

Agile Infrastructure Monitoring

pedro.andrade@cern.ch

CERN IT/CF

CHEP 2013

14th October 2013

Motivation

- Several **independent monitoring activities** in CERN IT
- Combination of data from different groups necessary
- Understanding performance became more important
- Move to a virtualised dynamic infrastructure

Challenges

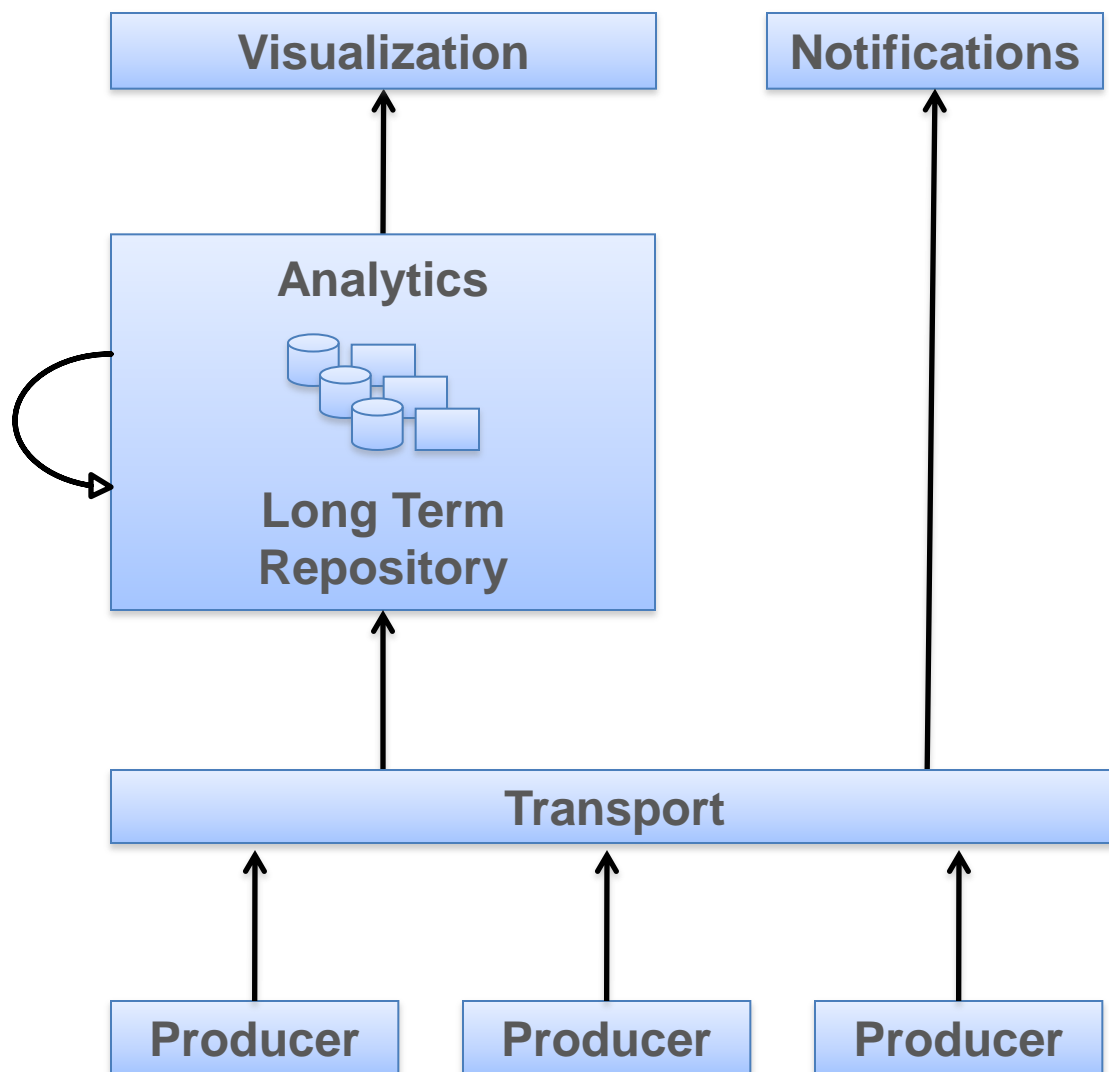
- Implement a **shared architecture** and **common tool-chain**
- Delivered under a common collaborative effort

