Storage Update in WLCG and at CERN

Dirk Duellmann

(material: O. Keeble, A. Peters, X. Espinal and others)

CERN IT, storage group

WLCG
Worldwide LHC Computing Grid

WLCG Data Steering Group - First Outcomes

- The group has collected and categorised input
  - presented conclusions at the WLCG workshop in Manchester
  - organised
    - data session at last WLCG workshop
    - pre-GDB on object stores & archival storage
  - initiated wg to optimise tape usage across WLCG
- Full details, agendas, mandate etc
  - at https://twiki.cern.ch/twiki/bin/view/LCG/WLCGDataSteeringGroup

Infrastructure diversity

- Cloud storage, object stores
- HPC Data management
- Colocation with other communities
- Adapting workflow to resources
  - Quality of Service

QoS - gains in parameter space?

- Role of tape beyond archive
- Archives backed other than by tape
- Tuneable reliability
- Cache hierarchies

Common strategy

- Can a common vision be described?
- Common White Paper
    - reference model for iterative refinement
- Sustainability and collaboration
    - Identify higher-level services are candidates for common projects
    - eg SKA, Globus project

New analysis models

- Access paradigms
    - Interfaces to storage systems
        - Queries, hints.
    - Event level access
- Data formats

Globus Toolkit EOL

- Univ. Chicago will end supporting the Globus Toolkit in January 2018
    - https://github.com/globus/globus-toolkit/blob/globus_6_branch/support-changes.md
- Maintenance and security fixes till end of 2018
    - GridFTP, GSI, MyProxy
- Main users met and agreed on joint support effort for mid-term
    - WLCG, OSG, EUDAT, PRACE, EGI, DPM,dCache
- Long term: initiative to find production quality alternatives (e.g. httpd/xrootd for GridFTP, Oath2 for GSI)

Infrastructure Architecture

- Multi-site storage
- Regional federations
- Caches
- Diskless sites
- -> Networking

Operations

- Storage accounting, space reporting
- Monitoring
- Topology system, CRIC

Authentication & Authorisation Infrastructures (AAI) and Federated Identity

Ensure that data management systems, and storage in particular, is ready for any new authentication and authorisation infrastructure

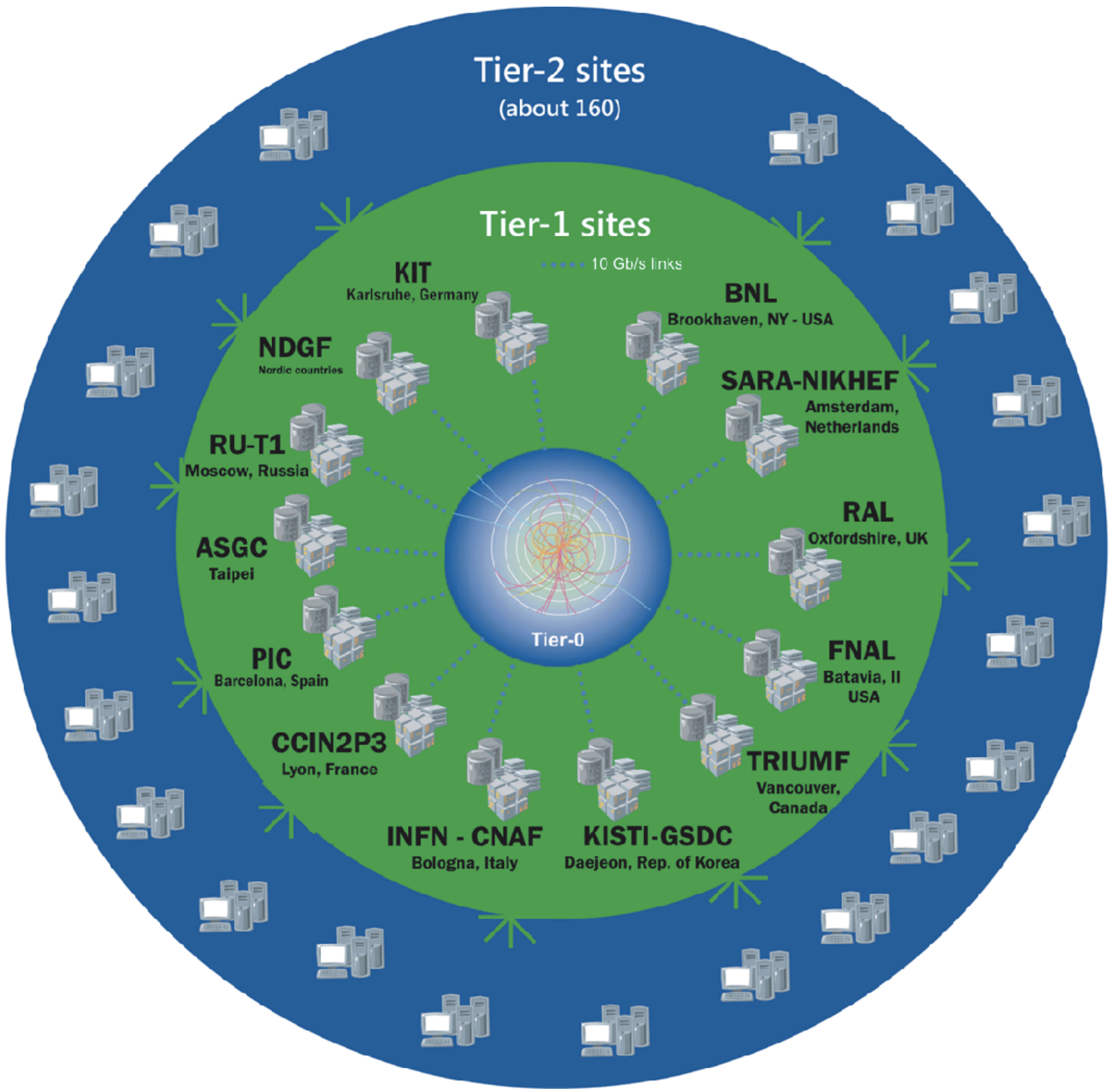
- eduGAIN
- OpenID Connect
- Web/non-web workflows
- GSI authentication

Hep Software Foundation (HSF)

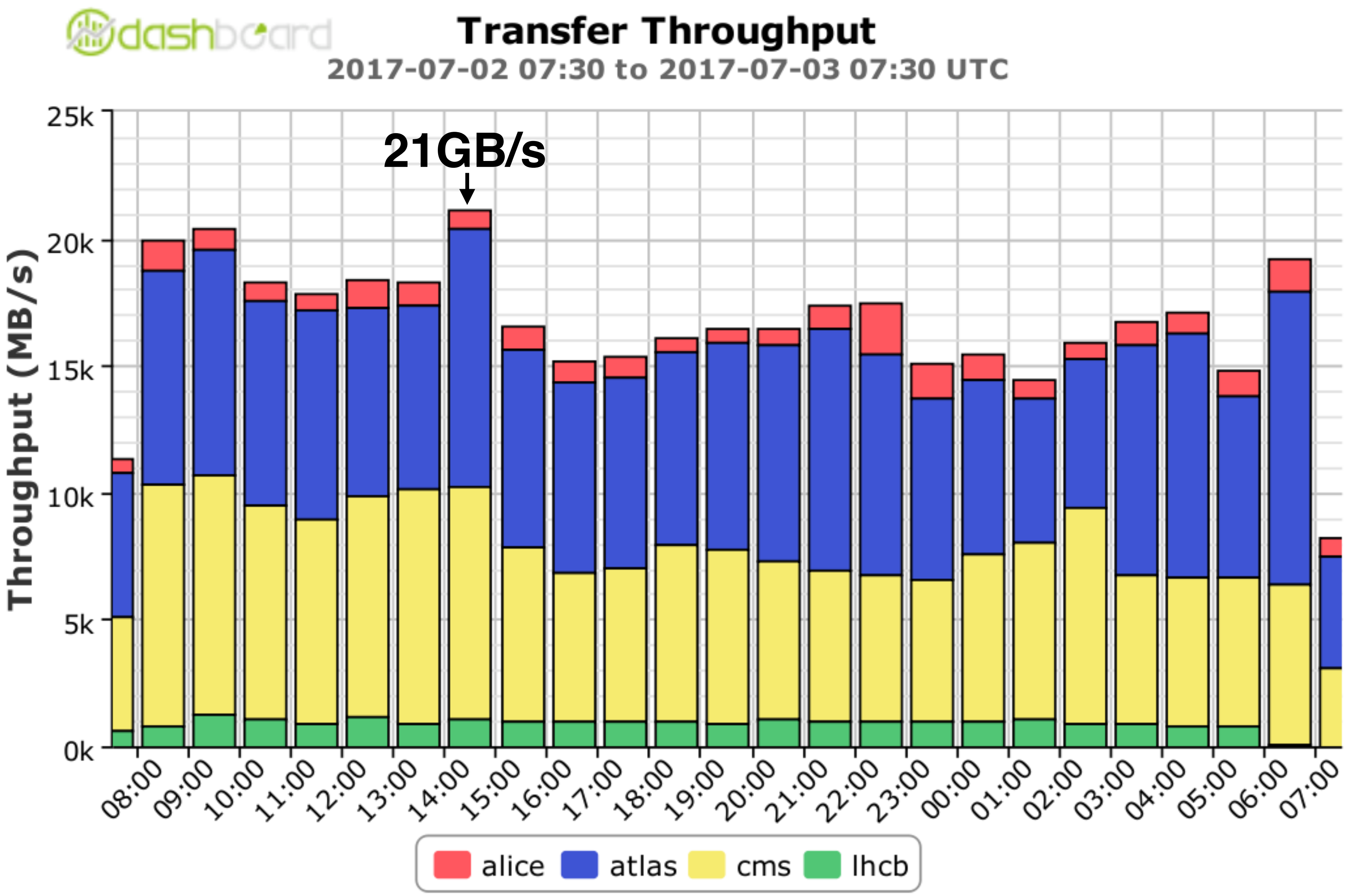
- The HSF is organising a "Community White Paper", designed to capture requirements and roadmap for HEP computing on an HL-LHC timescale.
  - http://hepsoftwarefoundation.org/activities/cwp.html
- Chapter on "Data Organisation, Management and Access" is in review
- Oriented towards analysis efficiency and experiment frameworks
  - Final version will include 1, 3 and 5 year roadmaps

http://wlcg.web.cern.ch/

WLCG
Worldwide LHC Computing Grid



IT-ST



Transfer Throughput
2017-07-02 07:30 to 2017-07-03 07:30 UTC
21GB/s
Throughput (MB/s)
25k
20k
15k
10k
5k
0k
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
22:00
23:00
00:00
01:00
02:00
03:00
04:00
05:00
06:00
07:00
alice
atlas
cms
lhcb

FTS3

IT-ST

File Transfer Service FTS3

http://fts3-service.web.cern.ch/

Grid oriented development **but crossing boundaries**

Third-party-copy file transfer **orchestrator**

Multiprotocol: **webdav/https**, **GridFTP**, **xroot**

Built-in **retry mechanism, checksuming**, real-time speed **optimization**, **scalability**

Web portal for end-users and web **monitoring**

Instrumental tool for the grid community:

>**30** active Virtual Organizations
Biggest VO (ATLAS): **51M** transfers, **15.6 PB**
All VOs in 30 days: **~35.4 PB and ~75M files**

IT-ST

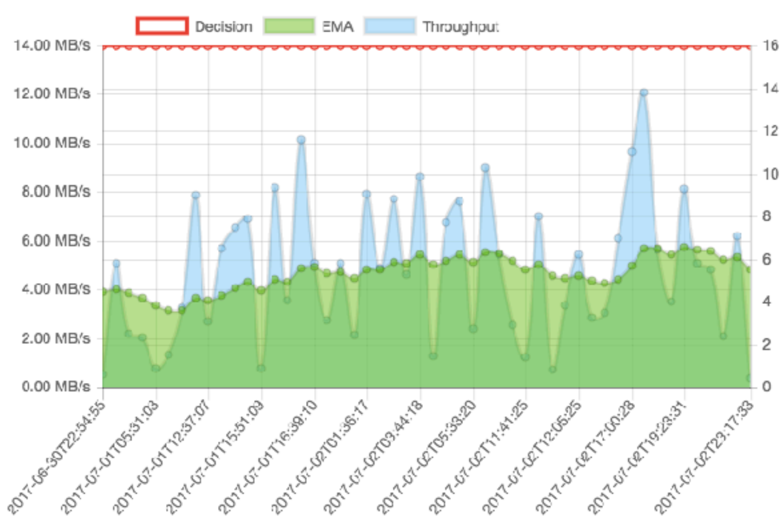
Continuous improvements: eg improvements to transfer optimiser,

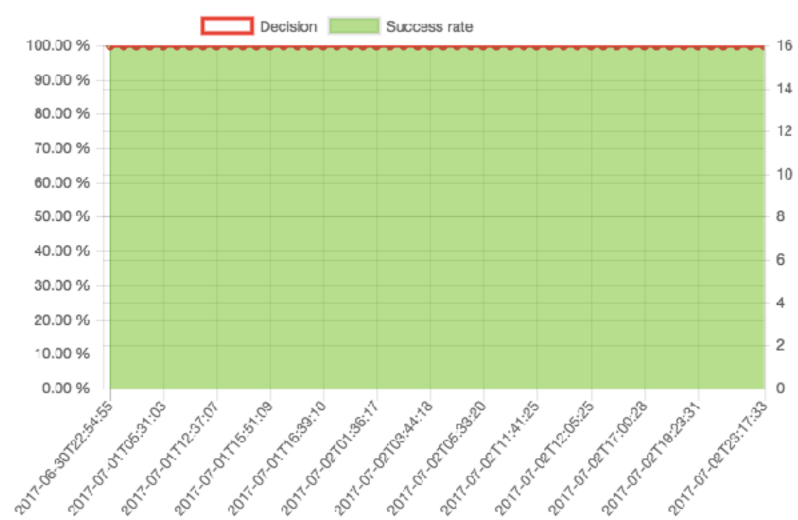
metrics collection to ES / Hadoop for optimisation of tape and transfer performance

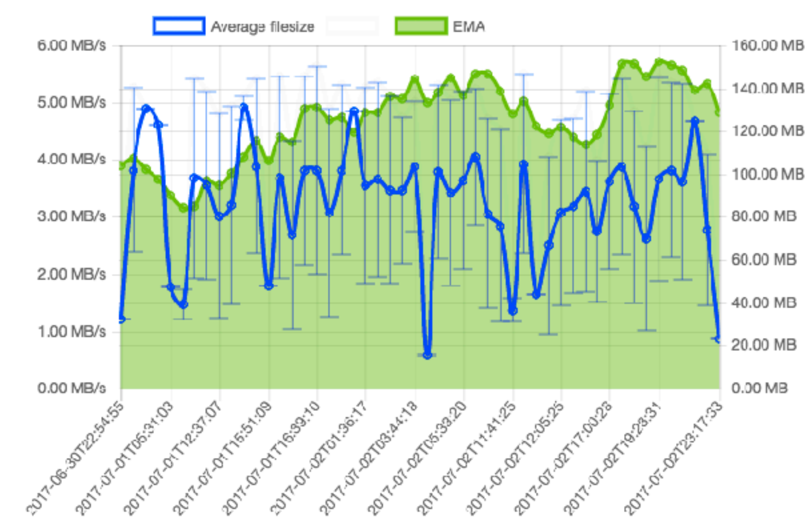
Grid oriented development **but crossing boundaries**

