
CERN

IT-ST: The Storage group

Alberto Pace on behalf of the IT-ST group

The Storage group

Mandate

- Ensure a coherent development and operation of storage services at CERN for all aspects of (physics) data
- Three goals
  - Storage
  - Distribution
  - Preservation

Three sections

- Analytics & Developments (IT-ST-AD)
  - Design and develop central storage services and their evolution. Operate and support the Analytics and FTS services
  - Leader: Dirk Duellmann

- File & Disk Operations (IT-ST-FDO)
  - Operate and support the storage and file system services for physics (AFS, CEPH, CERNBOX, CASTOR and EOS)
  - Leader Massimo Lamanna

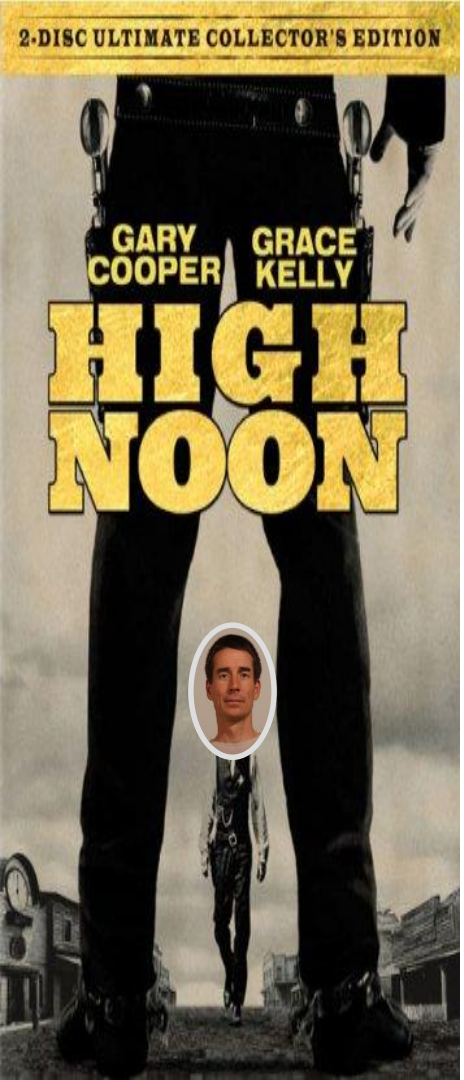
- Tape, Archives & Backups (IT-ST-TAB)
  - Design, operate and support the archive and backup services (includes robotics, drive and media, infrastructure for backup).
  - Leader: German Cancio Melia

IT-ST-FDO

Presented by Xavier Espinal

Storage Services for Physics and more

Storage Services for Physics and more
IT-ST-FDO

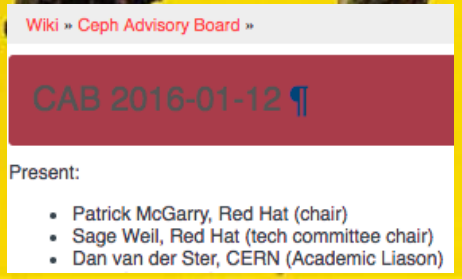


AFS (phase-out)

- Difficult task, but achievable
  - 520 TB (+16% this year vs 34% last year), 3.5 billions files (+16% vs +20% last year)
- We have solutions for nearly all current AFS usages
  - Successfully migrating **web/projects/personal files** to **EOS** and **software** to **/cvmfs**
- No hard line to “unplug” (Run3?)
- Starting a discussion with all relevant service managers concerning **home directories**:
  - lxplus, lxbatch, linux/win/Mac support, IT-CDA, IT-CM,…

Filer service

- The current Filer service solves the native NFS (posix) needs
- 80TB allocated and +30 projects:
  - Twiki, Puppet, MIC,LSG, gridCE, myproxy,gitlab, boinc, openshift, Indico, Inspire, etc.
- Currently running on a single NFS server over CEPH-RBD does not provide High Availablilty. We are evaluating possibilities to provide HA:
  - pNFS cluster
  - CEPH-FS



ceph cern

openstack

Beesly (5 PB + 433 TB, v0.94.9):
- Cinder (block storage volumes)
- Glance (images repo)
- Rados GW (object storage interface: S3, Swift)

Dwight (0.5 PB, v10.2.3):
- Preprod cluster for development (client side)
- Testing, upgrades and crazy ideas

Erin (4.2 PB, v10.2.3):
- New cluster for CASTOR: disk buffer/cache in front of tape drives

HPC

Flax (0.4PB, v.10.2.3 - early stage):
- Ceph-FS HPC cluster for QCD studies

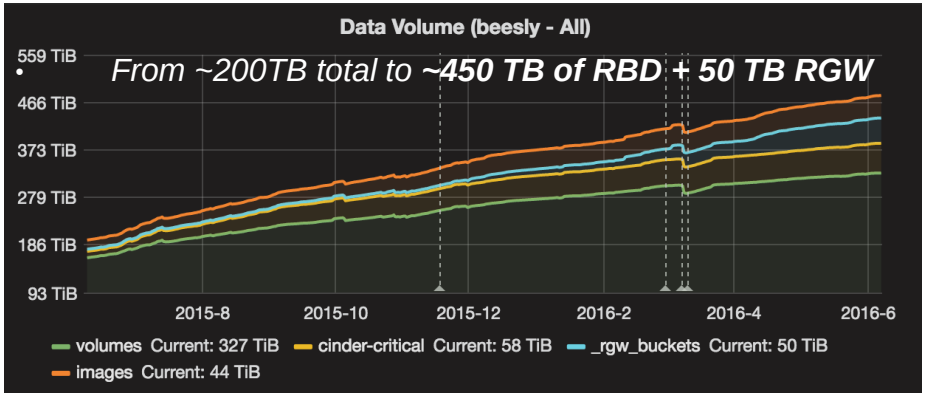
S3

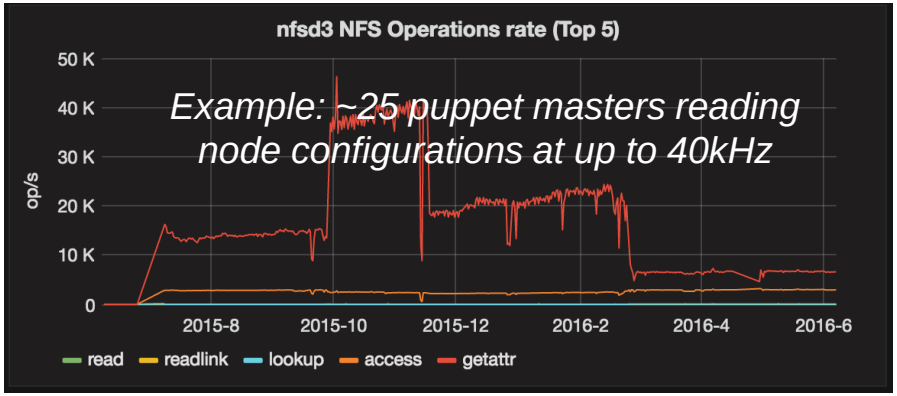
Gabe (1PB, v.10.2.3):
- New S3 Object Store IPV6 only

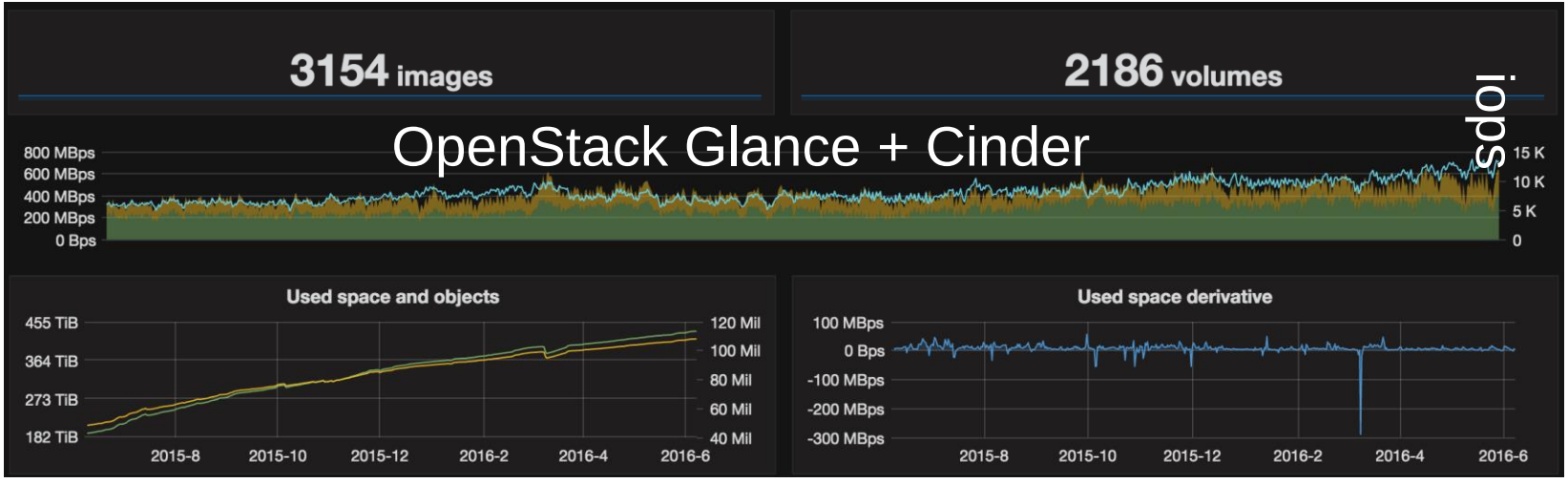
Bigbang (~30 PB, master):
- Playground for short term scale tests
- Usually when we receives new hardware