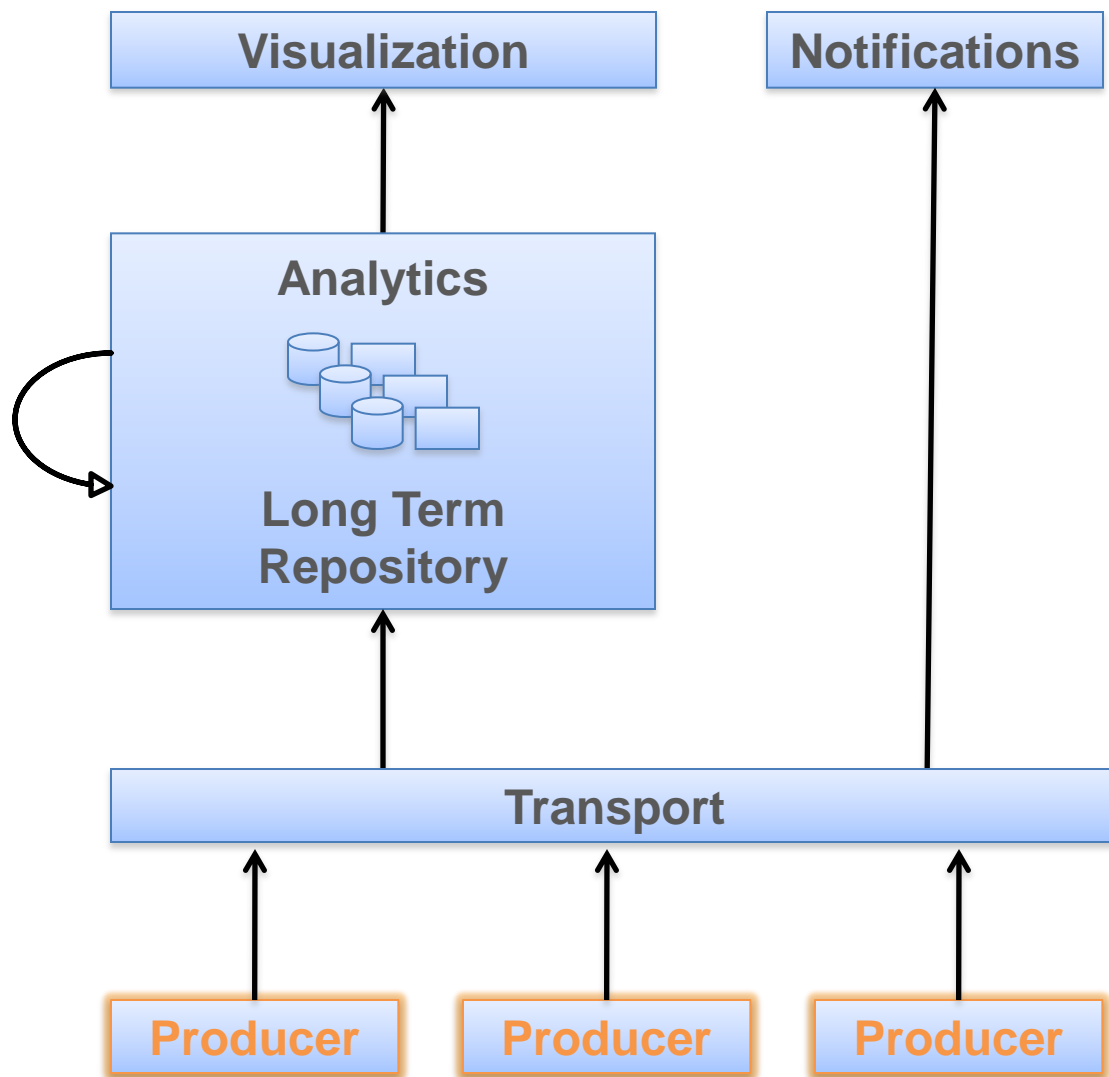
2012

- Monitoring team with contributions from few IT groups
- Architecture design and definition
- Several initial prototypes and studies

