XROOTD AND FEDERATED STORAGE MONITORING CURRENT STATUS AND ISSUES

A.Petrosyan, D.Oleynik, J.Andreeva

Creating federated data stores for the LHC
CC-IN2P3, Lyon, November 22, 2011

Outline

- Historical overview
- XRootD monitoring background
- Use cases and user categories
- Monitoring architecture of
  - Site
  - Federation
  - VO&WLCG
- Issues
- Conclusions

Historical overview

Tier 3 monitoring task force

Software suit to enable T3 sites monitoring, two layers

- T3mon site
  - Site infrastructure monitoring
  - Batch, storage software monitoring
  - Data representation via Ganglia monitoring system
- T3mon global
  - Monitor T3 site activity on global layer: data popularity, transfers
  - Data representation via Dashboard

Special part of project: XRootD monitoring for site, federation monitoring as a part of T3mon global

XRootD monitoring background

XRootD instrumented for monitoring

- Summary stream
  - Overview of site health, easy to configure and represent
- Detailed stream
  - Provides information about each single operation (authorization, staging, IO, etc.)
  - Complicated to use: decoding and aggregation needed

Every instance (server, redirector) uses same data transportation (UDP) and representation technology

Combination of info from summary and detailed allows to feed site and federation monitoring

XRootD monitoring metrics

- Following metrics provided and extracted:
  - File
    - Open/Close, Transferred volume, Read/Write
  - Username
  - Application
  - Trace
  - Client IP/name
  - Server IP/Name
- This metrics can be used as initial aggregation patterns

Use cases

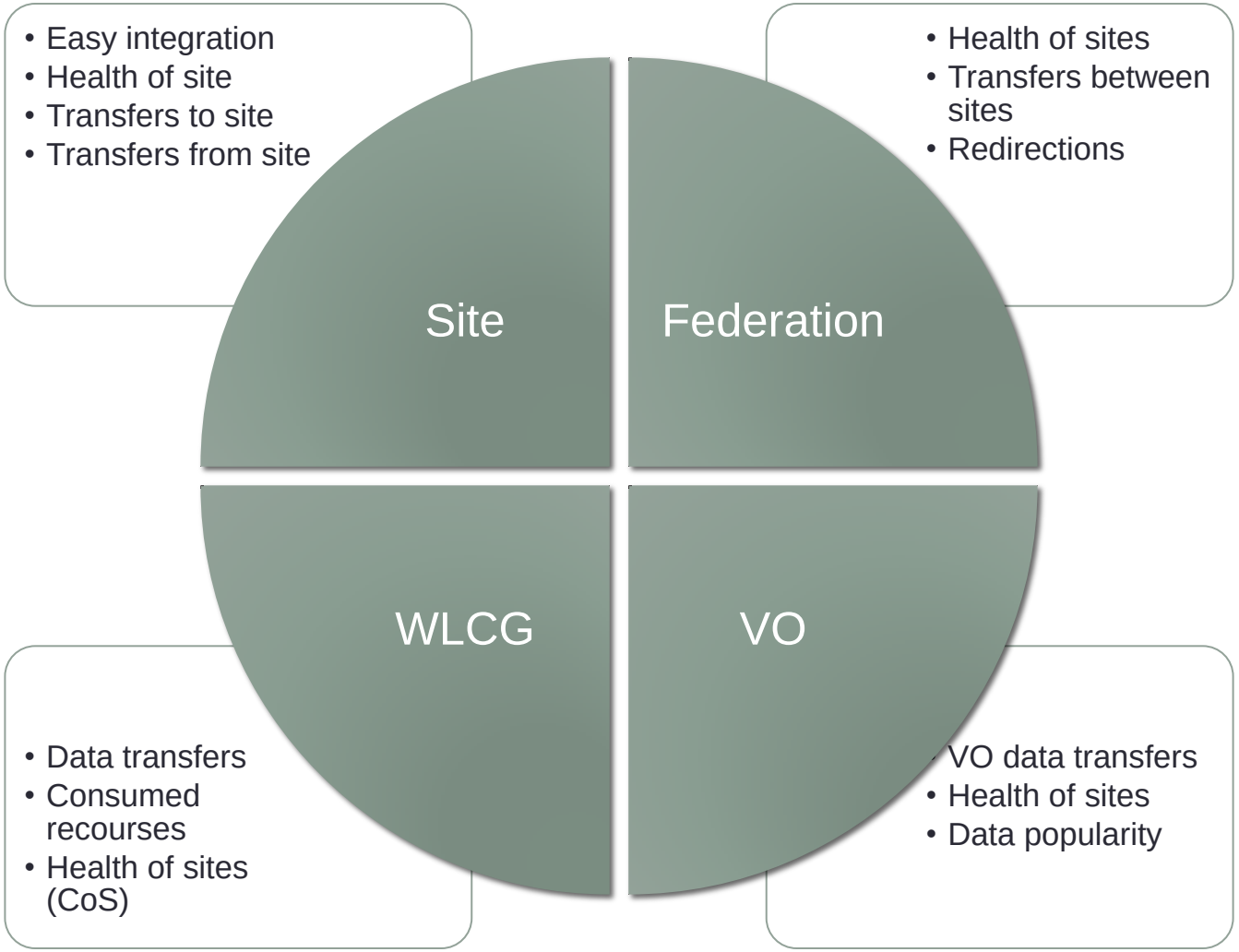
Monitoring of the XRootD transfers and data access are required at various levels:

