

# Monitoring Evolution

Alberto AIMAR, IT-CM-MM

# Outline

- Mandate
- Data Centres Monitoring
- Experiments Dashboards
- Architecture
- Plans
- Status
- Demo

# Monitoring

## Data Centre Monitoring

- Monitoring of DC at CERN and Wigner
- Hardware, operating system, and services
- Data Centres equipment (PDUs, temperature sensors, etc.)
- Used by service providers in IT, experiments

## Experiment Dashboards

- Sites availability, data transfers, job information, reports
- Used by WLCG, experiments, sites and users

**Both hosted by CERN IT, in different teams**

# Mandate

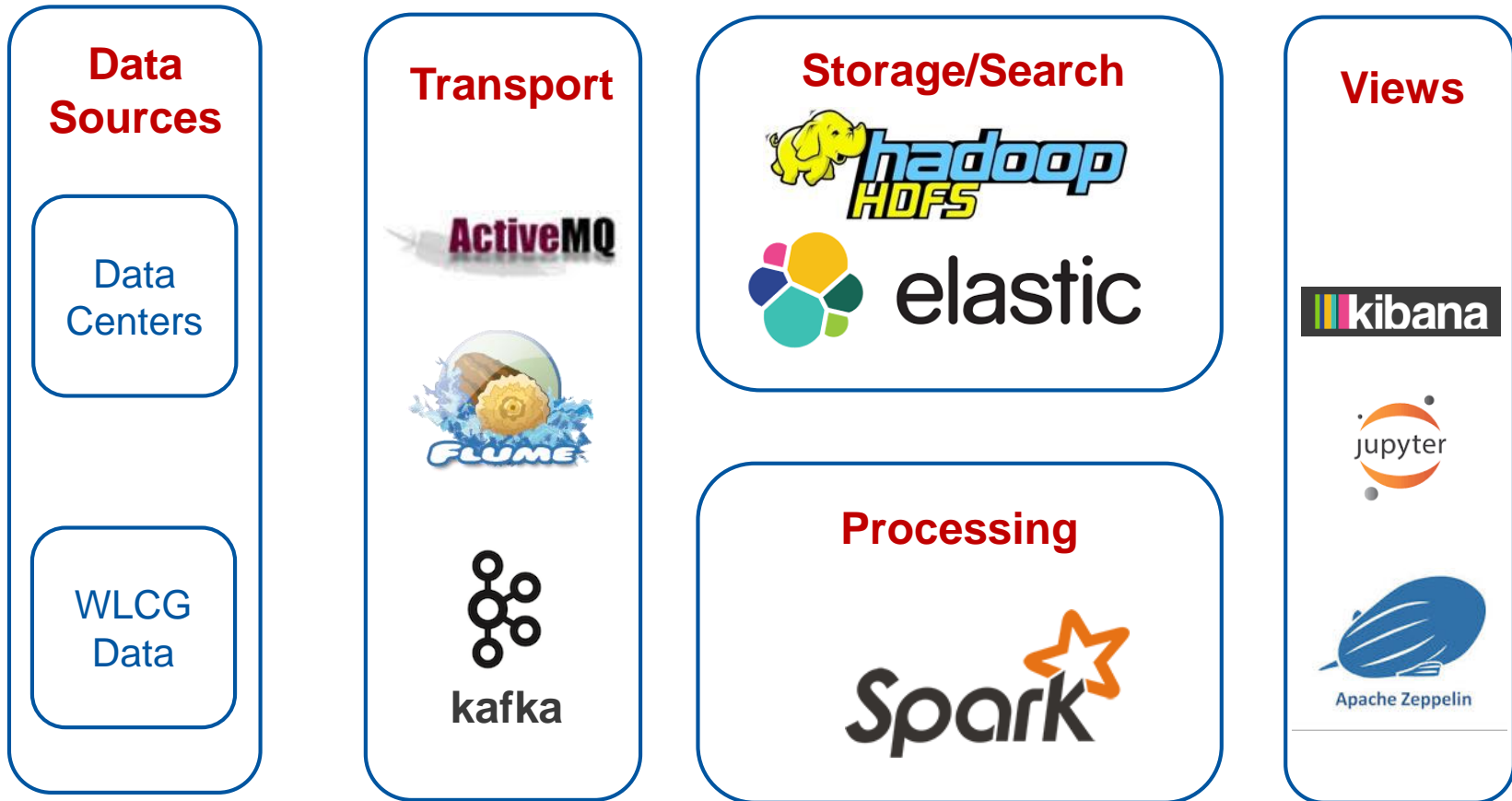
## Focus for 2016

- Regroup monitoring activities hosted by CERN/IT (Data Centres, Experiment Dashboards, ETF, HammerCloud, etc)
- Continue existing services
- Uniform with CERN IT practices
- Management of services, communication, tools (e.g. GGUS and SNOW tickets)

