



XRootD Monitoring Report

A.Beche

D.Giordano



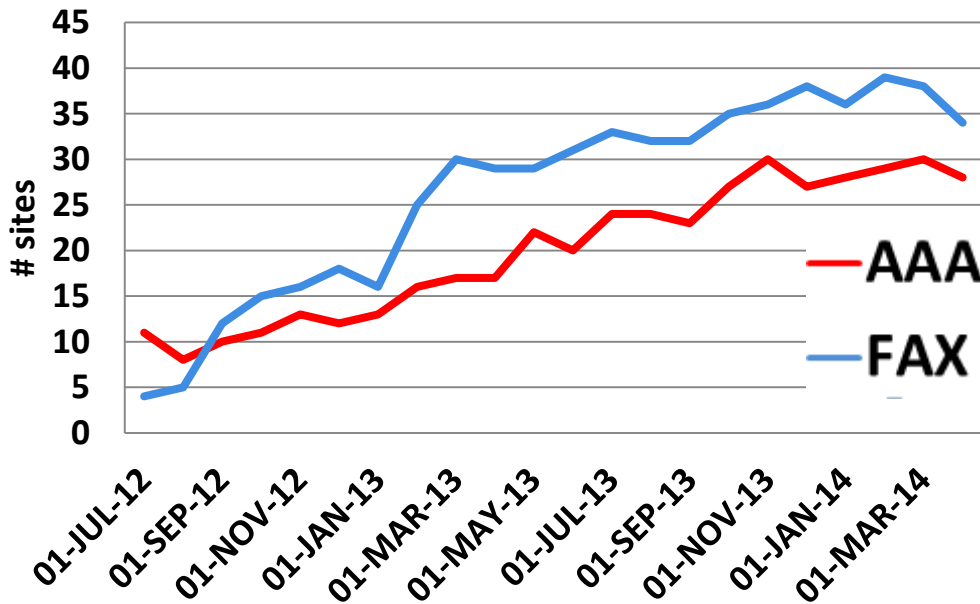
Outlines

- Talk 1: XRootD Monitoring Dashboard
 - Context
 - Dataflow and deployment model
 - Database: storage & aggregation
 - User interface & use cases
 - Open issues & future work
 - Summary
- Talk 2: Beyond XRootD monitoring
 - HTTP/WebDAV integration

