Monitoring of large-scale federated data storage
XRootD and beyond

Alexandre Beche[1], Julia Andreeva[1], Domenico Giordano[1], Pablo Saiz[1], David Tuckett[1], Daniel Diéguez Arias[1], Sergey Belov[2], Matevz Tadel[3], Ilija Vukotic[4], Danila Oleynik[2], Artem Petrosyan[2]

[1]CERN IT-SDC, [2]JINR, [3]UCSD, [4]UofC

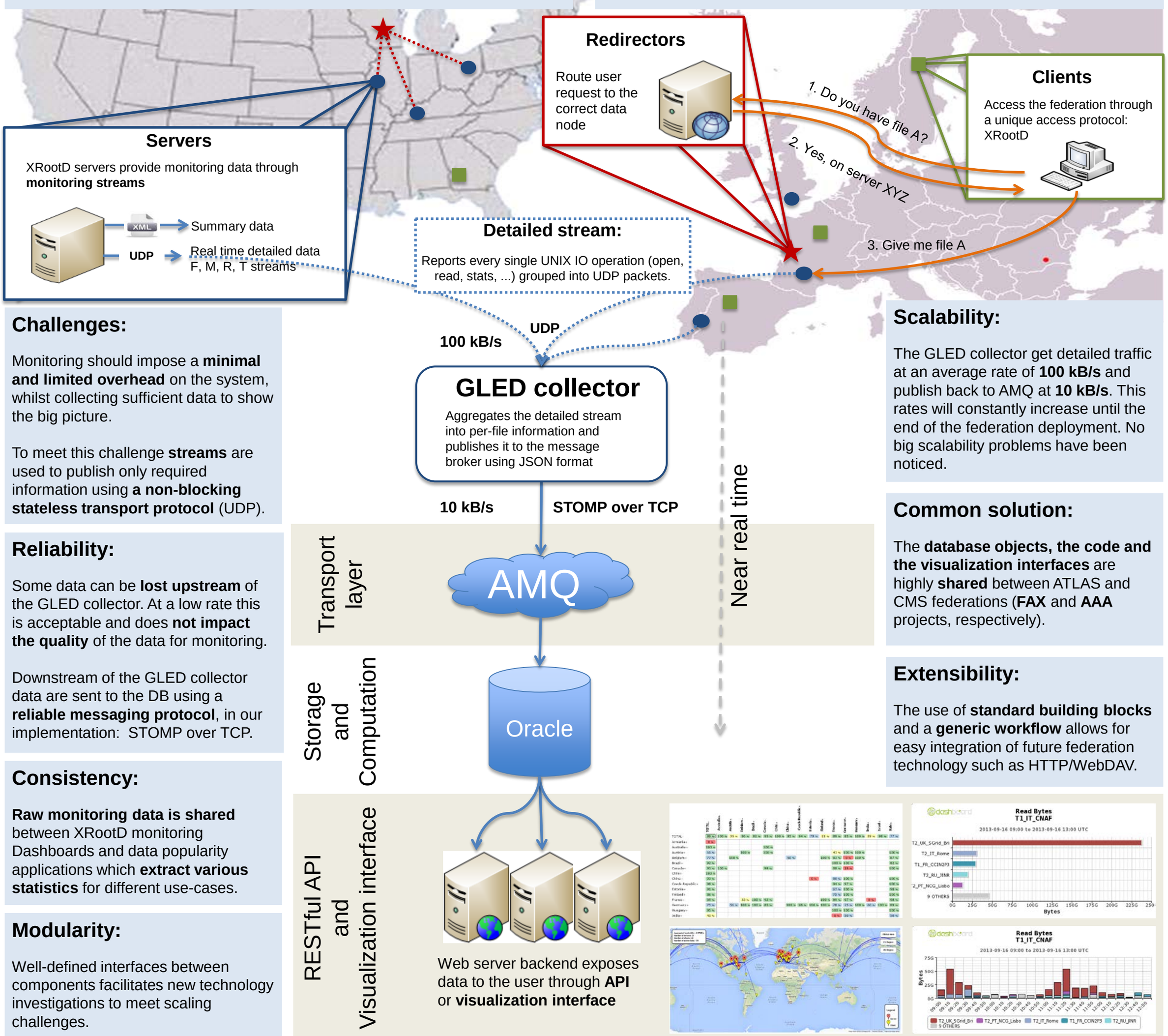
The computing models of the LHC experiments are gradually moving towards federated storage such as XRootD. Monitoring functionality should cover the status of all components on such a system and data mining of the collected monitoring data provides a deep insight into new patterns of usage of the storage resources. Handling of the monitoring flow generated by these systems has to be well optimized and visualization interface complete and concise.

Federation

Aggregation of storage systems of any kind into a **global namespace** using a single access protocol to provide **read-only access to worldwide replicated data** via virtual entry points : **redirectors.**

Why monitoring?

- Identify access patterns and propose data placement strategies
- Estimate data traffic and understand data flows
- Understand server usage
- And much more !



Challenges:

Monitoring should impose a **minimal and limited overhead** on the system, whilst collecting sufficient data to show the big picture.

To meet this challenge **streams** are used to publish only required information using **a non-blocking stateless transport protocol** (UDP).

Reliability:

Some data can be **lost upstream** of the GLED collector. At a low rate this is acceptable and does **not impact the quality** of the data for monitoring.

Downstream of the GLED collector data are sent to the DB using a **reliable messaging protocol**, in our implementation: STOMP over TCP.

Consistency:

Raw monitoring data is shared between XRootD monitoring Dashboards and data popularity applications which **extract various statistics** for different use-cases.

Modularity:

Well-defined interfaces between components facilitates new technology investigations to meet scaling challenges.

Web server backend exposes data to the user through **API** or **visualization interface**

Scalability:

The GLED collector get detailed traffic at an average rate of **100 kB/s** and publish back to AMQ at **10 kB/s**. This rates will constantly increase until the end of the federation deployment. No big scalability problems have been noticed.

Common solution:

The **database objects, the code and the visualization interfaces** are highly **shared** between ATLAS and CMS federations (**FAX** and **AAA** projects, respectively).

Extensibility:

The use of **standard building blocks** and a **generic workflow** allows for easy integration of future federation technology such as HTTP/WebDAV.

For more information:
https://twiki.cern.ch/twiki/bin/view/LCG/XrootdMonitoring
Contact: dashboard-support.cern.ch

LARGE HADRON COLLIDER ONLINE WORLDWIDE GRID COMPUTING

WLCG