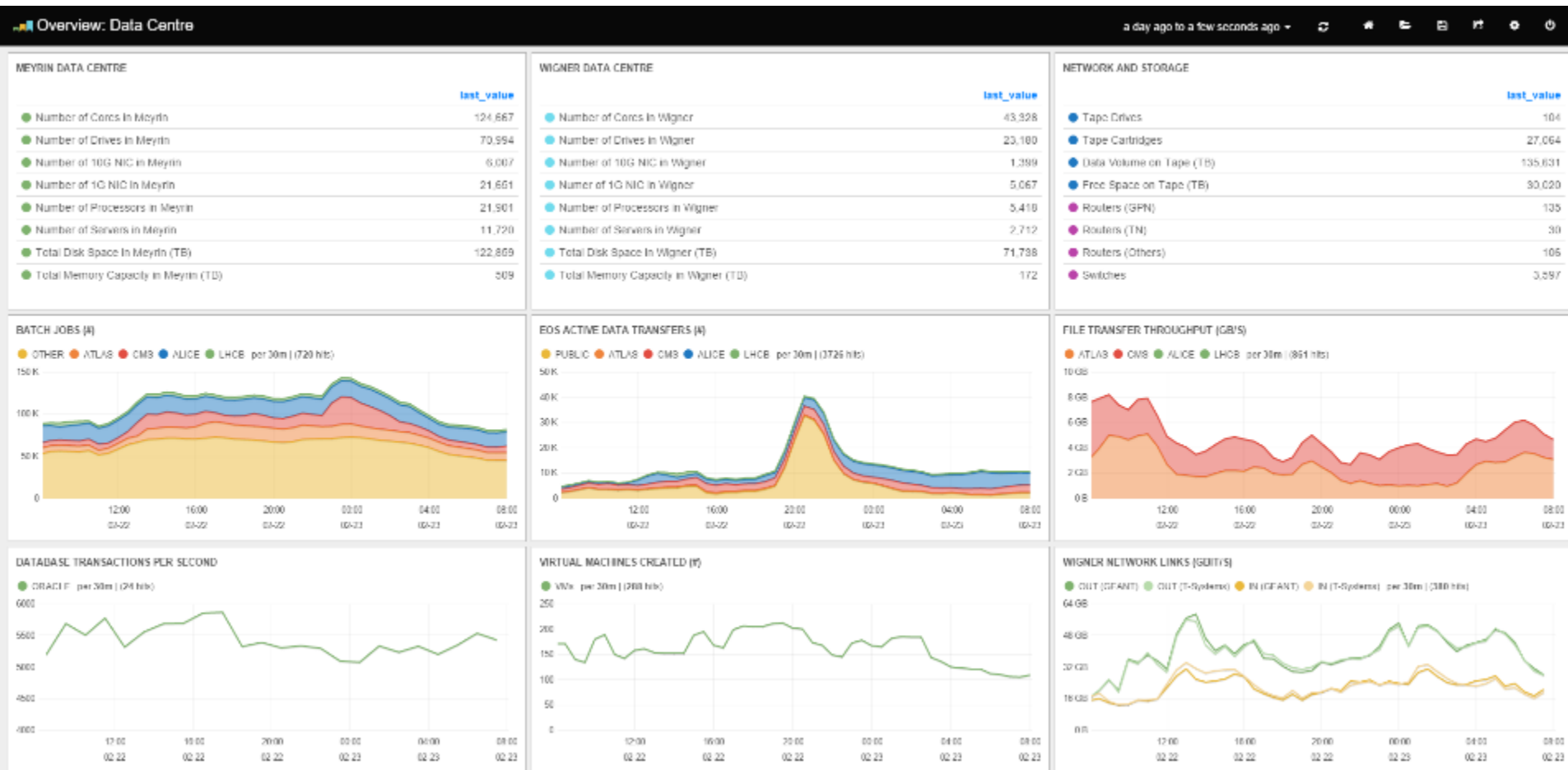
## Starting with

- Merge Data Centres and Experiment Dashboards monitoring technologies
- Review existing monitoring usage and needs (IT, WLCG, etc)
- Investigate new technologies
- Unchanged support while collecting user feedback