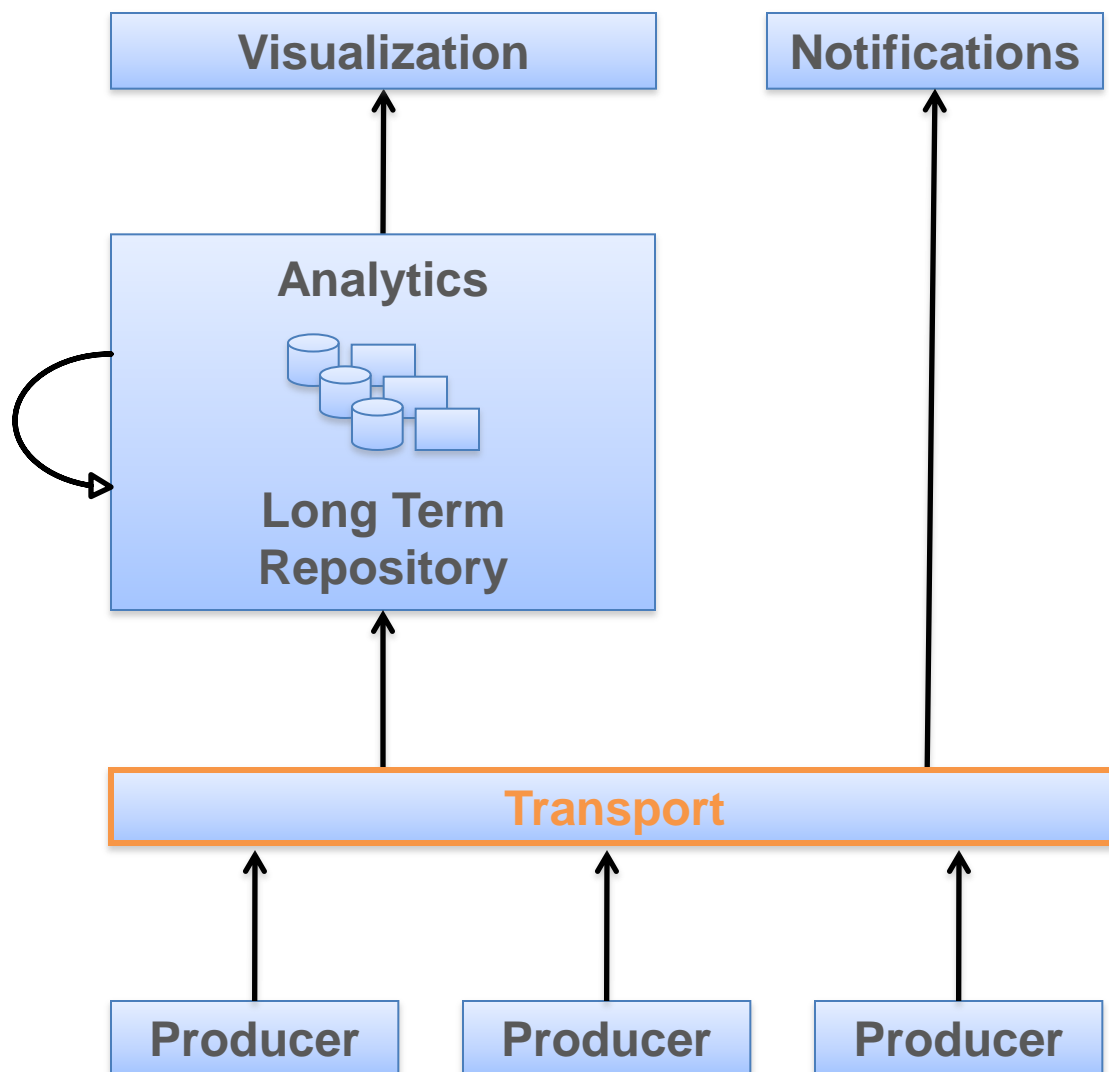
2013

- Core monitoring team (in IT/CF since March 2013)
- Close collaboration with other IT groups
- Implementation and deployment of solutions