openstack

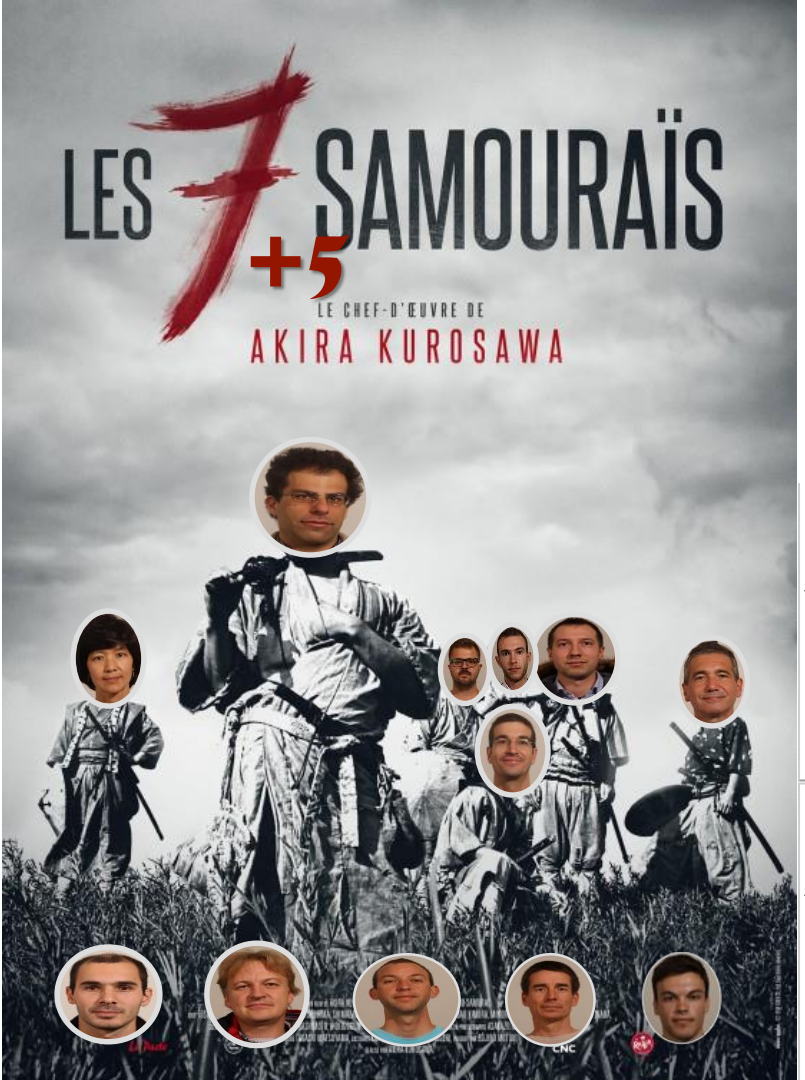
ceph cern







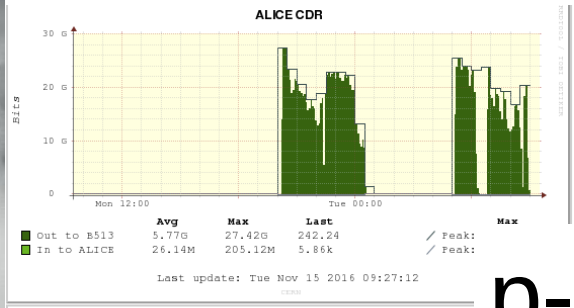
CERN

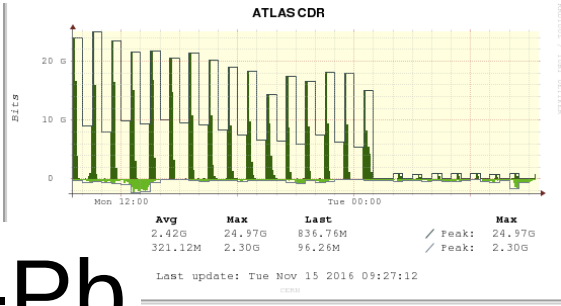


Large Scale Storage for LHC data

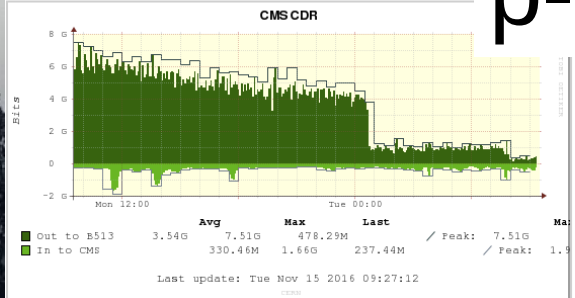
CASTOR
CERN Advanced STORage manager

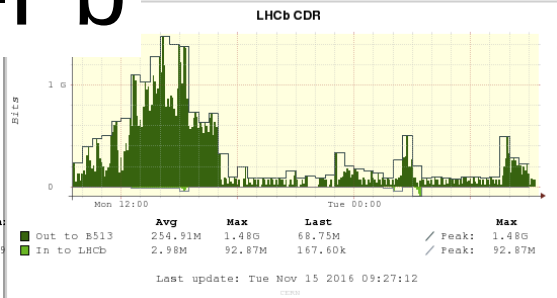
EOS



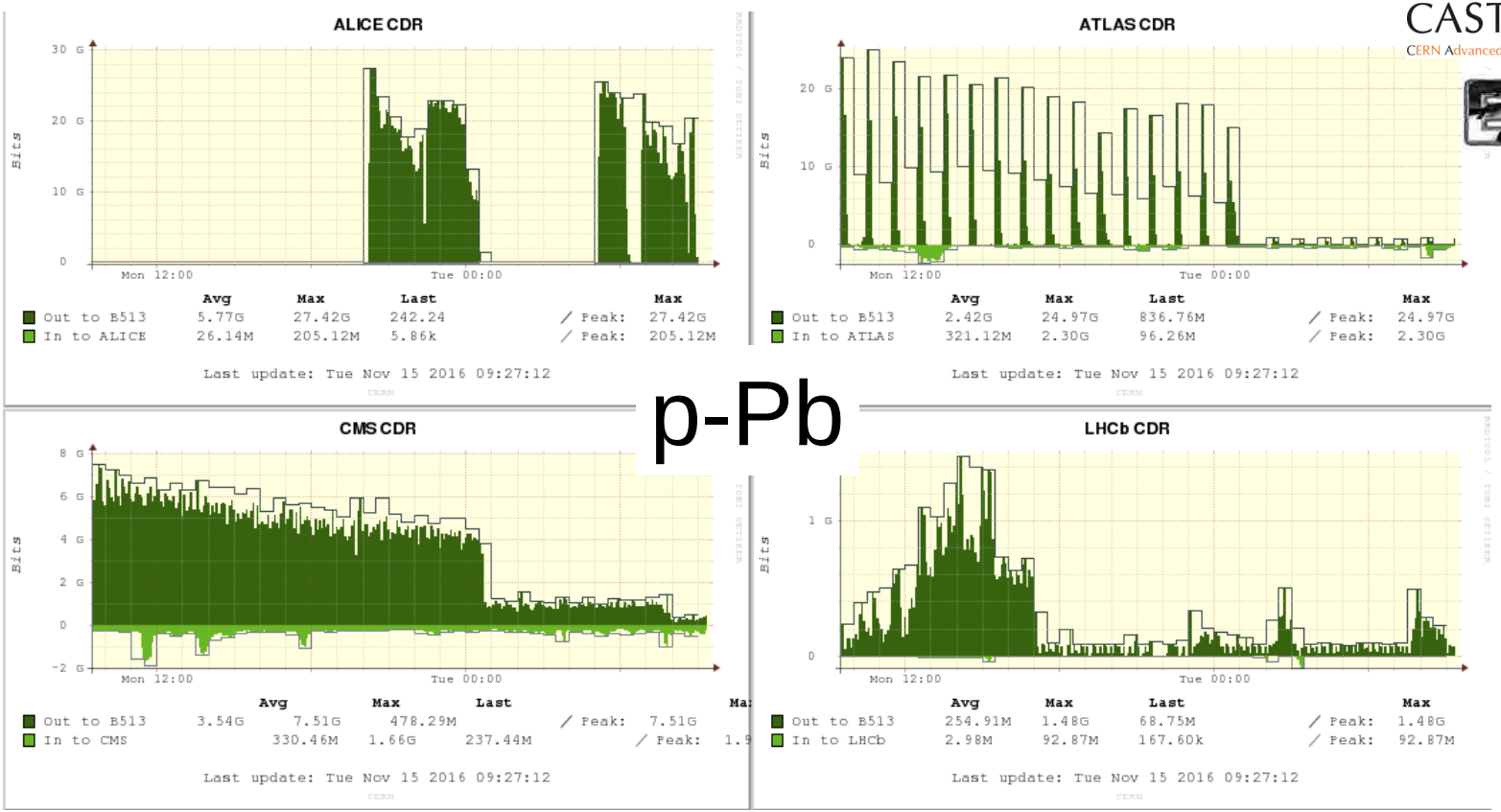


p-Pb





Large Scale Storage for LHC data



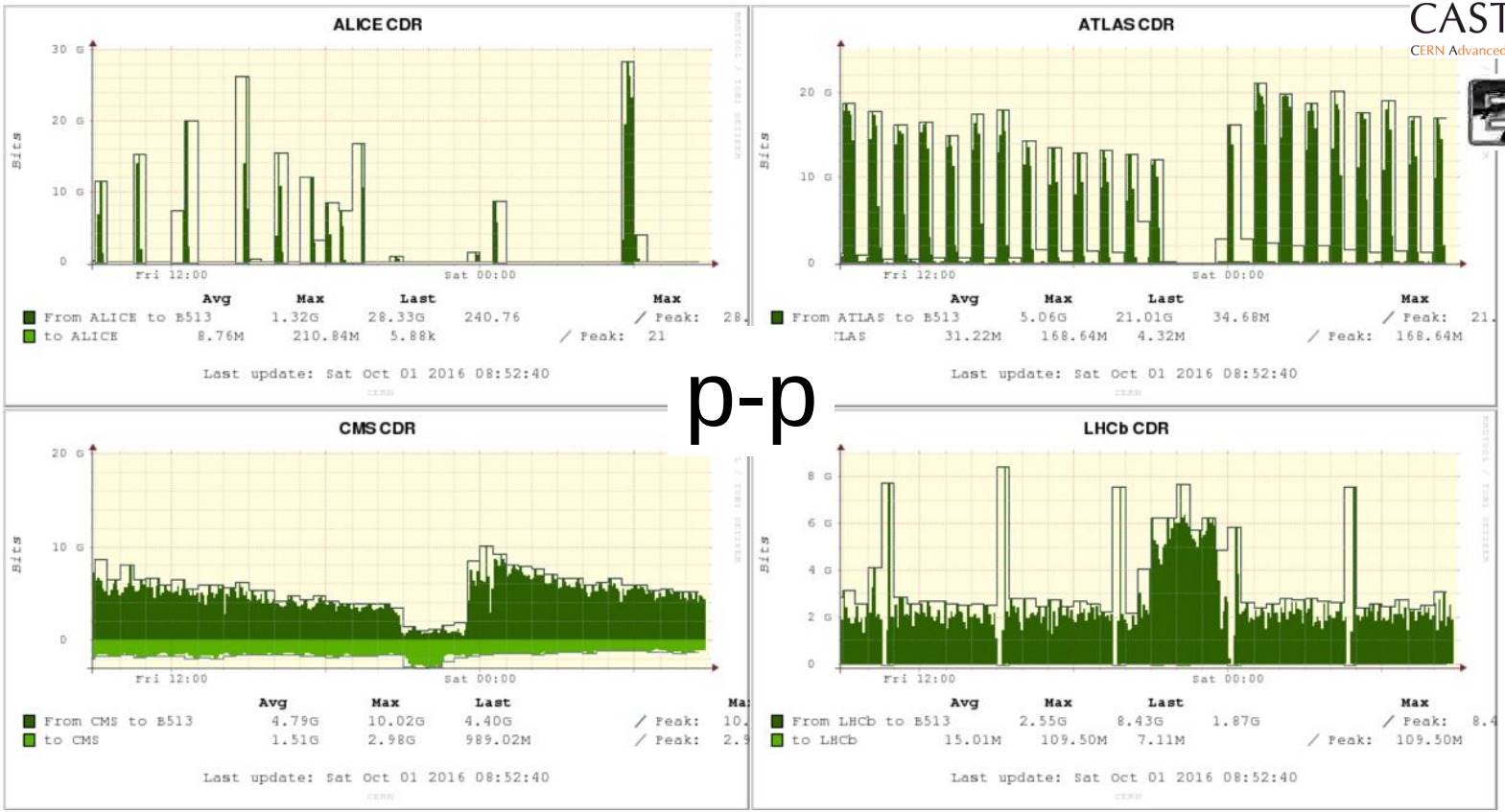
CASTOR
CERN Advanced STORage manager

EOS

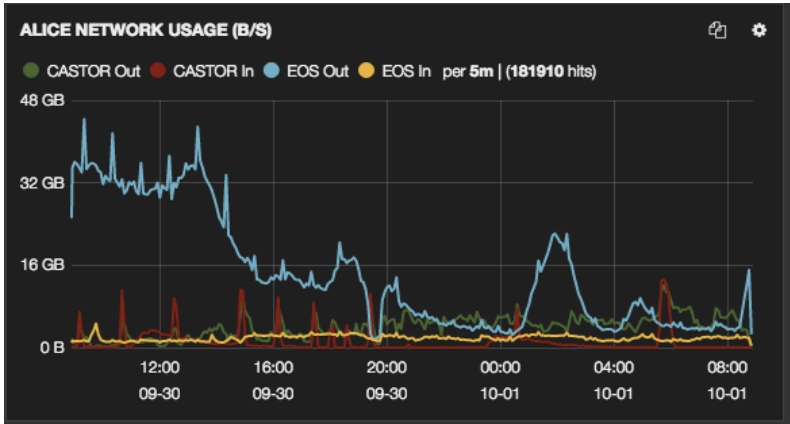
Large Scale Storage for LHC data

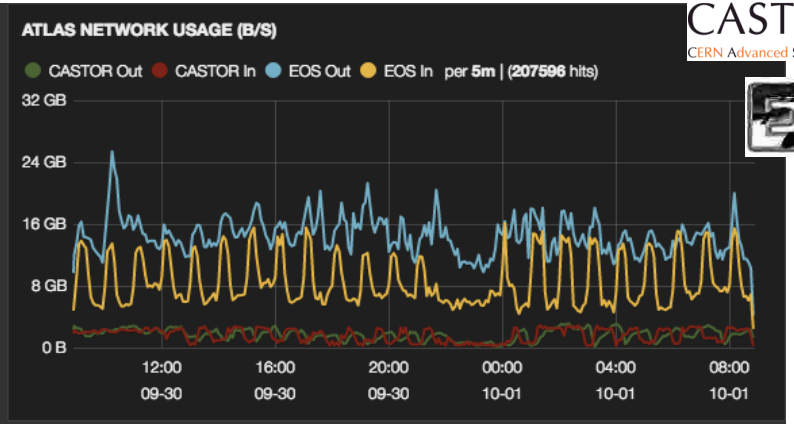
CASTOR
CERN Advanced STORage manager

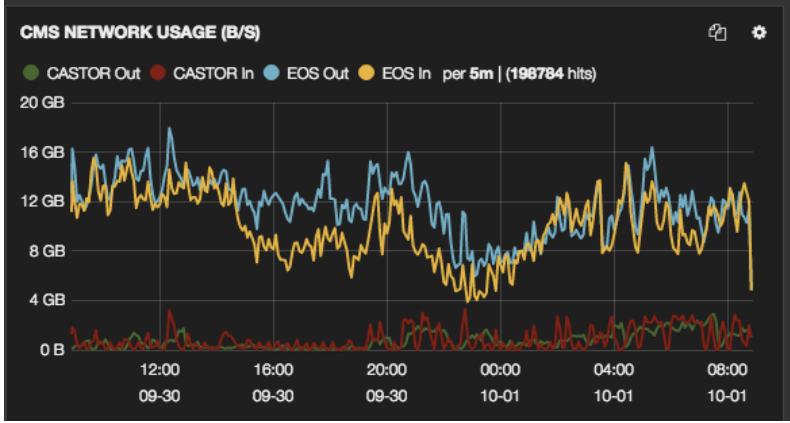
EOS

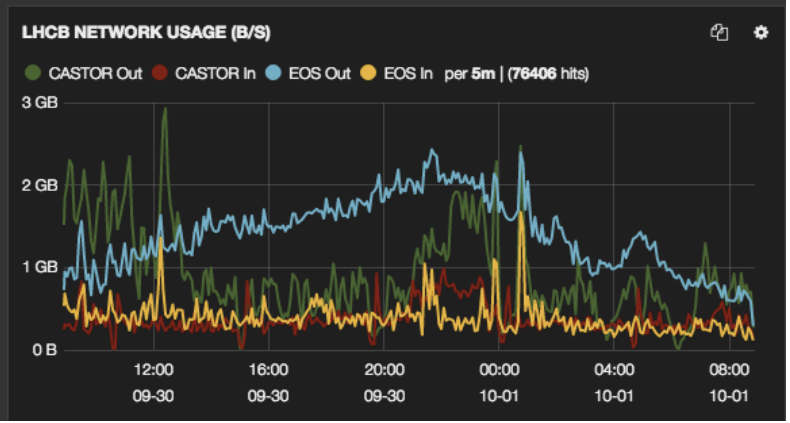


Large Scale Storage for LHC data







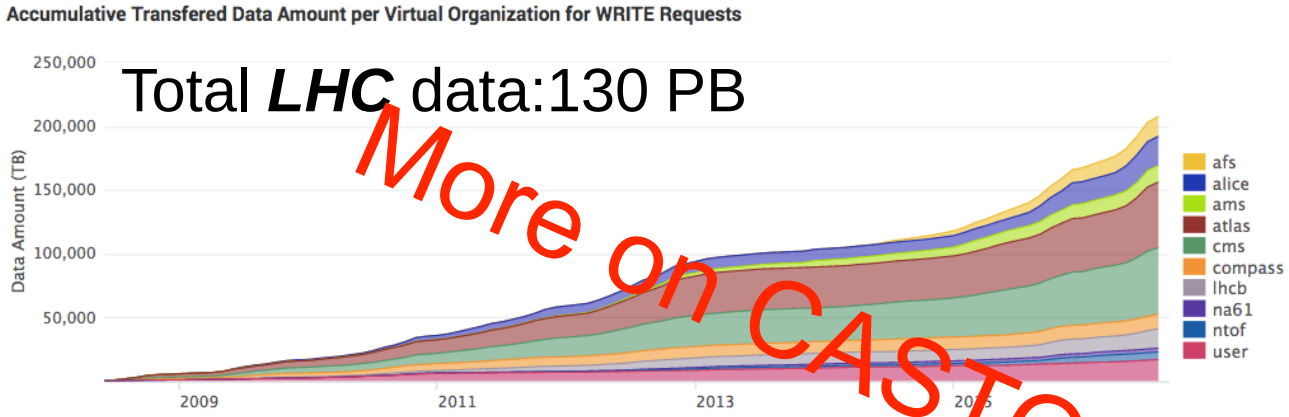


CASTOR
CERN Advanced STORage manager

EOS

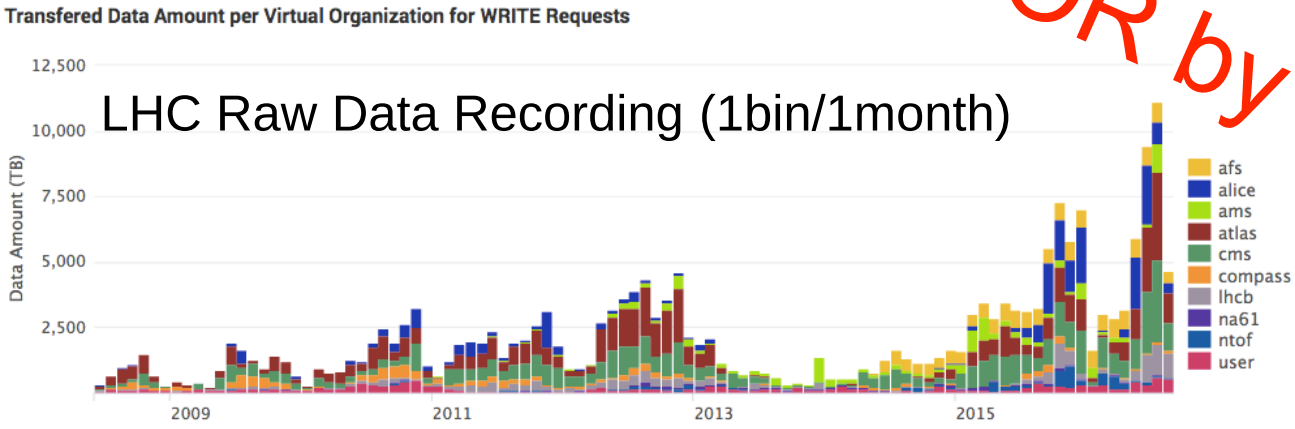
Large Scale Storage for LHC data

CASTOR
CERN Advanced STORage manager



Total ***LHC*** data:130 PB

More on CASTOR by Julien



LHC Raw Data Recording (1bin/1month)

