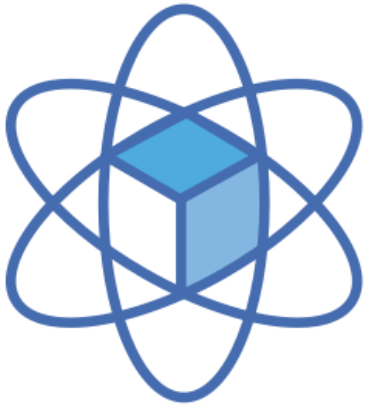
CERN

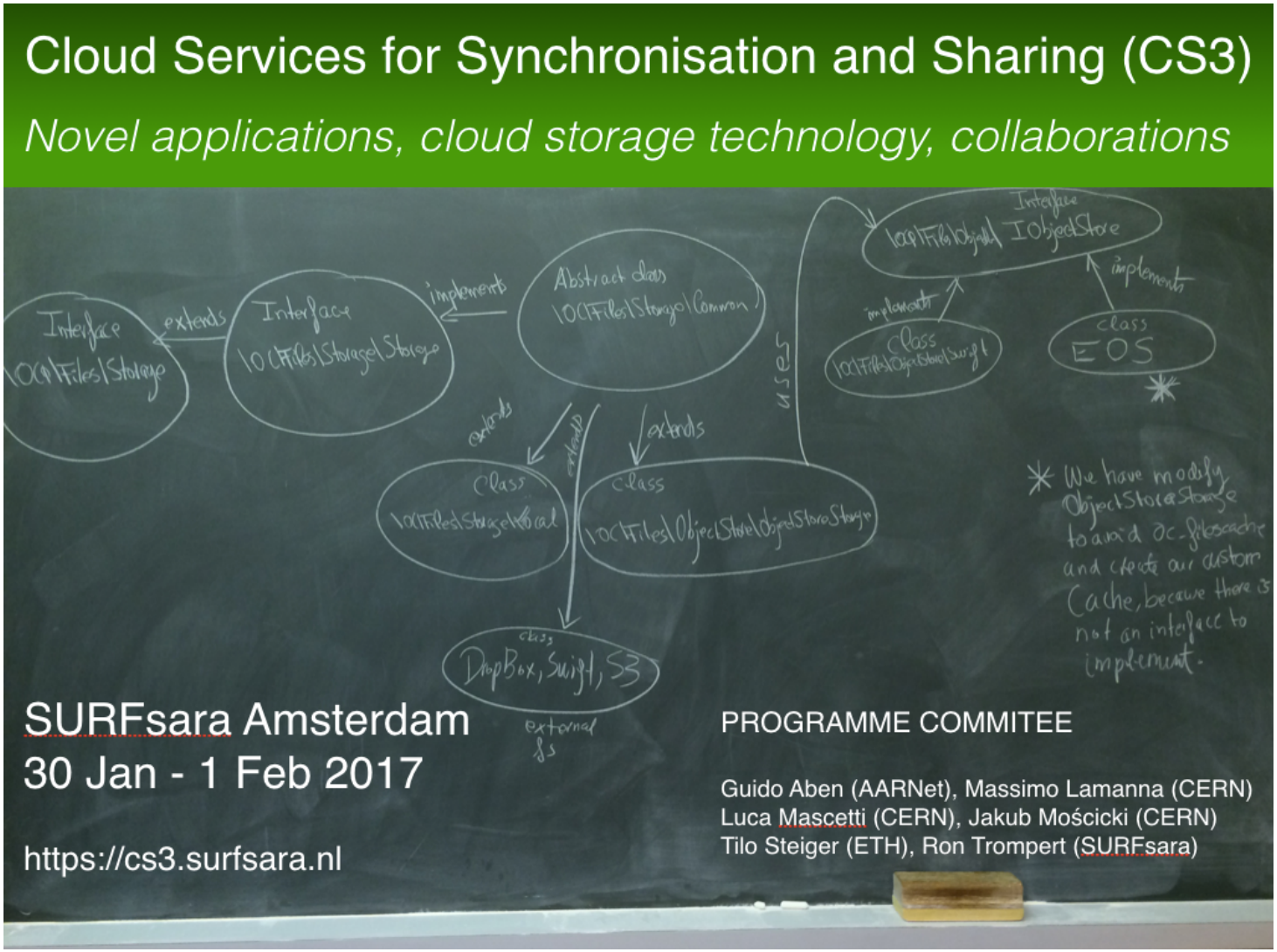


CERNBox

Storage Operations at CERN

Luca Mascetti
CERN - IT Storage

Just back from CS3 in Amsterdam



Dropbox
APIWISE
pydio
PUT YOUR DATA IN ORBIT
ownCloud
luckycloud
AGH
DDN STORAGE
360 ICT
SwiftStack
PowerFolder
ZIP IT Projecten
IBM
NetApp
Nextcloud
Seafile
CSC
cybera
UNINETT
TH BINGEN
University of Applied Sciences
JG|U
UNIVERSITAS MASARYKIANA BRUNENSIS
ONEDATA
ENEA
HUMBOLDT-UNIVERSITÄT ZU BERLIN
CERN
LINKÖPINGS UNIVERSITET
ktu kaunas university of technology 1922
cynny space
INFN Istituto Nazionale di Fisica Nucleare
aarnet
DESY
ETH Zürich
JRC EUROPEAN COMMISSION
GÉANT
CESNET
SURF SARA
UCL
LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
CYFRONET
UNIVERSITÉ DE STRASBOURG