XRootD federation monitoring

- Activity started during summer 2012
 - 4 sites for FAX, 11 for AAA

Number of sites reporting



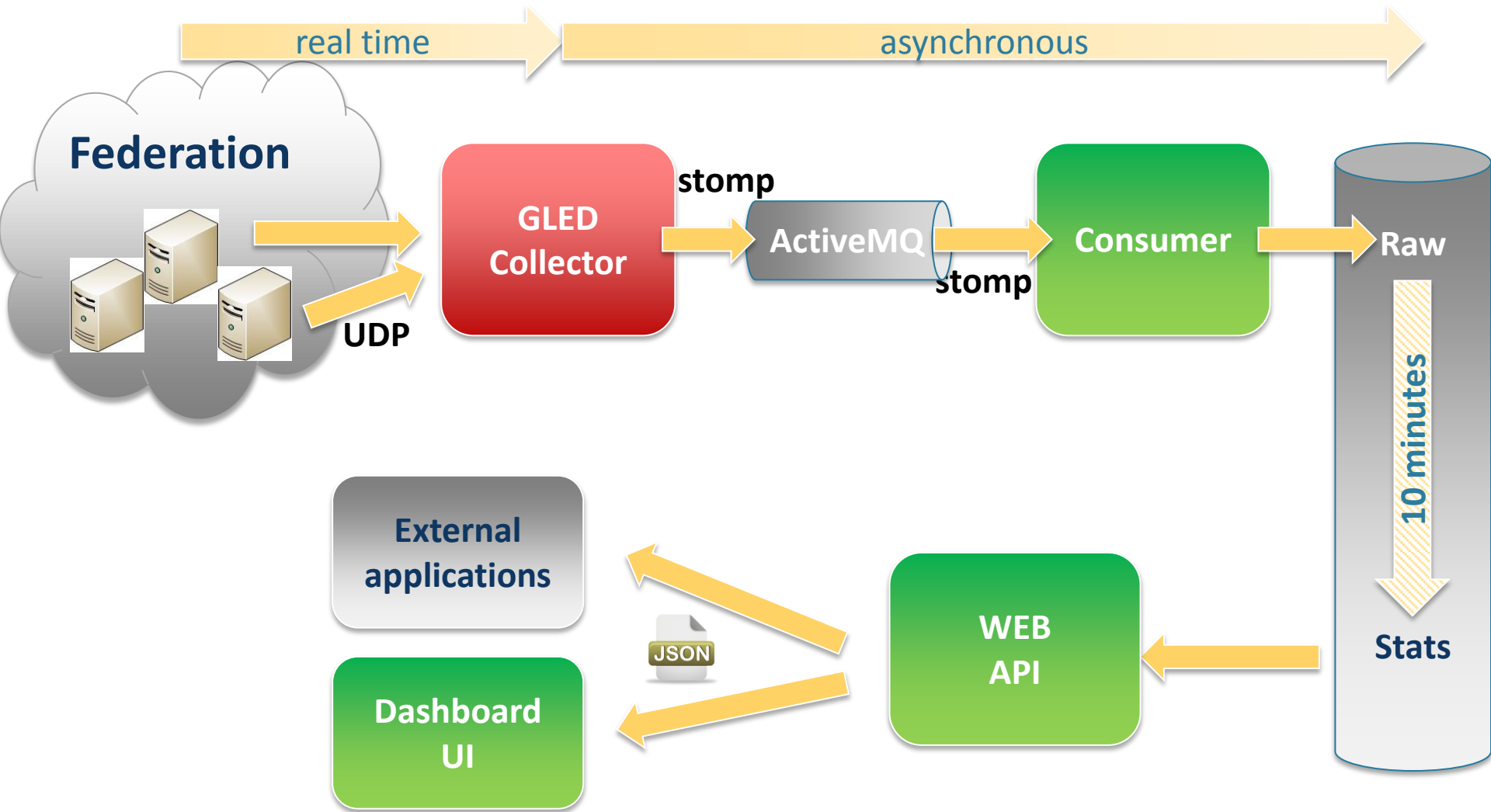
Monitoring data increased accordingly

	July 2012	March 2014
AAA	606k	43M
FAX	15k	22M

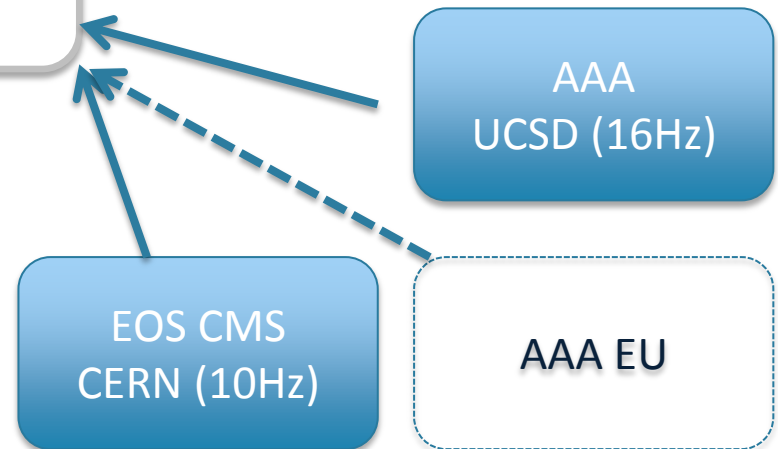
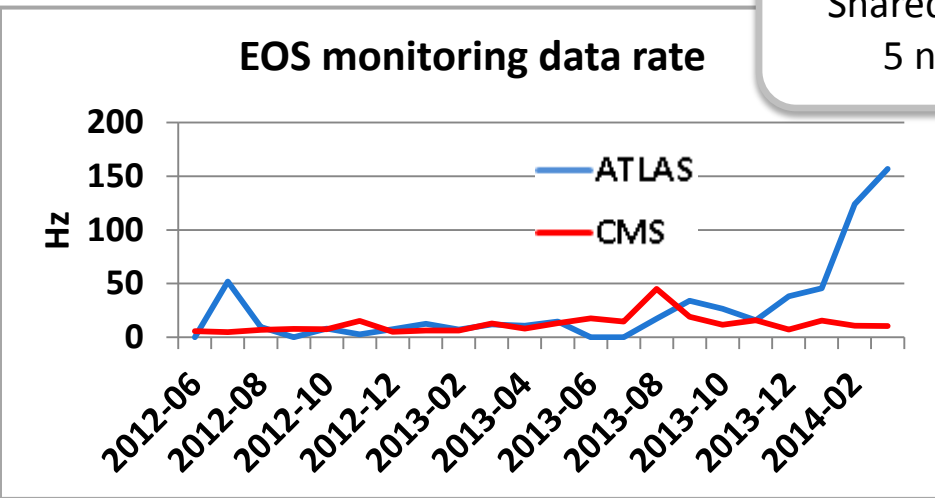
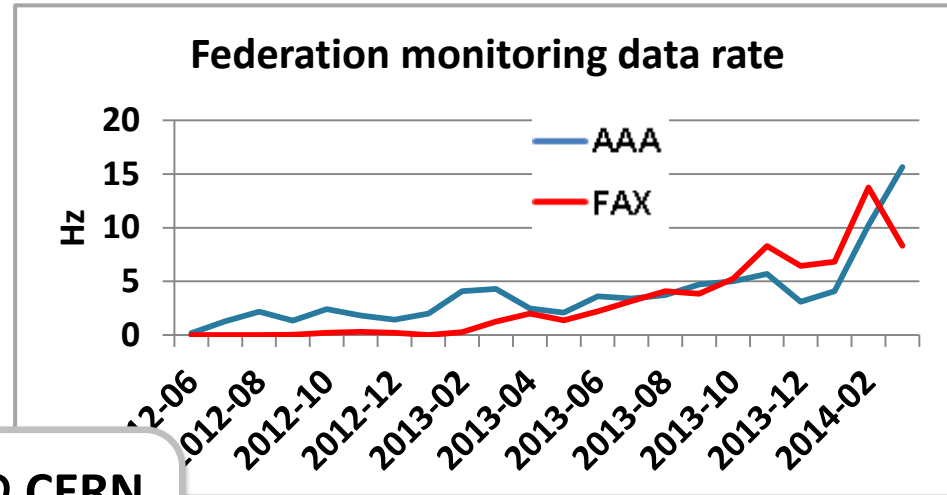
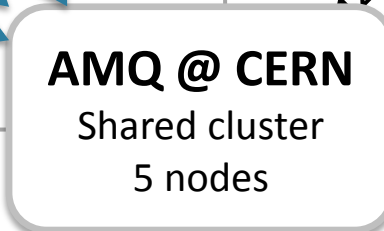
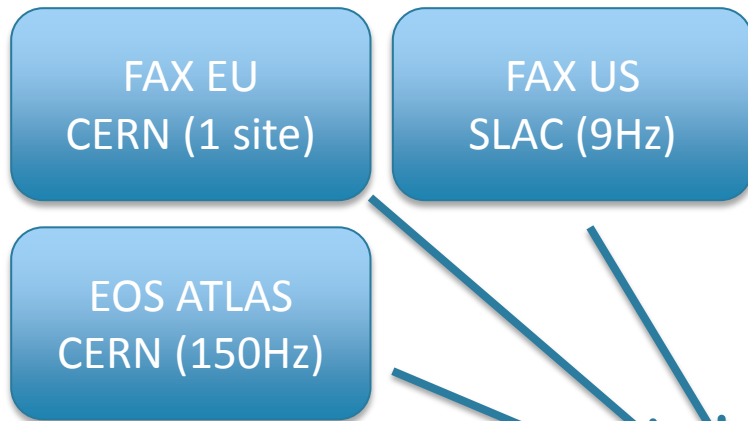
Why monitoring ?

- Understand data flows to estimate data traffic
- Provide information for efficient operations
- Identify access patterns and propose data placement strategies

XRootD monitoring dataflow



GLEED Deployment model

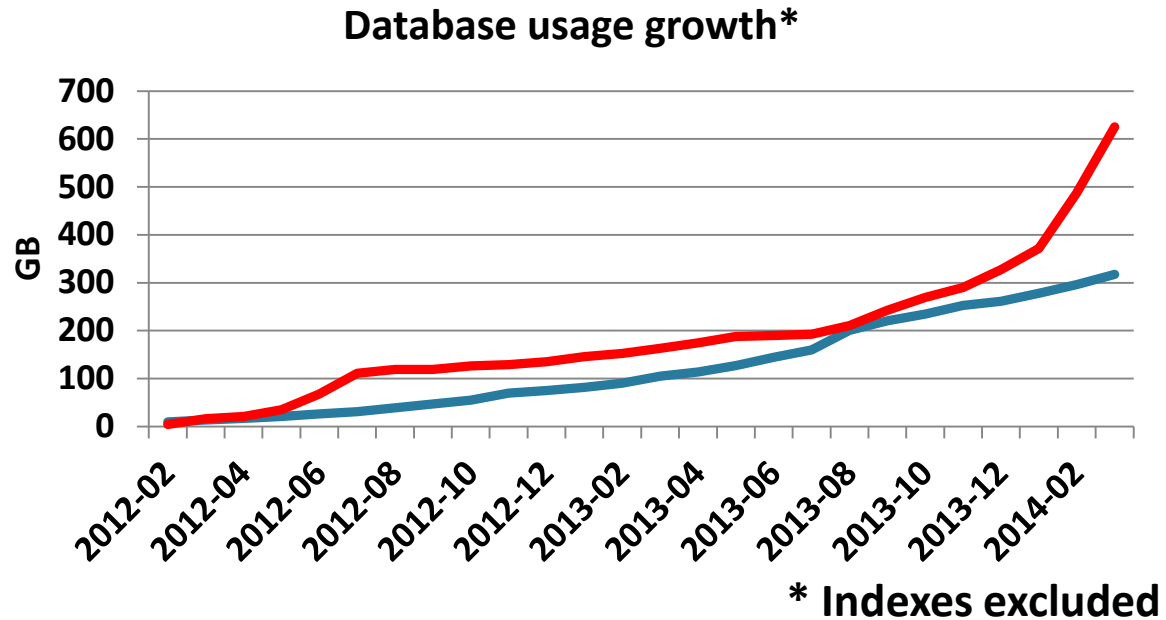
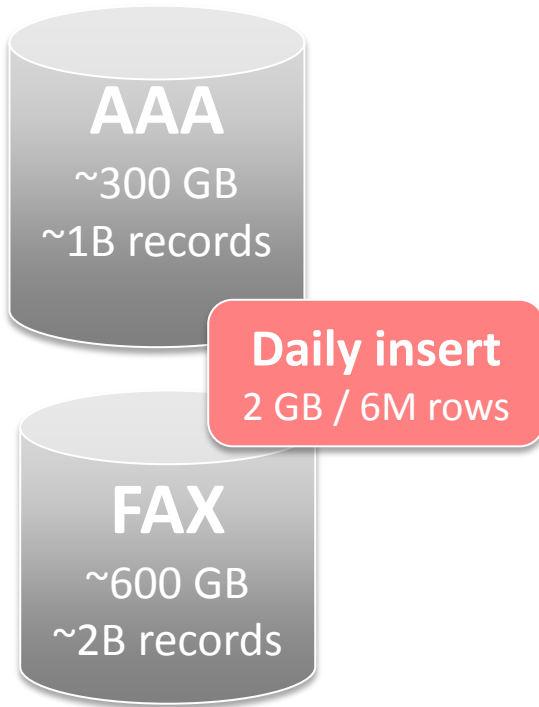