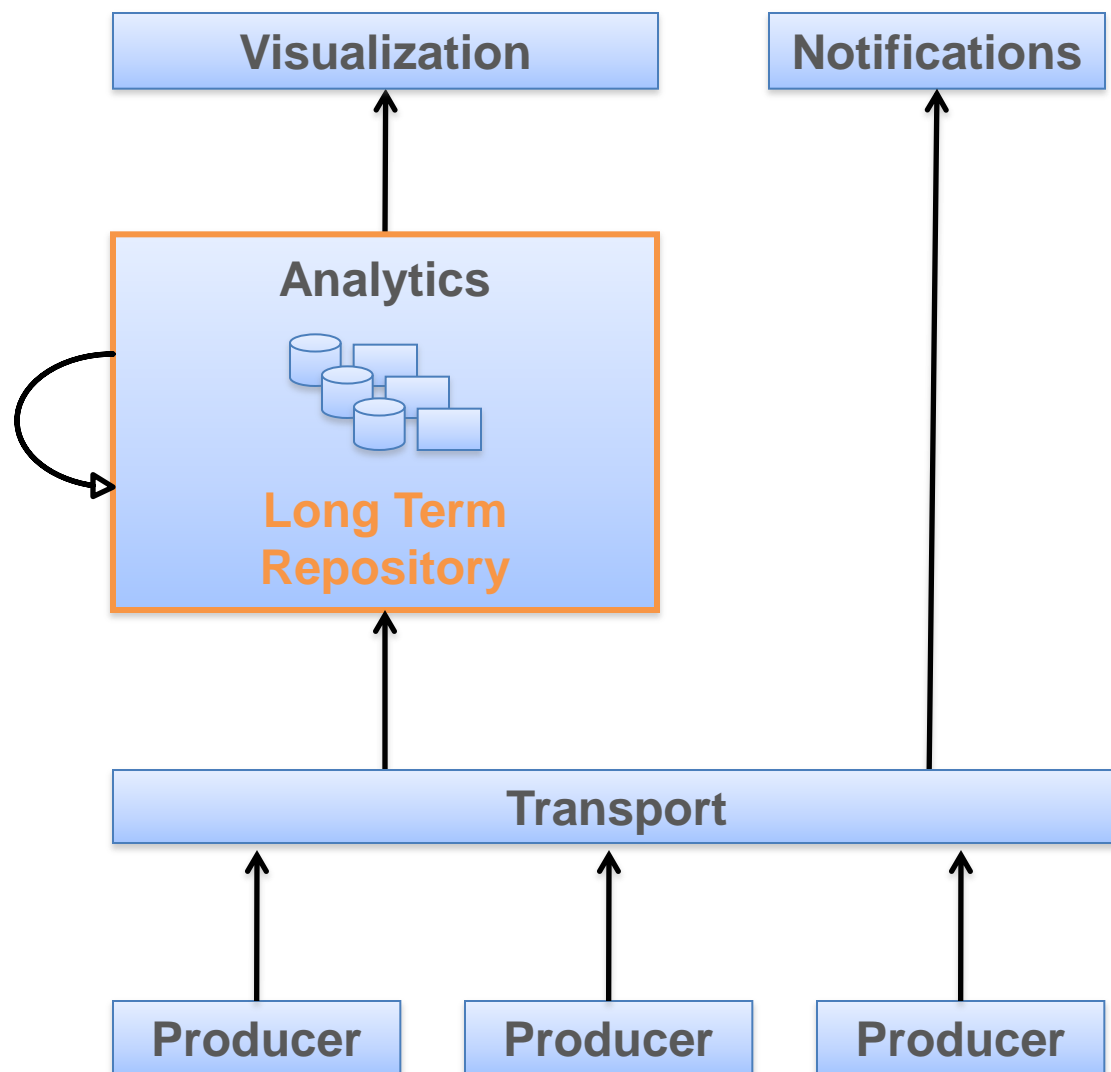




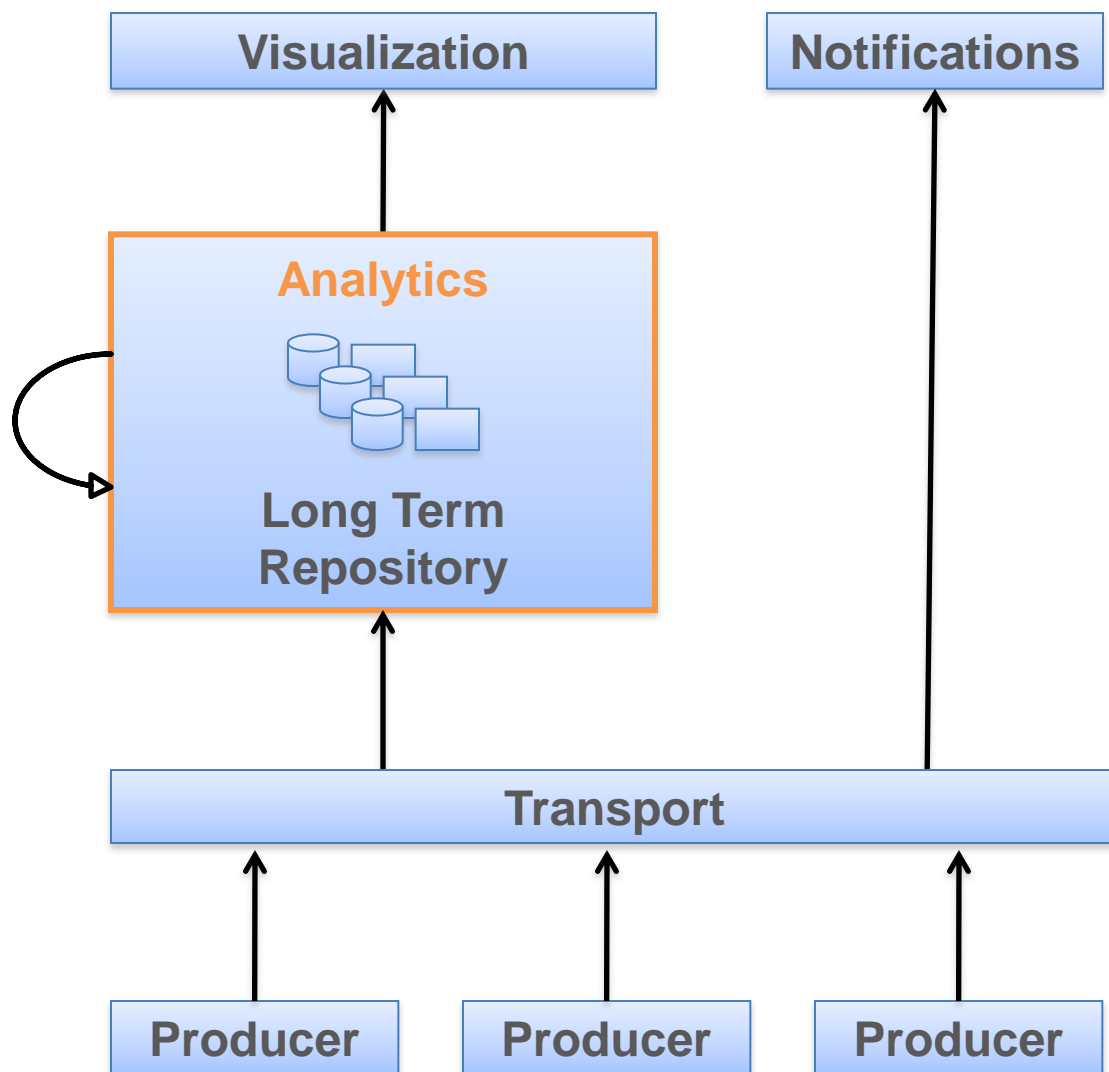
- Integrate data from multiple producers
- Support legacy producers