- Site (administrators)
- Federation (federation administrators)
- VO (VO managers)
- WLCG

Different data should be presented on every level

Different requirement for different user categories

Monitoring consumers & requirements

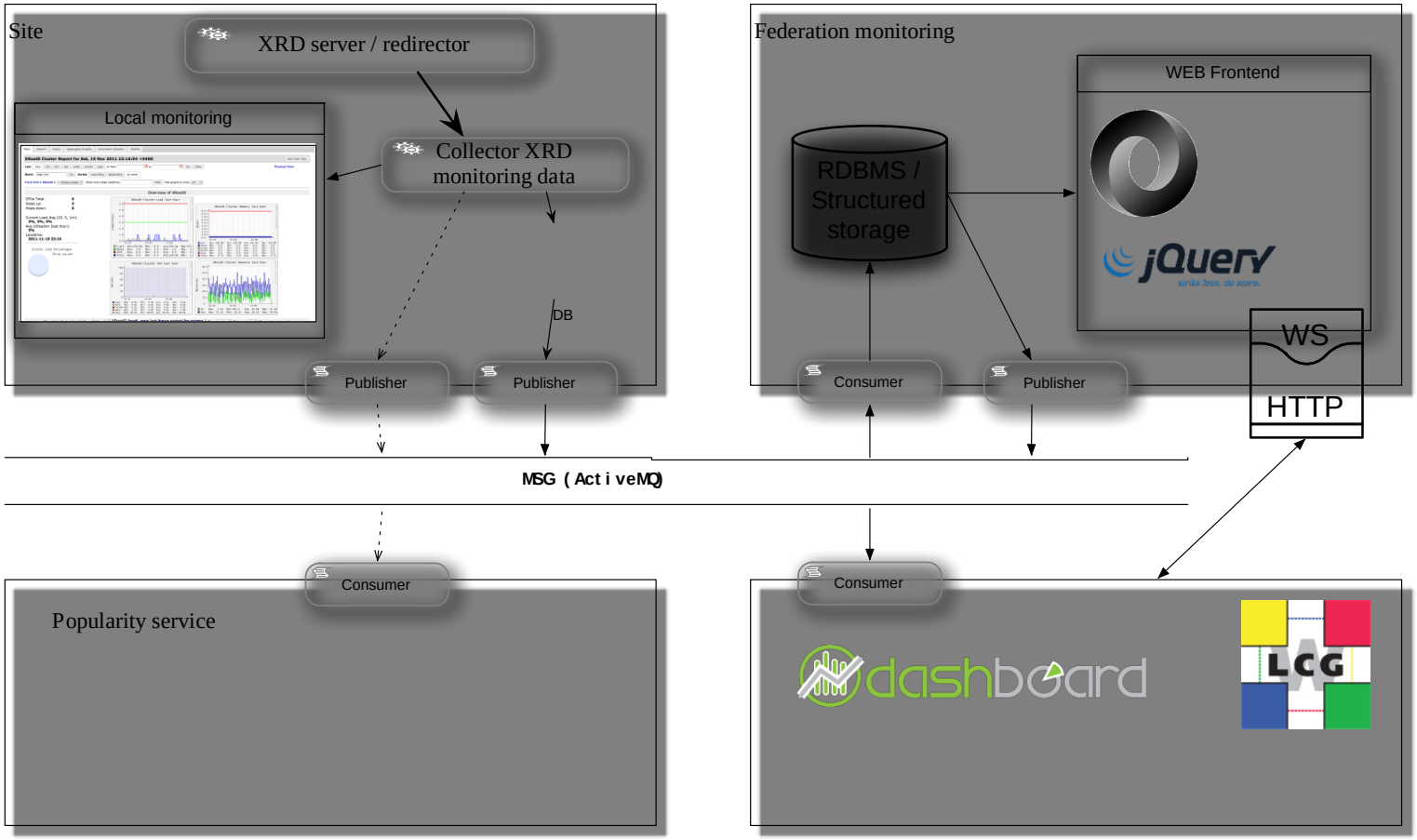


Architecture 1/2

- Information collected from servers/redirectors
- All top level metrics are calculated or aggregated based on data coming from servers/redirectors

- *At the moment we do not have good estimation of how much data has to be handled on the federation level (depends on number of sites, type of messages, granularity, etc.), but it is clear that* ***solution should be scalable***