Consolidated dataflow

- Two usage of these raw data:
 - Dashboard monitoring
 - XRootD popularity

- Now share the same database:
 - Storage optimization
 - Consistency guaranteed

Database



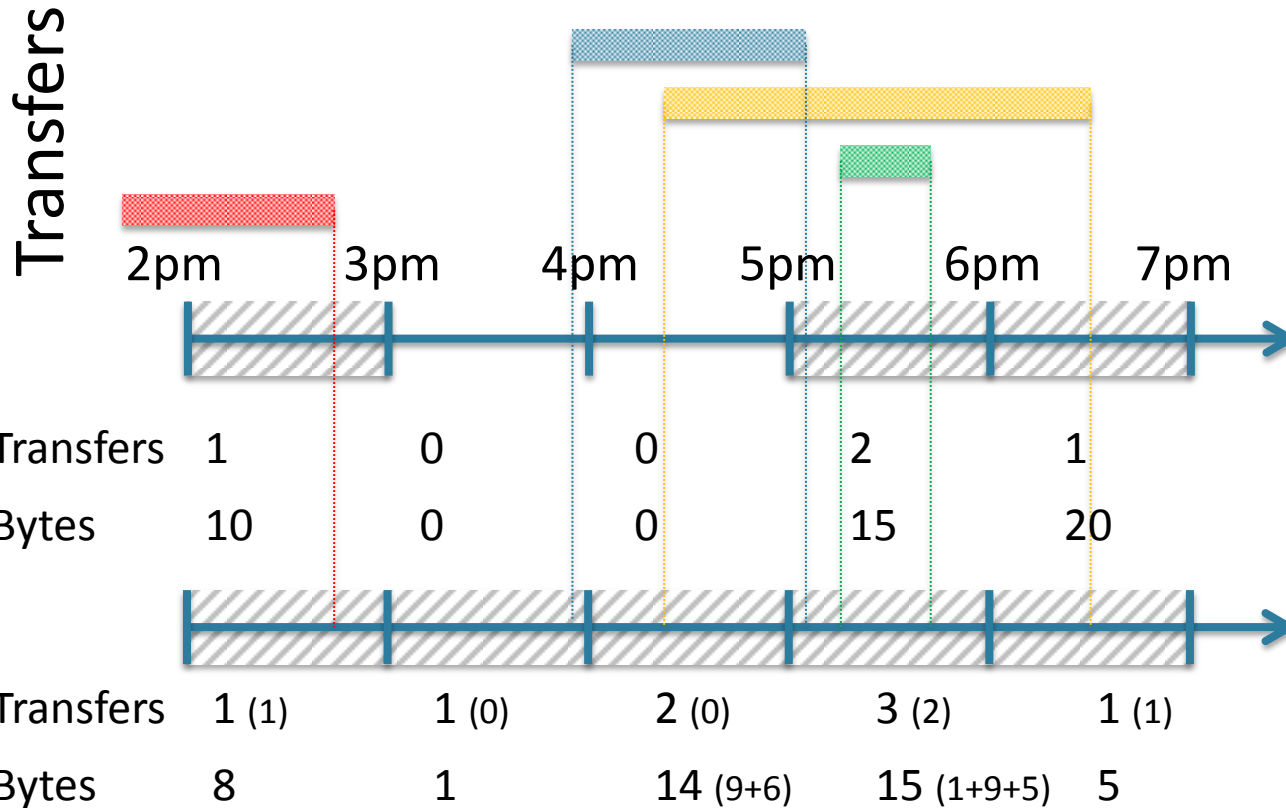
■ Storage

- Raw, statistics, metadata
- Tables daily **partitioned**, no global indexes

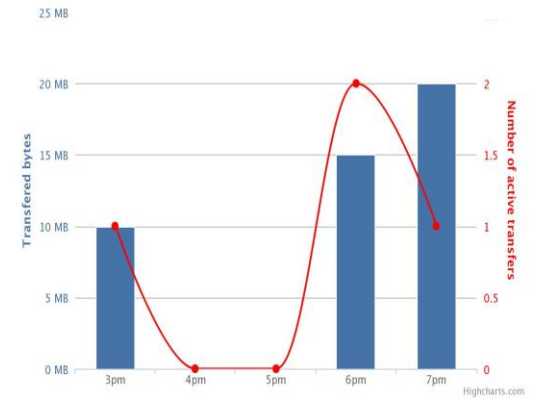
Database

- Raw data aggregation:
 - Done using PL/SQL procedures
 - Events are **unordered**
 - **Stateless: Full re-computation** of touched bins each time
 - Compute stats from raw data in 10 min bins
 - Aggregate 10 min stats in daily bins

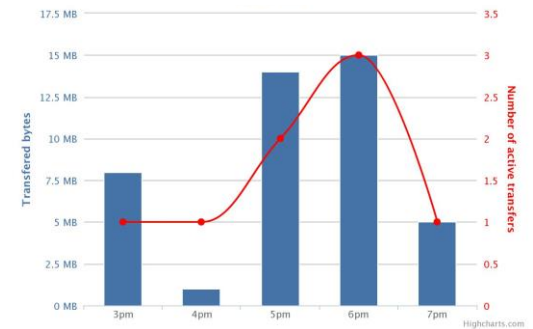
Aggregation methods



Easy method



Adopted method



Visualization Interface



FAX Monitoring (RC.34)

latest statistics update: 2013-11-27T09:10:00.03819Z

TRANSFER PLOTS (2013-11-27 05:20 to 2013-11-27 09:20 UTC SLIDING)

PLOT: GROUPING ▾ TYPE ▾ SERIES ▾ SIZE ▾ STYLE ▾ BIN: SIZE ▾ FORMAT ▾ STEP ▾

Summary

Matrix Transfer Plots Access Plots Site Statistics Custom Ranking Plots Site History Access Pattern MAP (Experimental)

Common options

Interval
Last 4 hours

Access type
Remote access

Transfer VS Reading
Reading
Copy

Client - Server options

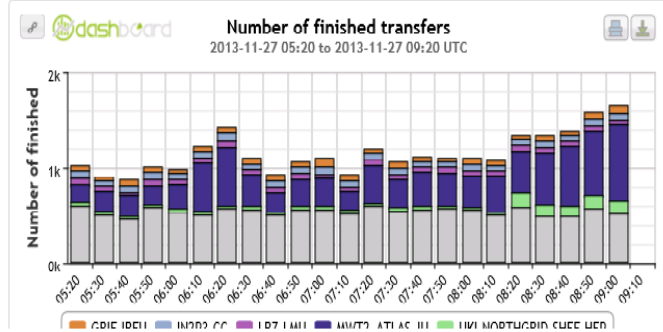
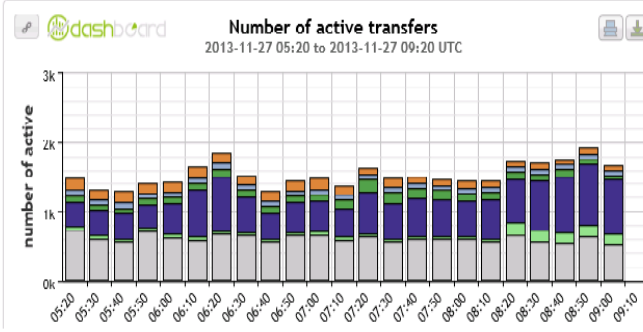
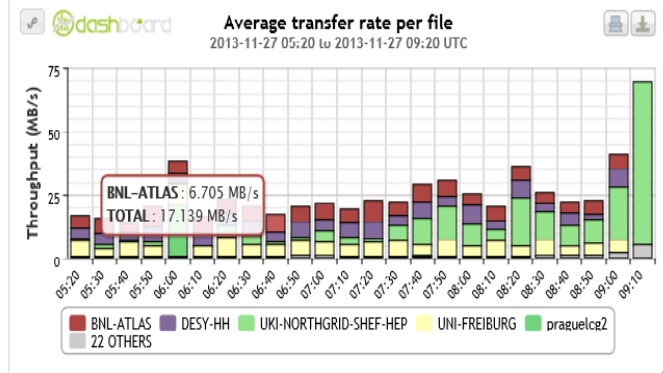
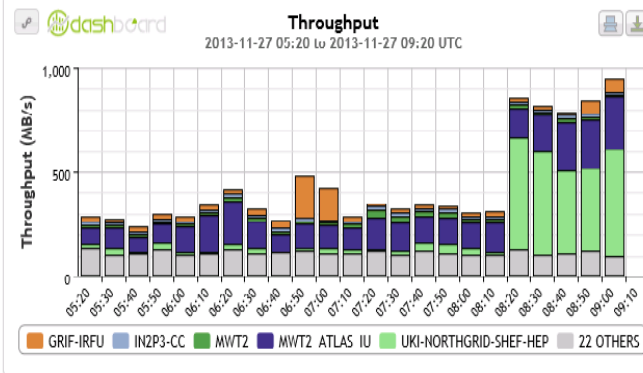
Server:
Grouping: SITE