NRF National Research Foundation
SAAO South African Astronomical Observatory
Consortium GARR
SWITCH
MAX-PLANCK-GESELLSCHAFT
JÜLICH FORSCHUNGSZENTRUM
WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
BROOKHAVEN NATIONAL LABORATORY

130 people attending the event

Outline

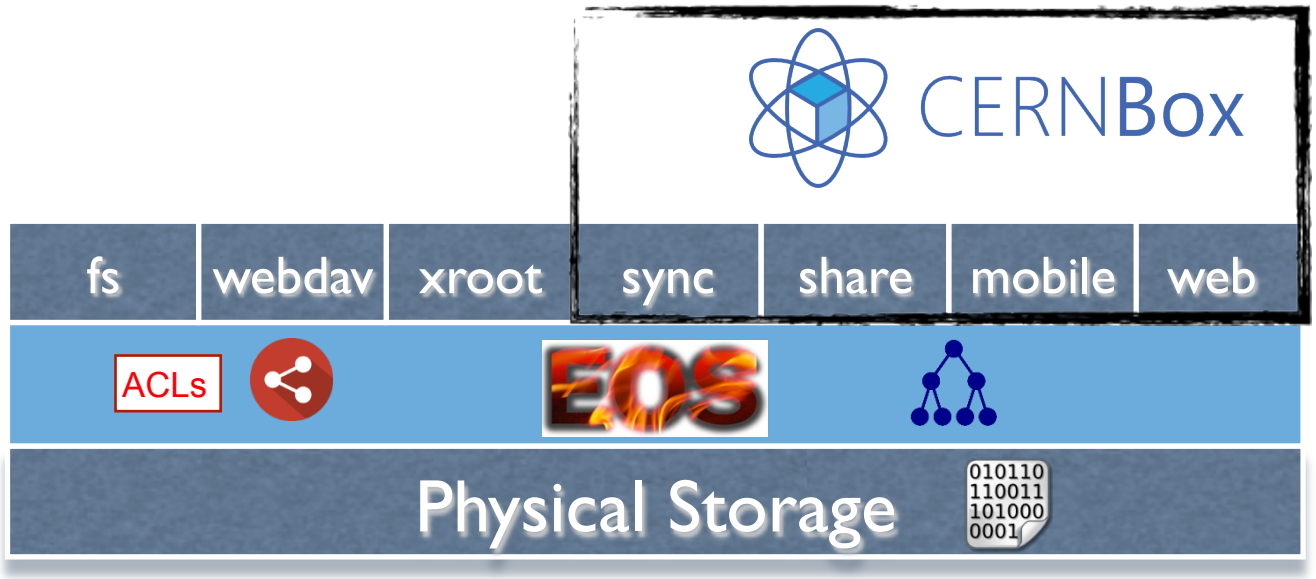
- Description and Architecture
- User Community and Service Numbers
- New features and ongoing work
- R&D Global EOS
- Overview and Summary

Description and Architecture

What is CERNBox ?

CERNBox provides a cloud synchronisation service

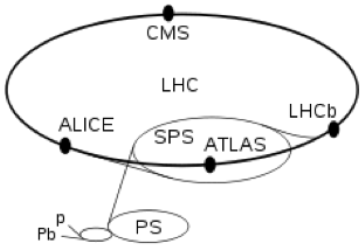
- Available for all CERN users (1TB/user)
- Synchronise files (data at CERN) and offline data access
- Easy way to share with other users
- All major platforms supported
- Based on **ownCloud integrated with EOS**