Biggest physics-repo worldwide :

180PB and **500M** files

Evolved towards a **Tape oriented** system during LS1

CEPH backend in production (Alice, Repack)

Cold by definition: high throughput, high latency

Future evolution for a pluggable tape backend, see CTA (Julien)

Large Scale Storage for LHC data

+1200 nodes

+45000 disks

EB era

Easily scalable (#disk #servers)
Performant and manageable
LHC main storage platform

850M files

150PB 80% usage volume

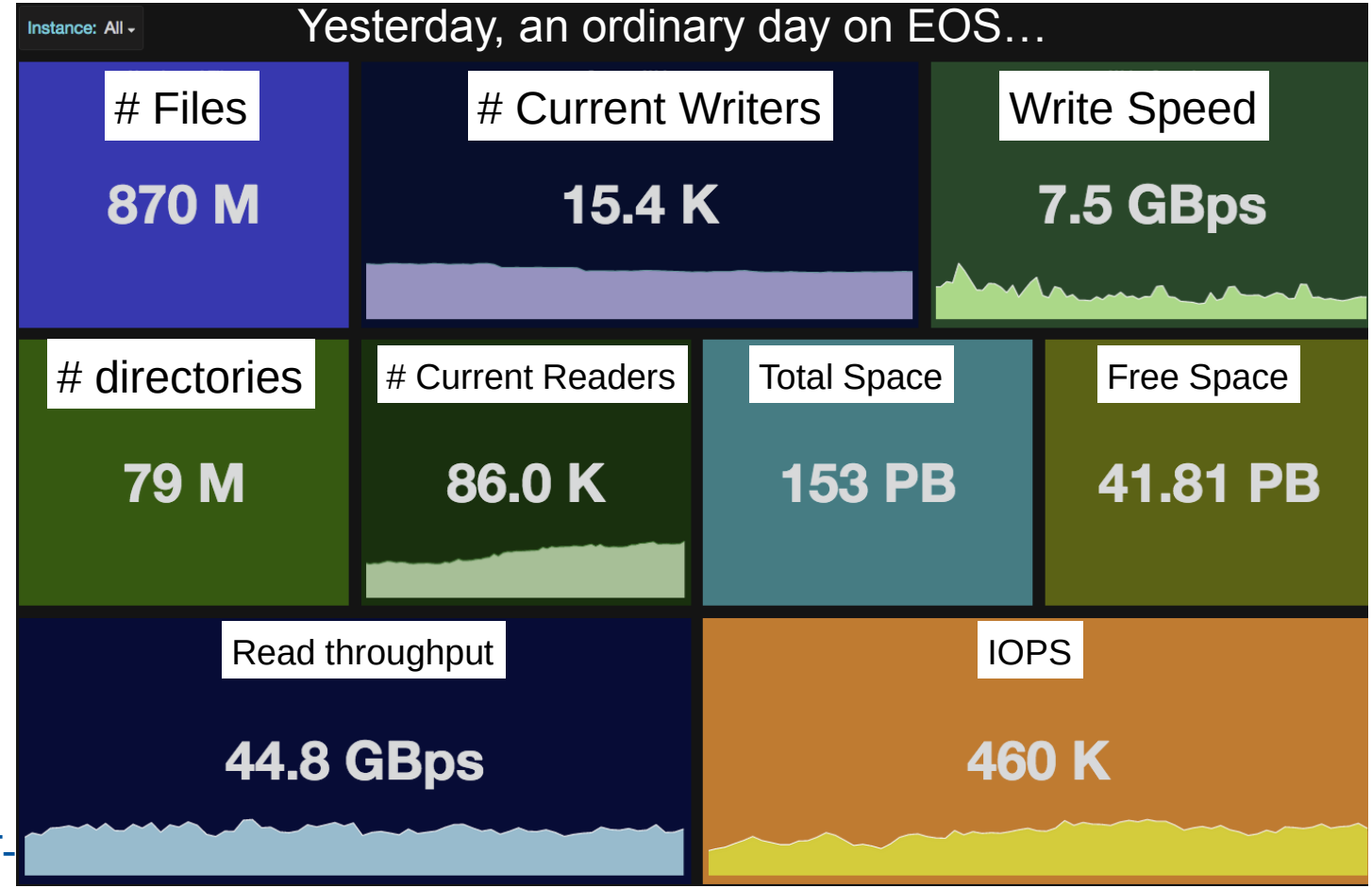
Concurrency peaks observed at

+25000 Writes

+100000 Reads

Throughput peaks at

+80GB/s Reads


Instance: All
Yesterday, an ordinary day on EOS…
Files
870 M
Current Writers
15.4 K
Write Speed
7.5 GBps
directories
79 M
Current Readers
86.0 K
Total Space
153 PB
Free Space
41.81 PB
Read throughput
44.8 GBps
IOPS
460 K

Large Scale Storage for (not only) LHC data

CERNBox

Users	6200 (Δ^{7d} +60)
#files	125M (Δ^{7d} +250K)
#dirs	15M
Quota	1TB/user
Used Space	310TB
Deployed Space	1.5PB

Collaborate

Share

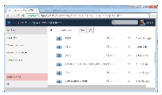
Offline work

Community storage

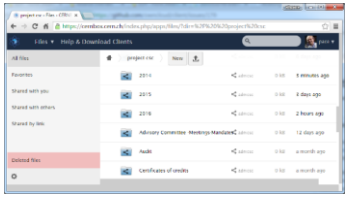
Sync

SWAN

Multiplatform / Multiuser



(web access)



Sharing with External collaborators (web access)

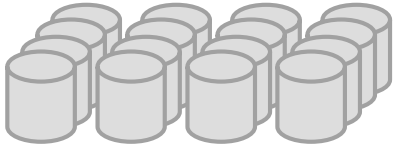
CERNBOX web interface

sync client

Choose what to sync

(offline access)

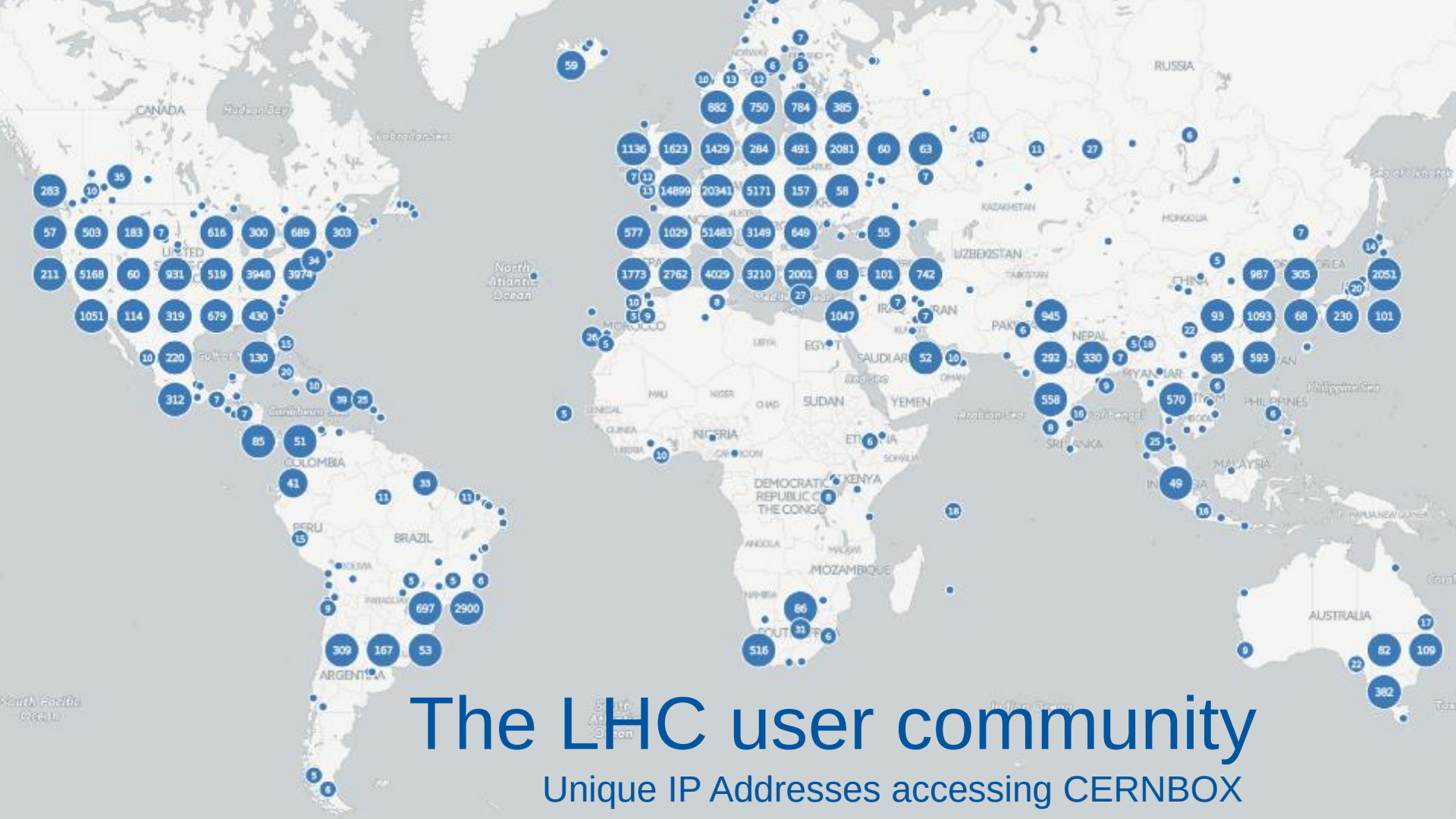
Same or Different user



Disk pools (EOS)

Online / offline File Sharing

(online access)



The LHC user community

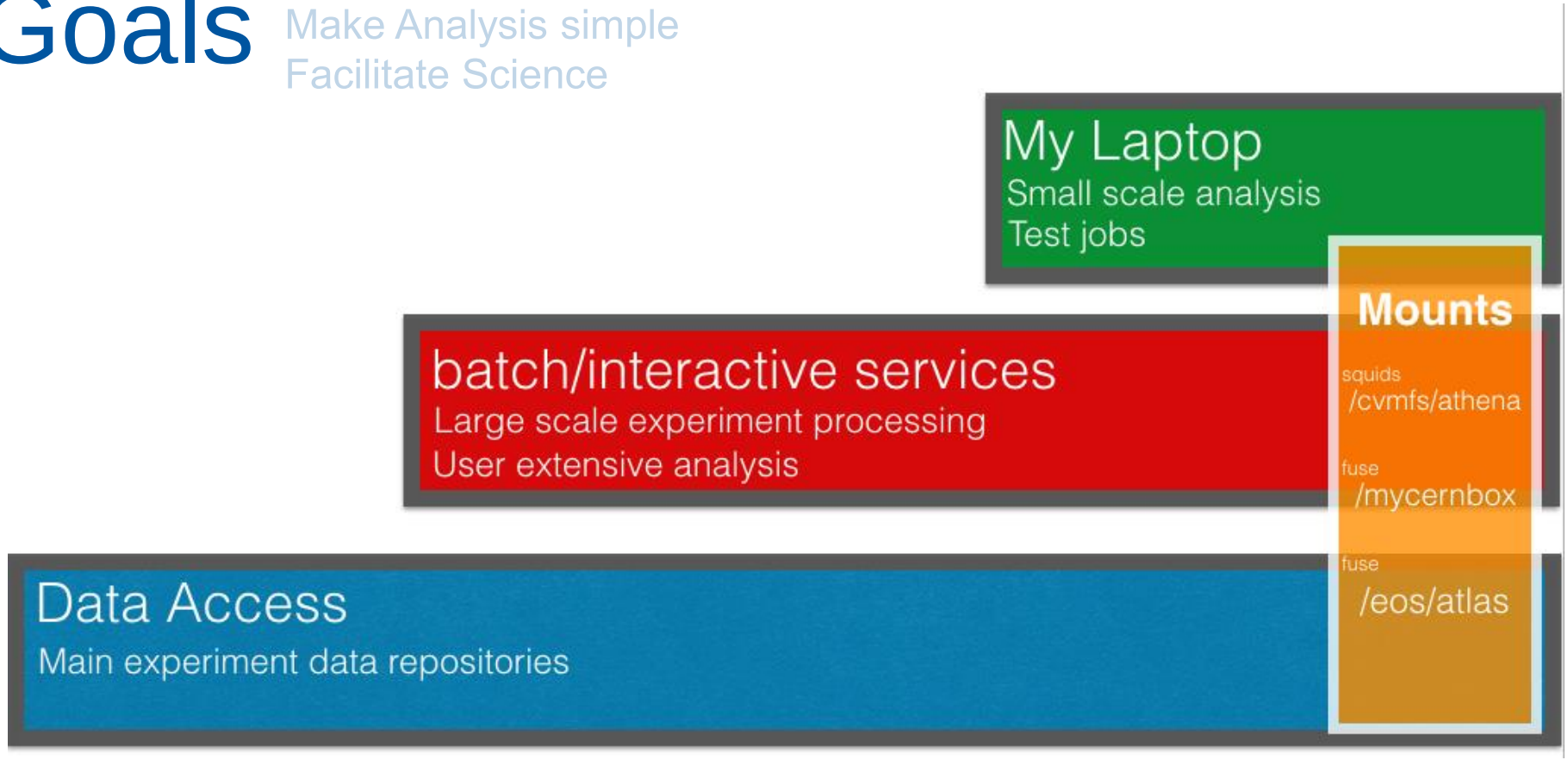
Unique IP Addresses accessing CERNBOX

Goals

Make data access easy
Make Analysis simple
Facilitate Science

Goals

Make data access easy
Make Analysis simple
Facilitate Science



Goals

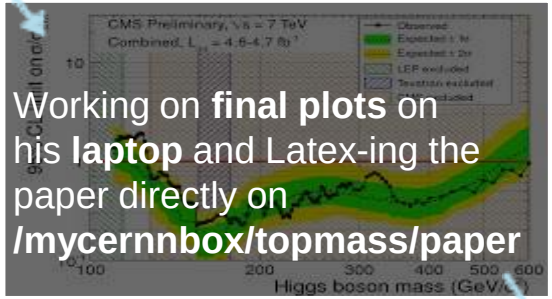
Make data access easy
Make Analysis simple
Facilitate Science