Clients:
Grouping: SITE

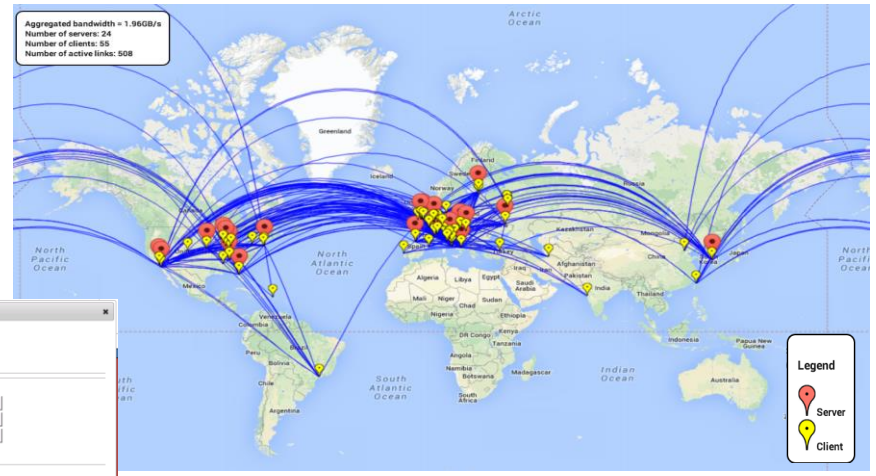
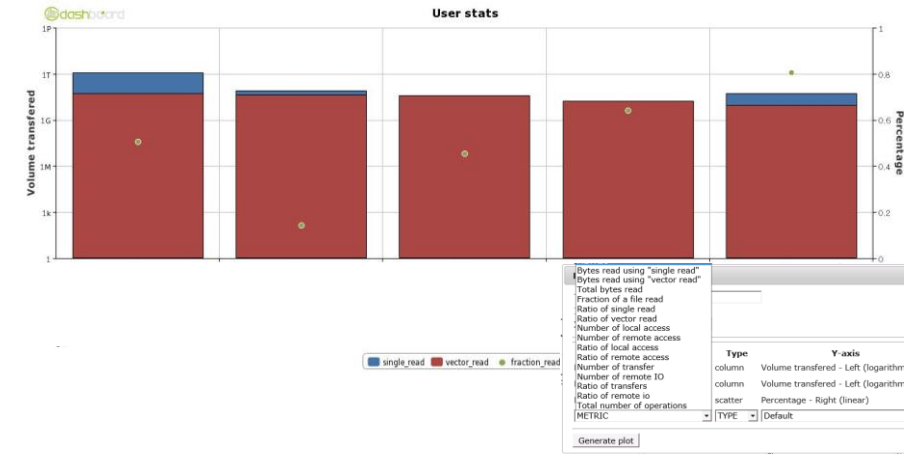
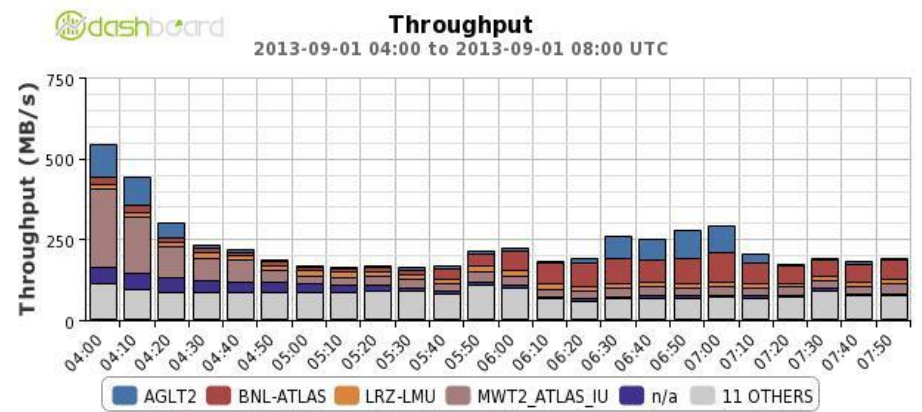
Users:

File:

- Interval
- Access type
- Transfer VS Reading
- Sources
- Destinations



Pre-defined set of views



Bytes read using "single read"
Bytes read using "vector read"
Total bytes read
Fraction of a file read
Ratio of single read
Ratio of vector read
Number of local access
Number of remote access
Ratio of local access
Ratio of remote access
Number of transfer
Number of remote IO
Ratio of transfers
Ratio of remote io
Total number of operations
METRIC

Type: column, Y-axis: Volume transferred - Left (logarithmic)

Type: column, Y-axis: Volume transferred - Left (logarithmic)

Type: scatter, Y-axis: Percentage - Right (linear)

METRIC: TYPE: Default

Buttons: Edit, Delete, Add, Reset, Generate plot

Matrix Example



Any Data, Any Time, Anywhere Monitoring (RC.34)

Latest statistics update: 2014-05-13 T08:10:00.072431

TRANSFER MATRIX (2014-05-13 07:20 to 2014-05-13 08:20 UTC SLIDING)

MAX CELLS

Summary

Common options

Interval
Last hour

Access type
Remote access

Transfer VS Reading
Reading

Client - Server options

Server:
Grouping:SITE

Clients:
Grouping:SITE

Users:

File:

Source - Destination options

Sources
Sites: T1_IT
Grouping:SITE

Interval

Access type

Transfer VS Reading

Sources

Destinations

Matrix Transfer Plots Access Plots Site Statistics Custom Ranking Plots Site History Access Pattern MAP (Experimental)

Volume
Throughput
Active
Finished

0 % 100 %

100 %

0 %

SOURCES

Displaying 1 of 1 sources and 5 of 5 destinations.

	TOTAL	T1_IT_CNAF
TOTAL	12 GB 3 MB/s 105	12 GB 3 MB/s 105
T1_IT_CNAF	0 kB 0 kB/s 21	0 kB 0 kB/s 21
T2_PK_NCP	3 GB 699 kB/s 12	3 GB 699 kB/s 12
T2_TP_METU	5 GB 1 MB/s 35	5 GB 1 MB/s 35
T2_UK_London_Brunel	99 MB 28 kB/s 3	99 MB 28 kB/s 3
T2_UK_SGrid_Bristol	5 GB 1 MB/s 34	5 GB 1 MB/s 34

DESTINATIONS

Matrix showing the remote IO CNAF served in the last hour

- # operations
- # bytes
- Averaged throughput



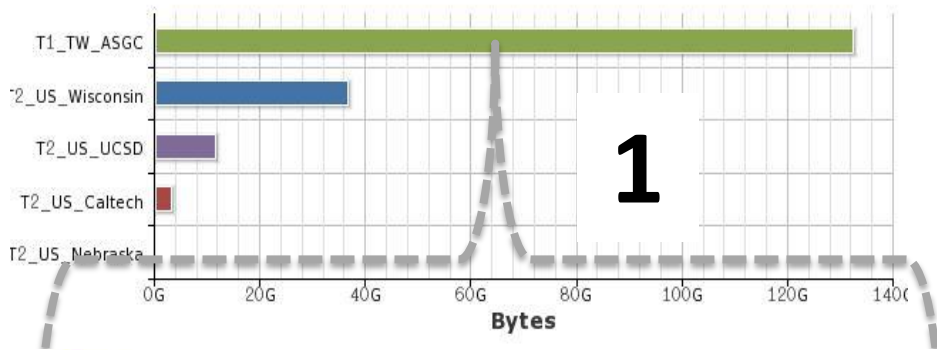
Use case example

Understand site access patterns



Read Bytes
T1_US_FNAL

2013-09-24 04:20 to 2013-09-24 08:20 UTC

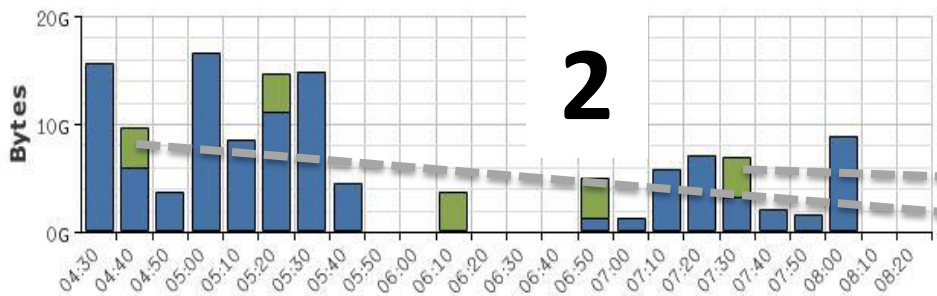


1. Which sites are reading from FNAL
2. Zoom to a specific site to understand which users are reading
3. Understand which files are read by a user