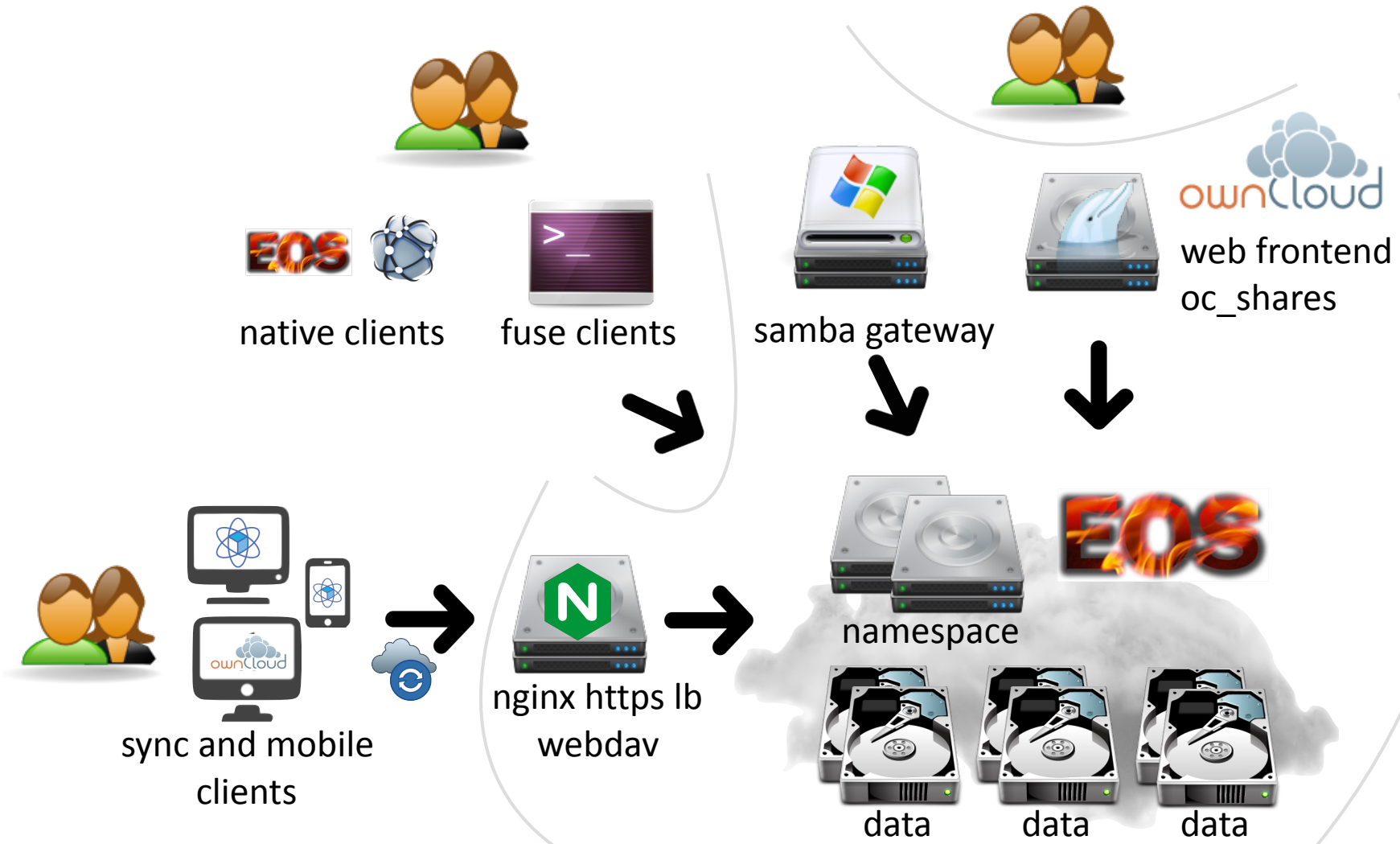
What is EOS ?

Files in the Namespace	Directories in the Namesapce	Used Space	Free Space
1.059 Bil	93.5 Mil	110.8 PB	37.8 PB

