Database Services

Katarzyna Dziedziniewicz-Wojcik
On behalf of IT-DB
Agenda

- Database service overview
- Status & plans
  - Oracle
  - Database On Demand
  - Hadoop
Databases

- 3 main pillars
  - Oracle
    - For critical, transactional load
    - Administered by experts (DBAs)
  - Database on-demand
    - Different database engines
      - MySQL, PostgreSQL, Oracle, InfluxDB
      - Instance owners have the DBA rights
  - Hadoop
    - For analysis where DBs are not suitable
At the heart of CERN, LHC and Experiment Operations

http://cern.ch/it-dep/db/
Oracle in numbers

• ~100 databases, most of them RAC
  • Mostly NAS storage plus some SAN with ASM
  • More than 600 TB of data files for production DBs in total
Oracle – short term plans

- Creation of a public database service supporting UNICODE
- Hardware migration of all experiment and project databases
  - New nodes 512GB of RAM
- Migration of all non-accelerator APEX installations to APEX 5
Oracle – long term

- Oracle 11.2 and 12.1
  - Supported until 2020 and 2021 respectively
- LS2 upgrade
  - Oracle 12.2 (recommended)
    - Release date 2016
- Main new feature
  - In-Memory database
    - Real-time responsiveness
On-Demand in numbers

- All databases – 415
  - 25% growth in last 6 months
- Production – 276
- By DB Type
  - MySQL: 321
  - PostgreSQL: 70
  - Oracle: 9
  - InfluxDB: 17 (pilot)
Time series databases

- Time series data workload assumptions
  - Billions of individual data points
  - High write and read throughput
  - Mostly an insert/append workload, very few updates/deletes
  - Large deletes to free up disk space
InfluxDB

- Pilot service, production in Q1 2017
- HTTP Write and Query API
- SQL like language to query aggregated data
- Schema-less
- Works with different visualization tools: Grafana, Chronograpf, …
- Authentication
- Data Retention policies
- Storage optimization & compression
InfluxDB use cases

• Cgroup stats for batch jobs
• Fifemon(scheduler and central manager of batch nodes)
• Gitlab
• Stats
  • Network
  • TSM
  • ActiveMQ
• Monitoring
  • LHCb
  • WLCG
  • DBoD
Databases – On-Demand

- MySQL CE 5.7
  - Q1 2017
  - To stay for next 3 years
- PostgreSQL 9.6
  - Q4 2016
- Oracle as for General Service
- Upgrades require database reboot
Status for On-Demand

- SSL for client access encryption and authentication ✔
- High availability
  - Replication ✔
  - App/DB Gateways (coming soon)
  - Clustered Databases (2017)
- **Instance Monitoring**
  - Currently studying a hybrid solution:
    - ElasticSearch+Packetbeat for database activity (SQL)
    - InfluxDB+telegraf for instance metrics
RDBMS vs Hadoop

RDBMS optimized for OLTP, Hadoop affordable at scale

The shared nothing architecture allows to scale for high capacity and throughput on commodity HW

Example of Oracle RAC deployed with shared storage
Main components of CERN-IT Hadoop service

- **Kafka**: Data streaming hub
- **Zookeeper**: Coordination
- **Flume**: Log data collector
- **Impala**: SQL
- **Spark**: Large scale data processing
- **Hive**: SQL
- **HBase**: NoSQL columnar store
- **YARN**: Cluster resource manager
- **MapReduce**: Data exchange with RDBMS

**HDFS**: Hadoop Distributed File System
Hadoop usage at CERN

- 3 production clusters
  - Aggregated capacity ~ 1000 cores, 3TB RAM, 6PB Storage (~2 effective PB)
- Main use-cases
  - IT Monitoring
  - ATLAS EventIndex
  - Accelerator Logging v2 (pilot phase)
Plans – Apache Kafka

- Distributed streaming platform
- Open source project started at LinkedIn in 2011
- Stores streams of data in distributed replicated cluster
  - Horizontally scalable (streams partitioning)
  - Fault tolerant (replication of partitions)
- Allows to access and process the streams of data in real-time
- Became an industry standard and important component of Hadoop ecosystem
Kafka as Hadoop enhancement

No data lost during downtime (scheduled and unscheduled) of a Hadoop cluster

Kafka buffers protects the recent data from being lost (before a daily HDFS snapshot can backed them up)

How Kafka enhances data pipelines on Hadoop

Stream Source

Events

Central data buffer

kafka

Flush periodically

Flush immediately

Real time stream processing

Big Files

HDFS

Batch processing

Indexed data

Fast data access

APACHE HBASE

DB Services
Apache Kafka as a service

- Core component for mission critical systems at CERN
  - CERN Accelerator Logging System v2
  - Industrial Controls and Data Acquisition systems
  - IT Monitoring & Security
- Production quality service needed
- Pilot service proposed
  - Joined effort of IT-DB and IT-CM groups
- Collecting inputs from the user communities ...
  - it-kafka-service@cern.ch
Hadoop data backup

- Introduction aimed at Q1 2017
- Cold, incremental copy of files every 24 hours
- Last 24 hours of stored in Kafka buffers
  - Data owner’s responsibility to insert them
Acknowledgments

• Colleagues from IT-DB
• Special thanks:
  • Zbigniew Baranowski
  • Ignacio Coterillo Coz
  • Prasanth Kothuri
  • Antonio Romero Marin