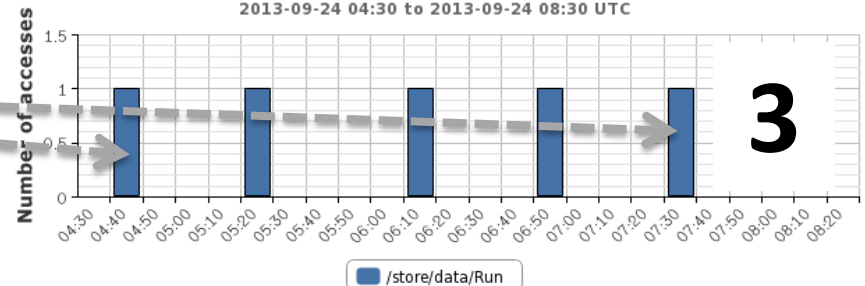
Read Bytes
T1_US_FNAL

2013-09-24 04:30 to 2013-09-24 08:30 UTC



Number of accesses
T1_US_FNAL

2013-09-24 04:30 to 2013-09-24 08:30 UTC



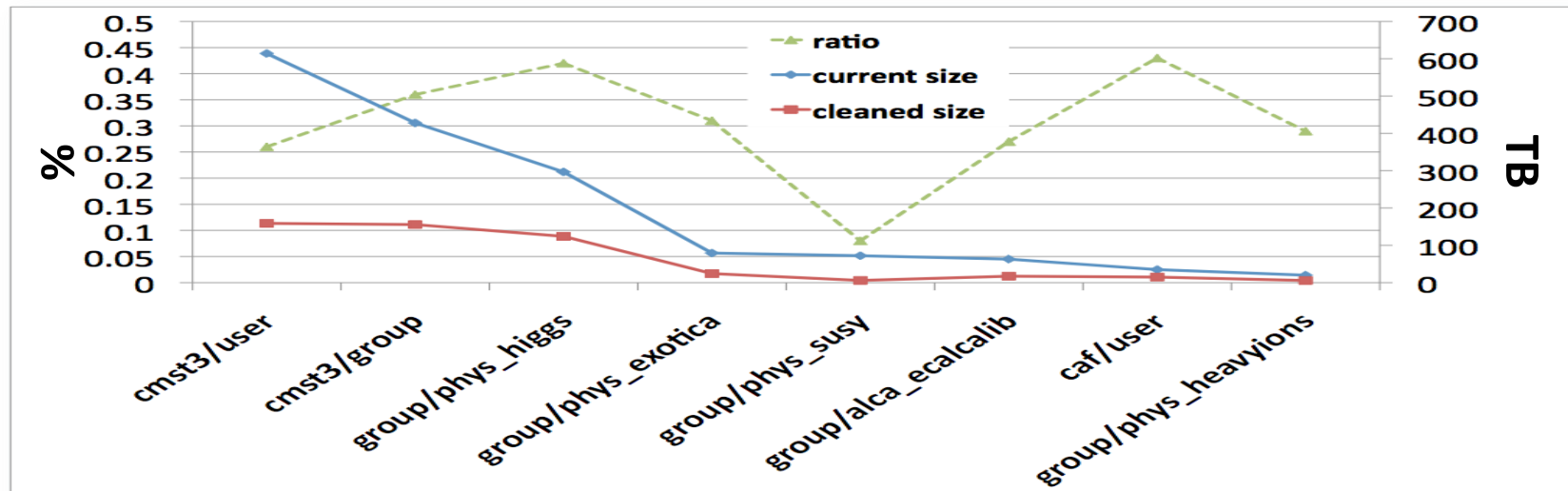
/store/data/Run

Data popularity

- XRootD monitoring provides information about file access patterns:
 - Including non official collections (ie: user files)
 - Contribute to simplify and make more efficient the usage of disk resources
- Popularity data analytics built on this information:
 - Adopted already for CMS-EOS
 - will be extended to full AAA

Archive recommendation for CMS-EOS

- Help to manage the disk space of EOS including user space
 - No central bookkeeping system



- Unused files: created > 4 months ago, no access in the last 3 months:
 - ~500 TB of space occupied and not used \Leftrightarrow 30% of total for these areas

Open issues

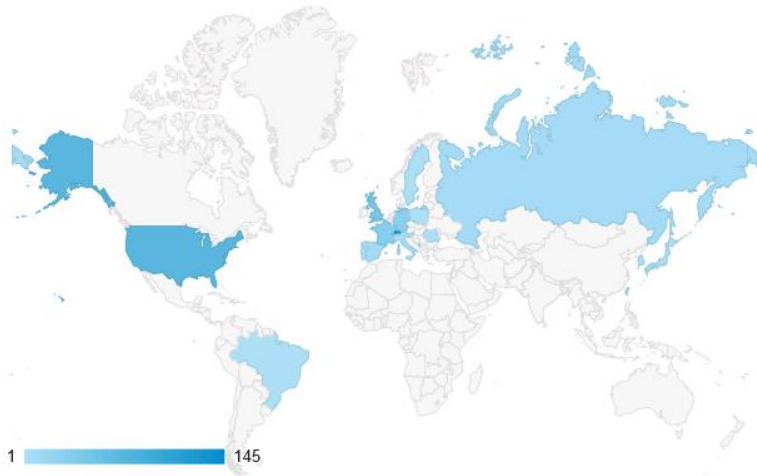
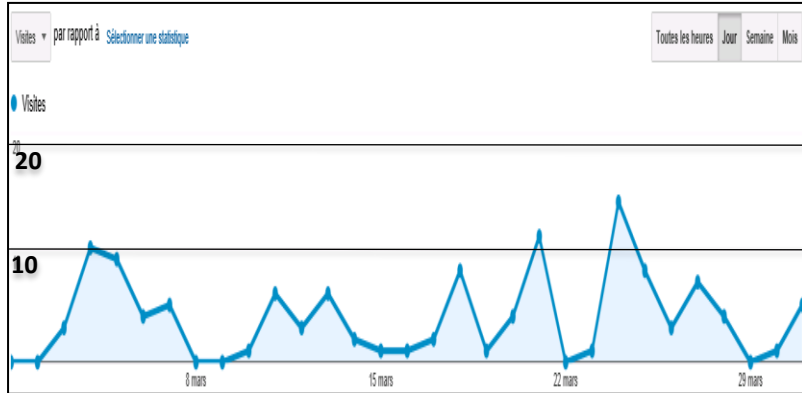
- Server should provide their site name.
 - CMS: only 5 sites (to be followed)
 - ATLAS: Done
- GLED Collector improvements:
 - Reliability of the service:
 - Recover time, can be long due to time difference
 - GLED should be operated as a production service
- Multi-VOs sites:
 - Discrimination will happened at GLED level

Future work

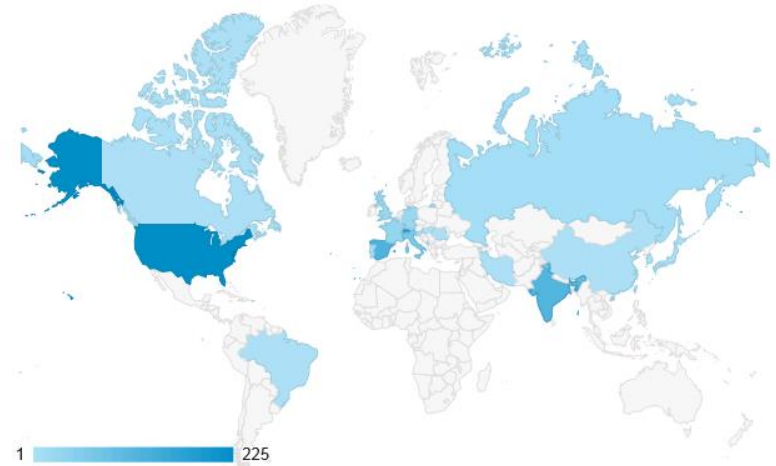
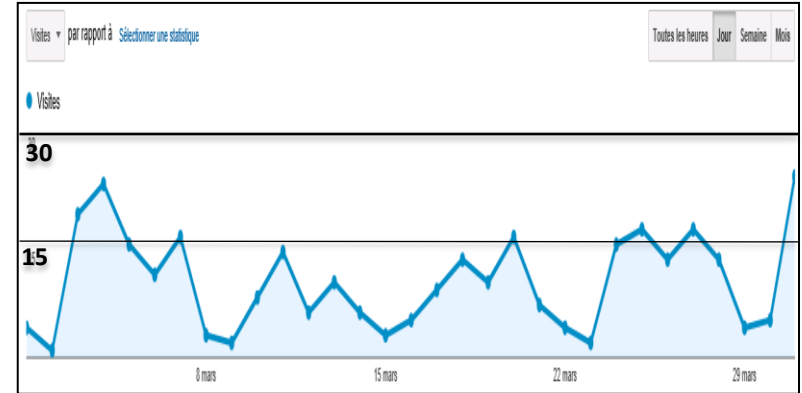
- Strong requirement from ATLAS to understand efficiency:
 - Need the concept of error / failure
 - How XRootD server could be instrumented to report it?
- Topology
 - Resolution will be based on the new “server site” field
- Data-mining activity (2 years of data ~ 1TB)