Physicist code: **topmass.**kumac
on his laptop on **/mycernbox**
and sync'd via **cernbox** client

Physicist identify an
interesting **dataset**
/eos/atlas/phys-top

Submit jobs to lxbatch/wlcg to
process the data
EOS Fuse: **/eos/atlas/phys-top**
EOS Fuse: **/mycernbox/topmass.**kumac
Experiment SW: **/cvmfs/athena**

Results (ntuples) aggregated on
/mycernbox/topmass are
synced on his laptop as the
↳ if desired
jobs are being completed



Working on **final plots** on
his **laptop** and Latex-ing the
paper directly on
/mycernnbox/topmass/paper


GLOBAL CONSERVATION LAWS AND MASSLESS PARTICLES*

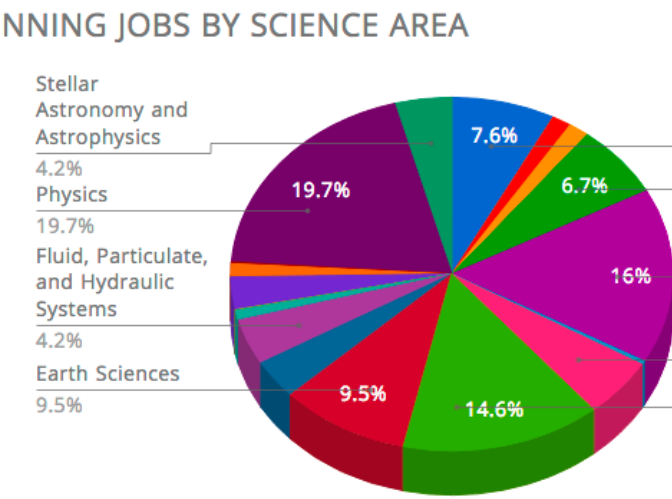
Share on-the-fly:
n-tuples
Final plots
Publication
via **/mycernbox**

Goals

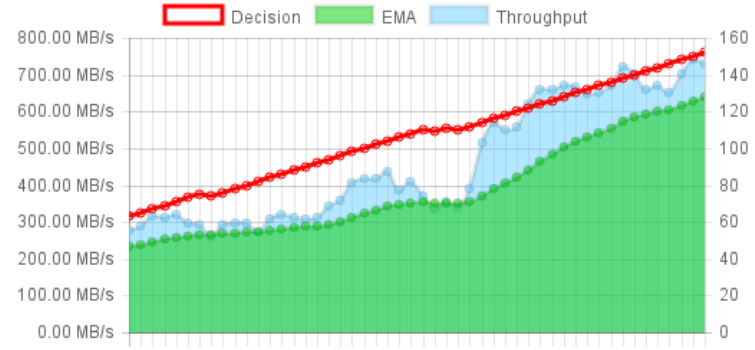
Make data access easy
Make Analysis simple
Facilitate Science

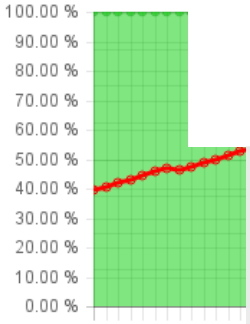
Goals

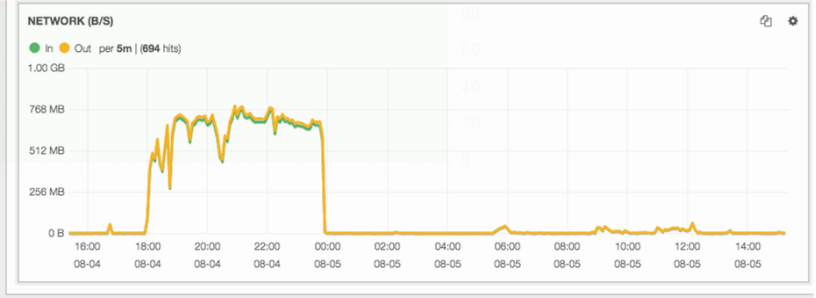
Make data access easy
Make Analysis simple
Facilitate Science



Details for srm://castorpublic.cern.ch → gsiftp://ie15.ncsa.illi







First | Previous | 1 | 2 | Next | Last

Timestamp	Decision	Running	Queue	Success rate (last 1min)	Throughput	EMA
2016-08-05T13:57:24	154	152	1898	100.00%	735.688 MB/s	648.032 MB/s

Goals

Make data access easy
Make Analysis simple
Facilitate Science

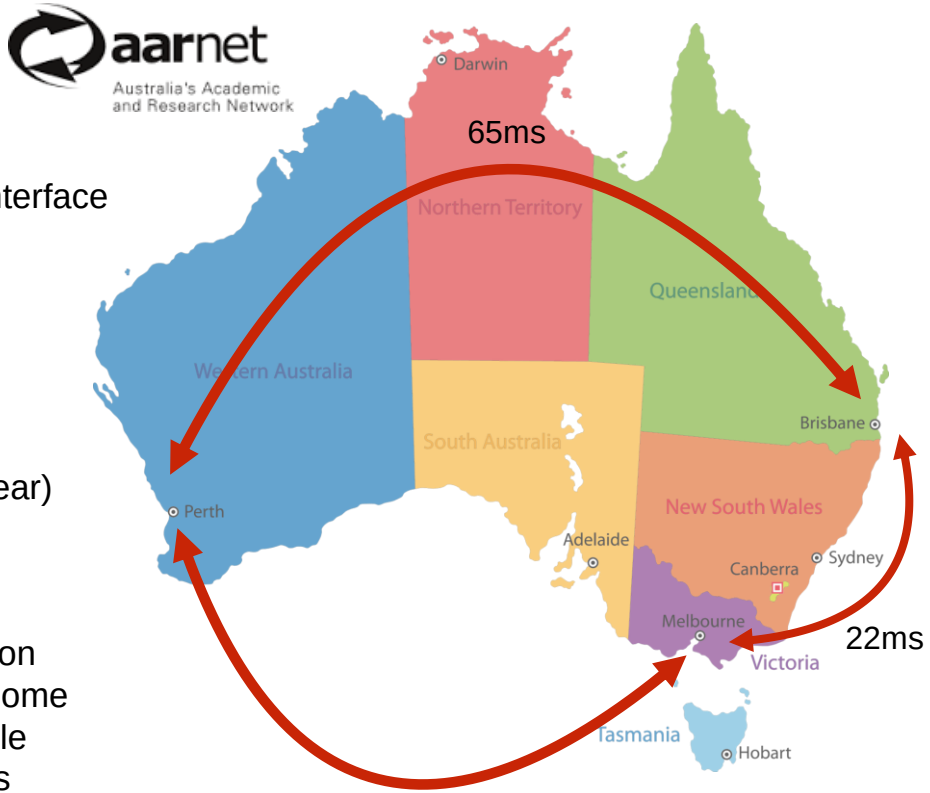
- Scale-out filesystem underneath the ownCloud app, using the eosd fuse interface for file IO

aspera RSYNC

- Geo-distributed setup: Brisbane, Melbourne, Perth
  - ~1PB (scale to ~20PB next year)

Australian National University

- Australian National University, in Acton Canberra: mirror archives of both genome sequences and open or freely available software distributed among three sites



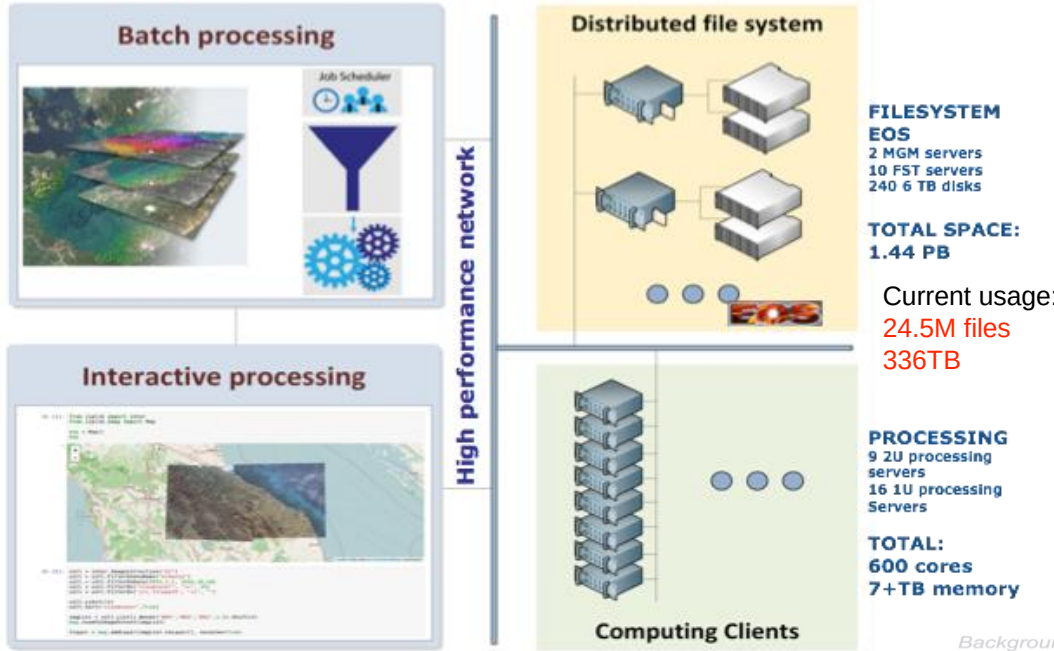
"This system is presently running 0.3.187, and has been so trouble free that I keep forgetting it's there." David Jericho -AARNetSolutions Architect

Background image: Shutterstock

Goals

Make data access easy
Make Analysis simple
Facilitate Science

Components of the *JRC Earth Observation Data Processing Platform (JEODPP)*



European Commission
Joint Research Centre

@V.Vasilev,F.Eyraud
(JRC)

Current usage:
24.5M files
336TB

Background image: Shutterstock

IT-ST-TAB

Presented by Julien Leduc

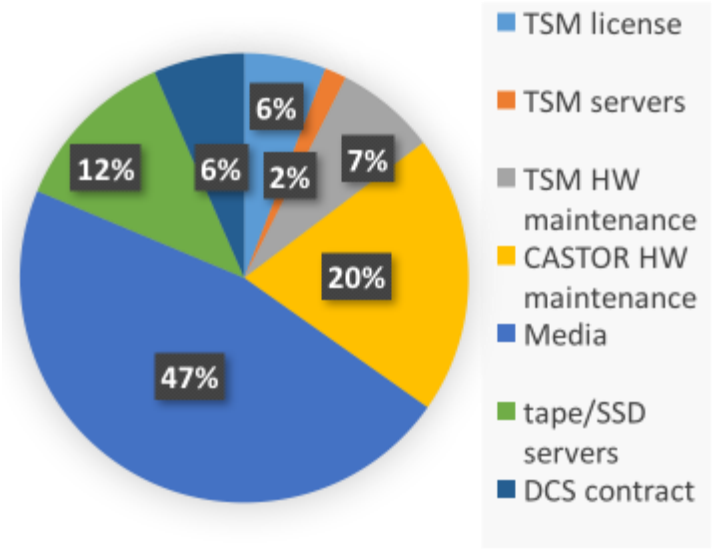
IT-ST-TAB

Storage Services for Physics and more

IT-ST-TAB

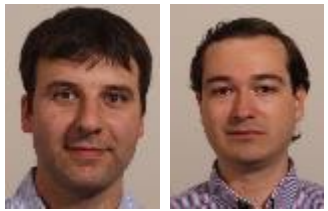
Tape, Archives and Backups (TAB)

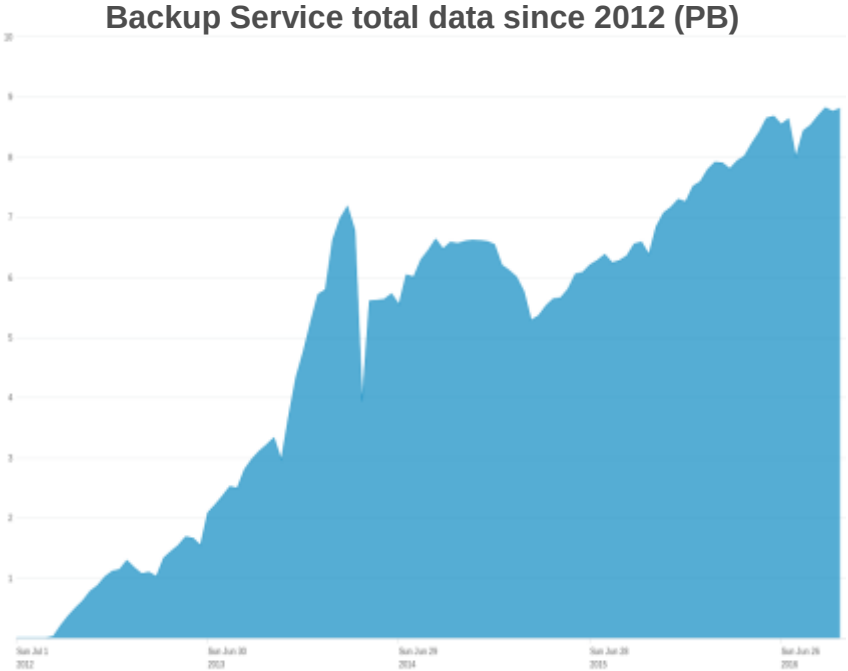
- Mandate
  - Design, operate and support the archive and backup services.
  - This includes the tape-based software backend for CASTOR, tape robotics, drive and media for physics, infrastructure for backup and restore of file servers and databases.
  - "Tape from A to Z"
- People
  - 7 staff + visitors + students
  - 1 external contract
- Budget
  - 1.7MCHF in 2016



Backup Service

- Operate, maintain and monitor CERN Backup infrastructure:
  - Critical service backups
  - High reliability, sustained daily traffic: 70TB/day
- Infrastructure:
  - 2 tape libraries (9k tapes)
  - 55 tape drives
  - 20 TSM servers