Architecture 2/2

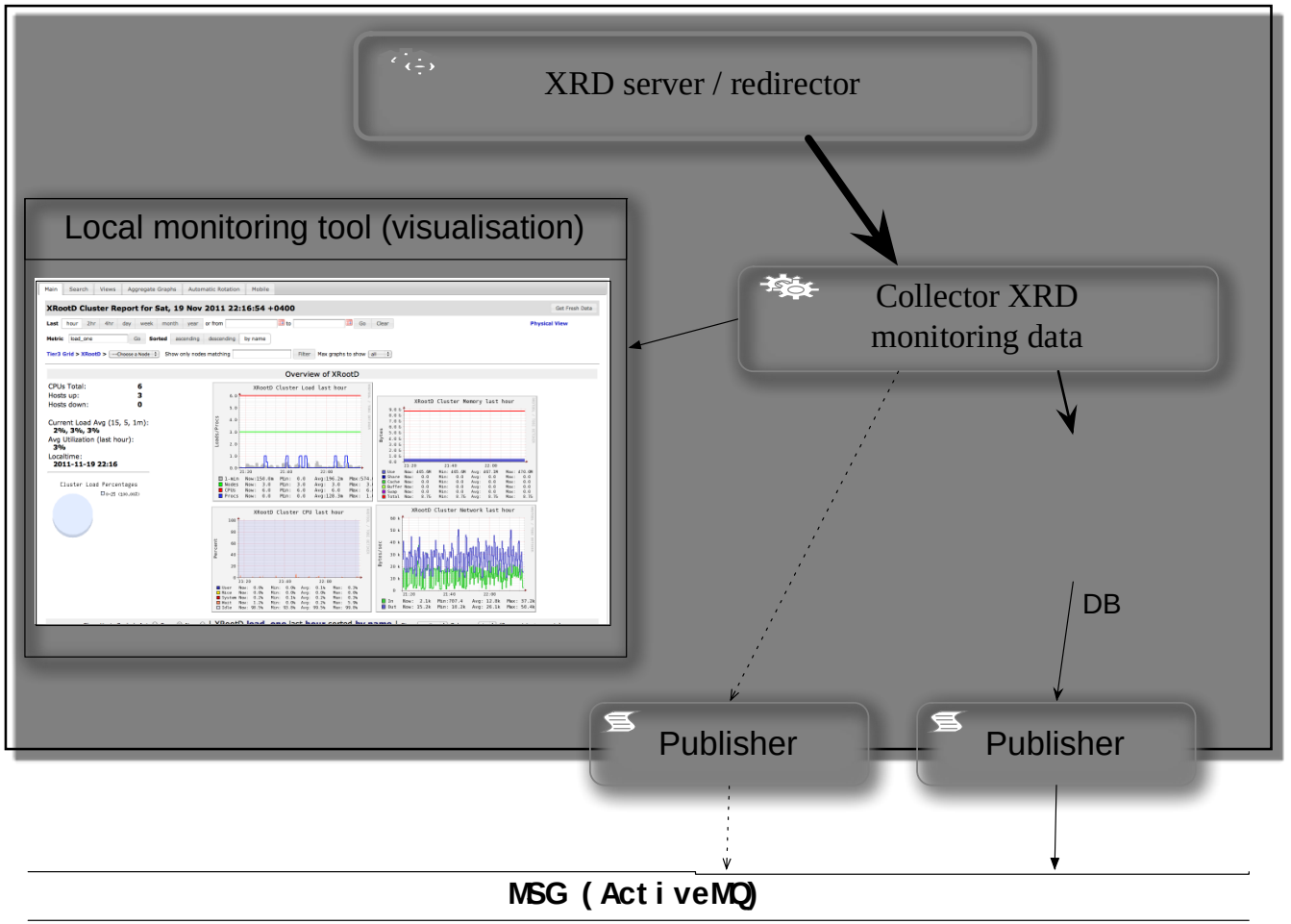


Architecture. Site monitoring 1/2

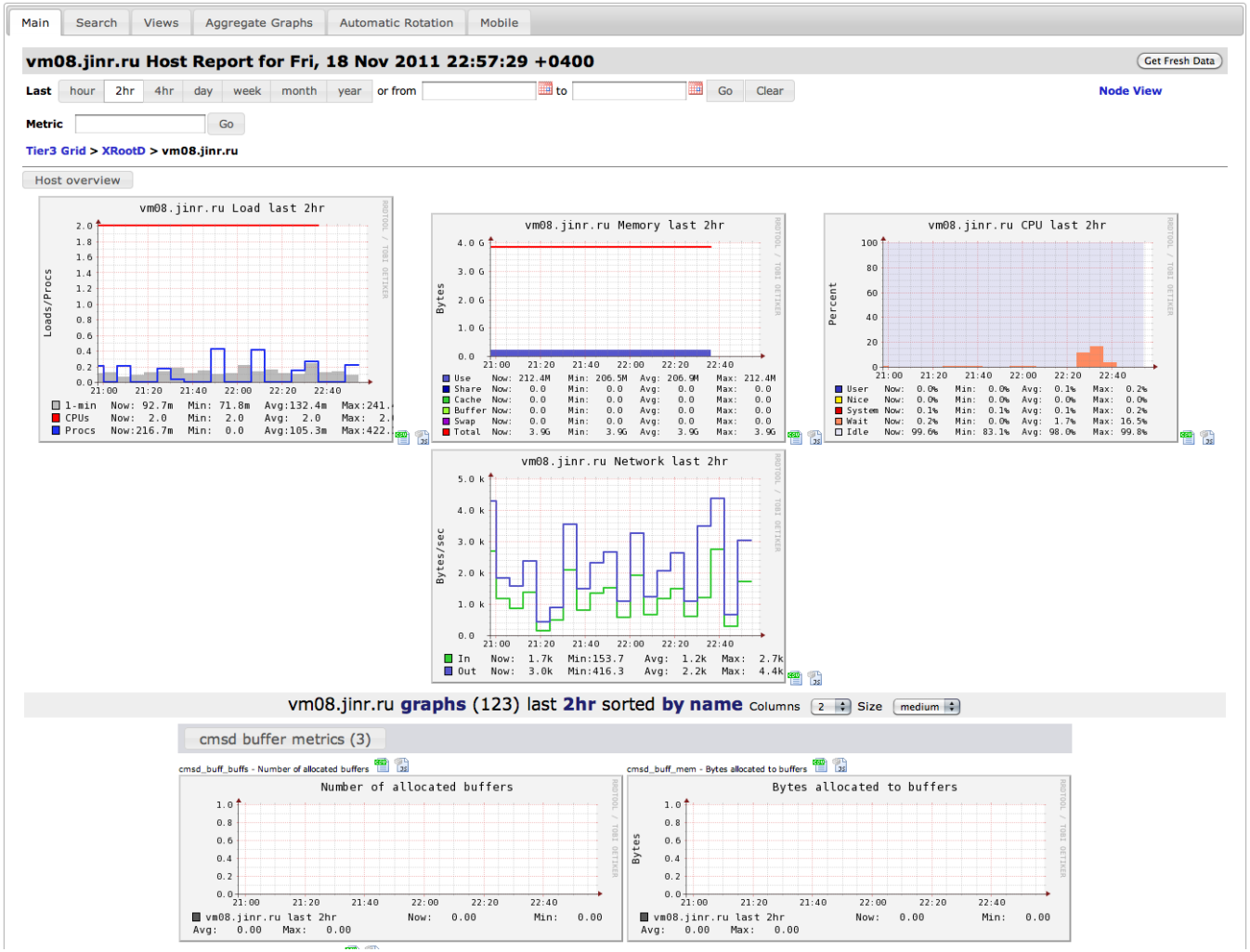
- Summary metrics are collected and represented through local fabric monitoring system
  - Ganglia solution is already implemented
- Python parser for extracting metrics from detailed stream is ready
- Plugin system for intercommunication with different backends:
  - PostgreSQL for T3 site
  - MSG for CMS popularity

  Both solutions have been implemented and are being tested
- Transfer info available only via detailed stream
  - Lot of data
  - Scalability & load tests needed