Application usage

FAX



AAA

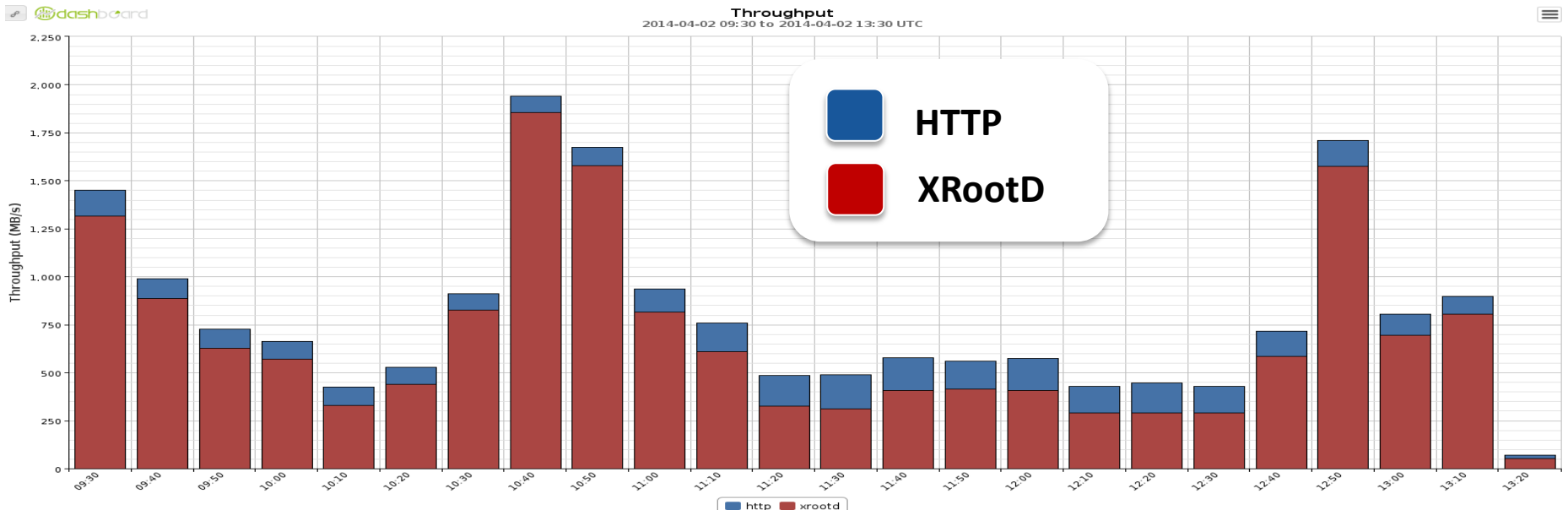


HTTP Federation is coming

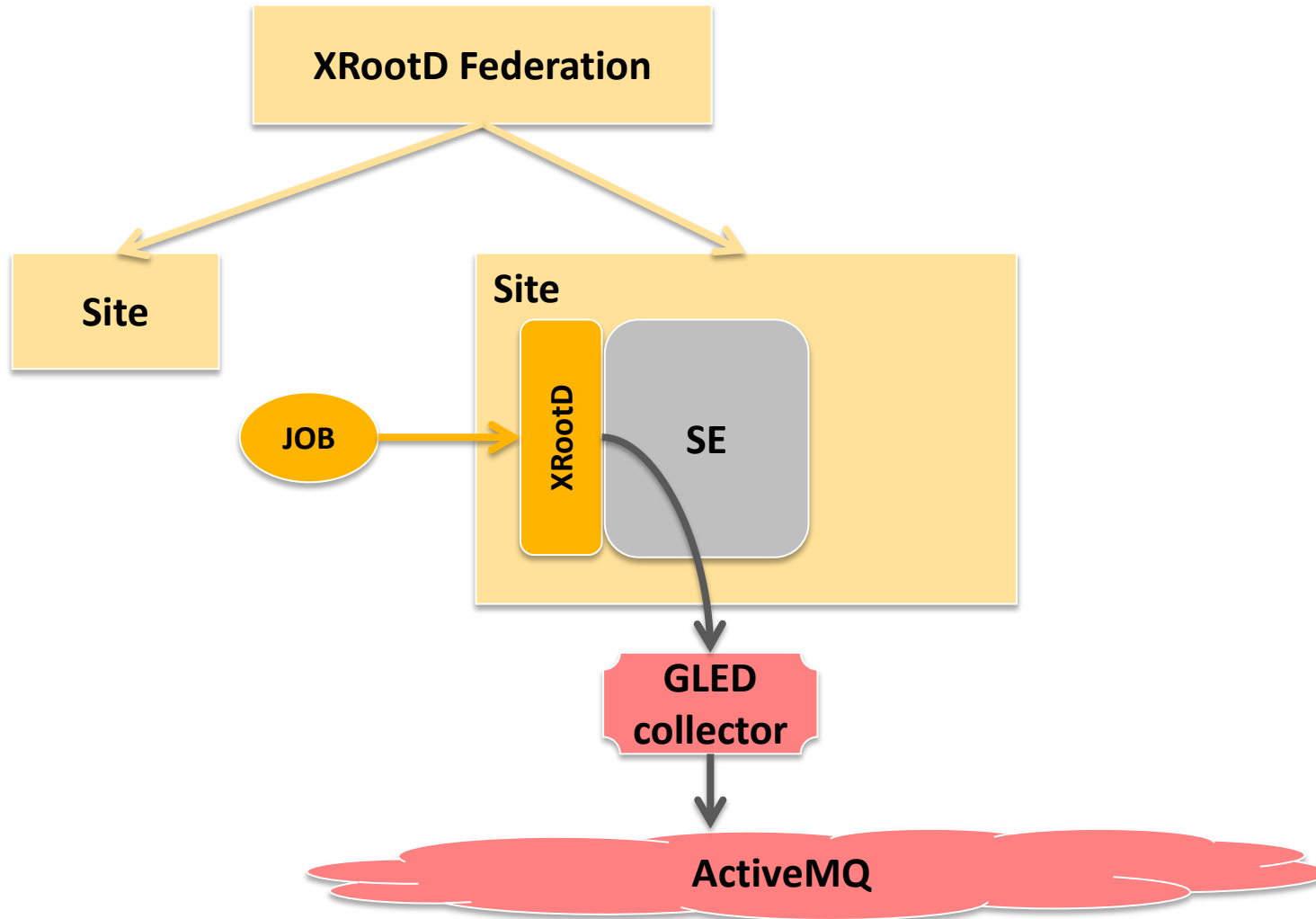
- HTTP protocol will be used in the future
 - XRootD servers can be accessed
 - See Fabrizio's presentation on xrdhttp
- Two kind of accesses:
 - Pure HTTP access (through Apache)
 - HTTP gate to XRootD server
- Can't be monitor in the same way