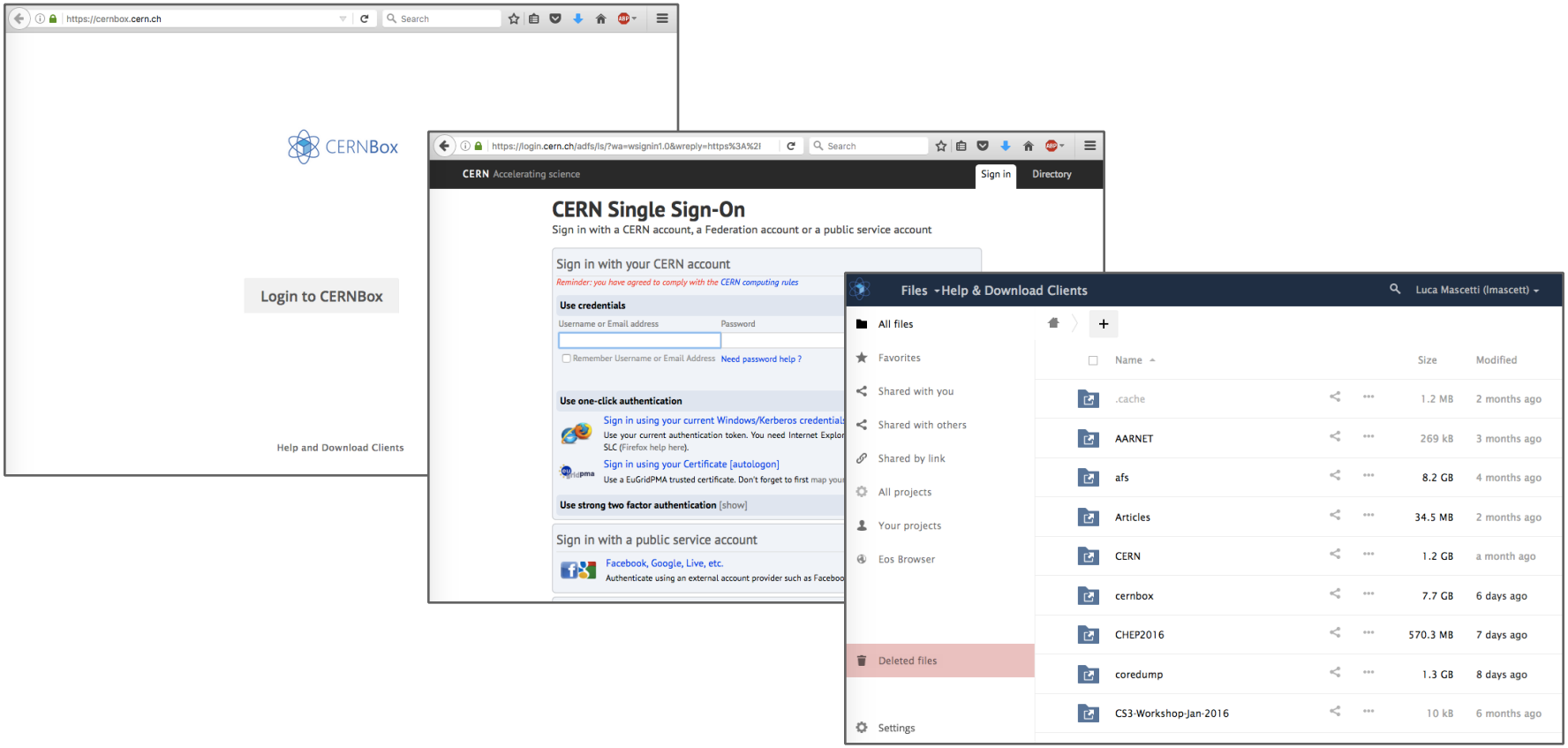
- CERNBox storage backend
- Disk Storage for all LHC and physics data
- Deployment across two computer centres