# Unified Monitoring Architecture



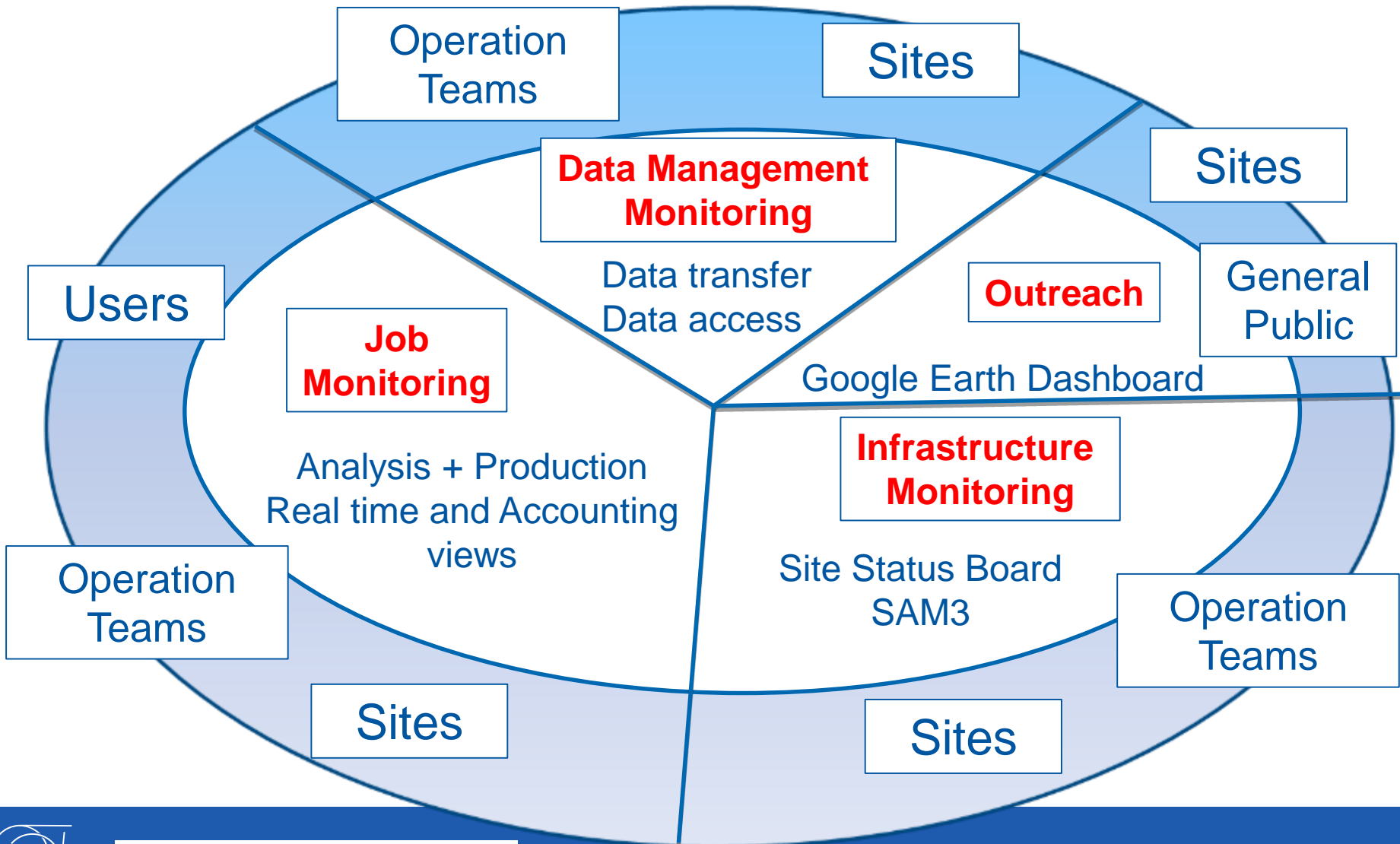
# Data Centres Monitoring



# Monitoring Technologies

Area	Services and Components	Technologies	Functions
Data Collectors	Metric Manager	CERN	Metrics registration
	Lemon Agent	CERN+Flume	Metrics producers (about 15000)
Transport	Gateway	Flume	Transport host metrics
	XSLS	Flume	Service metrics
	Messaging	Active MQ	Messaging of metrics
	River	Kafka	Streaming of metrics
Aggregation	Foz	Spark	Processing streamed metrics
Archive	HDFS	Hadoop	Long term storage
Displays	Meter	ElasticSearch + Kibana	Dashboard for metrics
	Timber	ElasticSearch+ Kibana	Dashboard for logs
	Meter Proxy	CERN	CLI and HTTP Interface to ES
Alerts	GNI	CERN	Alarms handling