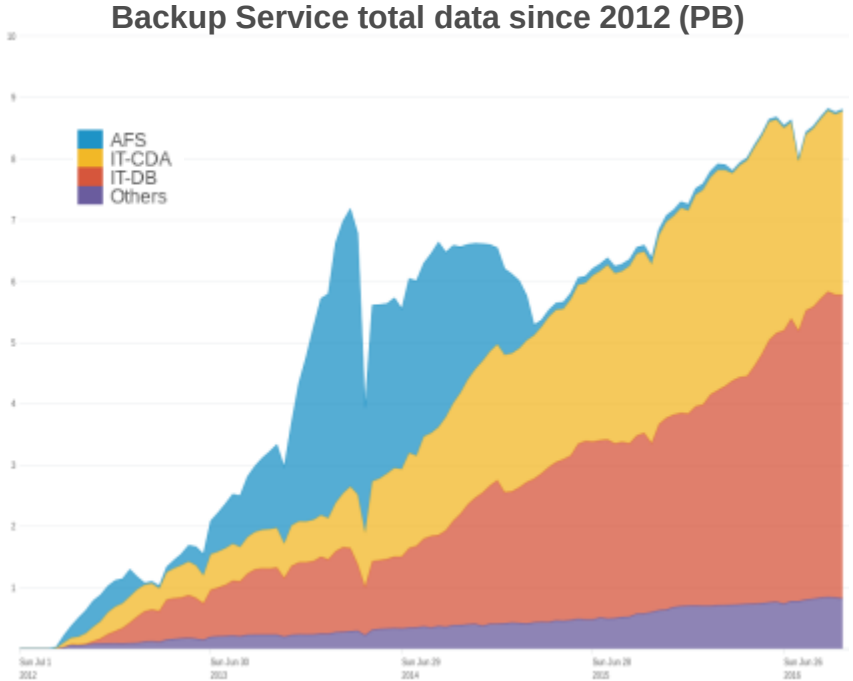




IBM Spectrum Protect

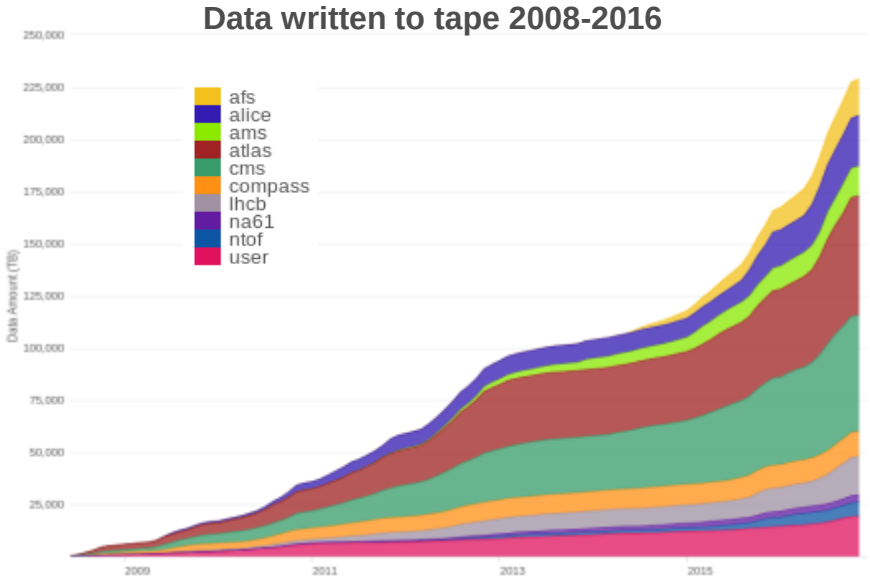
Backup Service

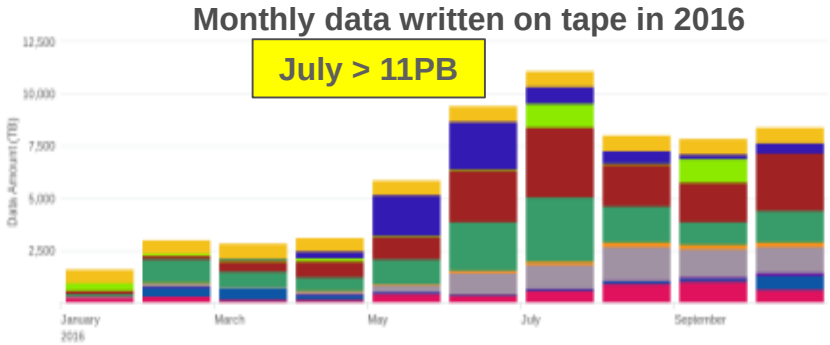
- TSM cost is proportional to backed up volume
  - Minimize backup volume in view of upcoming license retendering
- Ongoing efforts to limit annual growth rate: 24% (50% in the past):
  - Moved AFS backups to CASTOR
  - Worked with IT-CDA to remove redundant copies
  - Working with IT-DB to move Oracle backups out of TSM



Data Archiving

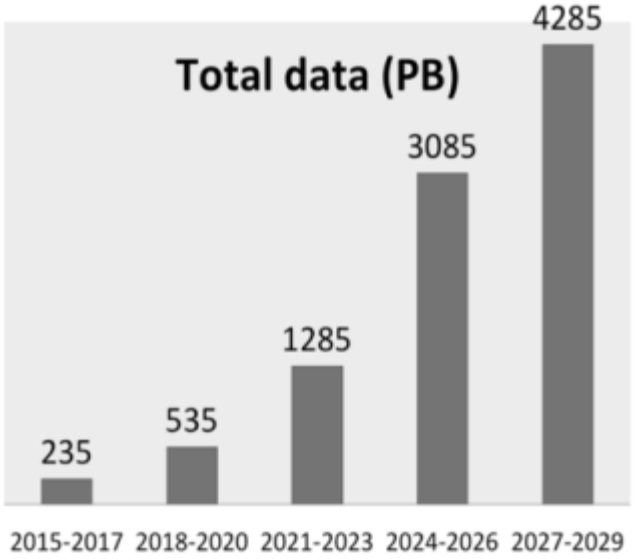
- Development, operation and monitoring of CERN's tape data archive:
  - Storage and long term (ad aeternum) archiving of CERN physics data
  - High throughput, high reliability, cost efficient
- Infrastructure:
  - 7 tape libraries, 83 tape drives / tape servers, 23k tapes
  - Current use: 180PB
  - Current capacity: 0.6 EB
  - Max throughput: 25GB/s





Data Archiving

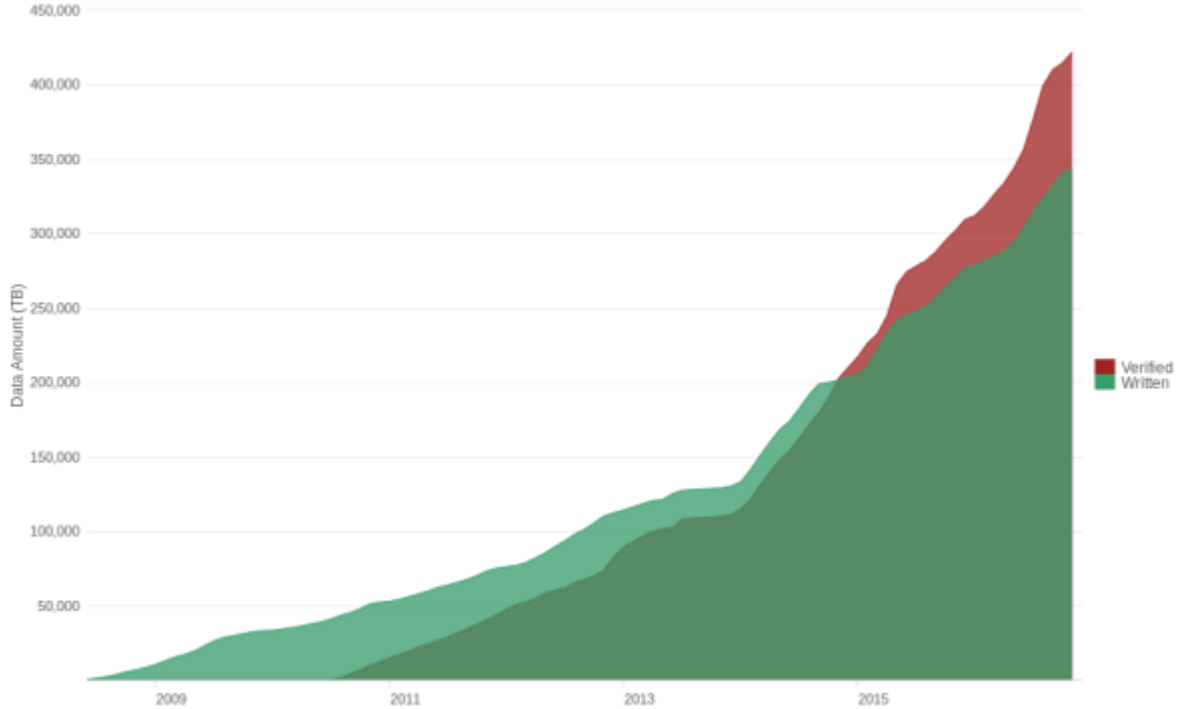
- Anticipating the future …
  - 2017-2018: 100PB/year
  - 2021++ : 150PB/year



- … and preserving the past
  - Protect against bit rot (data corruption, bit flips, environmental elements, media wear out and breakage …)
  - Migrate data across technology generations, avoiding obsolescence
    - Some of our data is 40 years old

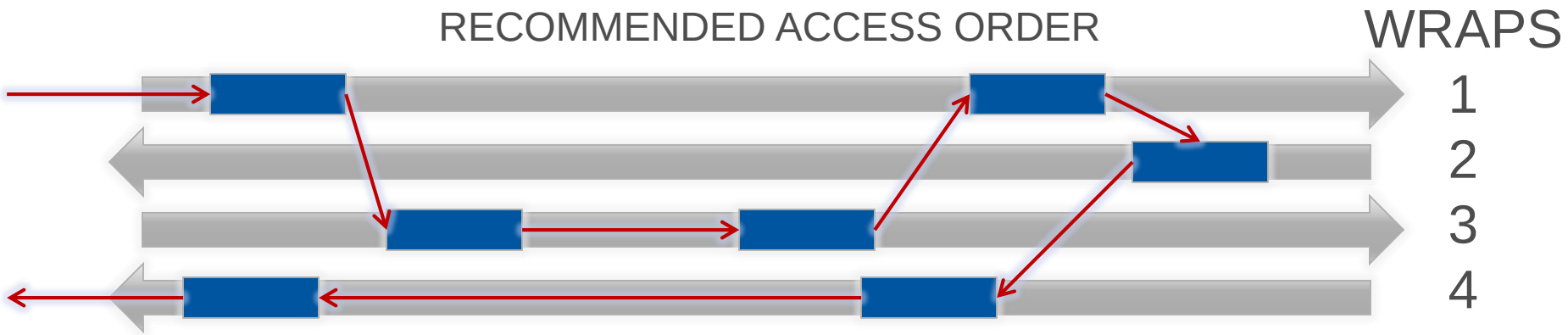
Data Archiving

- Systematic media verification



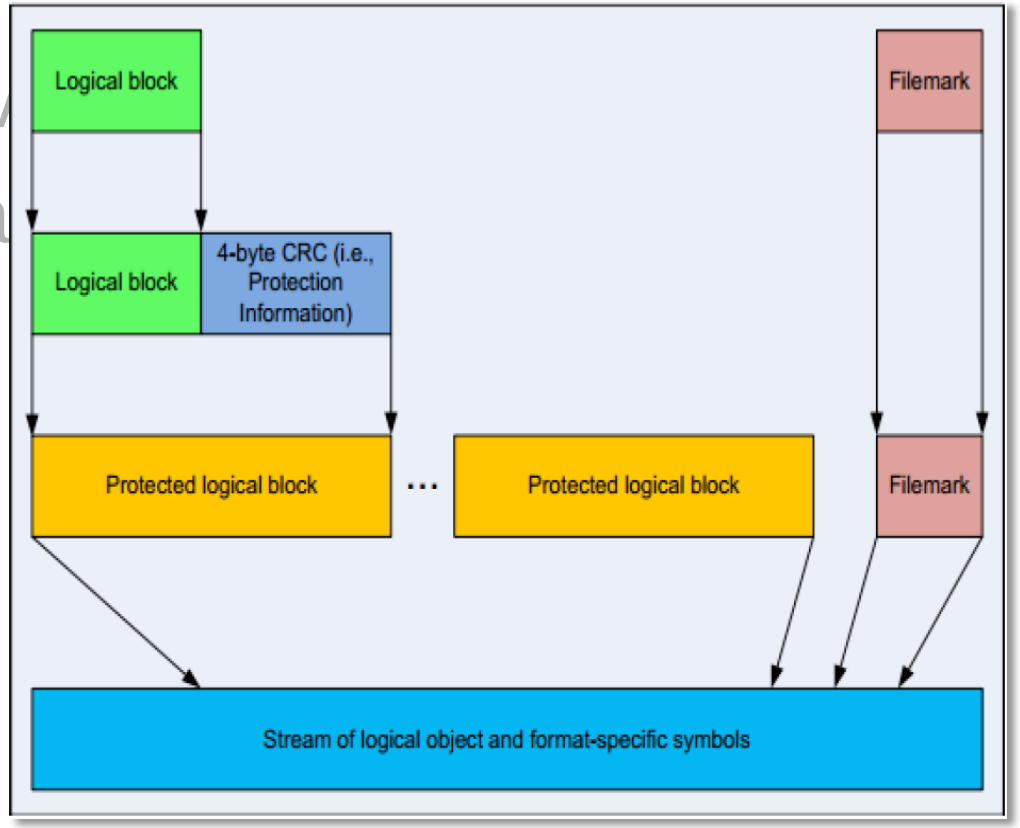
Data Archiving

- Systematic media verification
- Optimising media access



Data Archiving

- Systematic media v
- Optimizing media a
- Improve reliability



Data Archiving

- Systematic media verification
- Optimizing media access
- Im
- In ring

2016-01-26T05:46:03.594249+01:00 tpsrv220 tapeserverd[3335]:
LVL=Info TID=3350 MSG="Logging volume statistics"
firmwareVersion="460E" lifetimeBOTPasses="1486"
lifetimeMOTPasses="1556" lifetimeVolumeMounts="202"
lifetimeVolumeRecoveredReadErrors="167"
lifetimeVolumeRecoveredWriteErrors="30"
lifetimeVolumeUnrecoveredReadErrors="4"
lifetimeVolumeUnrecoveredWriteErrors="2"
validity="1" volumeManufacturingDate="20110603"

Not good..

Really bad!