Monitoring XRootD access protocol

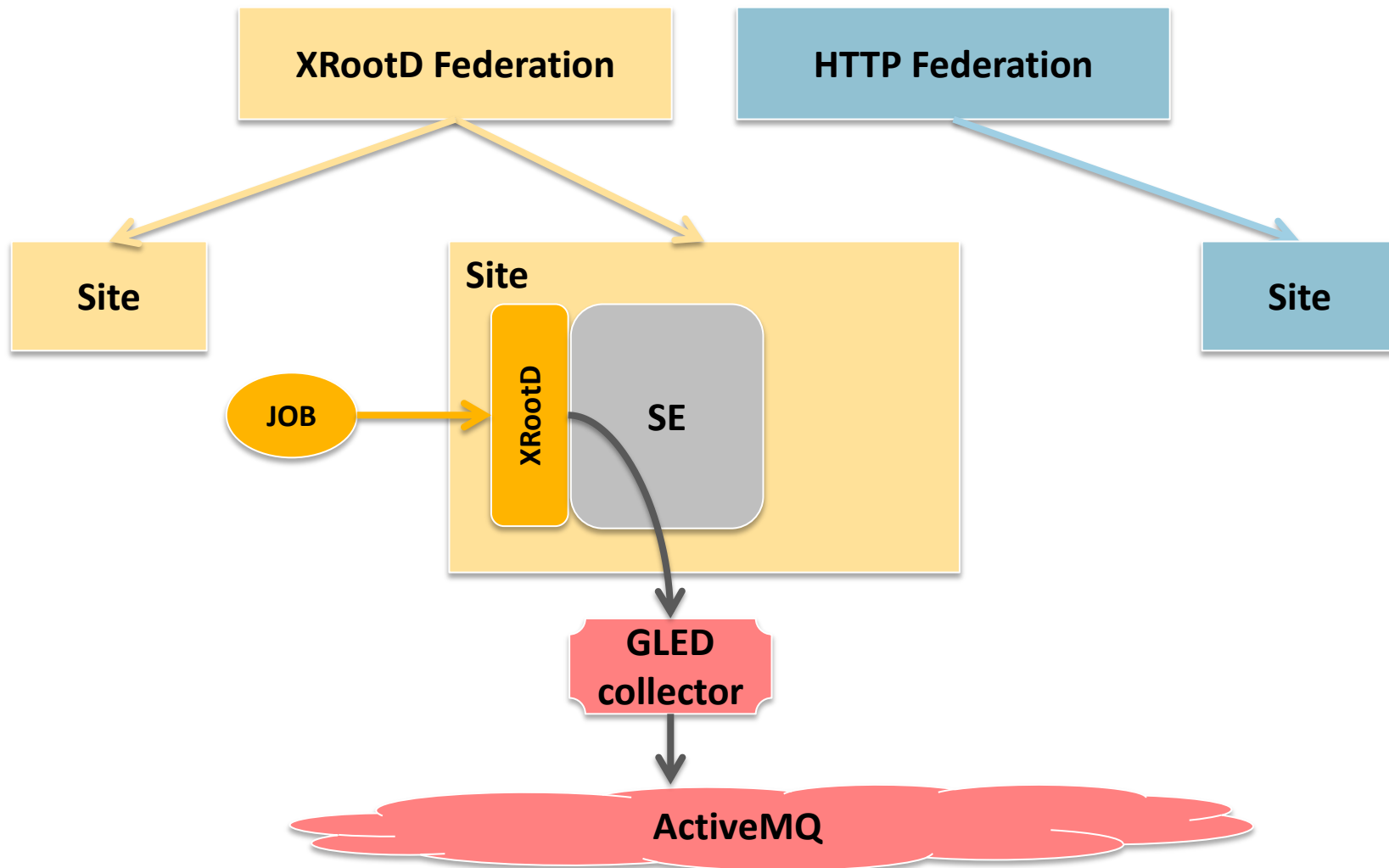
- XRootD 4 will now reports the user protocol:
 - All the monitoring chain needs to be updated
 - Dashboard DB and UI are fully ready