CERNBox Architecture

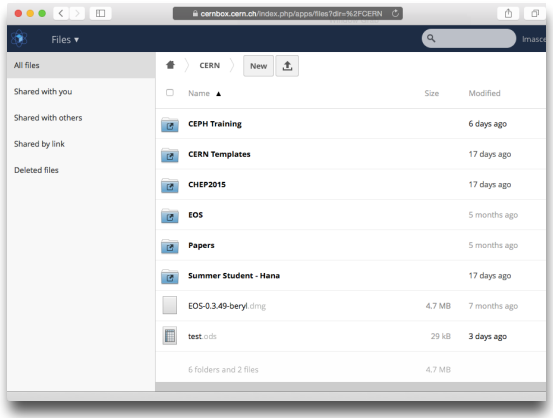


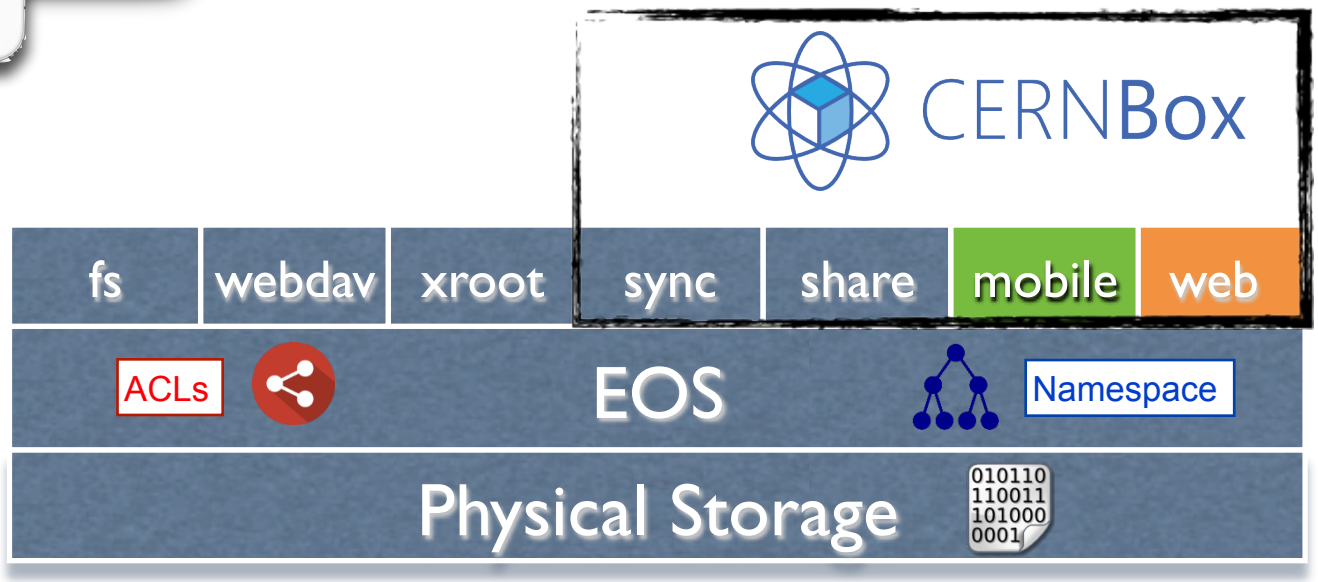
Single Sign On



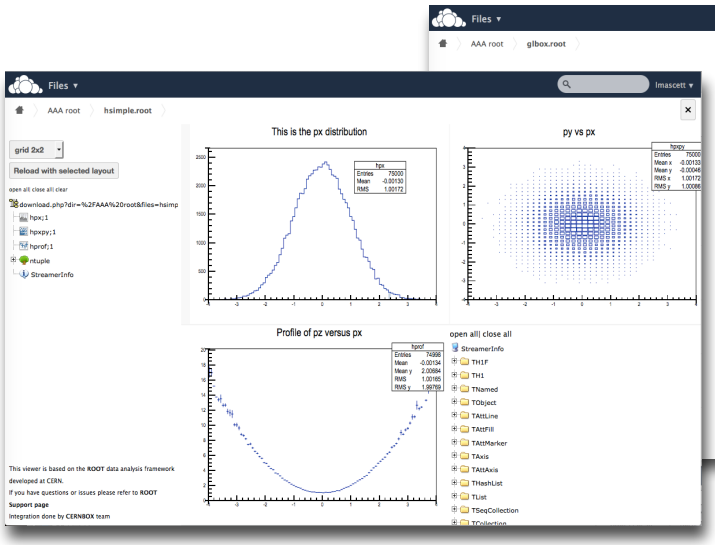
- SSO enabled to align with the other CERN web services
  - Desktop Sync Client continue using standard auth