Data Archiving

- Systematic media
- Optimizing media
- Improve reliability
- In depth performan
- Collaborations

IBM

ORACLE

CERN

IBM TS1150
tape drive test results

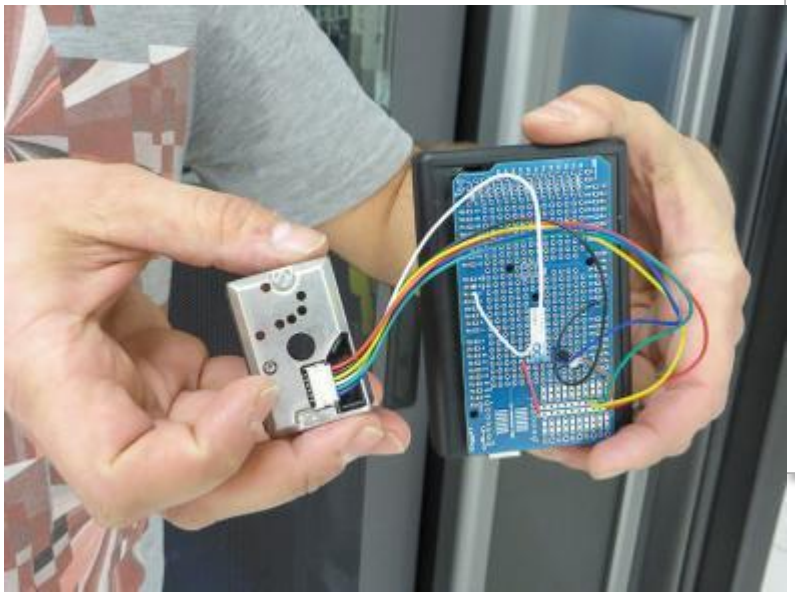
Vladi
IT De

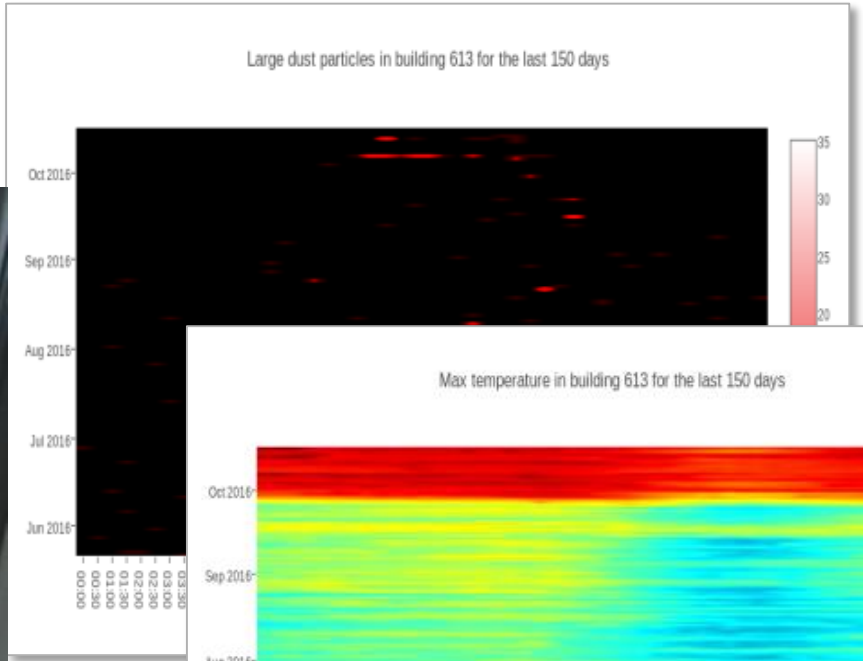
LTO/SpectraLogic testing

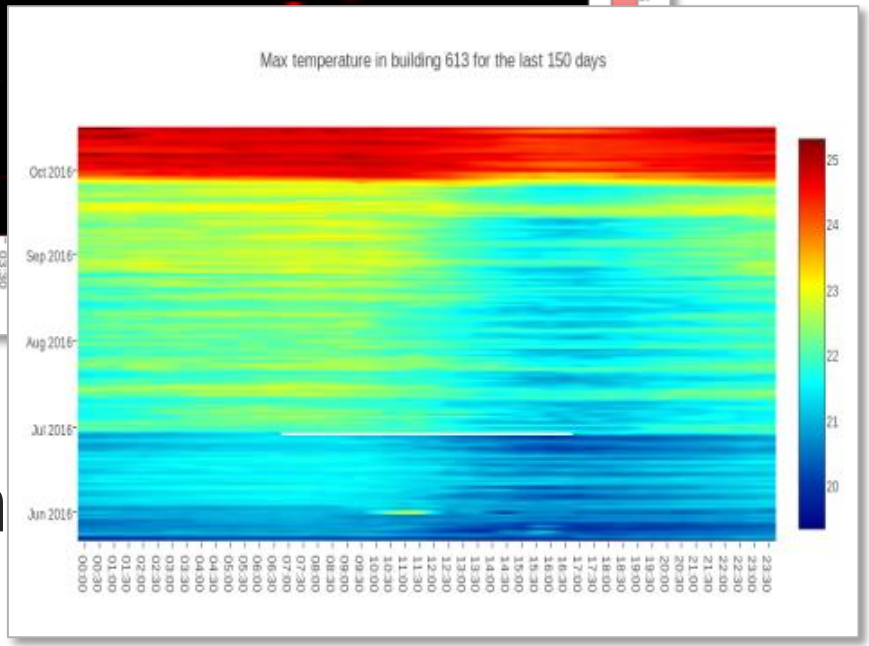
CERN IT Department

- "Small" test SpectraLogic library deployed for CASTOR integration during summer
  - LTO potentially interesting alternative/complement with reduction in tape mounts and read access; potentially useful for "dusty" archived data which is infrequently accessed
  - Necessary CASTOR adaptations coded and released
- Will test from November on a large SpectraLogic T-Finity (max 120 drives, 30K slots)
  - Test configuration: 5-10 LTO-5 drives, 1000 cartridges (1.5PB)
  - Loan free of charge except media
  - Field test drives, library, and vendor support

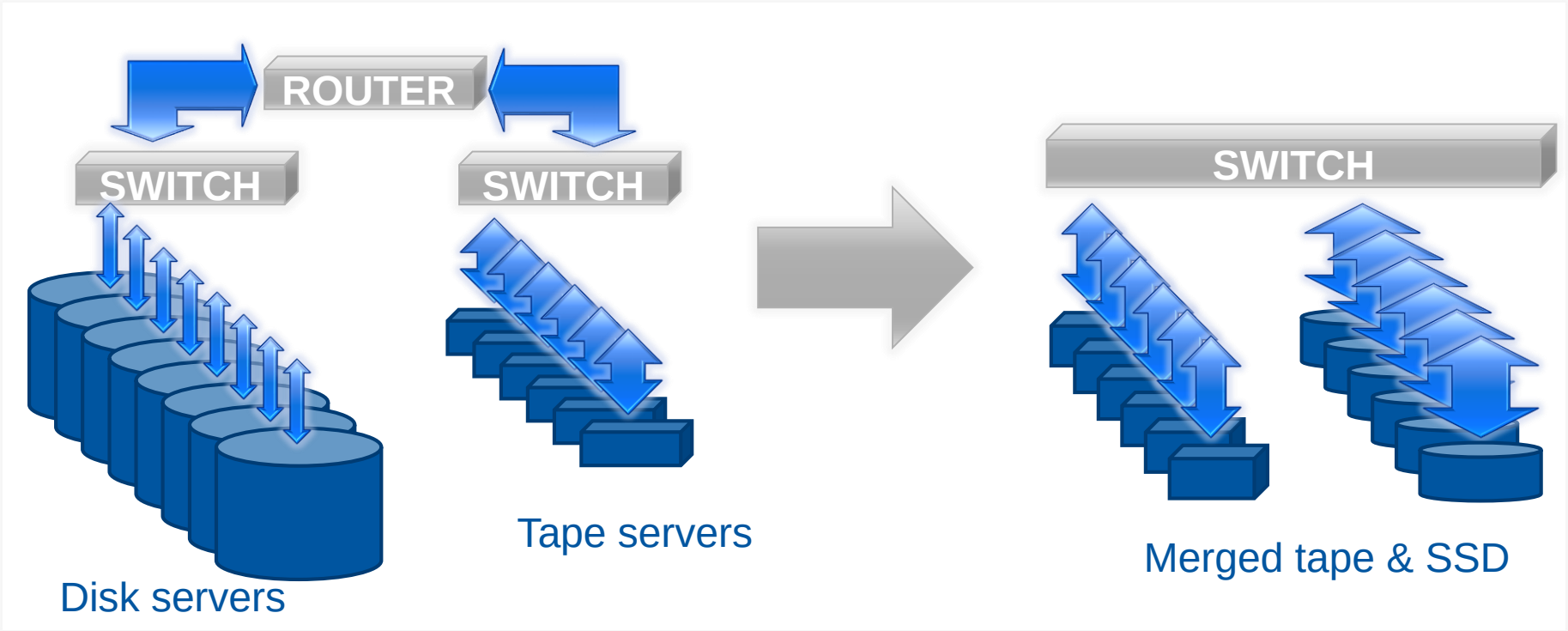
Data Archiving







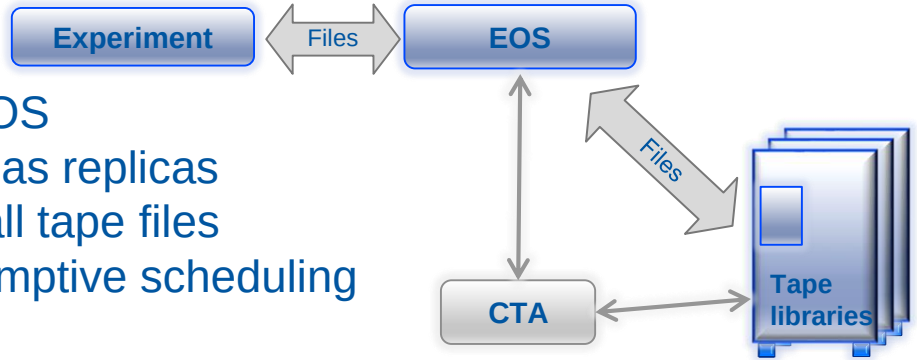
- Environmental protection



- Large scale media migration

Data Archiving: Evolution

- EOS + tape …
  - EOS is the strategic storage platform
  - Tape is the strategic long term archive medium
- EOS + tape = ❤️
  - Meet CTA : the CERN Tape Archive
  - Streamline data paths, software and infrastructure



- CTA is glued to the back of EOS
- EOS manages CTA tape files as replicas
- CTA contains a catalogue of all tape files
- CTA provides optimised, preemptive scheduling

Data Archiving: Evolution

- CTA Timeline
  - End 2016: First functional prototype release
  - April 2017: First release for additional copy use cases
  - 2018: Production-ready version

  - Easy migration path from CASTOR to EOS+CTA
    - Only metadata needs to be migrated
    - CASTOR tape format will be reused

IT-ST-AD

Presented by Andreas Peters

IT-ST-AD

STorage - Analytics & Developments

DPM
DynaFED

EOS
Open Source Storage

CERNBox

FTS3

CERN openlab

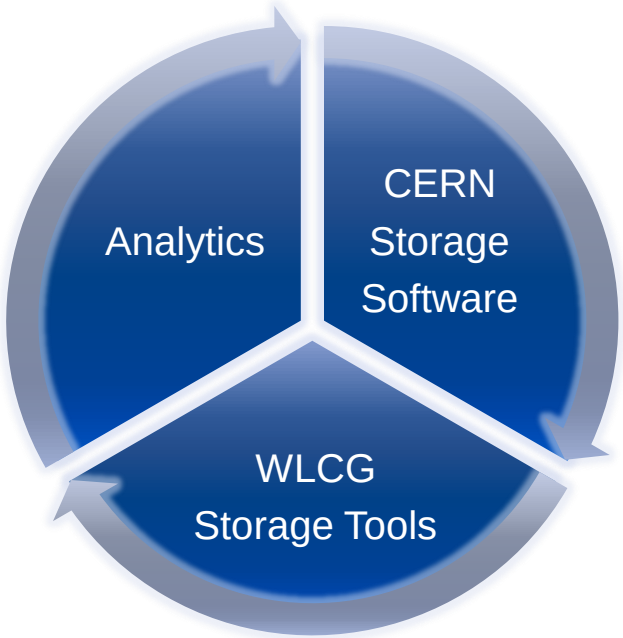
SEAGATE

XRootD

Storage Services for Physics and more

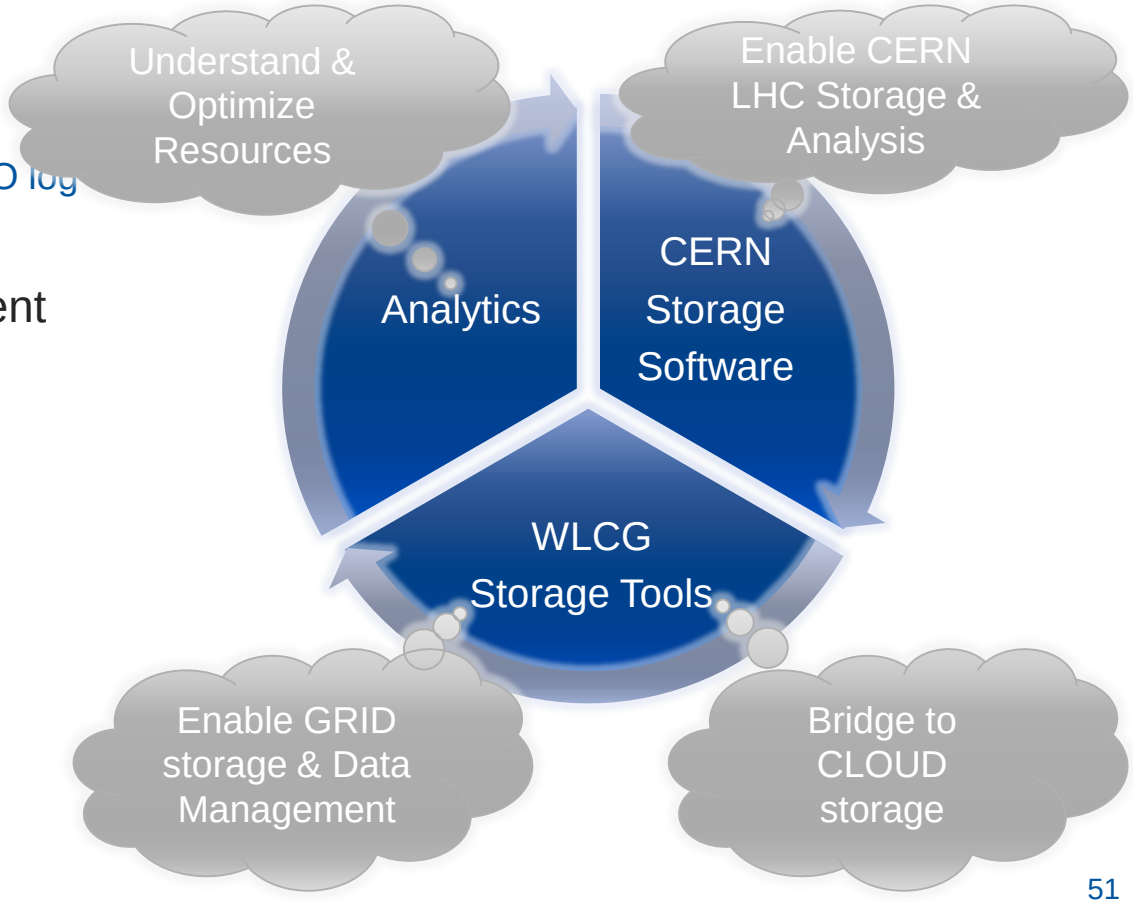
Section Activities

- Analytics
  - Activities/Working Group - CPU/IO log analysis
  - Hadoop Service
- Storage Software Development
  - **CERN**
    - CASTOR[XRootD]
    - EOS / CernBOX
    - XRootD [Client & Release Management]
  - **WLCG**
    - DPM
    - Dynafed
    - GFAL & DAVIX
    - FTS3 [Dev&Ops]



Section Activities

- Analytics
  - Activities/Working Group - CPU/IO log
  - Hadoop Service
- Storage Software Development
  - **CERN**
    - CASTOR[XRootD]
    - EOS / CernBOX
    - XRootD [Client & Release Management]
  - **WLCG**
    - DPM
    - Dynafed
    - GFAL & DAVIX
    - FTS3 [Dev&Ops]