Data acquisition **driven** by **FTS**

Details for gsiftp://na62primitive.cern.ch → srm://srm-public.cern.ch







First | Previous | 1 | 2 | 3 | Next | Last

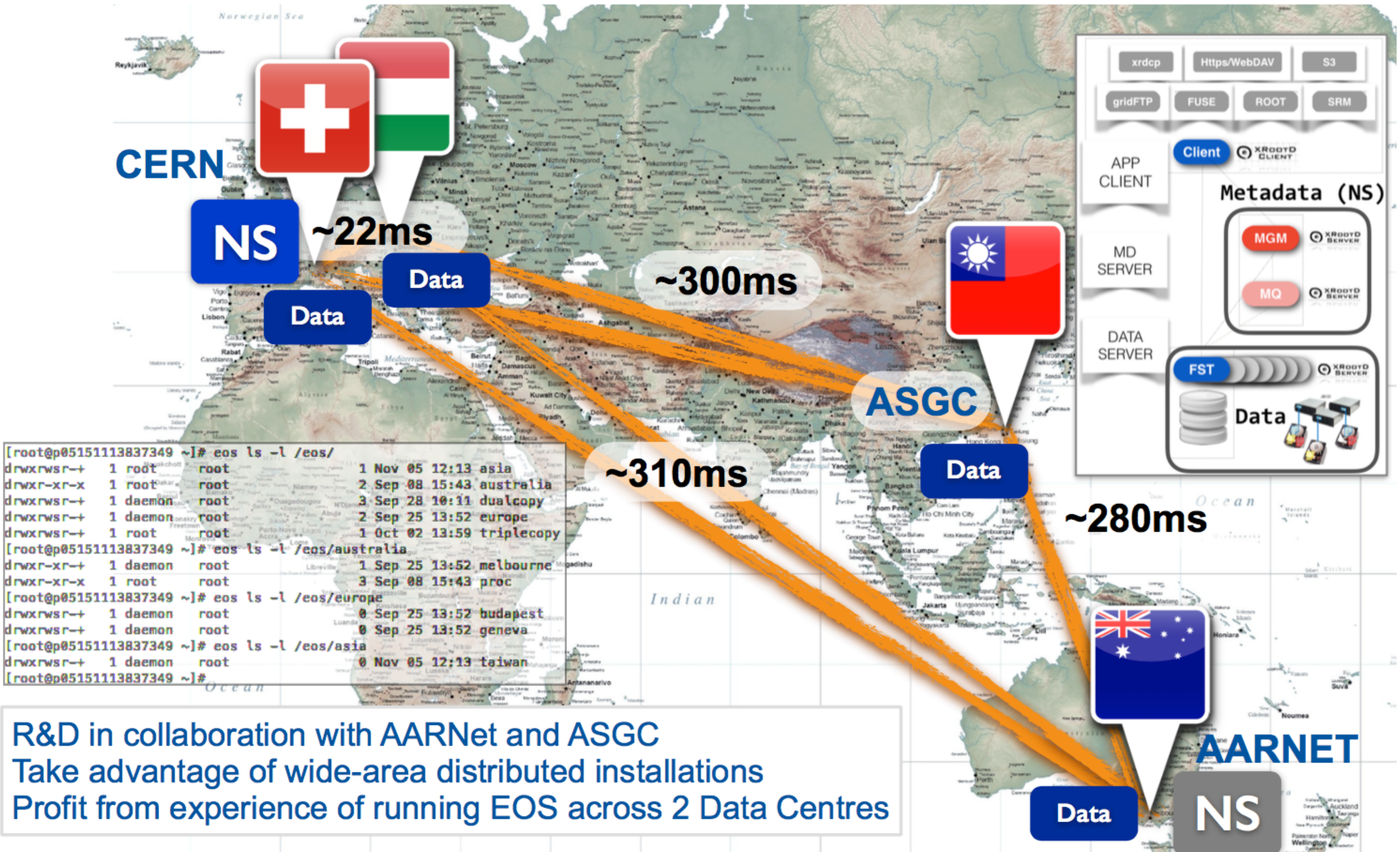
Timestamp	Decision	Running	Queue	Success rate (last 1min)	Throughput	EMA	Diff	Explanation
2017-07-02T23:17:33	16	5	0	100.00%	0.384 MB/s	4.827 MB/s	0	Good link efficiency, throughput deterioration. Too many streams
2017-07-02T21:01:32	16	1	0	100.00%	6.172 MB/s	5.321 MB/s	0	Good link efficiency, current average throughput is larger than the preceding average. Too many streams

IT-ST

Wide Area Federation - EOS

WAN-Testbed

Worldwide distributed storage system with extreme latencies - R&D prototype



R&D in collaboration with AARNet and ASGC
Take advantage of wide-area distributed installations
Profit from experience of running EOS across 2 Data Centres

Regional Clustering - EOS

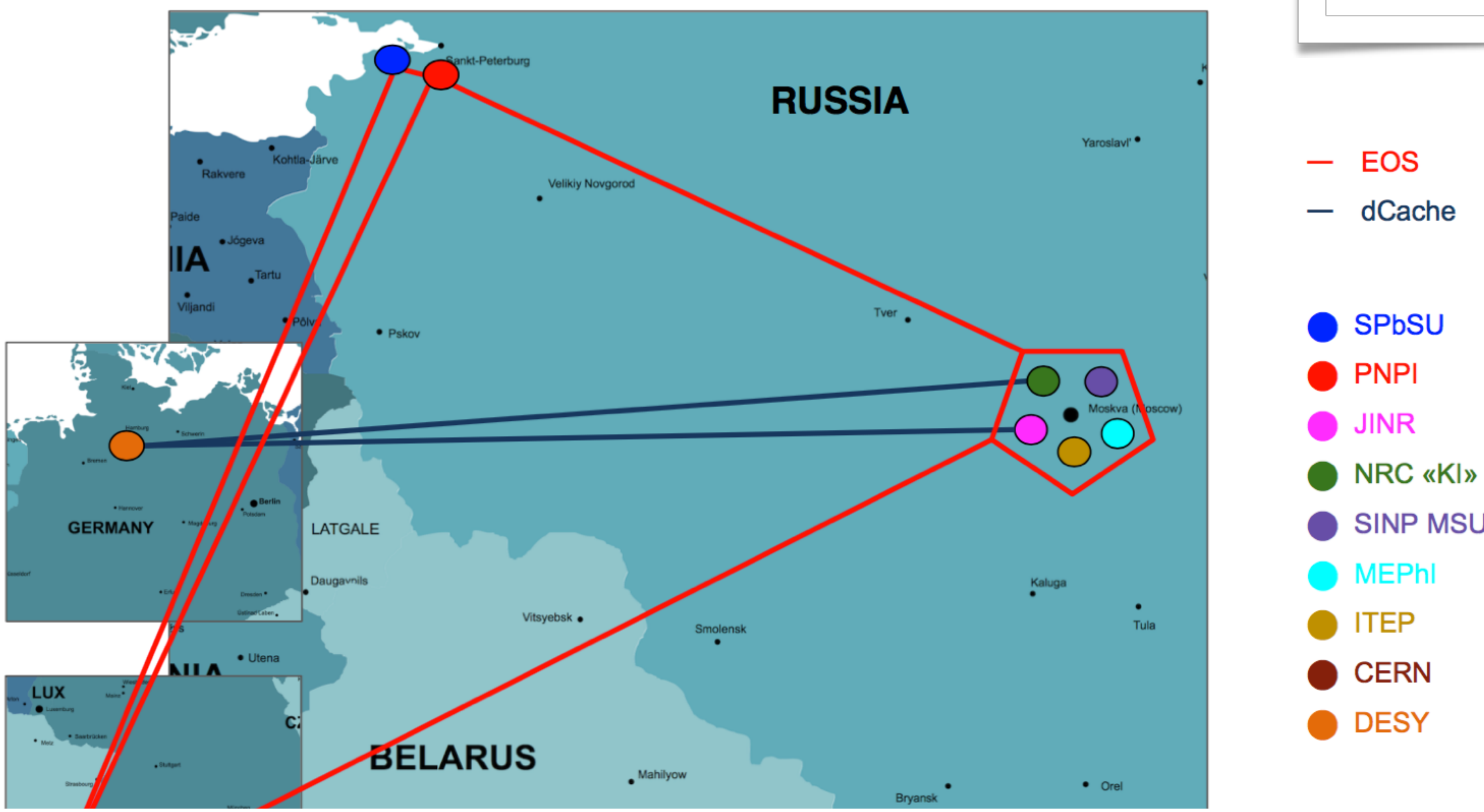
Examples

National Research Centre (NRC)
"Kurchatov Institute"

Big Data Technologies Laboratory
http://bigdatalab.nrcki.ru/

Federation topology

Andrey Kiryanov



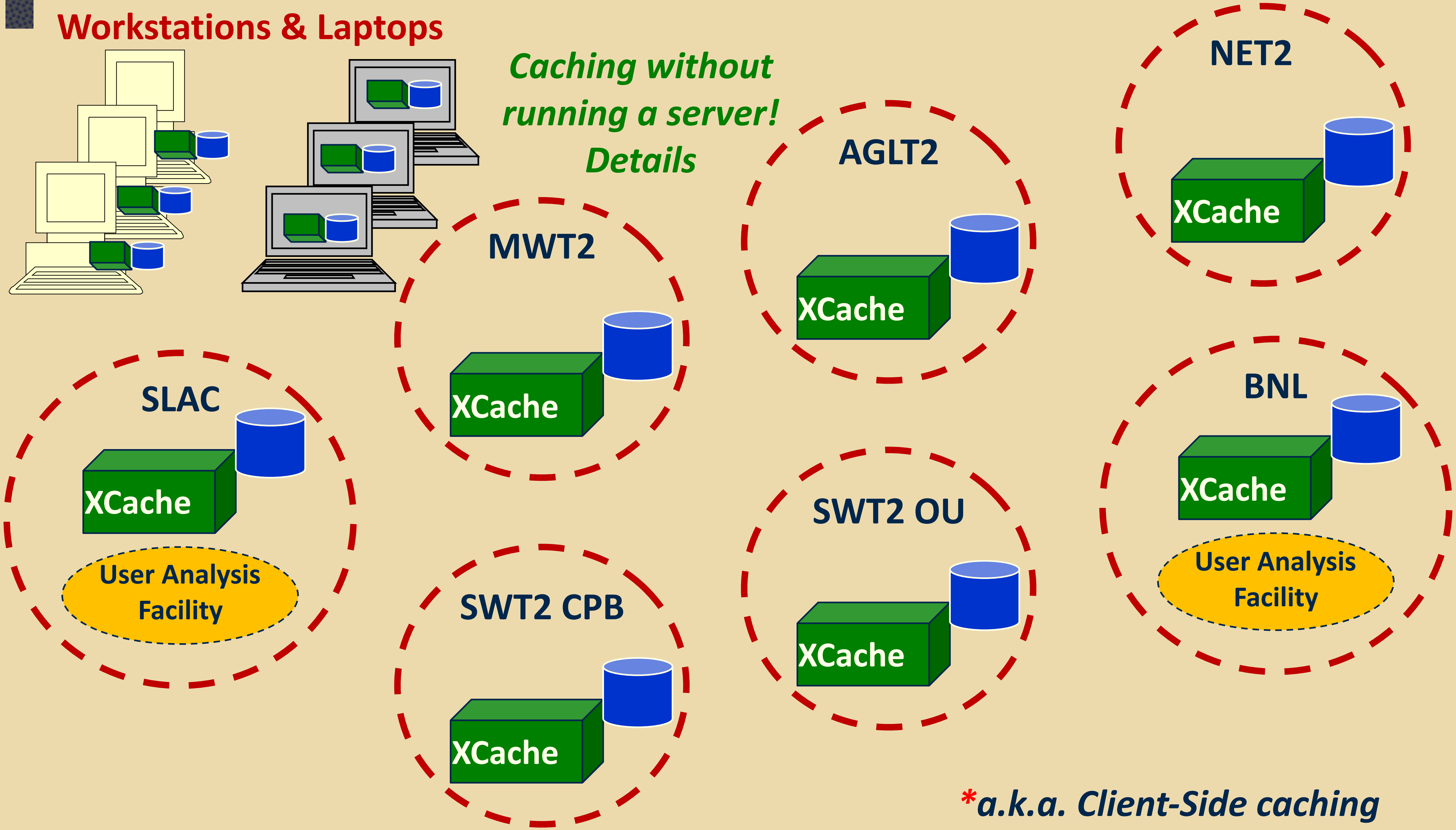
https://indico.cern.ch/event/595396/contributions/2532587/

See also eg: The Use of Proxy Caches for File Access in a Multi-Tier Grid Environment [CHEP 2010]

What is XCache?

- Infrastructure of distributed US caches
  - Deployed near computational resources to…
    - Reduce access latency
    - Keep hot data near where it's actually hot
    - Provide access to all ATLAS replicated data
      - With no changes to applications
    - Enhance Ceph random I/O performance
      - Where Ceph is deployed
- Via the XRootD Proxy Caching Server
  - In three formative phases

XCache Phase 3: adds serverless*



**a.k.a. Client-Side caching*

Phase 4?