Access Methods: Mobile & Web

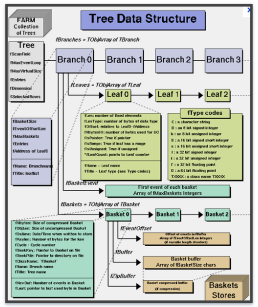


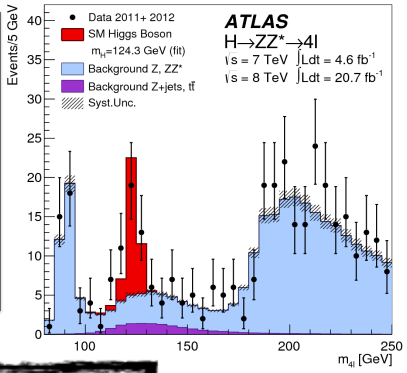


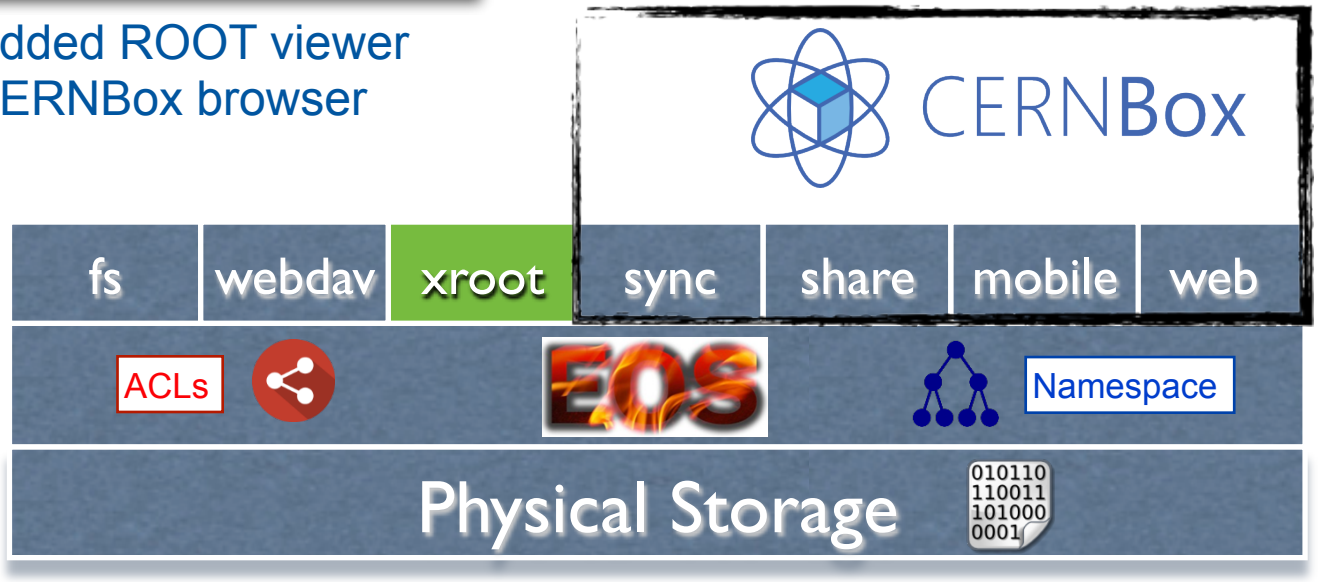
Access Methods: xroot & ROOT



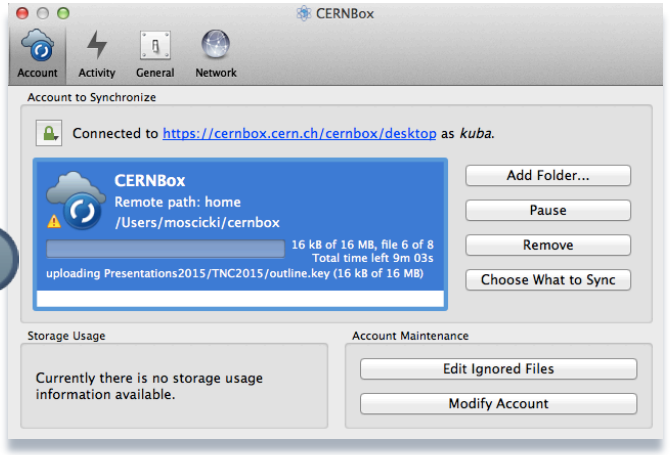
Embedded ROOT viewer
in CERNBox browser

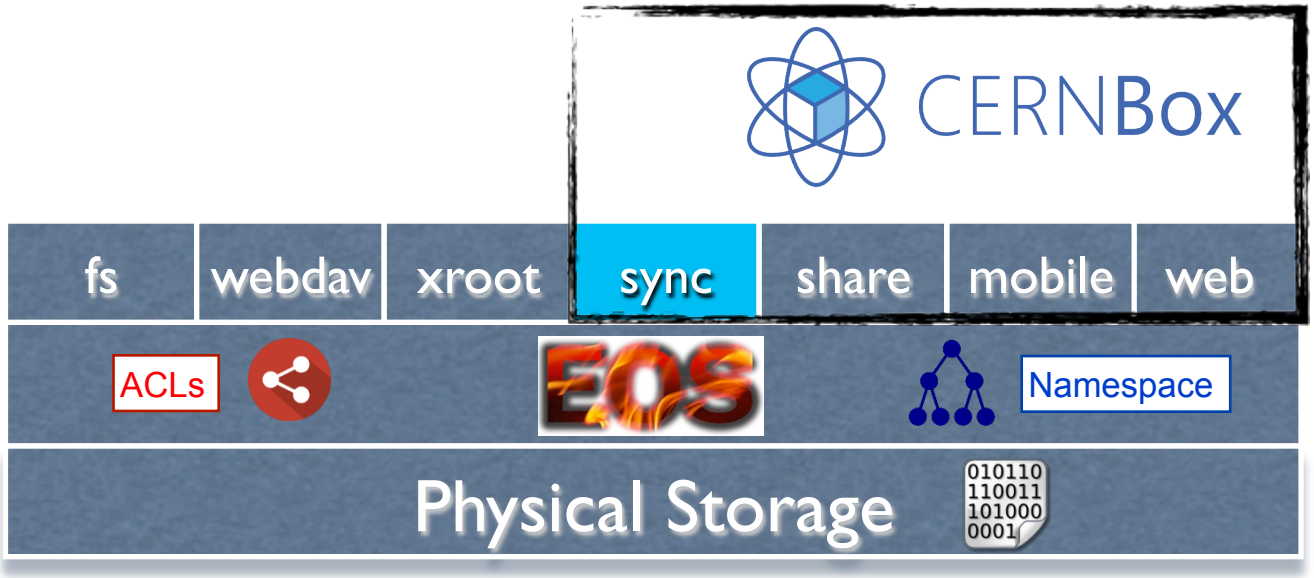




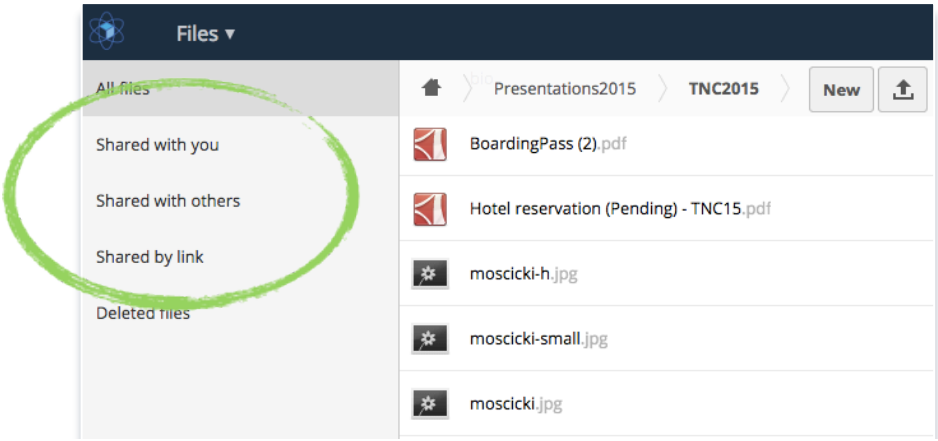


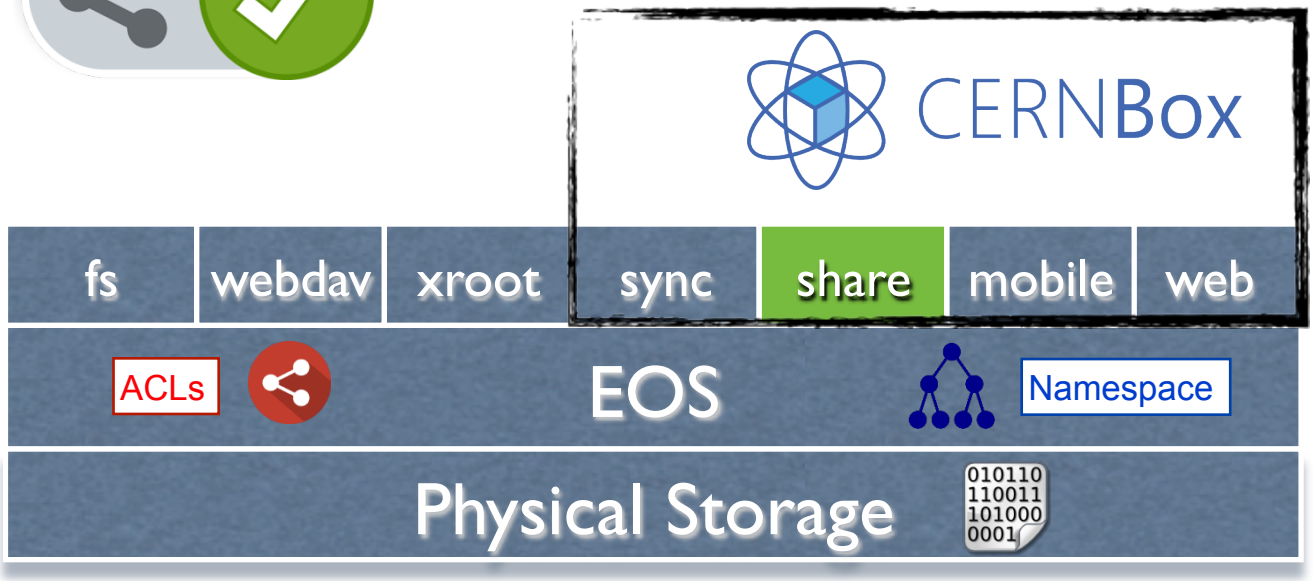
Access Methods: Sync





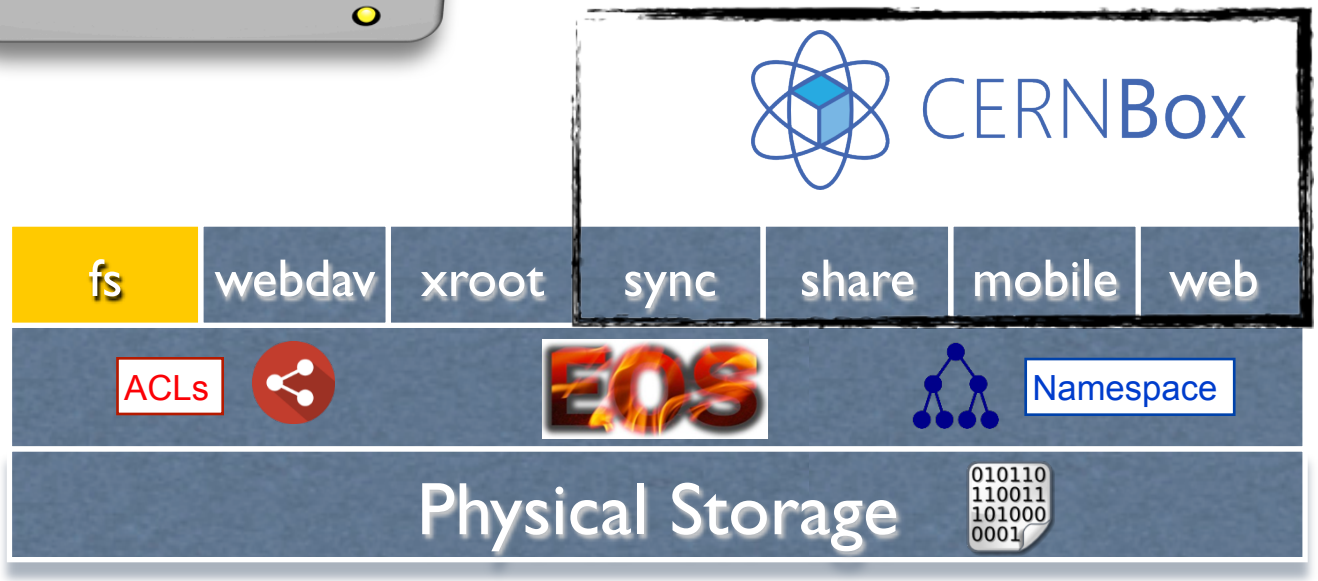