Section Members

Analytics

CERN Storage Development

WLCG Storage Development

Section Structure


PhD Christian Nieke
ANALYTICS

Analytics – Hadoop Service

3 production cluster – joint activity of IT-ST & IT-DB

	lxhadoop	analytix	hadalytic
Purpose	general, de facto Atlas	monitoring, log analysis	beams dep, new development
Nodes	20	38	14
Size (used) PB (%)	1.5 (60)	2.5 (66)	2 (55)
Files	35 M	8 M	6M
Jobs/week	2500	6600	2500

Highlights

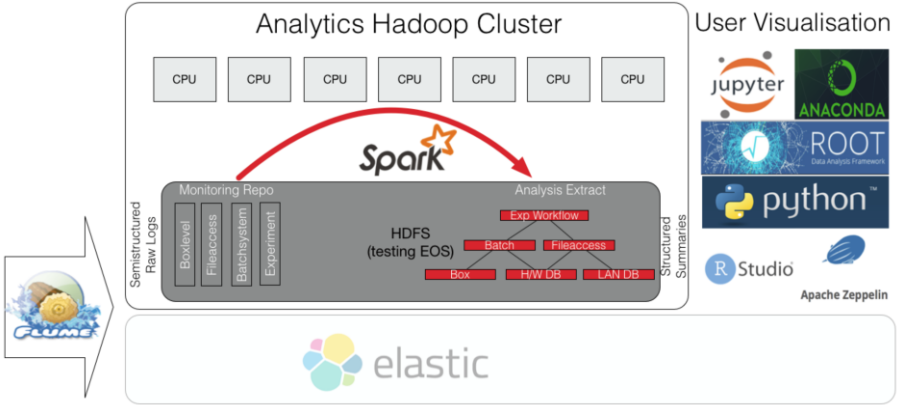
- Analytix repository for EOS log files
- **SWAN** hadoop integration
- CERNbox/EOS access for map/reduce and **Spark**
- HDFS **backup** to Castor

Future Challenge

- Hadoop on EOS disk server

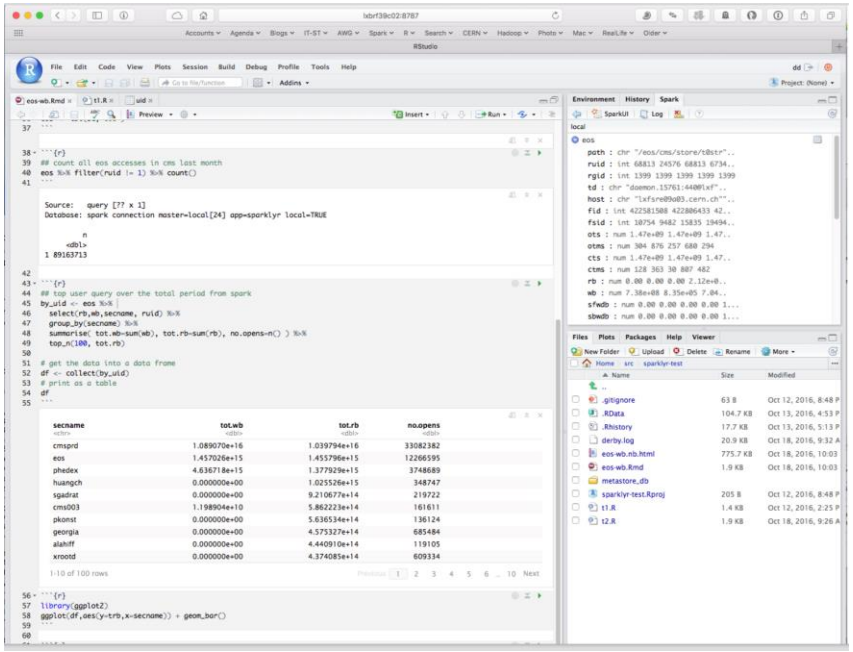
Analytics – Log Storage & Analysis

… feed log records via **IT monitoring chain**



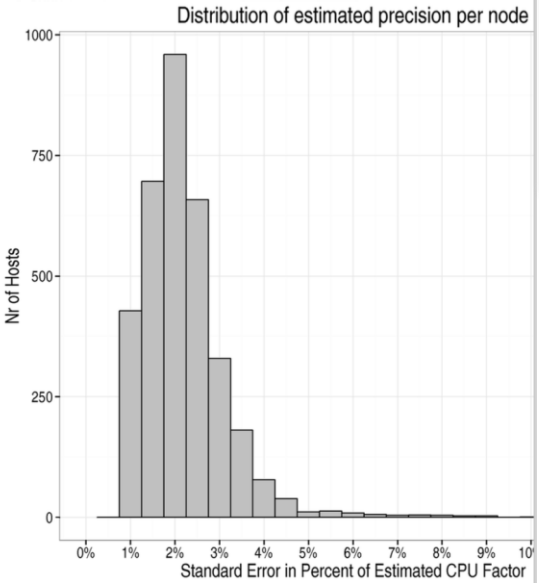
Hadoop & Friends – CHEP 2016

… **fast interactive** log data **access**



Analytics – Performance Projects

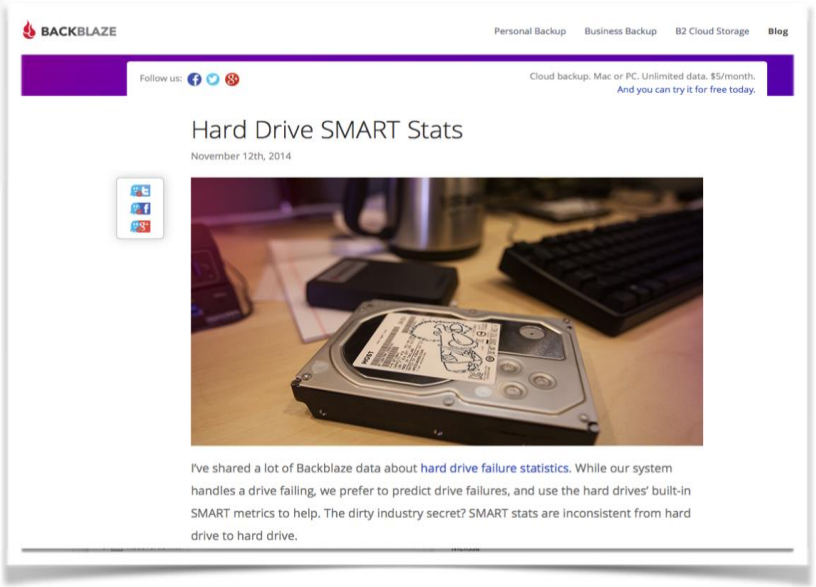
PhD Christian Nieke



Passive CPU benchmark

performance prediction error better than 5%

Study proposed by Dirk Düllmann and Darrell Long (University of California)

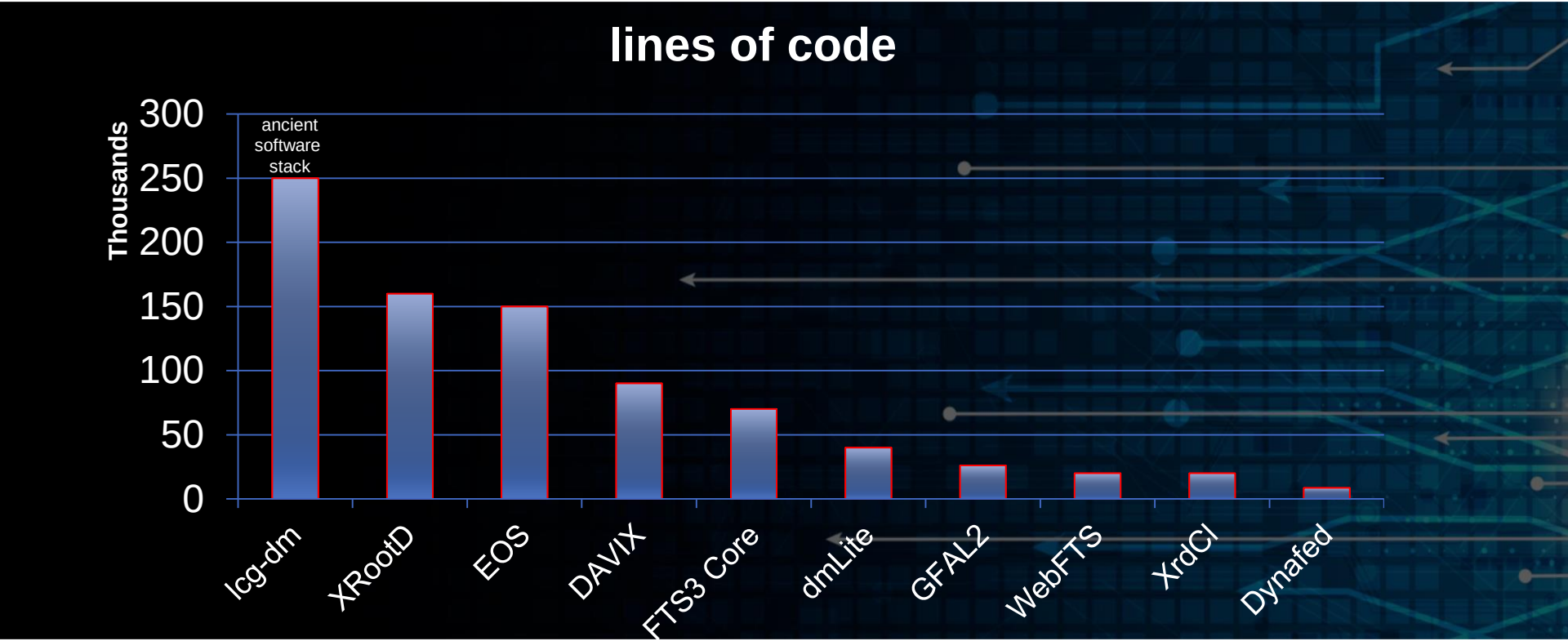


Harddrive failure analysis

Evaluate enterprise, consumer, network drives
For service planning & predictive maintenance


CERN
STORAGE SOFTWARE DEVELOPMENT

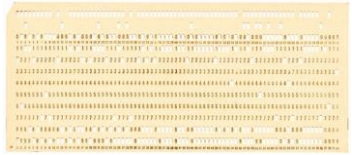
Open Source Software Projects



XRootD

What is that? What is that for?

- provides high performance, scalable fault tolerant access to data repositories
- main protocol for WLCG WAN access – latency optimized
  - >>99% of LAN data access at CERN via XRootD protocol
  - tight integration in ROOT framework
- provides a data oriented client-server C++ framework(plugin structure) - used by
  - EOS
  - LSST
- Participation via collaboration - core development SLAC



Our Responsibilities

- **client development** since 2010
- **release management & distribution**

Good to know …

- since 2010 XRootD **V3** – production (e.g. EOS)
- since 2014 XRootD **V4** IPV6, FileCache, HTTP(S) bridge, CEPH bridge production (e.g. CASTOR, lxplus, LHC exp.) wire protocol security added **V4.5**
- **V5** expected in 2017

Future Challenges…

- multi-source support, TPC[2]

GITHUB https://github.com/xrootd/xrootd.git
1000 source files – 160k lines of code

EOS

What is that?

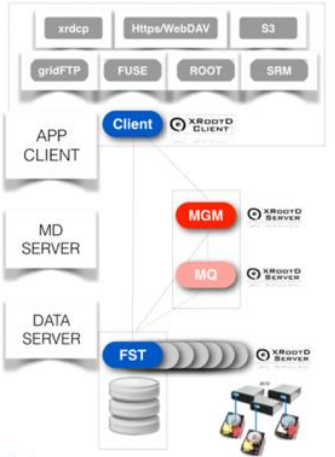
- Large scale data storage system (150 PB@CERN)
  - Multi protocol XRootD HTTP(S) WebDAV, FUSE
  - Strong Authentication KRB5 X509
  - Flexible provisioning and life-cycle management
  - Large user community features (limits, quotas)
- since 2010 developed in-house in storage group
- since 2011 production service at CERN
  - +11 HEP EOS sites not at CERN
    LBNL, ORNL, SUBATECH, IHEP, SASKE, FNAL, UNAM, MEPHI, SINICA, INFNTS, UNIVIE
- simple scalable architecture
  - Meta data *scale-up 0.5b files*
  - Data *scale-out 15k disks*
  - Messaging *1.5Mhz pipelined*

October 2016

150 PB disk space >**110 PB** used

- **1143** storage nodes

1y growth rates: +85% more files +50% more space used



EOS

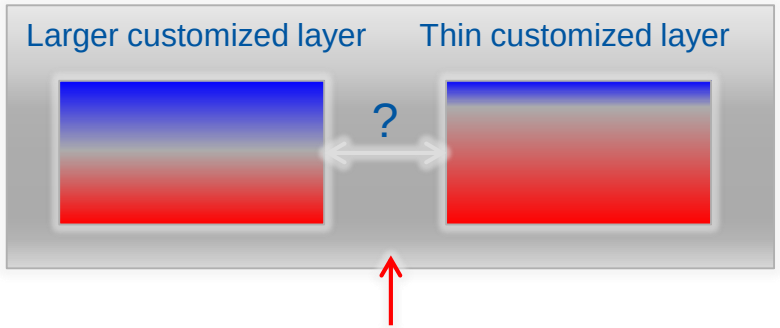
What is the current development strategy?