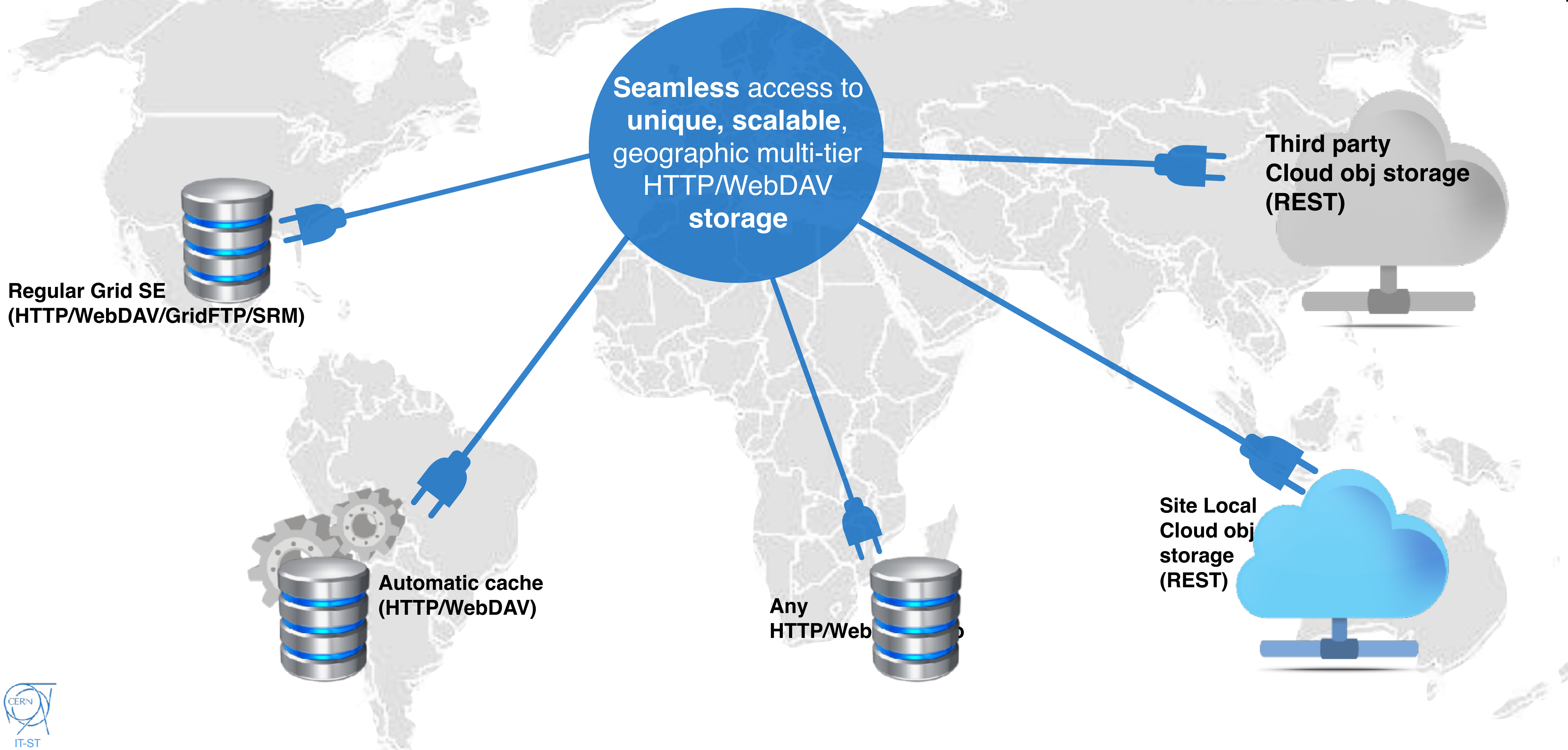
- Can **XCache** extend to…
  - DOE, NSF, and University HPC clusters?
  - Tier 3 clusters?
  - Shared analysis facilities?
  - Cloud-resident clusters?
- No technical reason preventing this!
  - Let's see where we are in Phase 2 or 3
    - Let's be deliberate to assure success!

Data Federations: DynaFed

http://lcgdm.web.cern.ch/dynafed-dynamic-federation-project

Browser-friendly realtime scalable **aggregator** of HTTP/WebDAV/S3/Azure **metadata sources**
Distributed heterogeneous **federation** of **storage** sites as a **single entity** without data movement

★ **Honours the possible volatility of the repo**



IT-ST

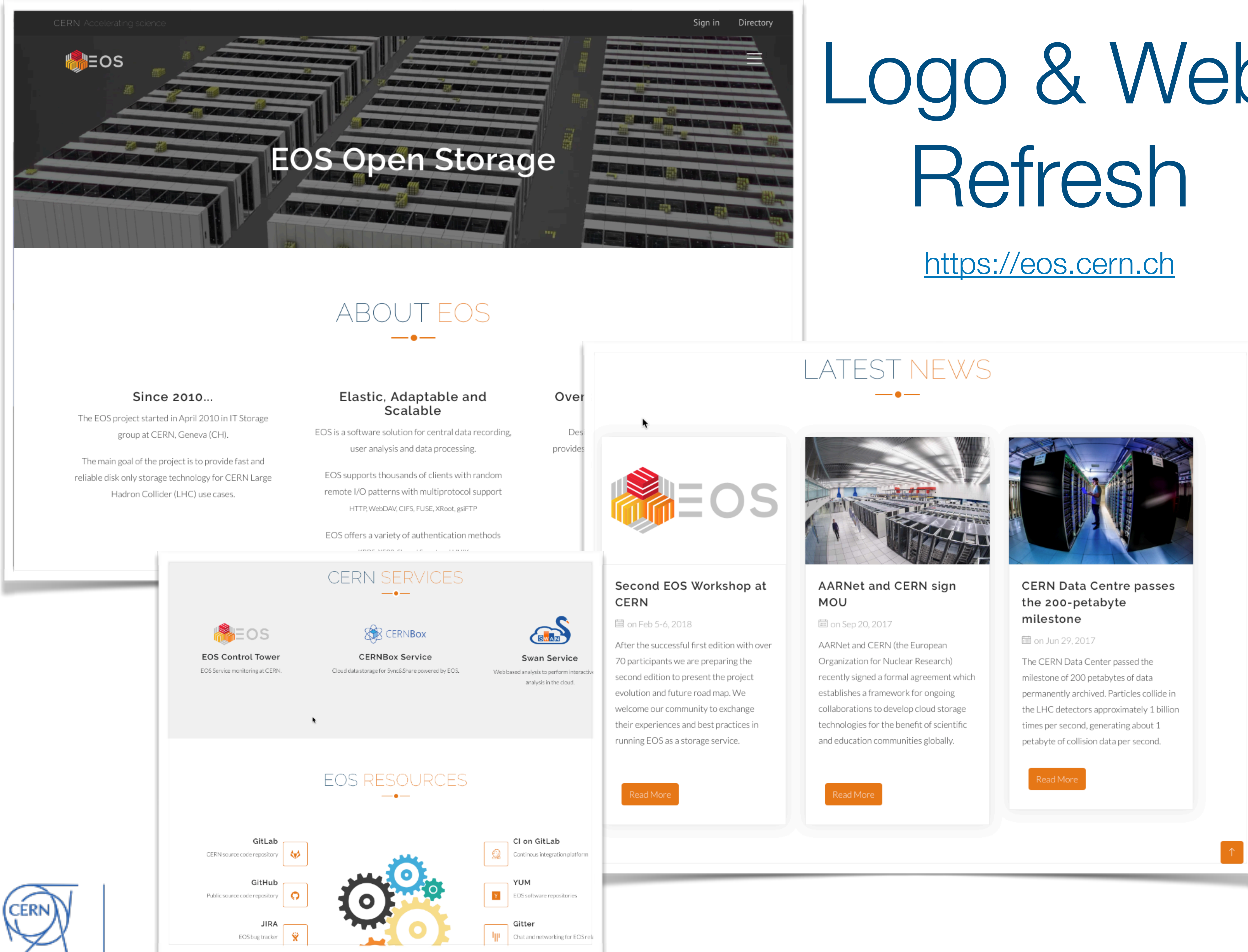
eXtreme Data Cloud (XDC)

- XDC project was funded in EINFRA-21-2017
- 2 year project dedicated to building scalable, federated data infrastructures
  - starting Q1/2018 latest
  - led by INFN
- partial continuation of Indigo DataCloud
- familiar technology providers participate
  - CERN, DESY, INFN

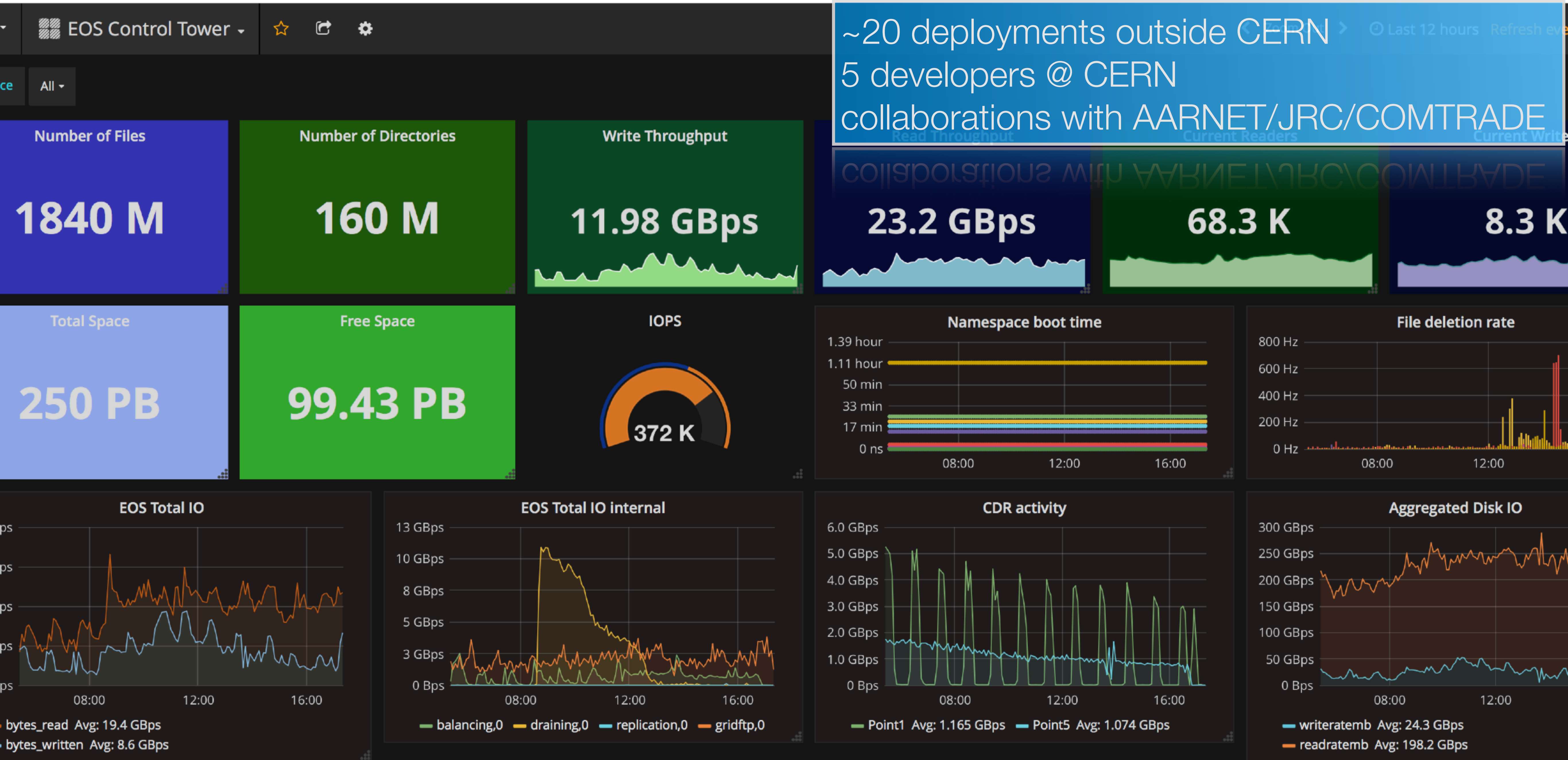
- CERN IT Storage group is actively involved in federation (Dynafed, EOS) and orchestration (FTS) topics

Logo & Web Refresh

https://eos.cern.ch



EOS at CERN

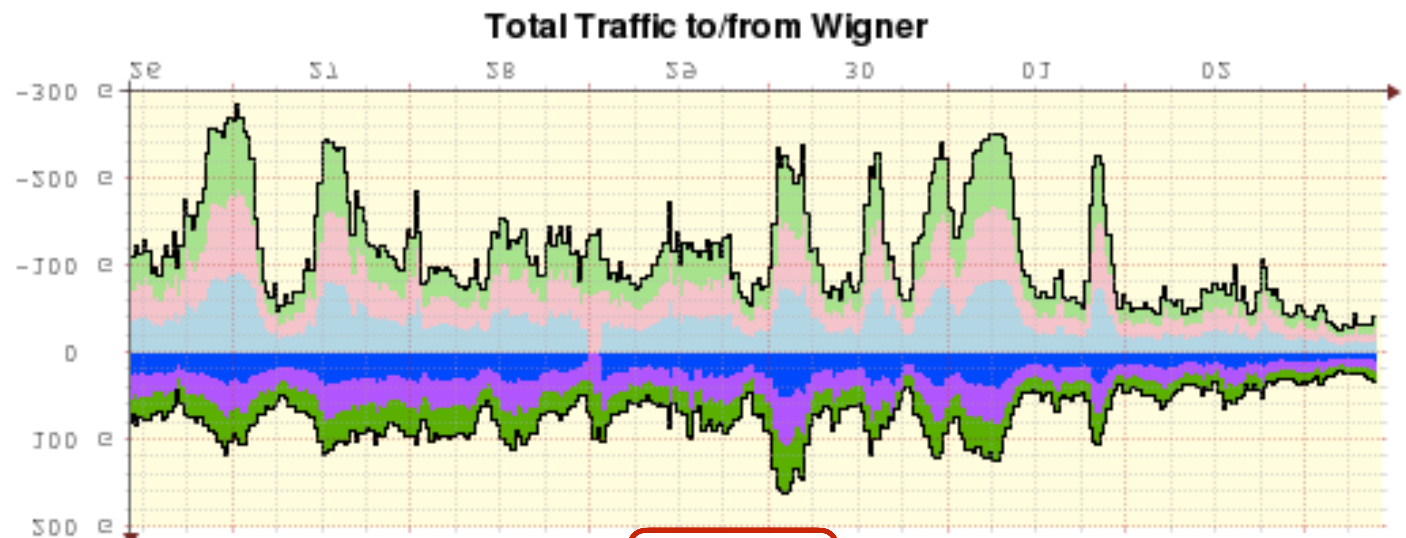


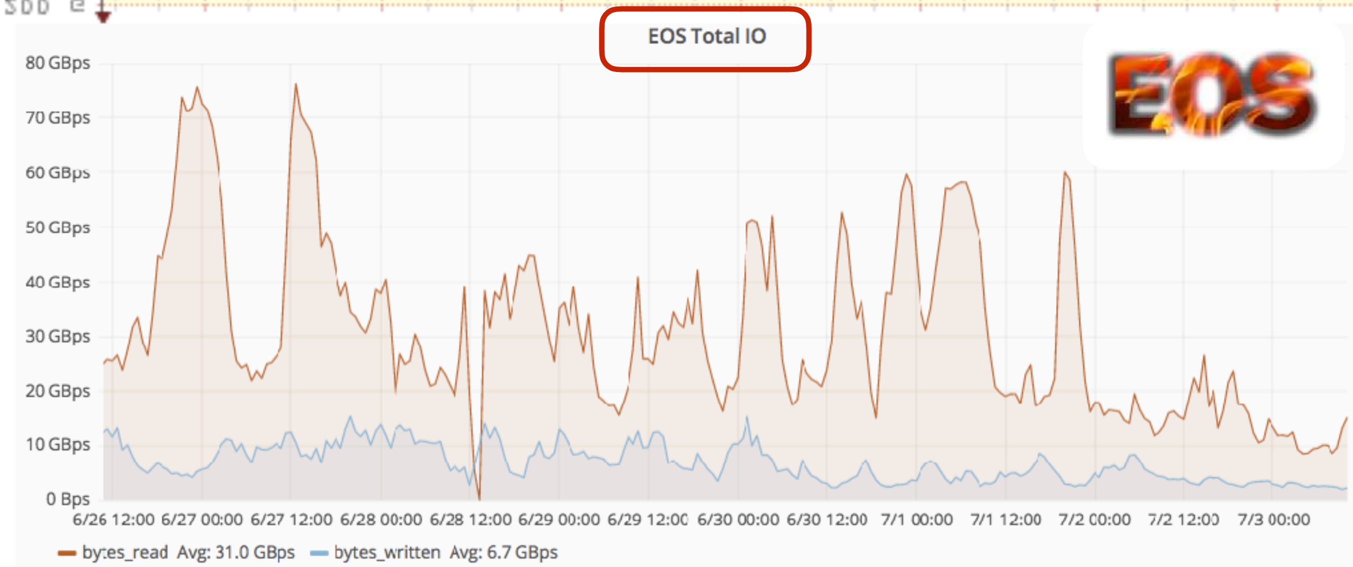
~20 deployments outside CERN
5 developers @ CERN
collaborations with AARNET/JRC/COMTRADE