# Experiment Dashboards

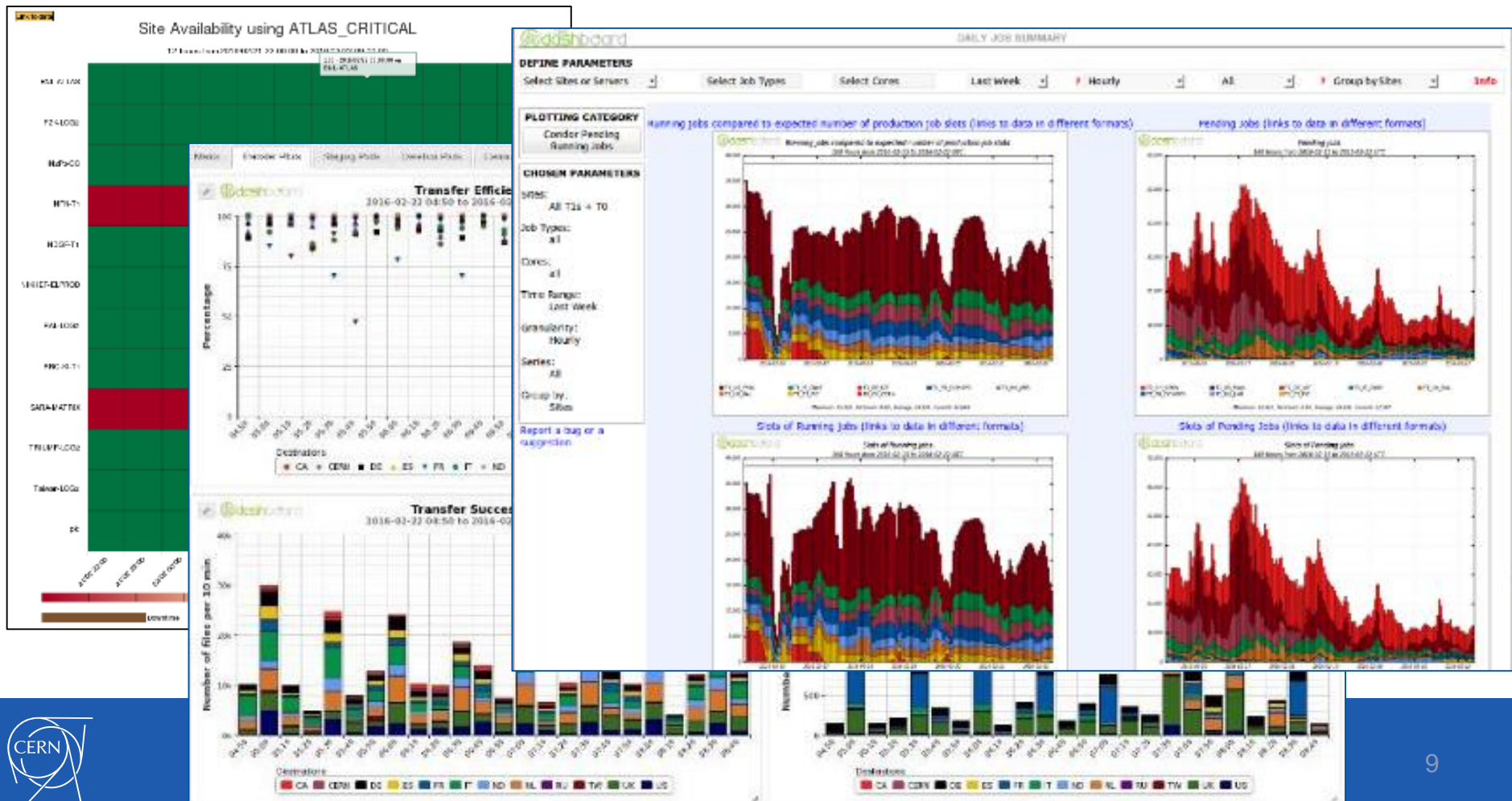


300-500 users per day



# Experiment Dashboards

- Job monitoring, sites availability, data management and transfers
- Used by experiments operation teams, sites, users, WLCG



# Current Monitoring

Data Centres Monitoring

Data mgmt and transfers

Job Monitoring

Infrastructure Monitoring

## Data Sources

- Metrics Manager
- Lemon Agent
- XSLS
- ATLAS Rucio
- FTS Servers
- DPM Servers
- XROOTD Servers
- CRAB2
- CRAB3
- WM Agent
- Farmout
- Grid Control
- CMS Connect
- PANDA WMS
- ProdSys
- Nagios
- VOFeed
- OIM
- GOCDB
- REBUS

## Transport

- Flume
- AMQ
- Kafka
- Flume
- AMQ
- GLED
- HTTP Collector
- SQL Collector
- MonaLISA Collector
- AMQ
- HTTP GET
- HTTP PUT

## Storage & Search

- HDFS
- ElasticSearch
- Oracle
- ElasticSearch
- HDFS
- Oracle
- ElasticSearch
- Oracle
- ElasticSearch

## Processing & Aggregation

- Spark
- Hadoop Jobs
- GNI
- Oracle PL/SQL
- ESPER
- Spark
- Oracle PL/SQL
- ES Queries
- ESPER

## Display & Reports

- Kibana
- Jupyter
- Zeppelin
- Dashboards (ED)
- Kibana
- Zeppelin
- Real Time (ED)
- Accounting (ED)
- API (ED)
- SSB (ED)
- SAM3 (ED)
- API (ED)

# Unified Monitoring

## Data Sources

Metrics Manager

Lemon Agent

XSLS

ATLAS Rucio

FTS Servers

DPM Servers

XROOTD Servers

CRAB2

CRAB3

WM Agent

Farmout

Grid Control

CMS Connect

PANDA WMS

ProdSys

Nagios

VOFeed

OIM

GOCDDB

REBUS

## Transport

Flume

AMQ

Kafka

## Storage & Search

Hadoop HDFS

ElasticSearch

*Other*