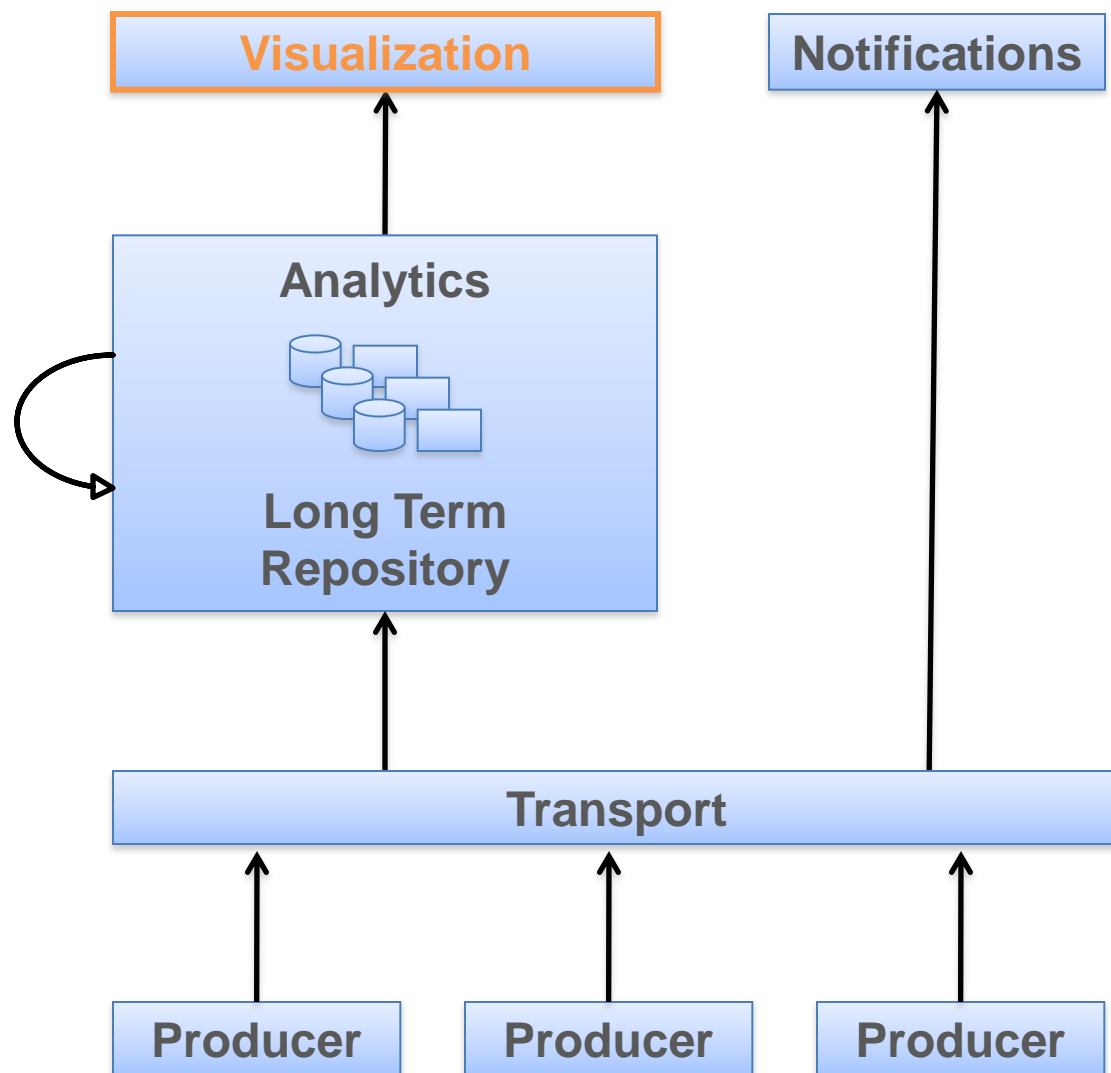
- Scalable transport to collect monitoring and operations data
- Easy integration with providers and consumers



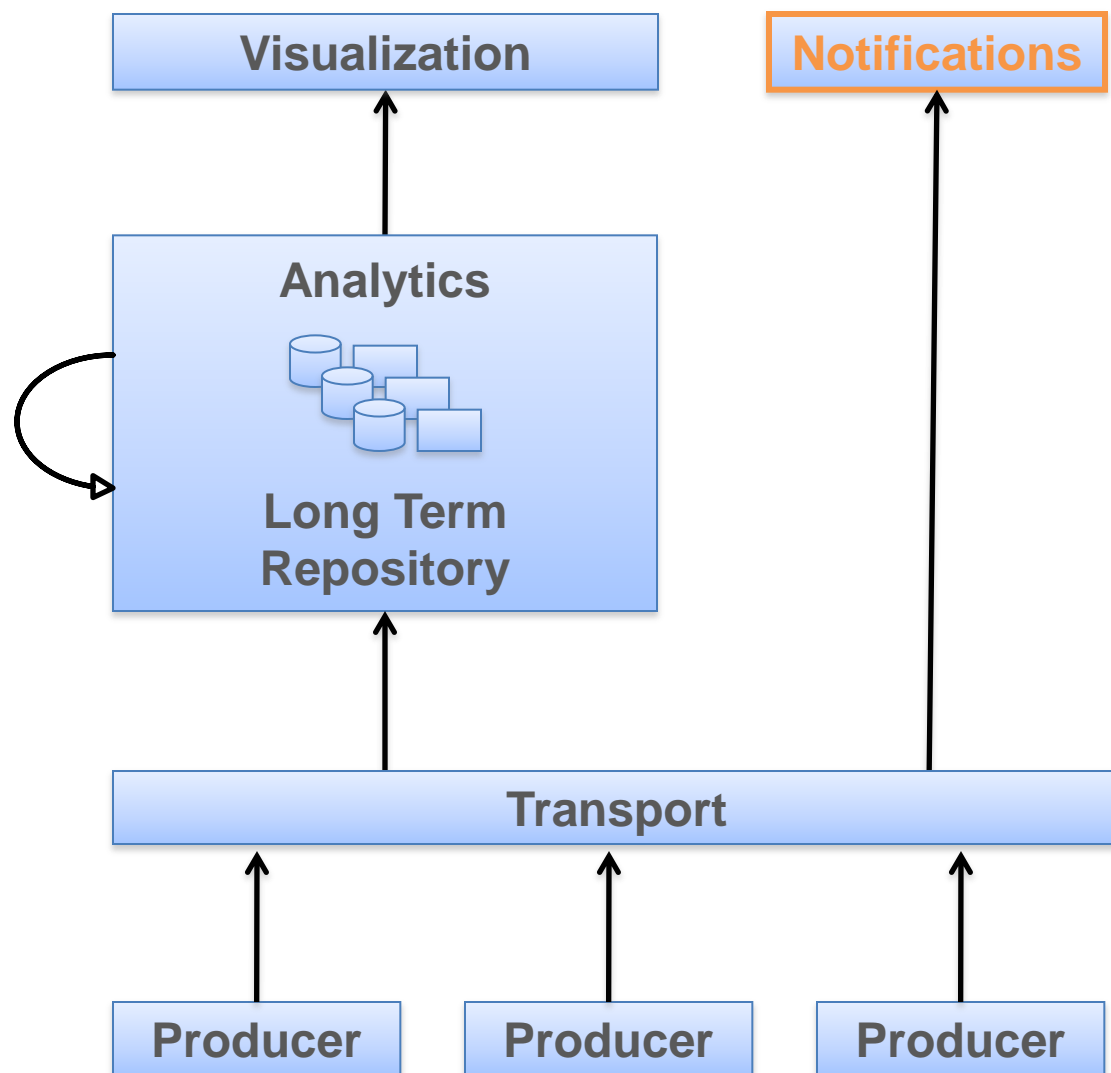
- Long term archival of collected data
- Offline processing of collected data
- Allow future data replay to other tools