Architecture. Site monitoring 2/2



XRootD monitoring through Ganglia



http://vm01.jinr.ru/ganglia/?c=XRootD&h=vm08.jinr.ru

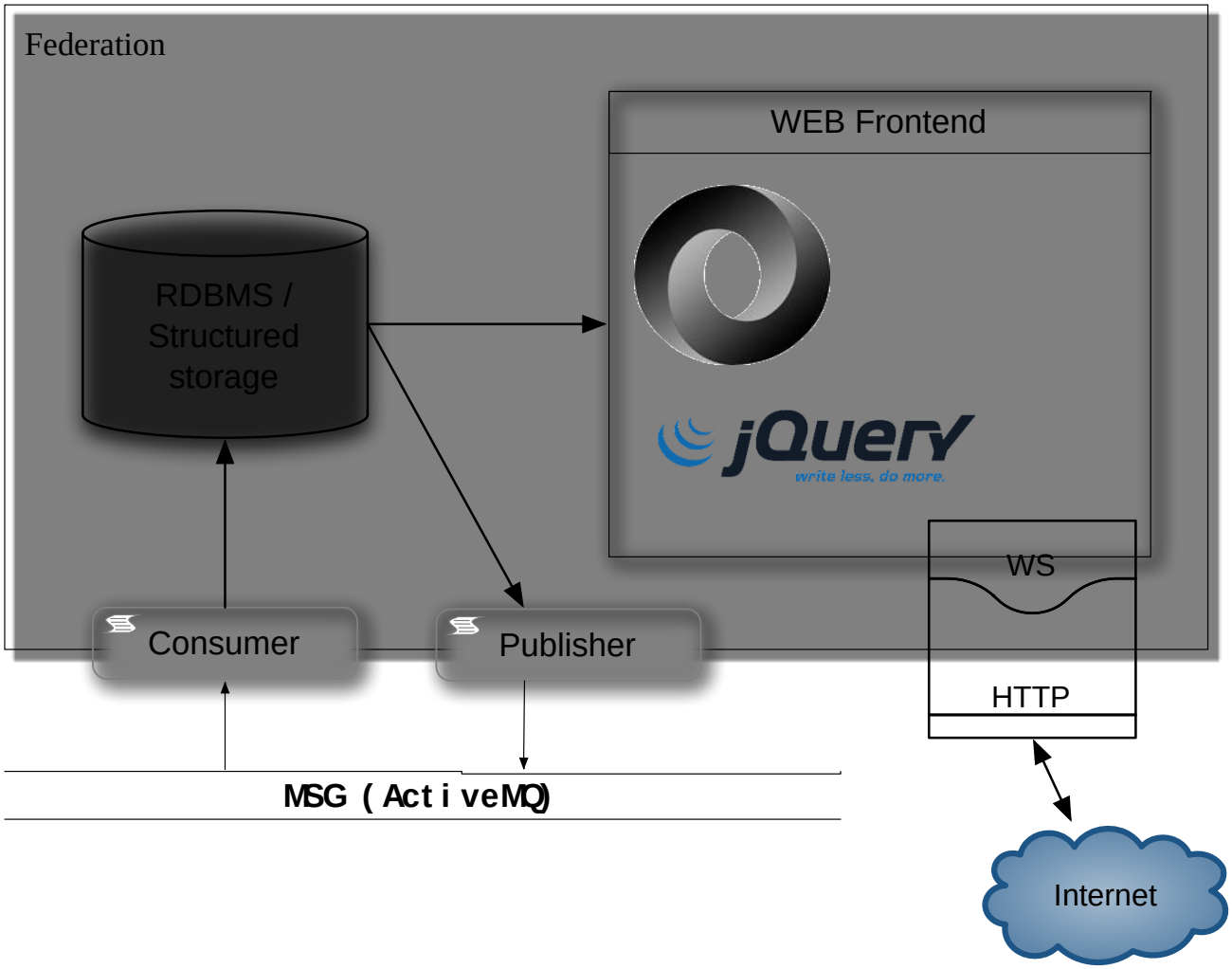
Architecture. Federation monitoring 1/3

- Collecting filtered list of metrics at the site (servers/redirectors)
  - All data generated by servers at site level
  - Each site must be declared in the WLCG&VO topology
- Transmitting data for aggregation (MSG/ActiveMQ) to central storage
  - Scalability – queues with asynchronous operations and multiple consumers support

Architecture. Federation monitoring 2/3

- Central storage/Data aggregation alternatives:
  - RDBMS
    - Data normalization, logical complexity, high cost of any data structure change
  - Structured data storage (Hbase)
    - Scalable, flexible, fast access to data
- Presentation & Visualization
  - JSON
    - Easy data integration within application and between web applications
  - Django, jQuery
    - Python, AJAX