scrubbing 7.8 Exabyte/year

CH to HU

HU to CH





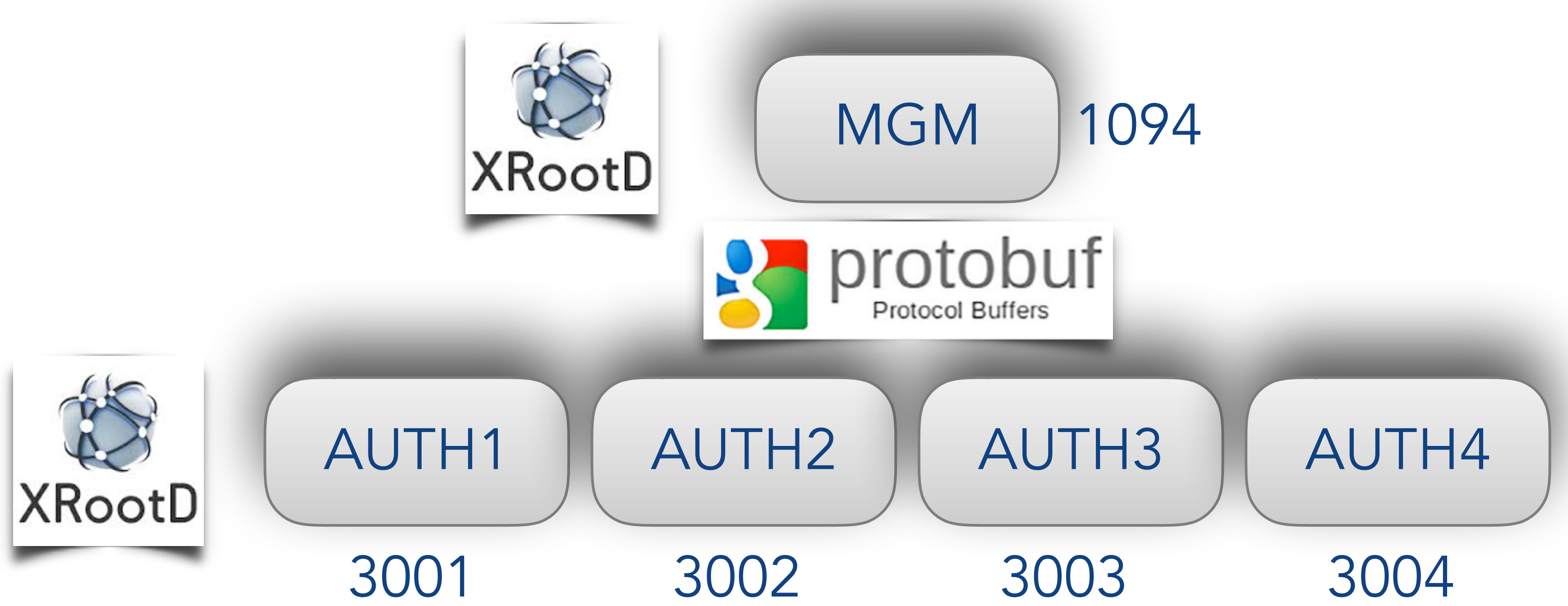
IT-ST

EOS R&D Storage Hardware

- CERN as several other places look into extra-large disk servers
  - Goal: further reduce the contribution of enclosure / server costs
  - Analysis suggests headroom in operational parameters
  - **8 x 24 x 6TB** disks connected to single front-end node [1.2 PB/node]
- Ongoing TCO evaluation
  - capacity/performance ratio
  - OS limitations handling 192 disks
  - RAID vs. ZRAID vs. Software EC
  - suitable network IF
  - CPU type

Improved Authentication Scalability

- G.Ganis boosted XRootD GSI plugin from 200 Hz to 1kHz handshakes
  - to avoid observed namespace performance limit in ATLAS & CMS instances
  - additional manpower coverage in this critical area
- deployed scale-out authentication service in CMS

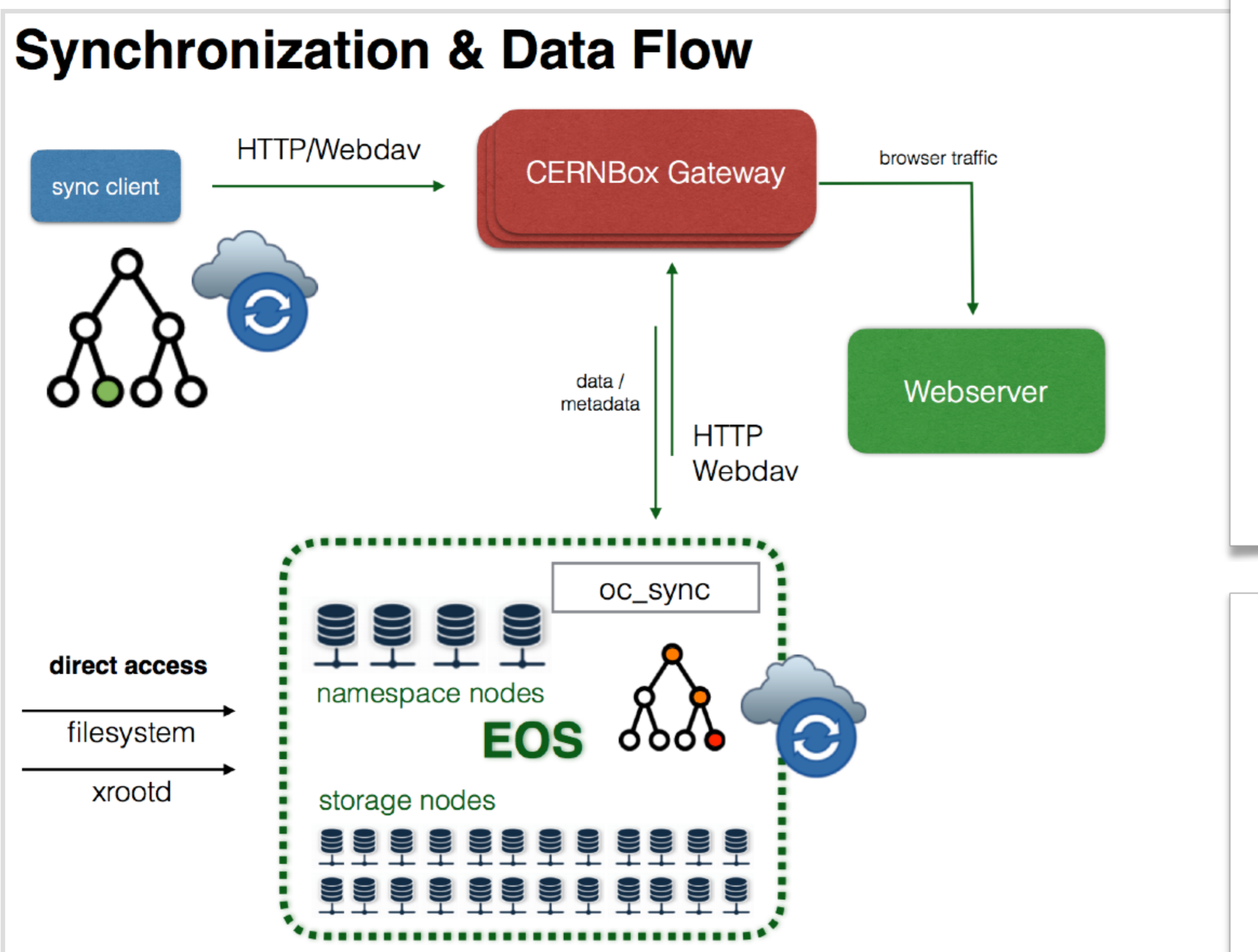


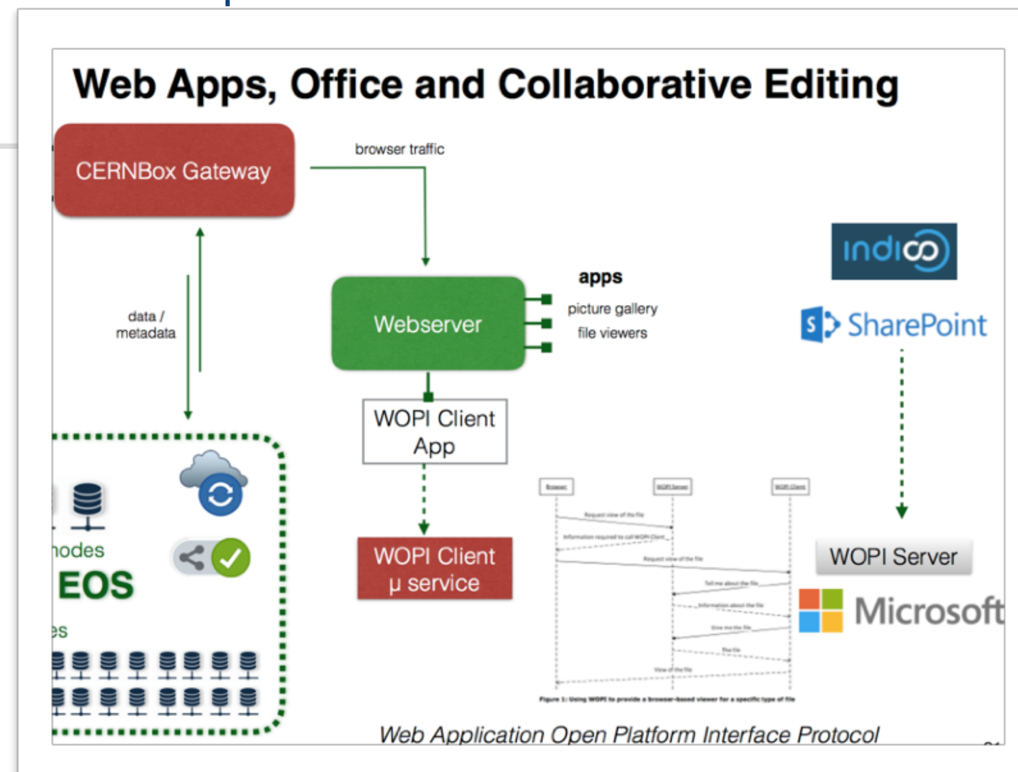

```
AUTHPROXY_0  tcp  --  anywhere            anywhere            statistic mode random probability 0.250000 /* 100 Authproxy probability routing */
AUTHPROXY_1  tcp  --  anywhere            anywhere            statistic mode random probability 0.333333 /* 101 Authproxy probability routing */
AUTHPROXY_2  tcp  --  anywhere            anywhere            statistic mode random probability 0.500000 /* 102 Authproxy probability routing */
AUTHPROXY_3  tcp  --  anywhere            anywhere            statistic mode random probability 1.000000 /* 103 Authproxy probability routing */
```

Simple stateful load-balancing 1094

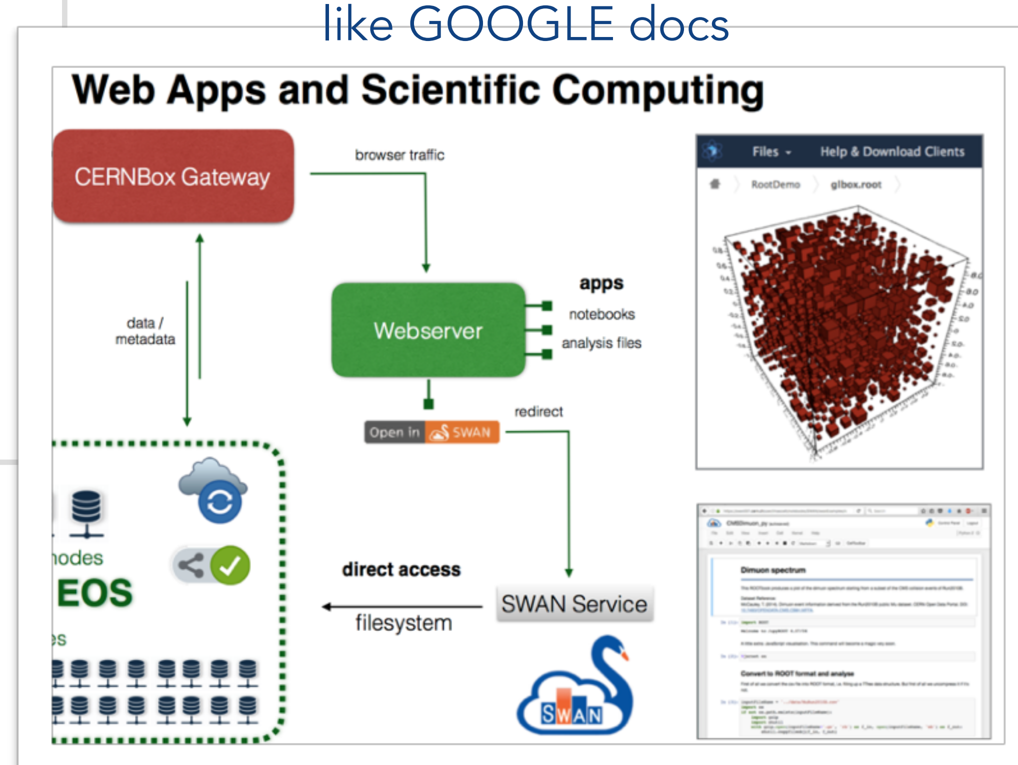
What is CERNBox

Sync & Share Platform = OpenSource Dropbox (OwnCloud)