- Limited data retention
- Real-time queries
- Easy to deploy and horizontal scalable



- Dynamic creation of dashboards
- Tailored for global and application specific views
- User friendly



- Quick and reliable delivery of alarms
- Delivery of notifications via multiple channels

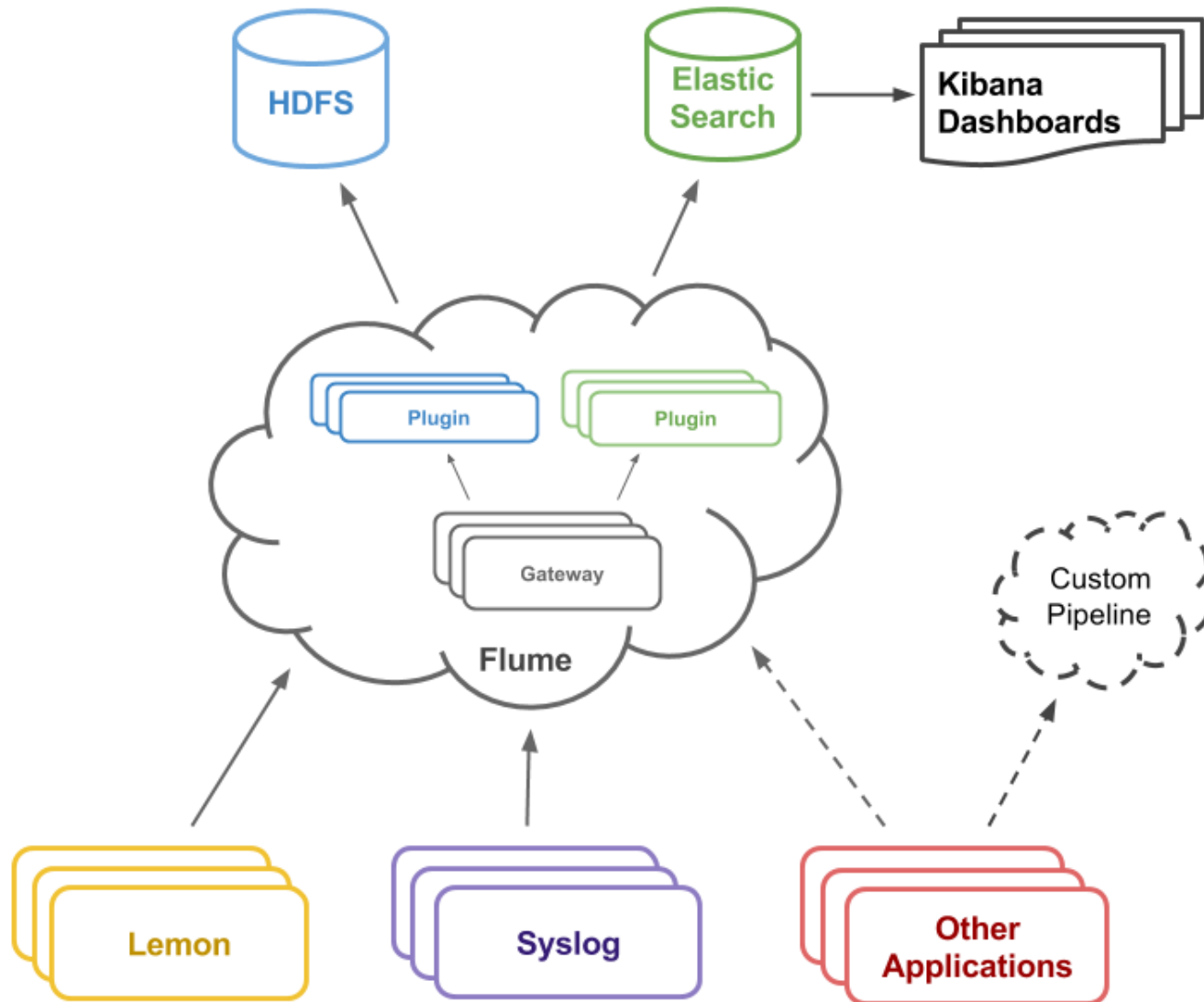
Adopt open source tools

- For each architecture block look outside for solutions
- Large adoption and strong community support
- Fast to adopt, test, and deliver (and easily replaceable)