- Evolve **flexible** storage infrastructure as ***scale-out solution*** towards exabyte scale
  - one storage – many views on data – analysis/analytics service integration
  - suitable for LAN & WAN deployments
  - enable cost optimized deployments

- Evolve **file-system** interface /eos
  - multi-platform access
    using standard protocol bridges CIFS, NFS, HTTP

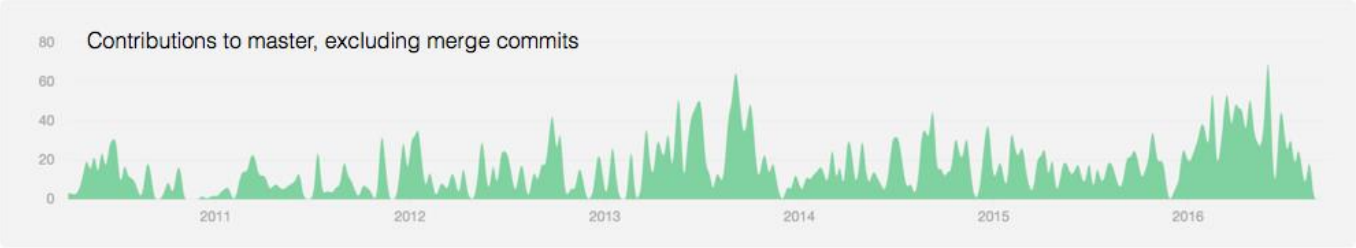
- Evolve scale-out **storage back-end** support
  - object disks, object stores, public cloud, cold storage

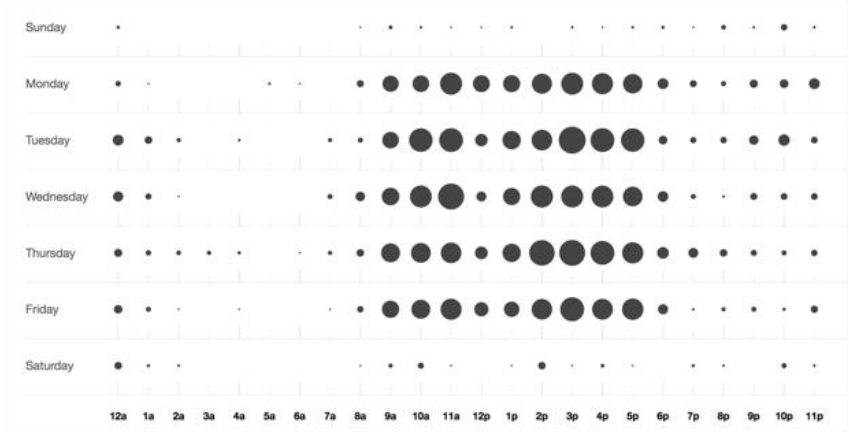
- Evolve **namespace scalability**
  - scale-out implementation



- **Review** regularly break-point between in-house and community open-source layers

EOS Development





GITHUB punch card

master: 181k LOC
production: 150k LOC
excluding external projects
jerasure/sqlite/gf-complete

70 tags in production branch
34 tags in master branch

8 contributors

+now several external contributors
AARNET + COMTRADE

GITHUB https://github.com/cern-eos/eos

EOS Challenges

1 Meta data *scale-up* ⇒ ***scale-out***

- *in-memory namespace ⇒ in-memory scale-up namespace cache + persistency in scale-out KV store*

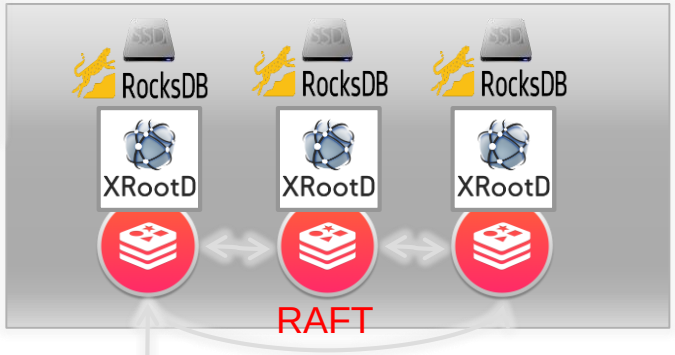
1st key project: EOS Citrine & QuarkDB

2 *Remote access APIs* ⇒ ***filesystem API***

2nd key project: EOS FUSE rewrite

EOS Scale-Out Namespace

QuarkDB
- built on open source software/ideas



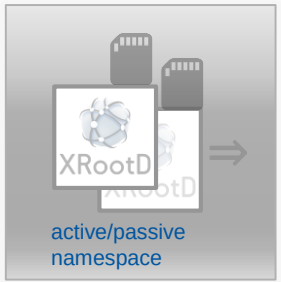
Facebook KV library for SSD persistency

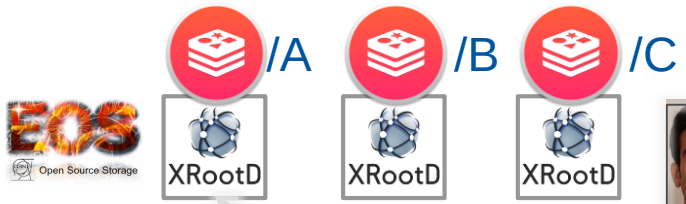
Redis XRootD plug-in

Redis protocol KV,SET,HASH

REDIS

current EOS





meta data server with namespace cache & plug-in for KV backend
⇒ stateless front-end architecture

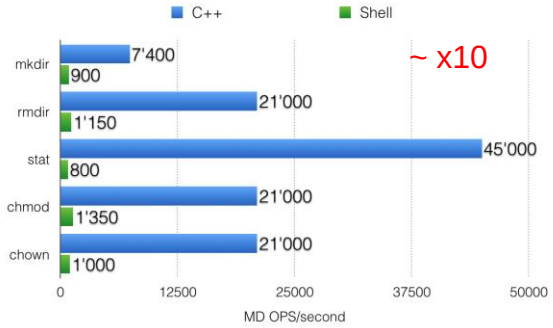
1st Key Project

EOS FUSE improvements

Open Source Storage

2nd Key Project

- FUSE was added as an external layer to an existing server API [path-inode translations]
  - 1st implementation 2011 clone of XRootD FUSE xrootdfs C-interface
  - 2nd implementation 2016 introducing many latency optimizations and pure C++ gained 10x + more POSIX
    - meta-data consistency very difficult/race conditions without server side changes – MD updates still synchronous
  - 3rd implementation on the way gain 10x + POSIX
    - client leases, async meta-data pipelining
    - optional persistent meta-data & data cache
    - libfuse2 ⇒ libfuse3 21k seq IOPS ⇒ 500k seq IOPS [wb cache]




WLCG
STORAGE DEVELOPMENT
WLCG

DPM

What is that?

- CERN **D**isk **P**ool **M**anager used in many GRID sites (outside CERN)
- originally storage software fork from **CASTOR1** code developed since 16 years
- *the* Tier-2 GRID storage solution – deployed at 160 sites world-wide providing 70 PB of storage space
- First refactoring included dmLite stack as plug-in architecture in 2011
- Latest release **1.9** includes now DOME component 2016
  - allows to to remove old legacy code & daemons ⇒ simplify and consolidate the architecture
    - better modularization
    - single internal communication protocol
  - SRM less operation
  - quota, space reporting, caching hooks
- Caching laboratory for ‘unmanaged’ Tier-2 storage, low-effort deployment, long-term consolidation
- DPM Workshop Paris https://indico.cern.ch/event/559673/

Key Components to Cloud Storage

- **DAVIX**

  - WebDAV/S3 C++ client – integrated in ROOT
  - developed since 2012 – focus on long term stability
    - latest features
      - compatible with Azure & AWS v4 authentication
      - performance optimizations by range coalescing and request parallelization

- **GFAL**

  - Grid file access library
    - provides protocol abstraction layer [⇒ FTS]

  https://gitlab.cern.ch/dmc/gfal2

- **DynaFED**

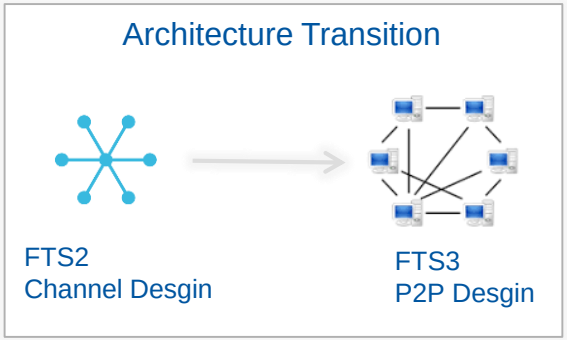
  - Provides in-memory cached namespace federation of multiple HTTP enabled storage systems
  - Provides a Grid⇒Cloud bridging functionality translating X509 identities to signed URLs using DAVIX

FTS^3

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

What is that?

- 3rd generation of file transfer service building core platform for scheduled LHC data transfers
  - developed since a 2001 at CERN
  - 1st production release 2013
  - key component in WLCG
  - transfers hundreds of peta bytes per year worldwide
  - Latest release 3.5
    - Transfer Optimization based on network performance analysis
    - Decision visualization



Future Challenges

- Boost Scalability of database centric design
  - Initiated R&D project **Flutter** trying more radical redesign solutions
- HPC integration (dual credential support)
- Alarms & Containerization

GITLAB https://gitlab.cern.ch/groups/fts
500 source files – 67k lines of code

FTS Optimizer Visualization

Details for srm://ccsrm.in2p3.fr → srm://atlassrm-fzk.gridka.de

First | Previous | 1 | Next | Last

Timestamp	Decision	Running	Queue	Success rate (last 1min)	Throughput	EMA	Diff	Explanation
2016-10-18T11:06:2	13	21	3940	100.00%	48.359 MB/s	48.426 MB/s	-1	Good link efficiency, throughput deterioration
2016-10-18T11:05:1	14	14	3948	100.00%	89.522 MB/s	48.434 MB/s	1	Good link efficiency, current average throughput is larger than the preceding average
2016-10-18T11:04:0	13	14	3968	100.00%	42.308 MB/s	43.869 MB/s	-1	Good link efficiency, throughput deterioration
2016-10-18T11:02:5	14	13	3986	100.00%	70.434 MB/s	44.042 MB/s	1	Good link efficiency, current average throughput is larger than the preceding average
2016-10-18T11:01:4	13	14	4006	99.00%	57.463 MB/s	41.109 MB/s	1	Good link efficiency, current average throughput is larger than the preceding average

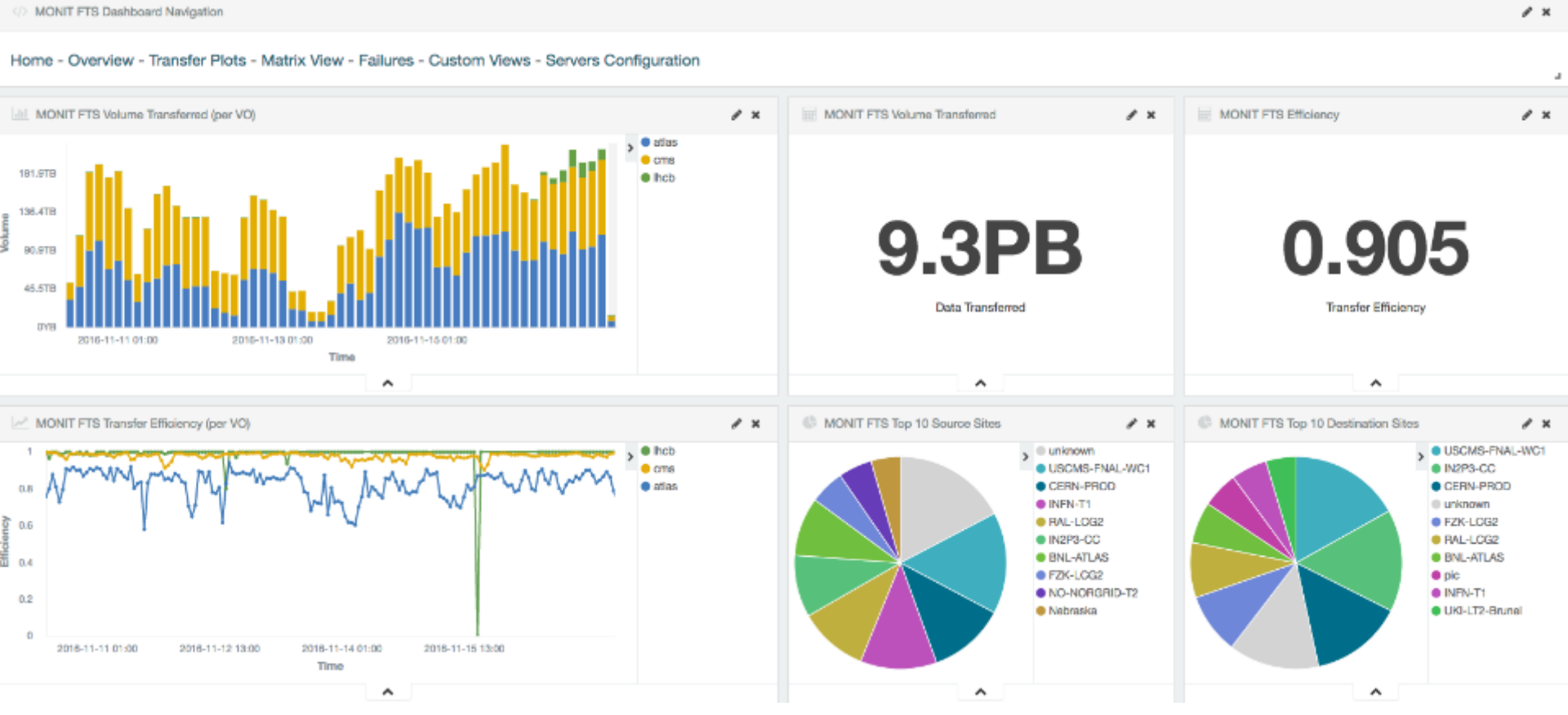
Optimizer Decision Visualization

Optimizer Status Explanation

FTS Operations

- AD responsible for operations since January ‘16
  - 2 cluster production & pilot
  - Moved to CC7
  - IPV6 enabled
  - Service Monitoring **https://monit.cern.ch/app/kibana#/dashboard/_project-FTS-Service-Level**
  - Transfer Log archival in EOS on the way

FTS @ CERN last week stats



WebFTS & LastMile

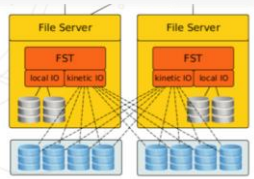
https://webfts.cern.ch

- WebFTS
  - end-user transfer portal
  - targets small VOs
- LastMile
  - end-point less FTS transfers
    - transfer data to/from workstations/laptops
      run a client – not a server

IT-ST-AD Collaborations

COMTRADE — EOS white-paper – now core software contributor

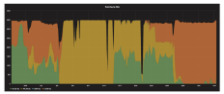
EOS OpenKinetic interface



European Commission Joint Research Centre — EOS collaboration

aarnet Australia's Academic and Research Network — EOS/CERNbox collaboration

Open Telekom Cloud — DPM Grid storage in the cloud



Dynafed demonstrator — University of Victoria (CA), INFN and Belle-II
https://indico.cern.ch/event/394788

Data Intensive Cloud Research — PhD focused on modeling ATLAS performance

www.cern.ch

IT-Storage Plans for 2017

- EOS as the strategic storage platform
  - New, scalable name server (new REDIS-based implementation)
  - Improved FUSE to satisfy more requirements
- CERNBOX as the strategic service to access EOS storage
  - Offer migration paths out of AFS (target = EOS)
  - Start a discussion with relevant service managers on the home directory service
    - Define an architecture to host lxplus home directory
    - Improve Windows support, Samba service for Windows Clients
- CERN Tape Archive
  - First internal release interfacing EOS planned end 2016 using IBMLIB1 (test library)
  - Apr 2017: D1T1 and D0T1 workflows available in EOS
- TSM
  - New license to be negotiated, investigate alternative options for Oracle backups
- FTS: improvements on manageability and scalability
  - Files to/from laptops
- DPM: Consolidation
  - Going SRM-less with DPM
- CEPH the backend storage solution
  - Understand if CEPHFS can replace the Filer service

Strategic Vision for IT-ST

- Build a flexible storage infrastructure
  - Unlimited storage for LHC experiments, at exabyte scale
    - Disk pools (EOS on-demand reliability, on-demand performance)
    - CERN Tape Archive (high reliability, low cost, data preservation)
    - Backbone of the CERN cloud (CEPH block storage for Openstack and Filers)
    - Cluster(s) for analytics (provided with Hadoop)
  - A generic home directory infrastructure
    - Fuse mounts, NFS and SMB exports, (and HADOOP).
    - CERNBOX (Sync client and Web/HTTP/DAV Access)
- Maintain and enhance (grid) data management tools
  - To empower global scientific data workflows
    - Data transfers (FTS), data caches (DPM, CVMFS), …

www.cern.ch