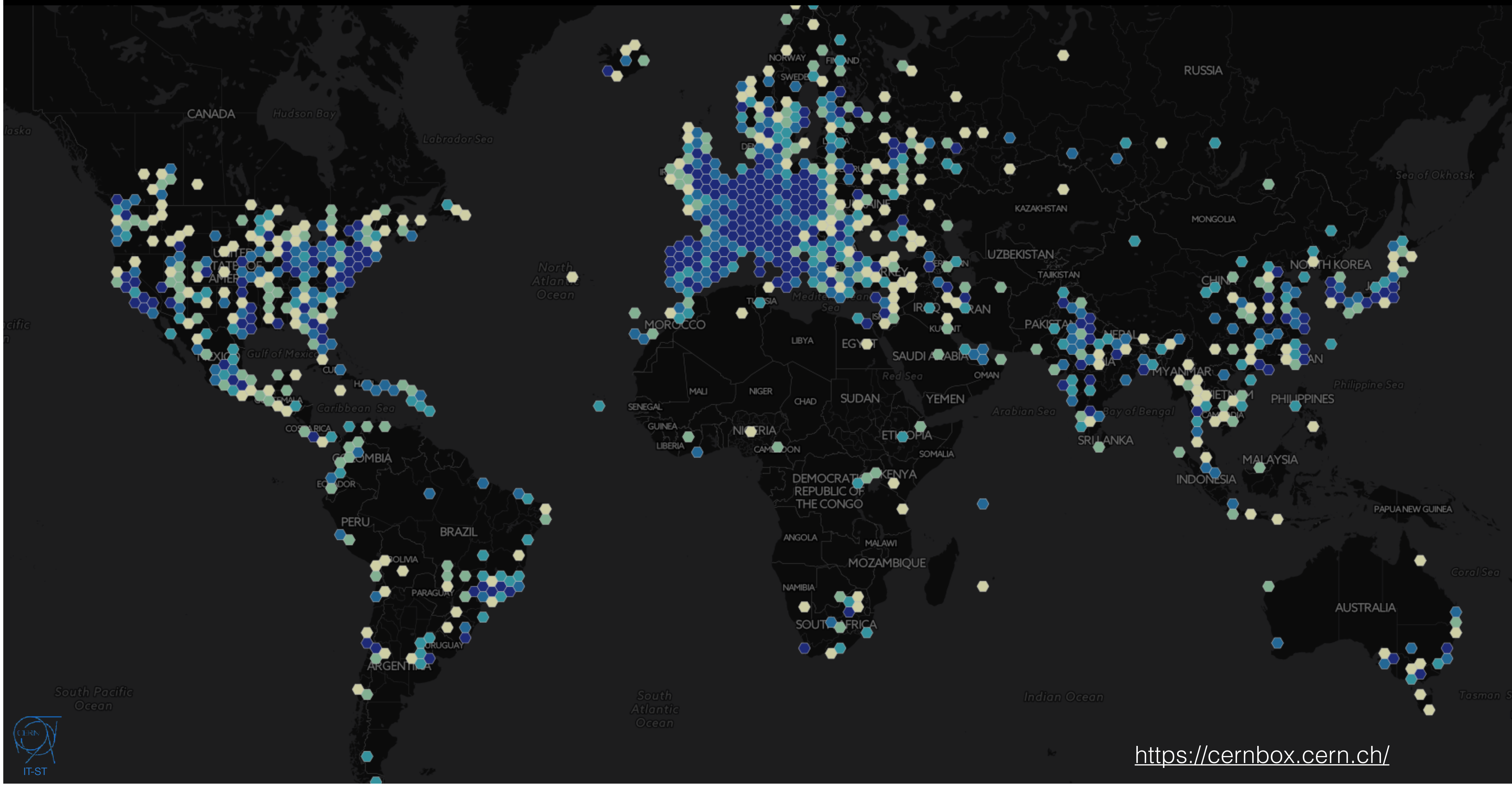


like GOOGLE docs



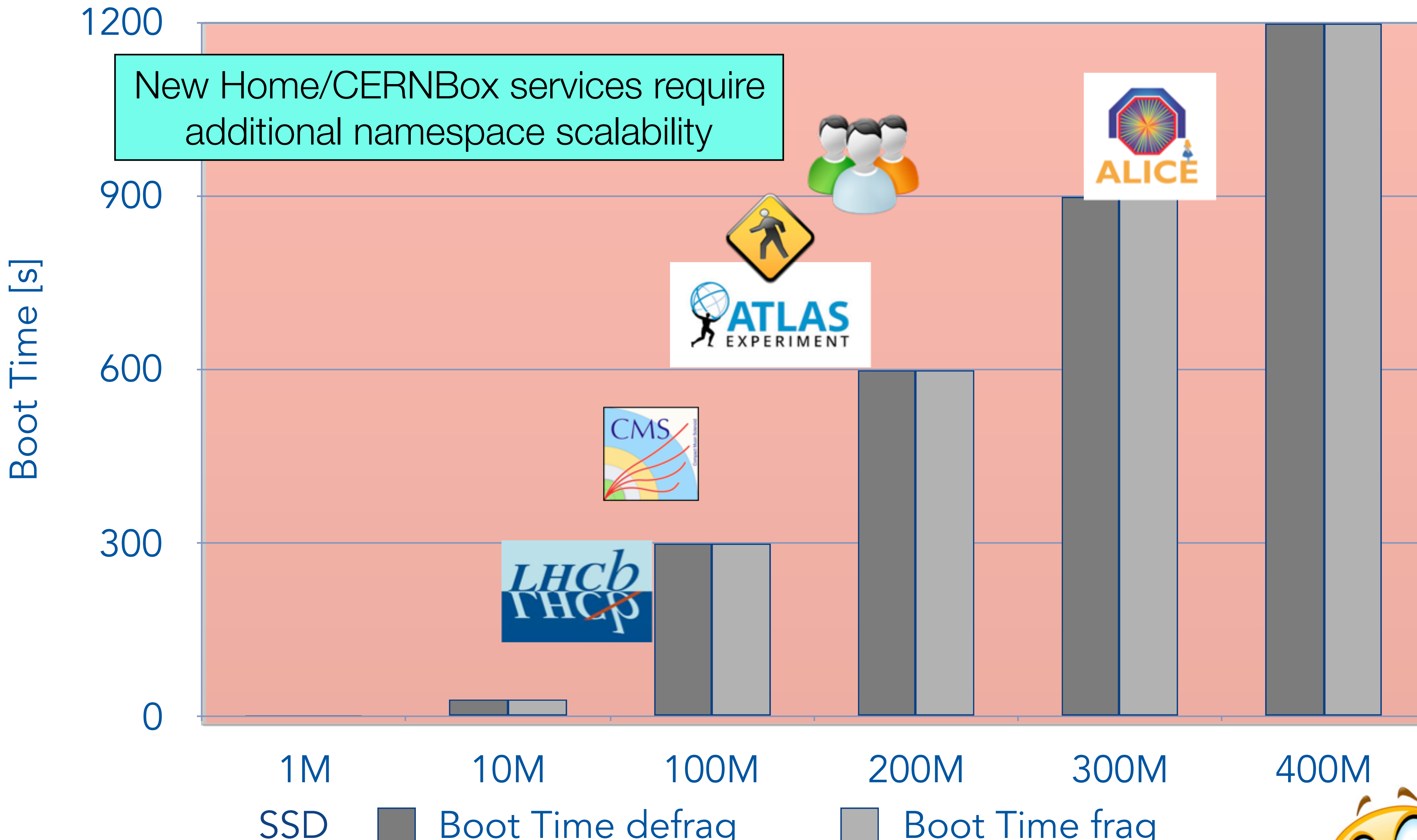
Jupyter Notebooks

CERNBOX flexible data access: store, sync and share



https://cernbox.cern.ch/

New Parallel Namespace Boot



2x - 6x faster boot

QuarkDB

HA Clustered Meta Data

https://gitlab.cern.ch/eos/quarkdb

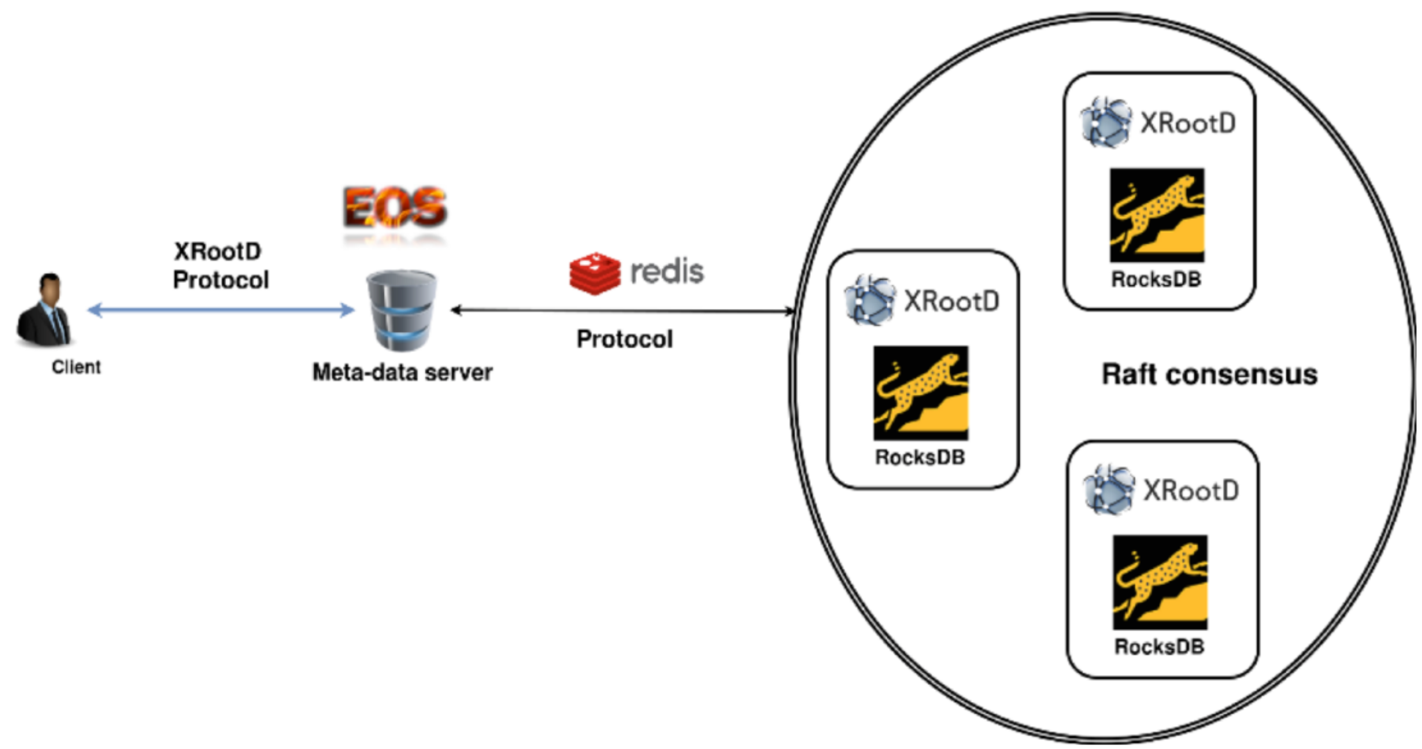
Georgios.Bitzes@cern.ch

Namespace on top of a datastore

- Requirements: **consistent** low latency, scalable, very high rate of writes per second
- EOS to replace AFS at CERN, hold network home directories
- Needs to be reasonably performant for tasks such as
  - interactive usage
  - compiling
  - untarring archives the size of the linux kernel

QuarkDB: a highly-available datastore

- Implement the minimum necessary to keep the system simple
  - QuarkDB runs as a plug-in to the XRootD server framework used by EOS
- A redis-like server on top of RocksDB
  - Support for a subset of the redis command-set: `HASH`, `STRING`, `SET` operations
- High availability through multiple **strongly-consistent** replicated nodes
  - Raft consensus algorithm to keep replicas in sync



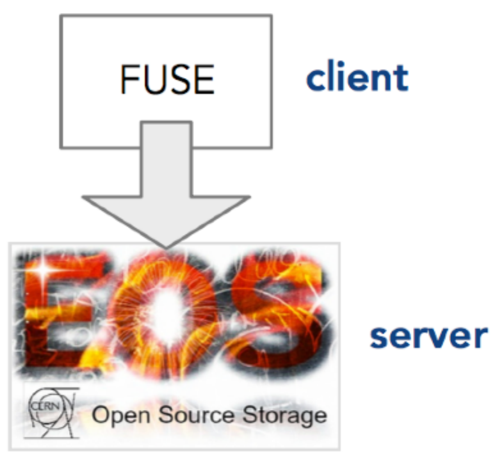
- **10k** lines of C++ (including tons of tests)
- Preliminary benchmarks: peak of **100khz** 200-byte writes, **300khz** reads (non-replicated mode)
- Replicated performance currently **10–15 khz** writes – plans to improve through automatic sharding

EOS FUSE Current Status

/eos mounted on lxplus and lxbatch nodes

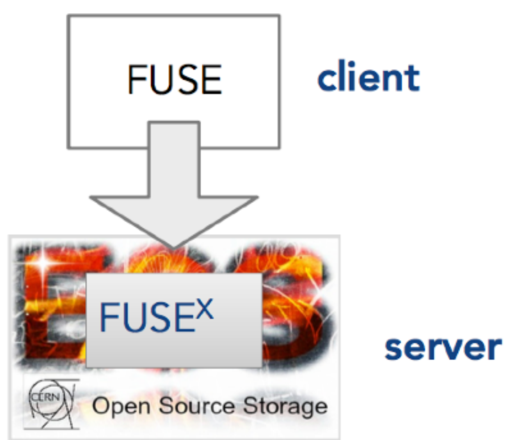
- encountered significant amount of problems and obstacles
  - consistency, stability and kerberos integration
- experience triggered clean rewrite of FUSE client
  - implementation of a filesystem = challenging task !

V2 implementation



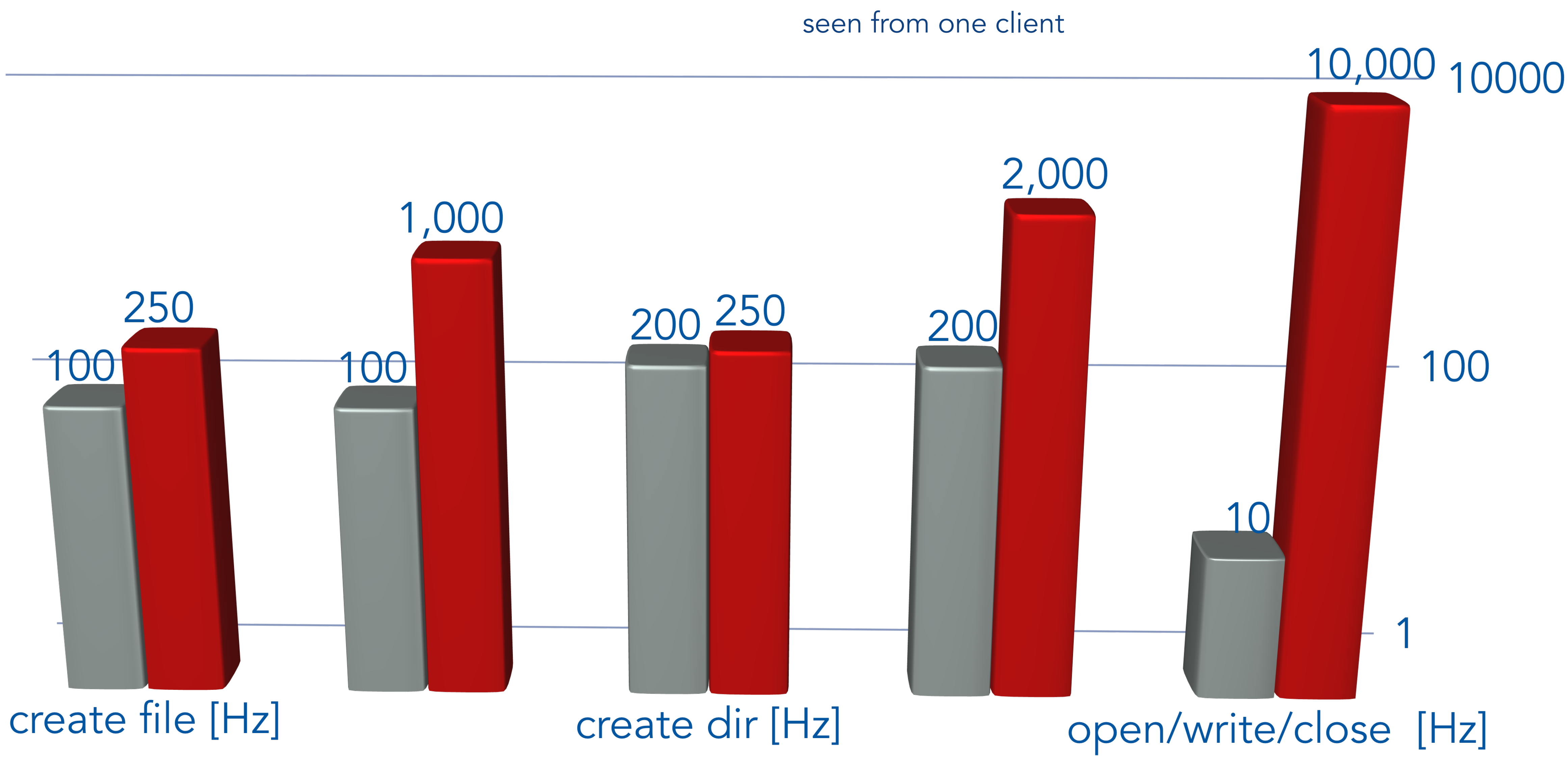
FUSE filesystem implemented as **pure client side** application without dedicated server side support.