Integrate with new CERN infrastructure

- AI project, OpenStack, Puppet, Roger, etc.

Focus on simple adoption (e.g. puppet modules)



Distributed service for collecting large amounts of data

- Robust and fault tolerant
- Horizontally scalable, multi-tier deployment
- Many ready to be used input and output plugins
 - Avro, Thrift, JMS, Syslog, HTTP, ES, HDFS, Custom, ...
- Java based, Apache license

Feedback

- Needs tuning to correctly size flume tiers
- Available plugins saved a lot of time



Distributed framework for large data sets processing
Distributed filesystem designed for commodity HW

- Suitable for applications with large data sets
- Cluster provided by other IT group (DSS)
- Data stored by cluster (might be revised)
- Daily jobs to aggregate data by month

Feedback

- Large analytics potential to explore
- Reliable external long term repository



Distributed RESTful search and analytics engine

- Real time acquisition, data is indexed in real time
- Automatically balanced shards and replicas
- Schema free, document oriented (JSON)
- Based on Lucene (full-featured IR library)

Feedback

- Easy to deploy and manage
- Robust and fast API
- Powerful query language (DSL)



Visualize time-stamped data from Elasticsearch

- Designed to analyse log, perfectly fits time stamped data
- No code, point & click to build your own dashboard
- Built with AngularJS (from google)

•

Feedback

- Easy to install and configure
- Very cool user interface
- Fits many use cases (e.g. text, metrics)
- Still limited feature set, but active growing community



General Notifications Infrastructure

- Based on ActiveMQ messaging broker
- Same monitoring producers (e.g. lemon)
- Multiple consumers
 - Snow consumer for ticket creation
 - Dashboard consumer for web application
 - No contact for node heartbeat checking



Feedback

- Efficient, direct ticket routing
- Flexible, easy to add more consumers (e.g. SMS)

Producers

- From all puppet-based data centre nodes
- Central monitoring
 - Computing Facilities ([lemon](#), [syslog](#))
- Application monitoring
 - OS & Infrastructure Services ([openstack](#))
 - Platform & Engineering Services ([batch lsf](#))
 - Security Team ([netlog](#), [snoopy](#))
 - Databases Services ([web apps](#))
 - Data and Storage Services ([castor logs](#))
 - Support for Distributed Computing ([wlcg monitoring](#))

Flume

- 10 aggregator nodes, 5 nodes to HDFS + 5 nodes to ES


HDFS

- ~500 TB cluster, 1.8 TB collected since mid July 2013

ElasticSearch

- 1 master node, 1 search node, 8 data nodes
- 90 days TTL, 10 shards/index, 2 replicas/shards
- Running ElasticSearch Kibana plugin



Lemon  Kibana 3 milestone 3


Controls 

5m 15m 1h **6h** 12h 24h 2d 7d 30d

Relative | Absolute | Since | Auto-refresh

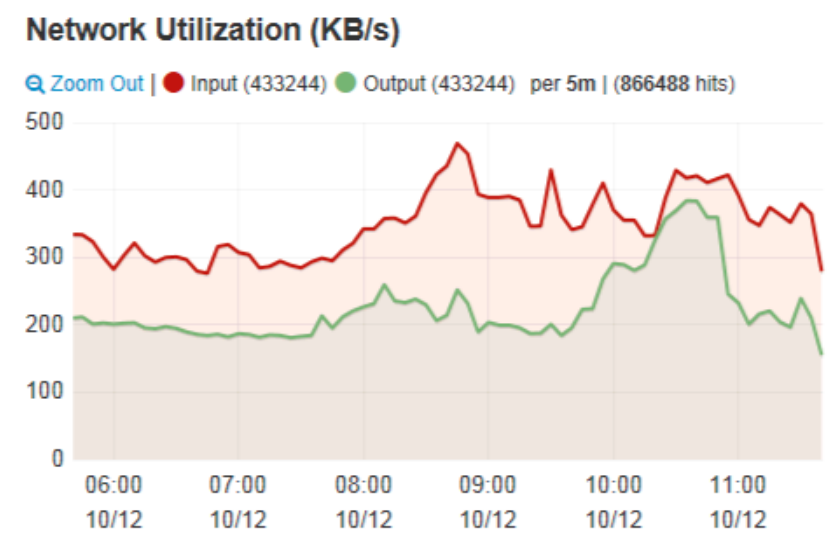
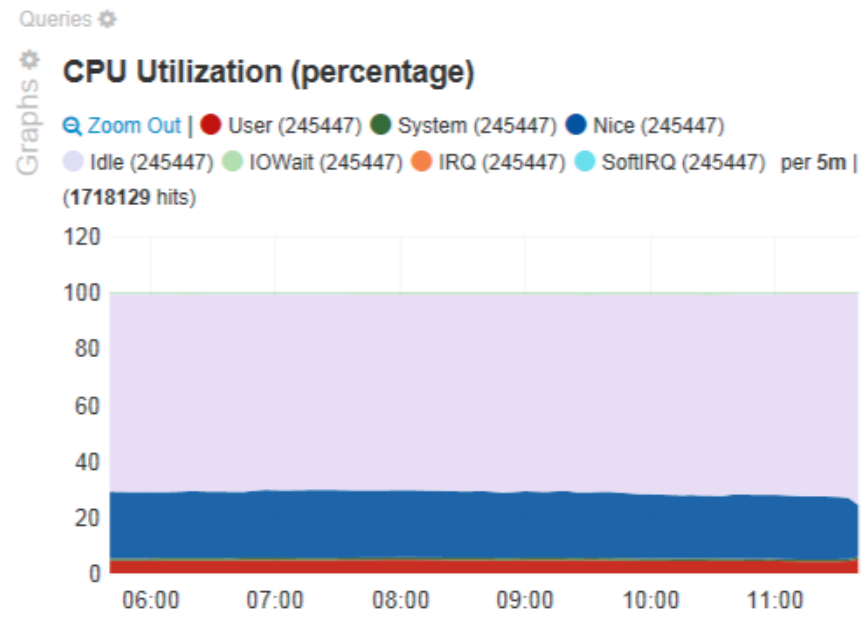
Dashboard Control