Architecture. Federation monitoring 3/3

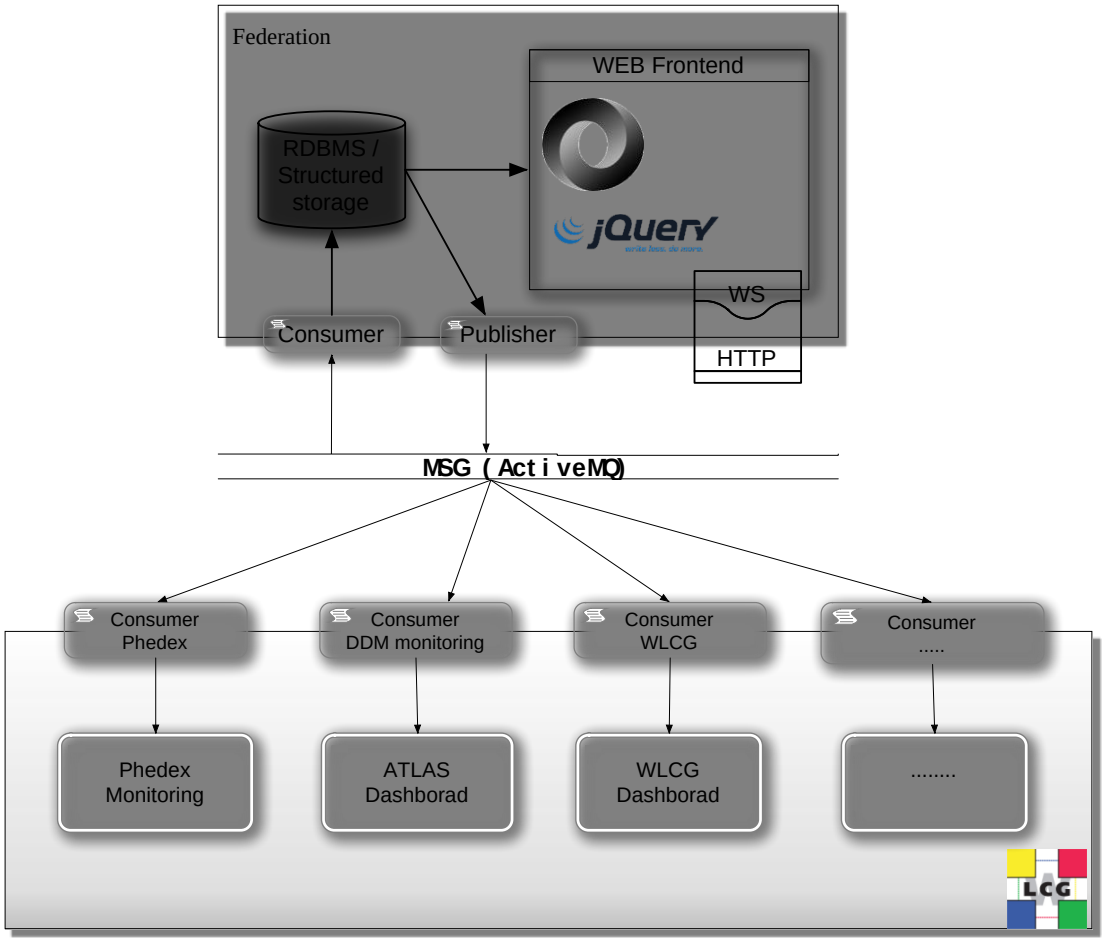


Architecture. VO&WLCG monitoring

- Data collected from XRootD federations
  - Federation must be declared in the WLCG&VO topology
  - Applications must provide VO mark of user for transfer at the site level
- Transfer of monitoring data

  (MSG/ActiveMQ/WebService)
- Integration with Dashboard
  - Data formats to be agreed

Architecture. VO monitoring



Issues

- Federation topology
  - Site and federation declaration in the WLCG&VO topology
- Metrics
  - Information is available, list of metrics for federation/VO/WLCG level to be agreed
- Initial aggregation patterns

Conclusions

- Monitoring architecture for XRootD data transfers on the local and global levels is being prototyped
- Collection and publishing of basic metrics at the site level is implemented and is being tested
- Technology and infrastructure for data transmission is in place
- Technology for data aggregation and visualization is defined