V3 implementation



Dedicated server-side support providing a fully asynchronous server->client communication, leases, locks, file inlining, local meta-data and data caching

New FUSE - first performance impressions



SQLITE use case

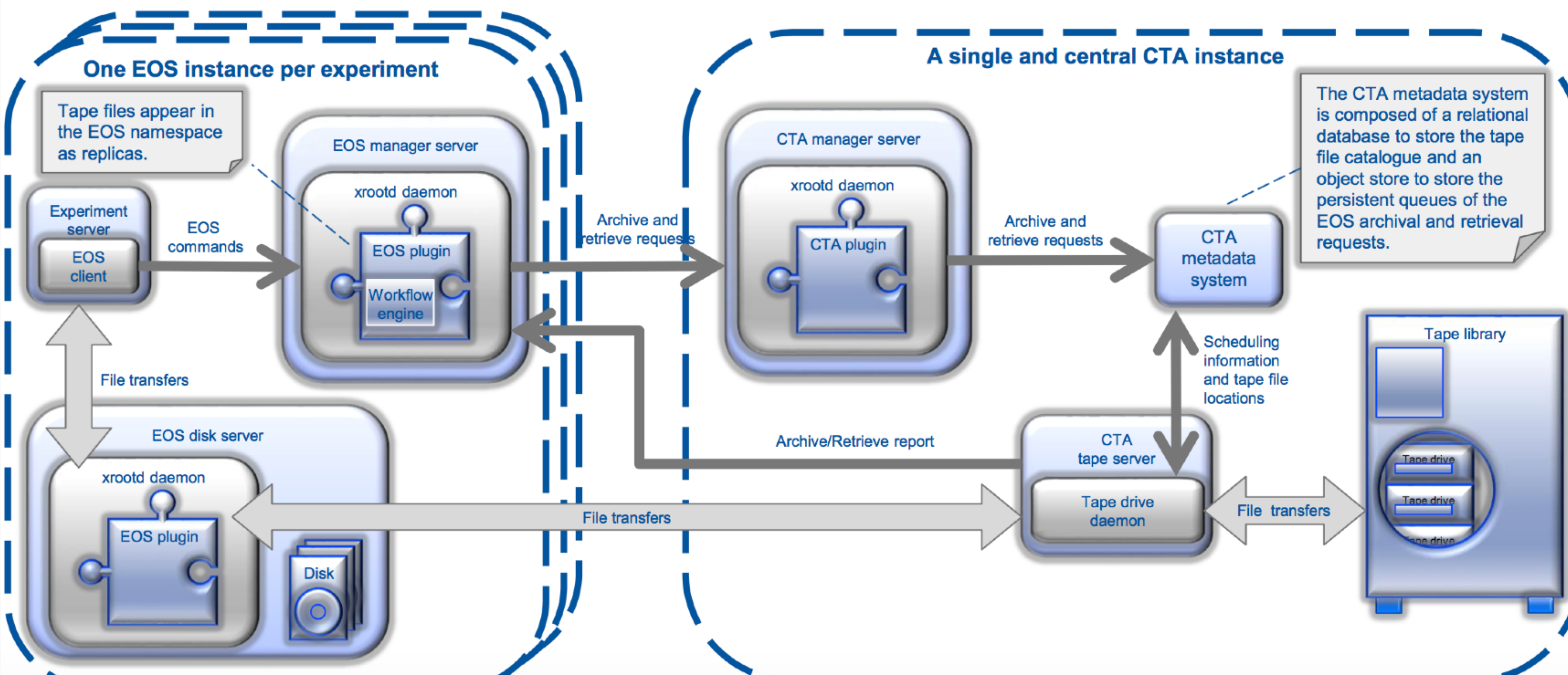
- optional client-side meta-data cache in REDIS
- configurable client-side data cache & file journal - leverage SSDs on batch nodes
- nfs4 exports via kernel nfsd

EOS & CTA server communicate via protocol buffer bus

EOS XRootD ⇄ protobuf Protocol Buffers ⇄ XRootD CTA

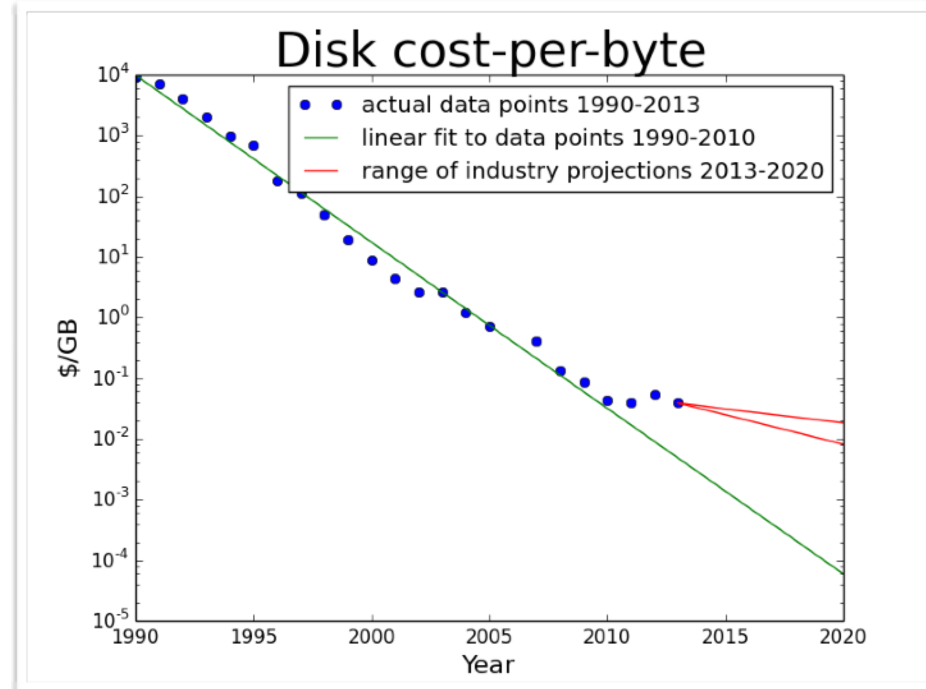
EOS+CTA architecture

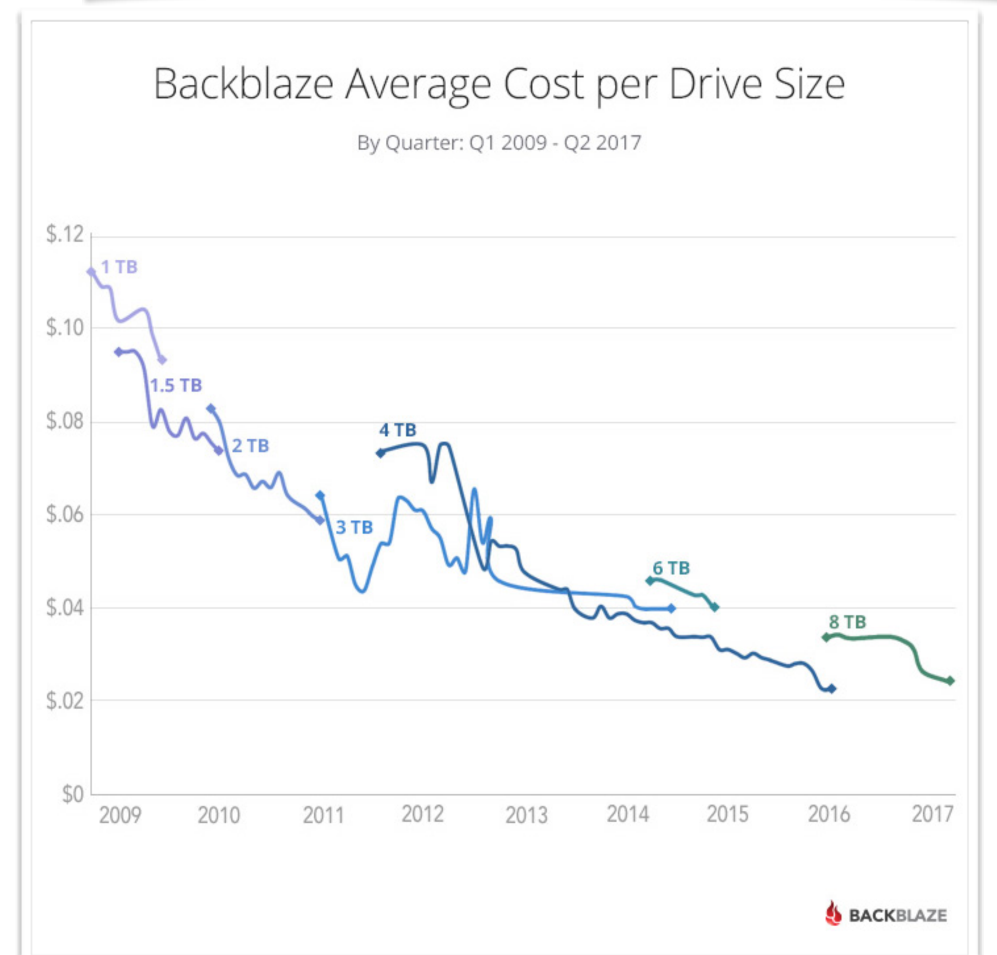
- CTA is integrated with EOS: all user interaction via EOS
- CTA tape files appear in the EOS namespace as file replicas
- CTA contains an internal flat catalogue of all tape files



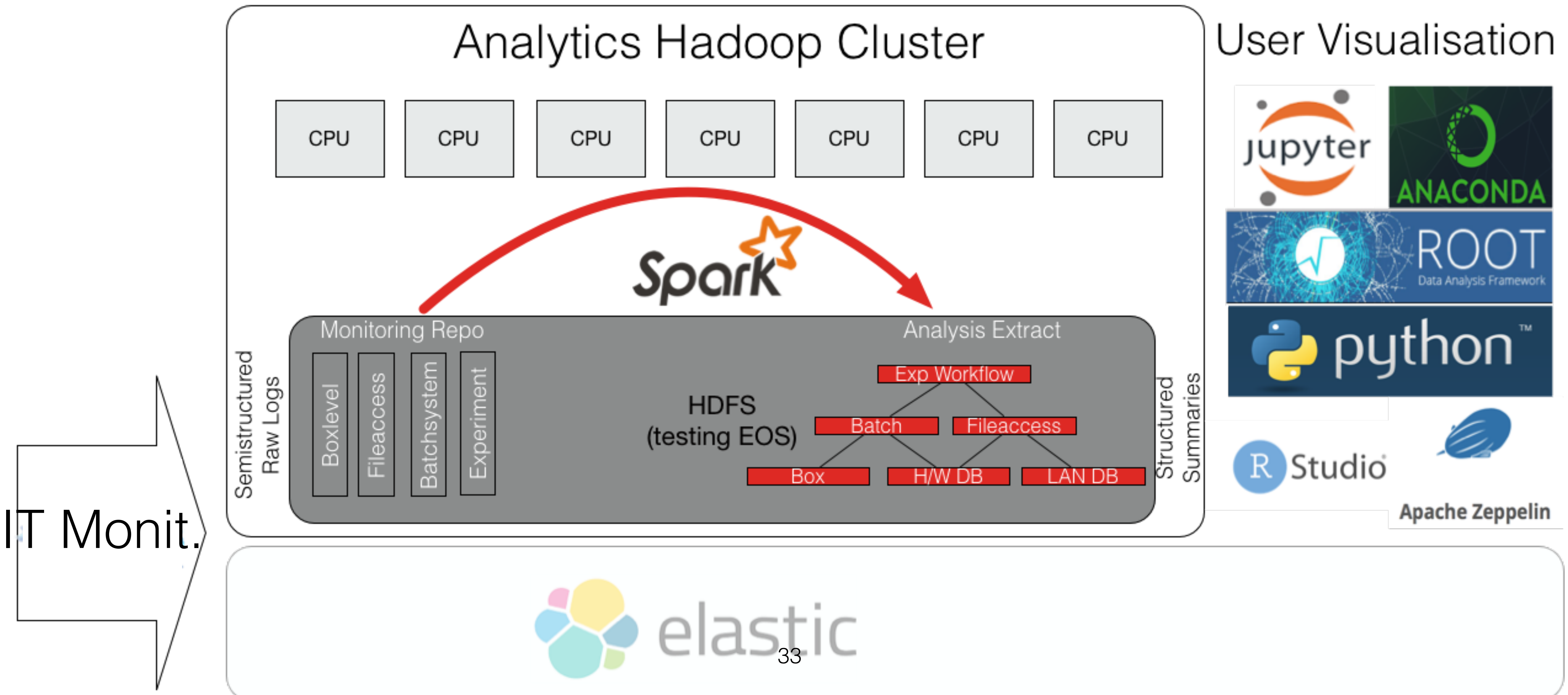
Why Analytics?