Filters 

<p><u>field must</u>   </p> <p>field : @fields.submitter_cluster query : "aimon/flume/gw/dev"</p>	<p><u>field must</u>   </p> <p>field : @fields.environment query : "production"</p>	<p><u>field must</u>   </p> <p>field : @fields.entity query : ""</p>	<p><u>time must</u>   </p> <p>field : @timestamp from : "2013-10-12T03:40:35.374Z" to : "2013-10-12T09:40:35.374Z"</p>
--	--	---	---



Several interesting **technologies tested** and deployed

Full workflow deployed for concrete monitoring needs

Verified by **different monitoring** producers

Improve monitoring tools under a **common effort**

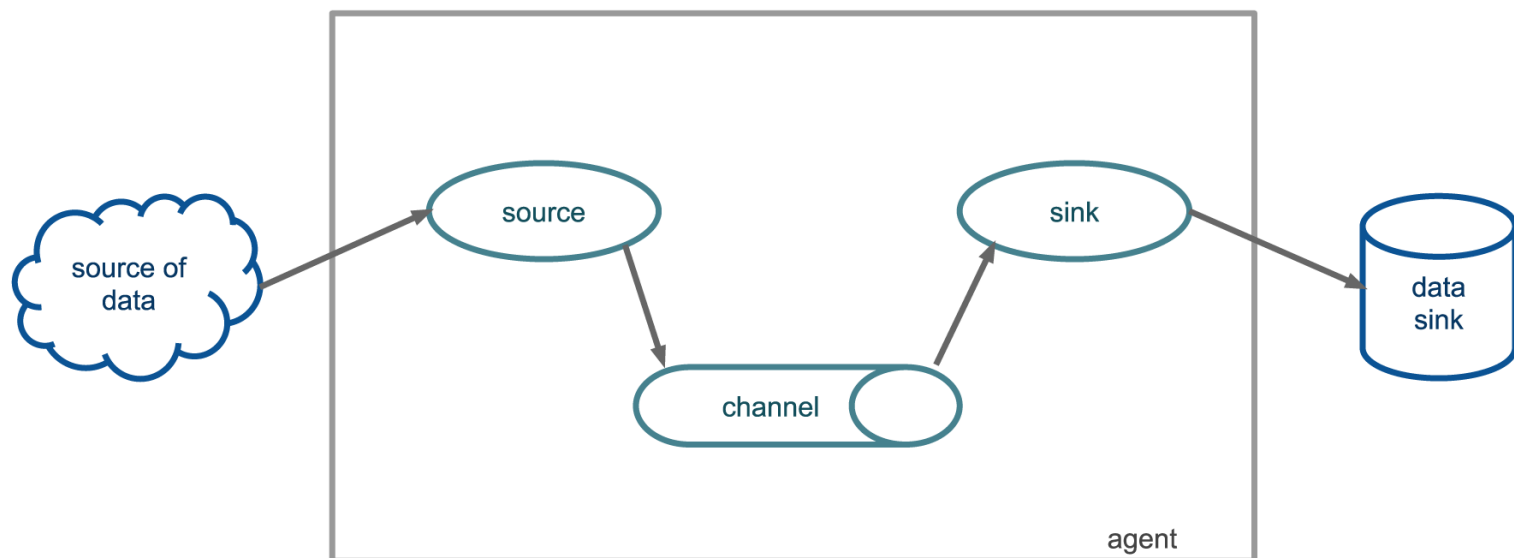
Thank you !!

Questions ??

itmon-team@cern.ch

<http://cern.ch/itmon>

Backup Slides



Sources

- Avro, Thrift, JMS, Syslog, HTTP, Custom, ...

Sinks

- Avro, Thrift, ES, Hadoop HDFS, Custom, ...

