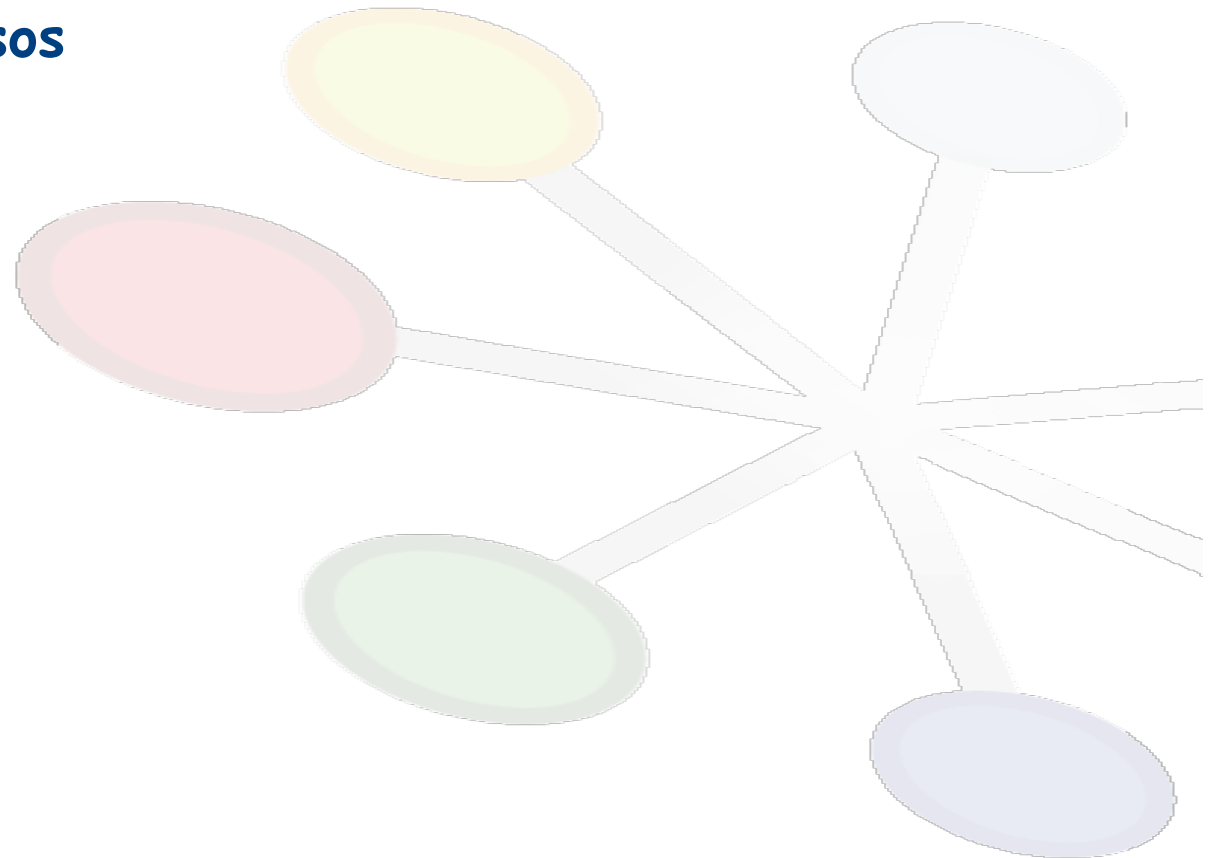
○ ELK for monitoring the web server (pre-production)

LHCb Offline DBs





- More monitoring data
- Centralized Monitoring based on Elasticsearch
- Hadoop for keeping the data for longer period
- InfluxDB for Mesos





- We plan to continue using the current DBs for RUN3
- We will use more Elasticsearch for monitoring
- We are considering to introduce Hadoop and InfluxDB
- We are happy for the DBs support (Oracle, MySQL and Elasticsearch)
- Thanks to the database groups for providing scalable database services