- LHC data volume and complexity will increase
- Budgets are expected to stay constant
  - Moore's and Kryder's "law" are slowing down
  - Several disruptive changes ahead commercial clouds, disk->flash->NV memory
- Need to understand cost and be able to predict impact of changes
  - quantitatively instead of just qualitatively
  - absolute (not just relative) numbers

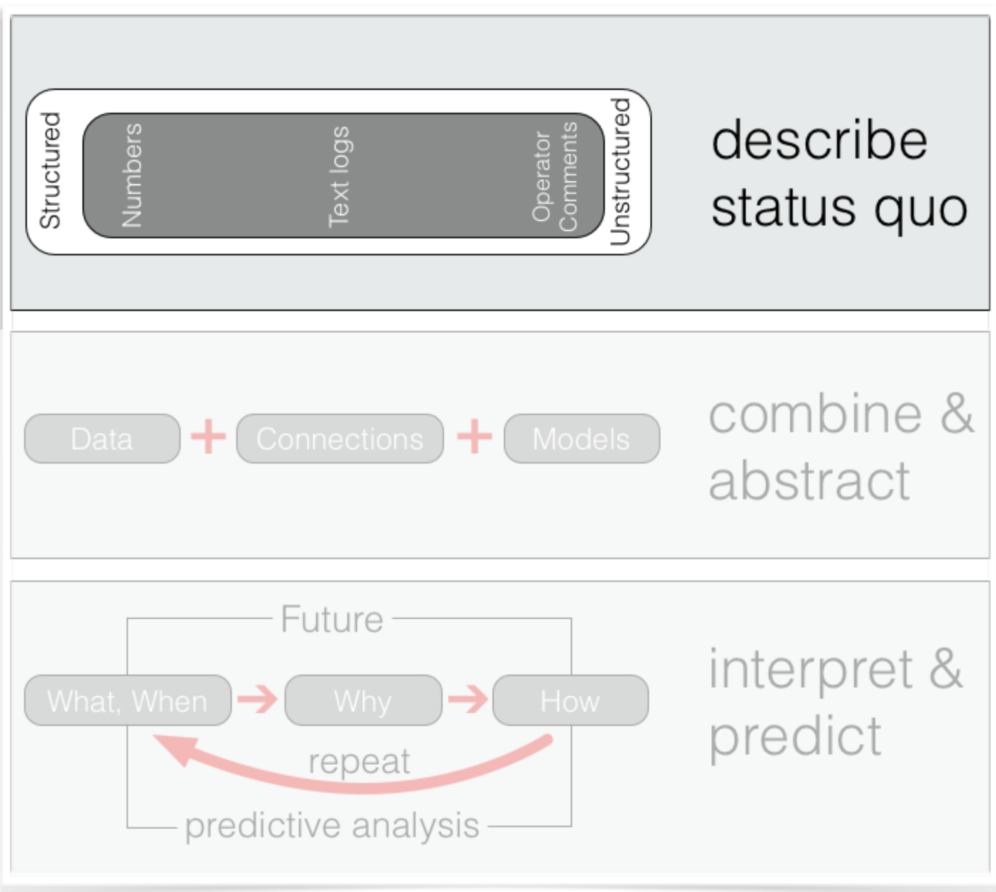




Analytics How?

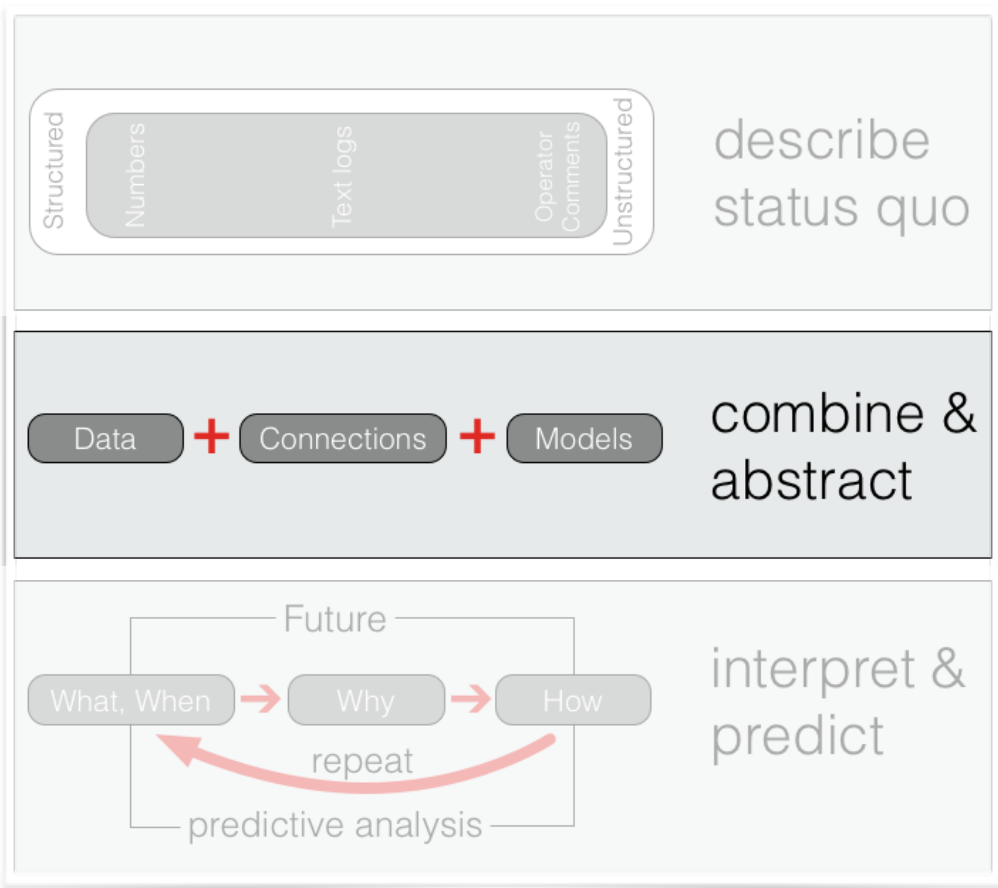


Metric Collection



- **Collection** via *IT monitoring project*
- select and summarise **relevant metrics**
    - Find & remove unexpected / unintended access patterns
- **To what level** can we trust our metrics & assumptions?
    - Evaluate data quality: eg accuracy, units(!)
        - data that has not been used quantitatively likely still has problems
- **Quantitative cross-checks with 2016 data**
    - CPU consumed in LSF (Condor analysis upcoming)
        - $\sum job_{cpu} \sim \sum sched_{cpu} \sim \sum host_{cpu}$ in rough agreement (within 15%) per step
    - Data from and to disk in EOS
        - $\sum$ disk I/O ~ $\sum$ user I/O + $\sum$ internal I/O in good agreement (within 8%)

Connecting Data

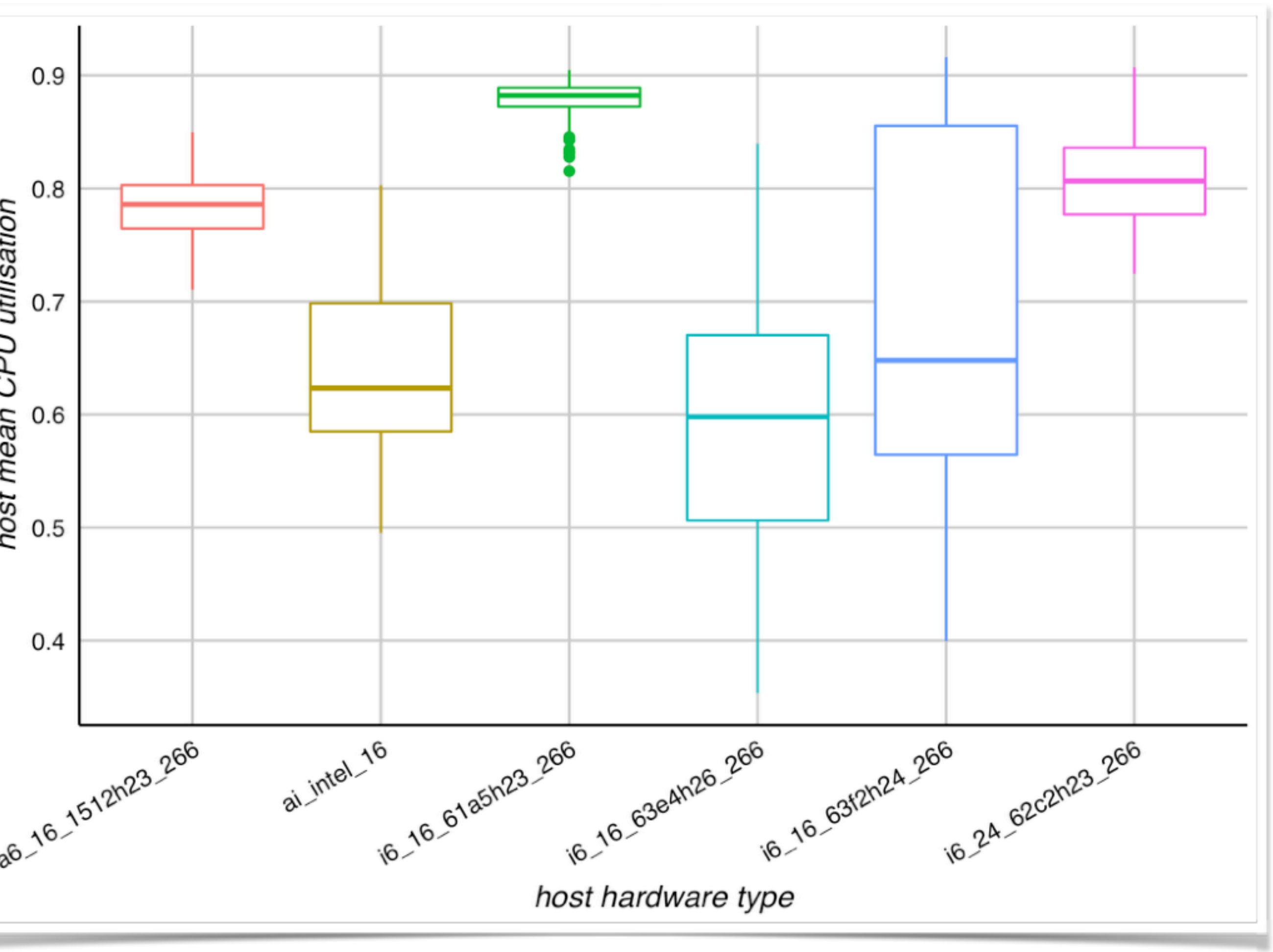


- Involved in several experiment performance studies
  - Starting point: why do users/service providers see:
    - **slow** file **access**? **inefficient CPU** usage?
    - differences: Wigner vs CERN, CERN vs T1, etc..
  - **where is** the **bottleneck? where should be?**
- Connected data from experiment, storage, batch
  - connected infrastructure data: LAN db, hardware db
  - enables correlation with location, hw type, HEPSPEC