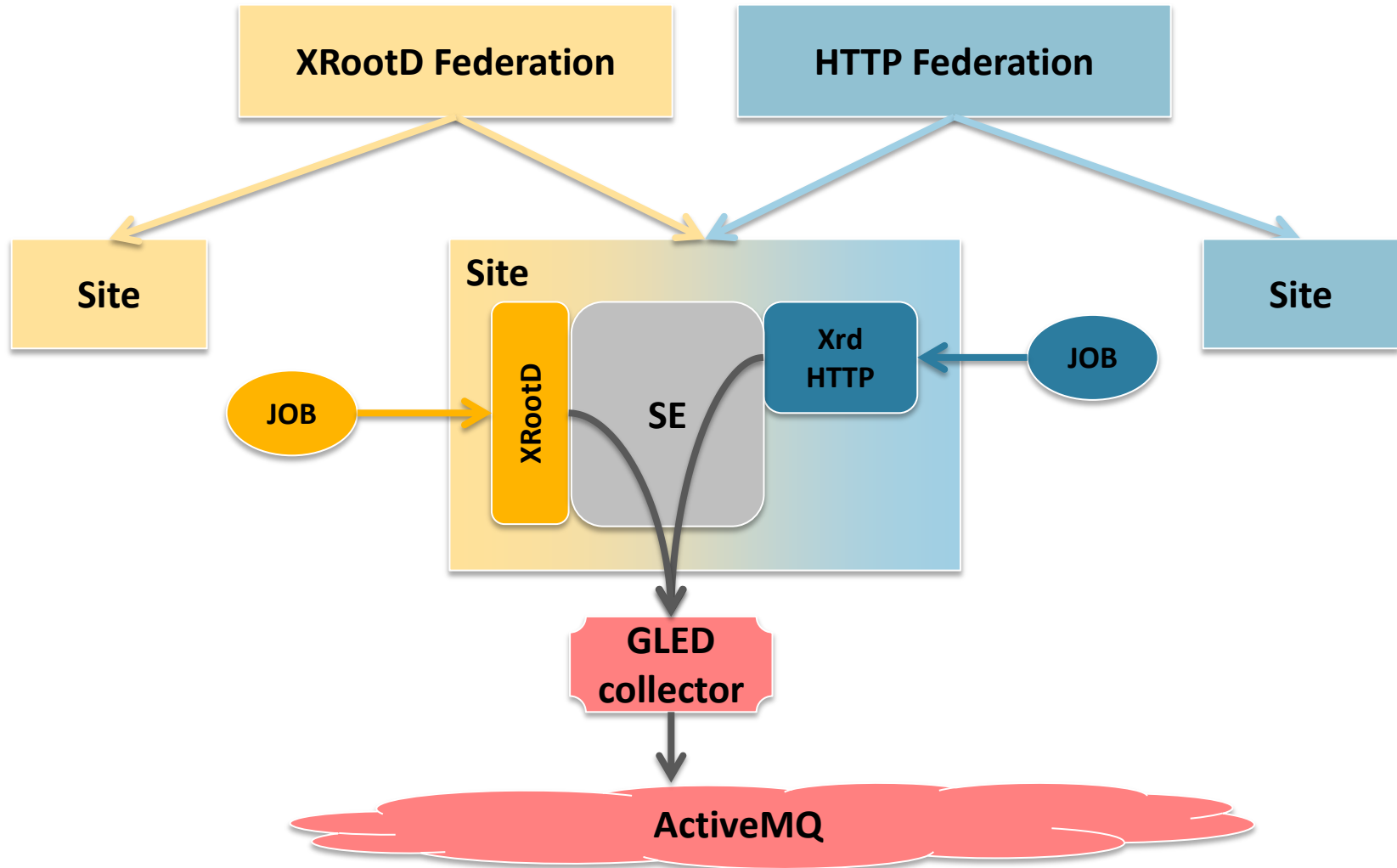
HTTP/WebDAV federation monitoring



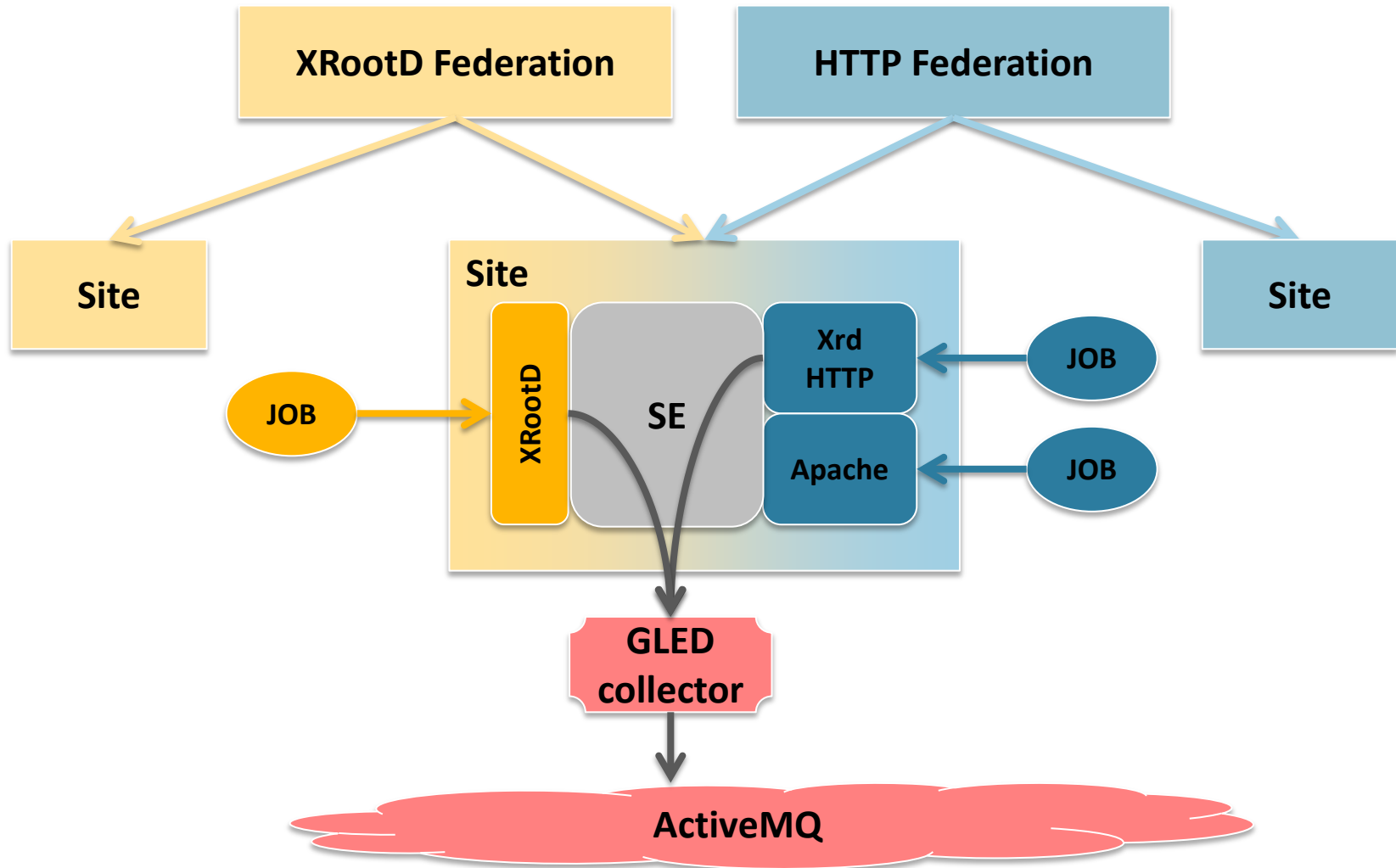
HTTP/WebDAV federation monitoring



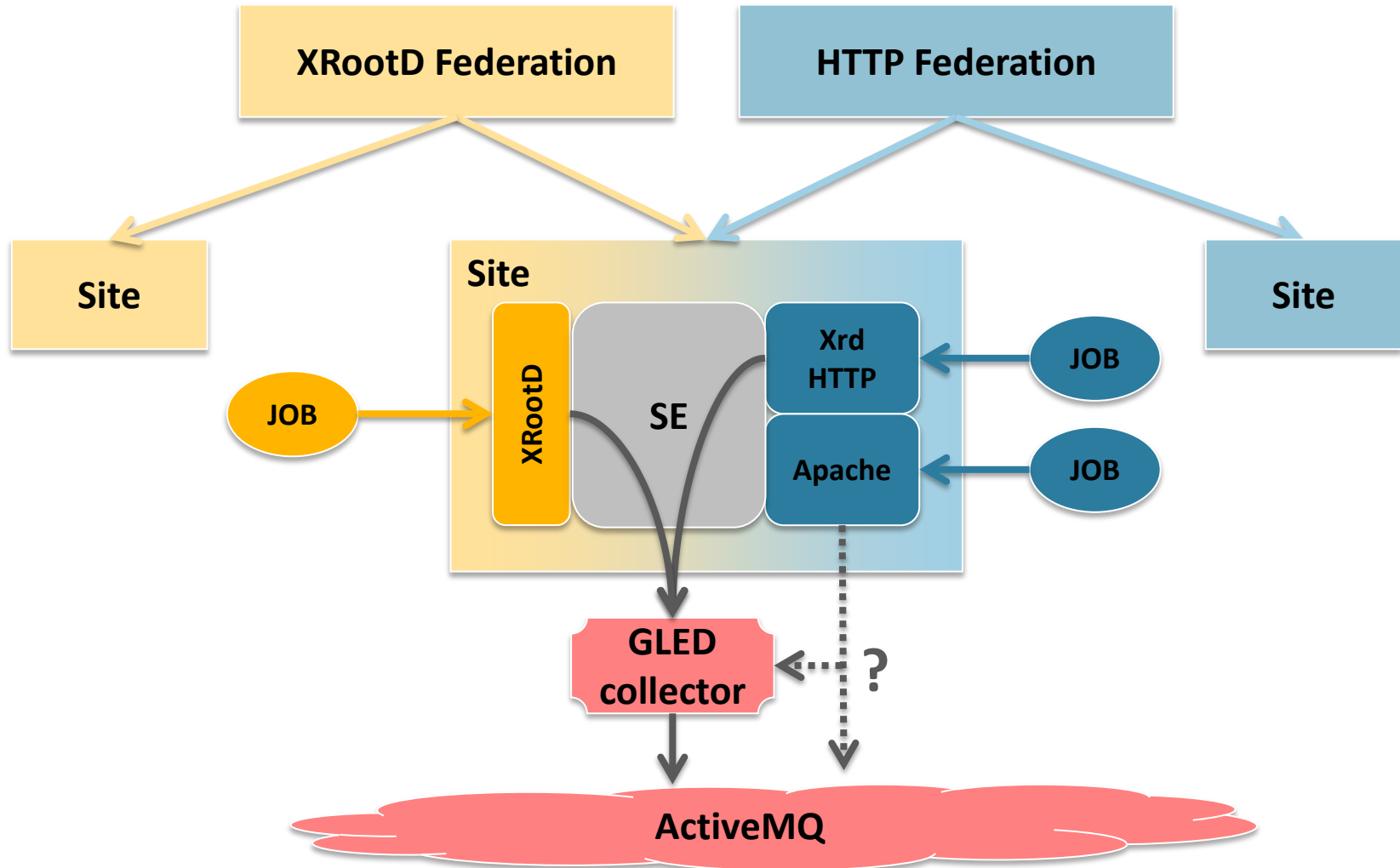
HTTP/WebDAV federation monitoring



HTTP/WebDAV federation monitoring



HTTP/WebDAV federation monitoring



Summary

- Lots of effort has been put in XRootD monitoring workflow and dashboard in the last 2 years
 - Reliable system achieved
 - Lots of use cases covered
- HTTP Monitoring already started
 - Will require a lot of effort to reach XRootD monitoring level
 - First prototype for pure HTTP monitoring will be ready by autumn thanks to DPM team

Credits

- Andreeva Julia
- Cons Lionel
- Giordano Domenico
- Saiz Pablo
- Tadel Matevz
- Tuckett David
- Vukotic Ilija
- The AAA and FAX deployment team
-

Useful links

- AAA Dashboard
 - <http://dashb-cms-xrootd-transfers.cern.ch>
- FAX Dashboard:
 - <http://dashb-atlas-xrootd-transfers.cern.ch>
- CHEP materials
 - <https://indico.cern.ch/abstractDisplay.py?abstractId=101&confId=214784>
 - <https://indico.cern.ch/getFile.py/access?contribId=94&sessionId=6&resId=0&materialId=slides&confId=214784>
 - <https://indico.cern.ch/getFile.py/access?contribId=265&sessionId=6&resId=1&materialId=slides&confId=214784>
- Xbrowse framework:
 - <https://twiki.cern.ch/twiki/bin/view/ArdaGrid/XbrowseFramework>

Thanks for your attention