## Processing & Aggregation

Spark

Hadoop Jobs

GNI

*Other*

## Display & Reports

Kibana

Jupyter

Zeppelin

*Other*

# Status

## Producers and Transport

- Moving all data via new transport (Flume, AMQ, Kafka)

## Storage and Search

- Data in ES and Hadoop

## Processing

- Doing aggregation and processing via Spark

## Display and reports

- Experimenting using only the standard features of ES, Kibana, Spark, Hadoop
- Introduce notebooks and data discovery

## General

- Selecting technologies, learning on the job, looking for expertise
- Evolve interfaces (e.g. dashboards for users, shifters, experts, managers)

# Plans

## **Unified architecture and technologies**

- Focus on migrating to common architecture
- Review the existing architecture and areas

## **Update to new technologies in several areas**

- Better performance and new versions with new features and major improvements
- Look into technologies as needed (collectd, Kafka, Grafana, etc.)
- Benefit from experience and feedback received from Experiments , WLCG and IT groups

## **Move to central services**

- Central service for ES is being created, InfluxDB for time series DBoD
- Continue to use central Hadoop services

## **Continue with standard operations and upgrades**

- At least for all 2016
- Make available the new monitoring platform, in parallel with the existing ones

# Demo

- Data in ElasticSearch
- FTS and Xrootd transfers data
- Examples of Dashboards
- Data discovery and error investigation
- Specific views for specific tasks