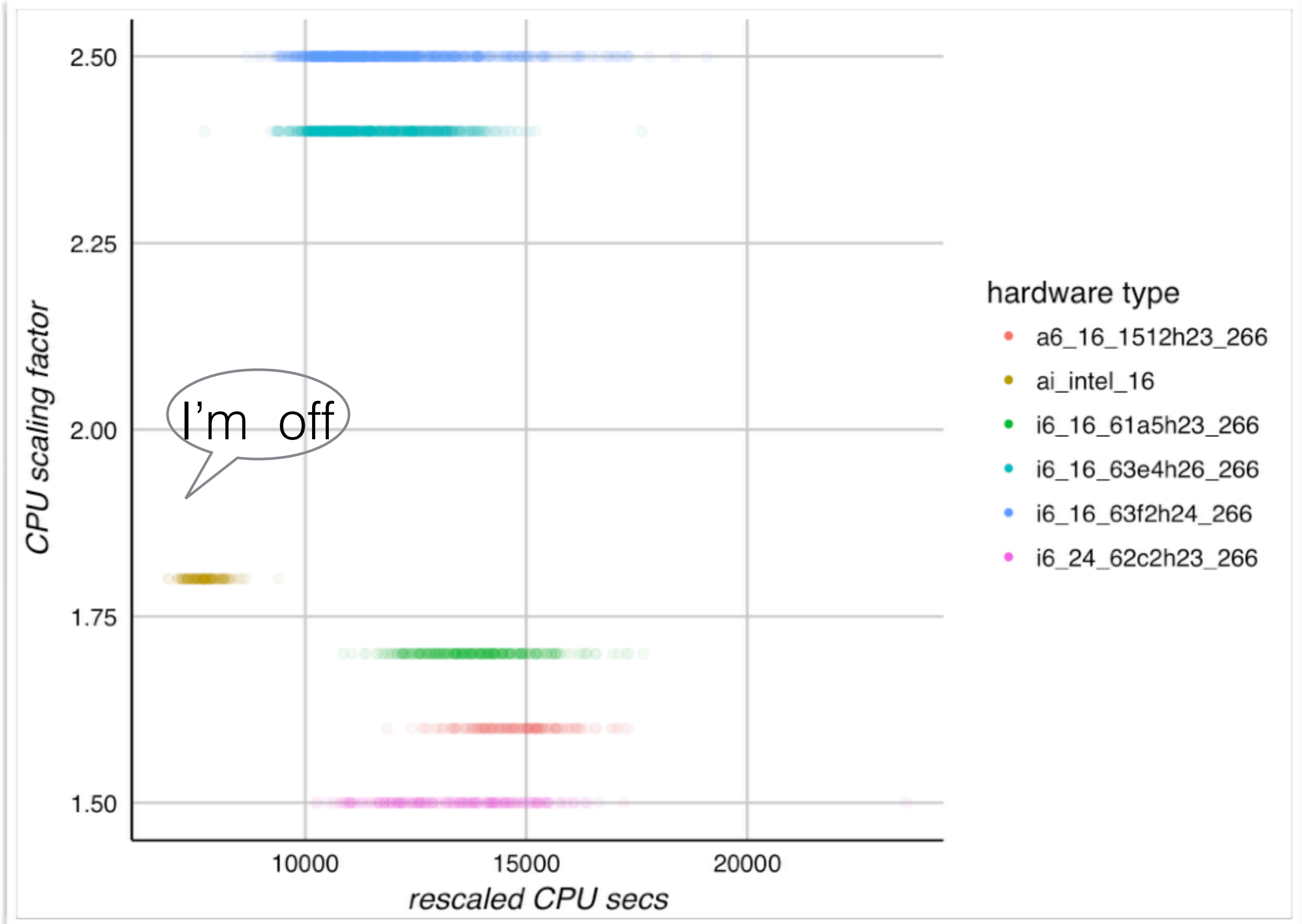
Examples: One production task

data from Alice production in 2016

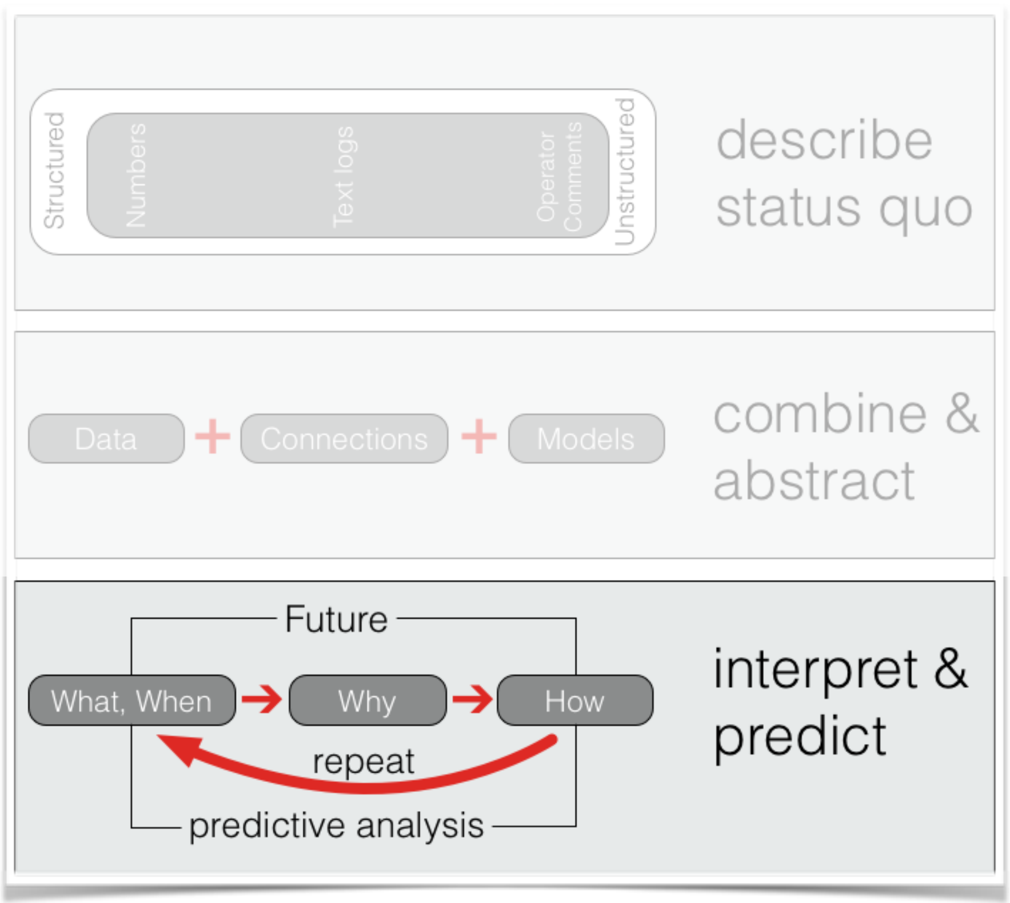
CPU “Efficiency” versus H/W types



CPU Calibration check



Model Predictions

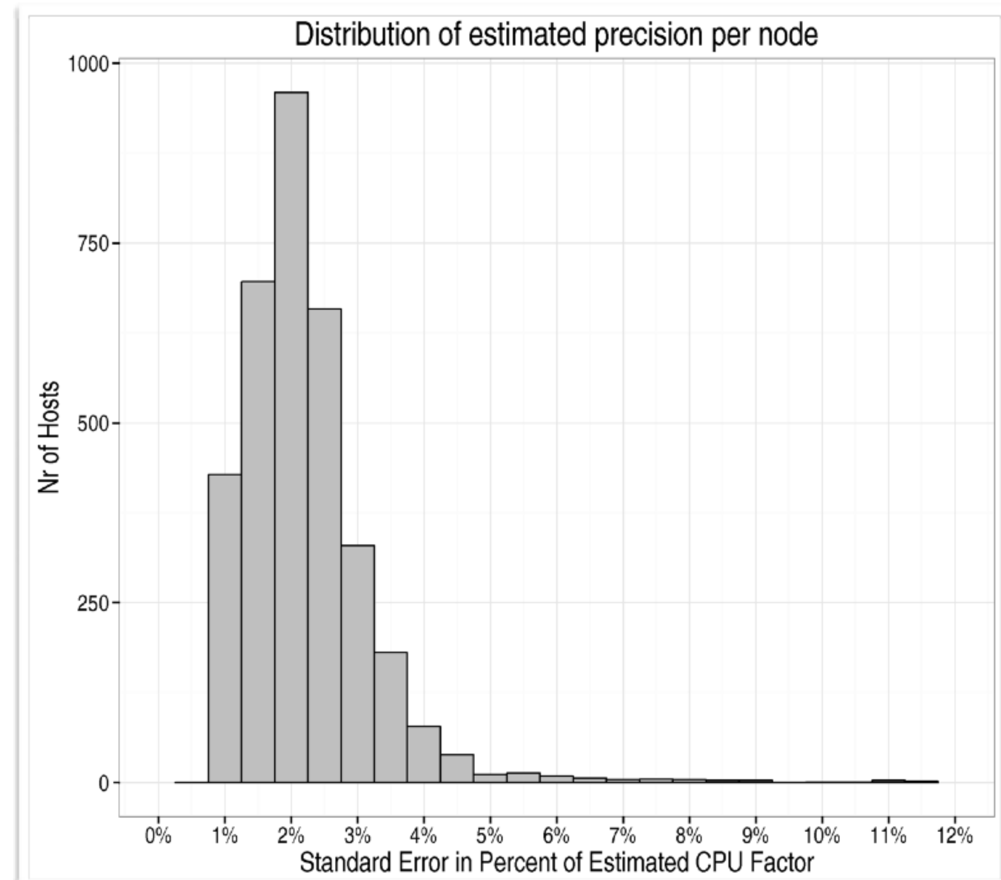


- Answer via predictive models:
  can we construct a more performant system for the same price?
- Simplest case: CPU-bound jobs
  - **CPU & RAM speed** => MC throughput
- More balanced case:
  - need to consider:
    CPU, local I/O, LAN I/O, WAN I/O, network speeds

Passive Benchmark

PhD C. Nieke, TU Braunschweig presented @ IEEE Cloud 2017

- Basic idea: use the real workload as ‘benchmark’
  - Assumption: jobs in a give task are equal
    - estimate rel. CPU performance per task by comparison of runtime
    - data from on existing monitoring logs (time and evt number)
- Advantages
  - No intrusion or overhead
  - Coverage of all used nodes
  - Prediction stays representative wrt changing workload
- Application
  - Observe performance during operation
  - Compare configurations by performance on the actual workload
- Accuracy & Precision
  - Experiment on LSF dataset: ATLAS and CMS, 3 months
  - Equal or better prediction of performance than HepSPEC06
  - Precision per node is below 5% error for 98% of nodes



Analysis of Disk Failures

Collaboration w SSRC, UC Santa Cruz and Backblaze

- Failures on CERN sample of 70 k disks (similar O(backblaze))

  1: failure impact on service performance

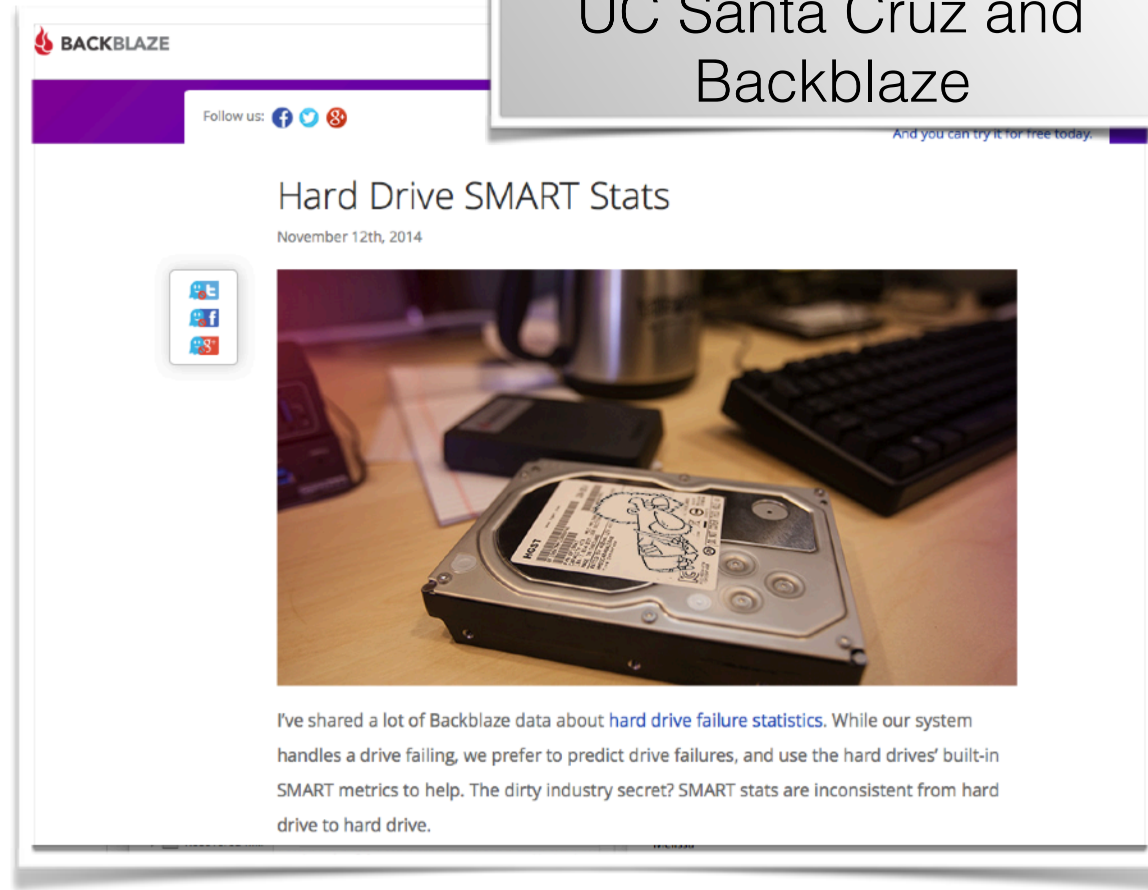
  2: comparison of enterprise and “consumer” disks

  between different use cases

  3: predictive maintenance

  following ML approach as in IBM study [KDD 2016] and [MSST 2017]

- Prototype of smart counter collection in place since 6 months
  - smart status and relevant counters
  - disk replacement totals - logs are becoming available
    - hardware repository: describes purchase, but not status quo
  - Focus on EOS cluster disks



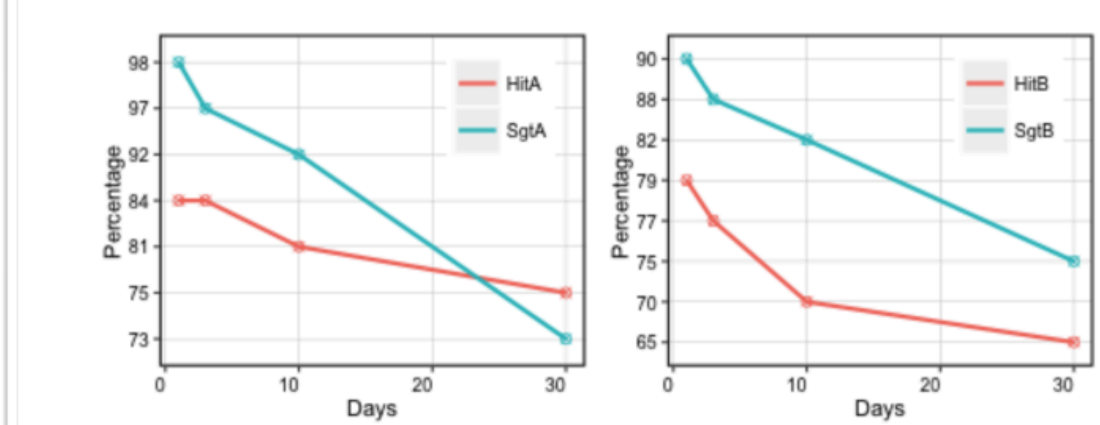
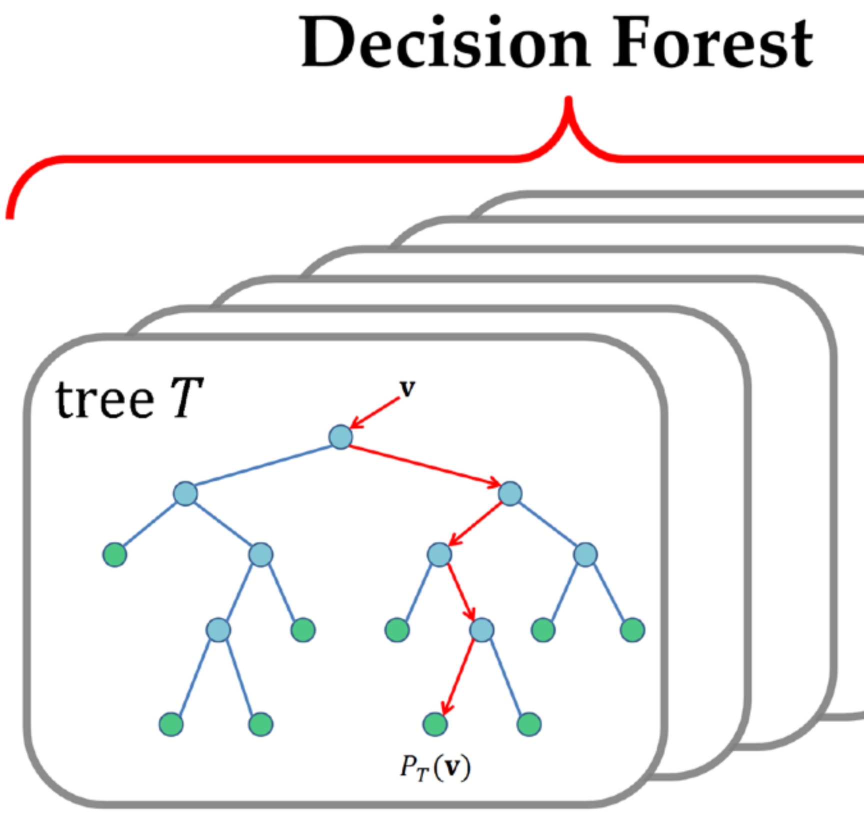


Figure 6: Percentage of disks correctly predicted as replaced on snapshots taken 1,3,10 and 30 days before the actual replacement event.

Batch job vs WN hardware classifier

- Can we automatically classify jobs wrt to ?
- Idea: a job is either a) CPU-bound, b) WN-I/O bound, c) server I/O-bound
  - depending of class, it profits from a) faster CPU, b) WN with SSD c) more file replicas on backend service
  - main metrics: experiment & task ID, WN I/O and LSF CPU stats, EOS I/O, CPU and disk type
- Convenient classifier are random forests
  - eg “party” R-package [http://party.R-forge.R-project.org]
  - more stable results and less manual effort than cut based approach across a larger number of different job types
- Classifier output can be used as optimisation hints for
  - file replication: eg these files (don’t) need additional replicas
  - job placement: eg these jobs (don’t) need a local SSD
  - hw planning



Summary

- Data steering group has identified focus topics to organise an active discussion across experiments, sites and technology providers
- CPW effort is documenting the main architectural options to achieve stringent storage requirements of upcoming LHC runs
  - Data technology is key factor in several CWP R&D proposals
- Update on storage developments at CERN
  - namespace availability and next generation FUSE binding are focus areas
- Infrastructure Analysis Working Group [meetings] and [twiki]
  - Metrics collection and analysis environment are now in place and have allowed first quantitative studies
  - Resource/budget pressure and increasingly available ML training are expected to provide an addition boost of this activity

CERN