Access Methods: Sharing



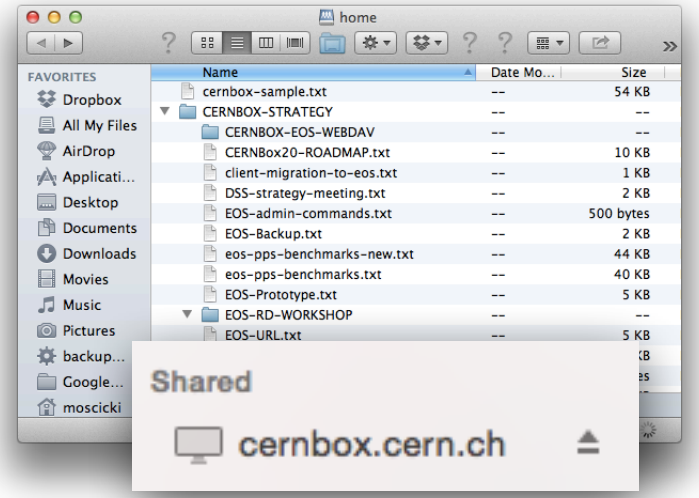


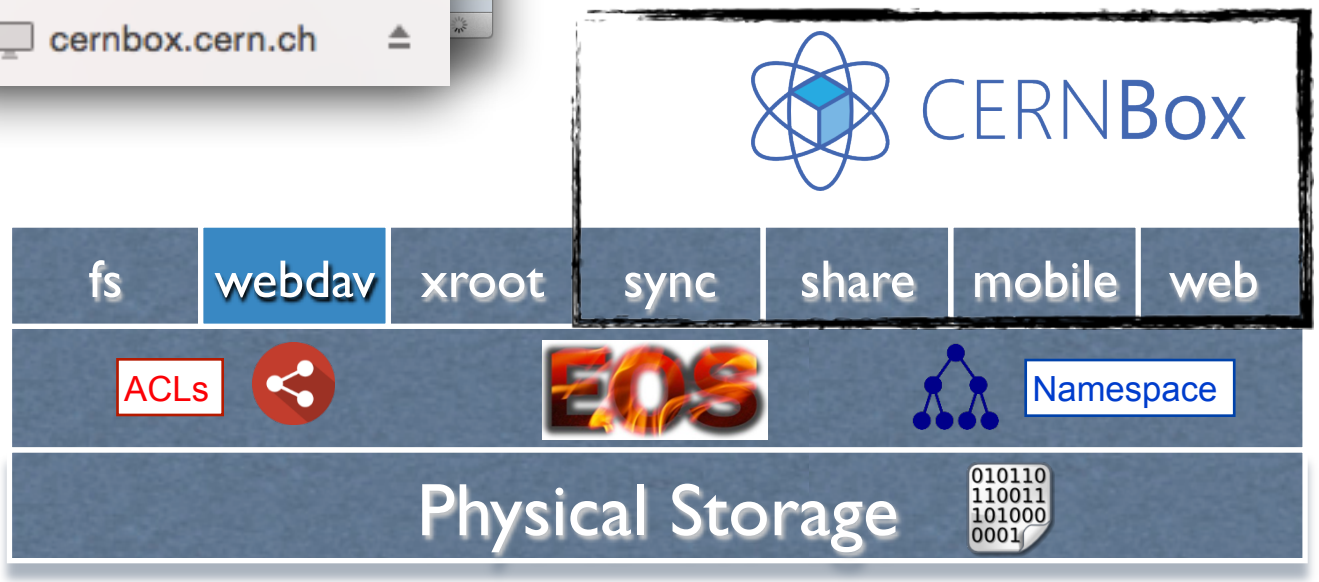
Access Methods: FUSE

```
[lmascett@lxplus2015 ~]#
[lmascett@lxplus2015 ~]# df -H -t fuse
Filesystem      Size  Used Avail Use% Mounted on
eosuser         506T   70T  437T  14% /eos/user
eosatlas         36P   17P   20P  45% /eos/atlas
eosalice         20P   11P  8.5P  57% /eos/alice
eoscms           28P   14P   15P  49% /eos/cms
eoslhcb          13P  7.6P  4.6P  63% /eos/lhcb
eospublic        16P  5.8P   11P  36% /eos/public
[lmascett@lxplus2015 ~]#
[lmascett@lxplus2015 ~]# ls -lc /eos/user/l/lmascett/
total 6644
drwx------. 1 lmascett c3       5 Dec 10 15:58 CERN
drwx------. 1 lmascett c3       0 Jan 26 18:18 debug
drwx------. 1 lmascett c3       0 Dec 11 09:43 download
drwx------. 1 lmascett c3       0 Oct 31 18:24 pdf
drwx------. 1 lmascett c3       1 Dec 11 09:44 personal
drwx------. 1 lmascett c3       8 Dec 10 12:11 pictures
```

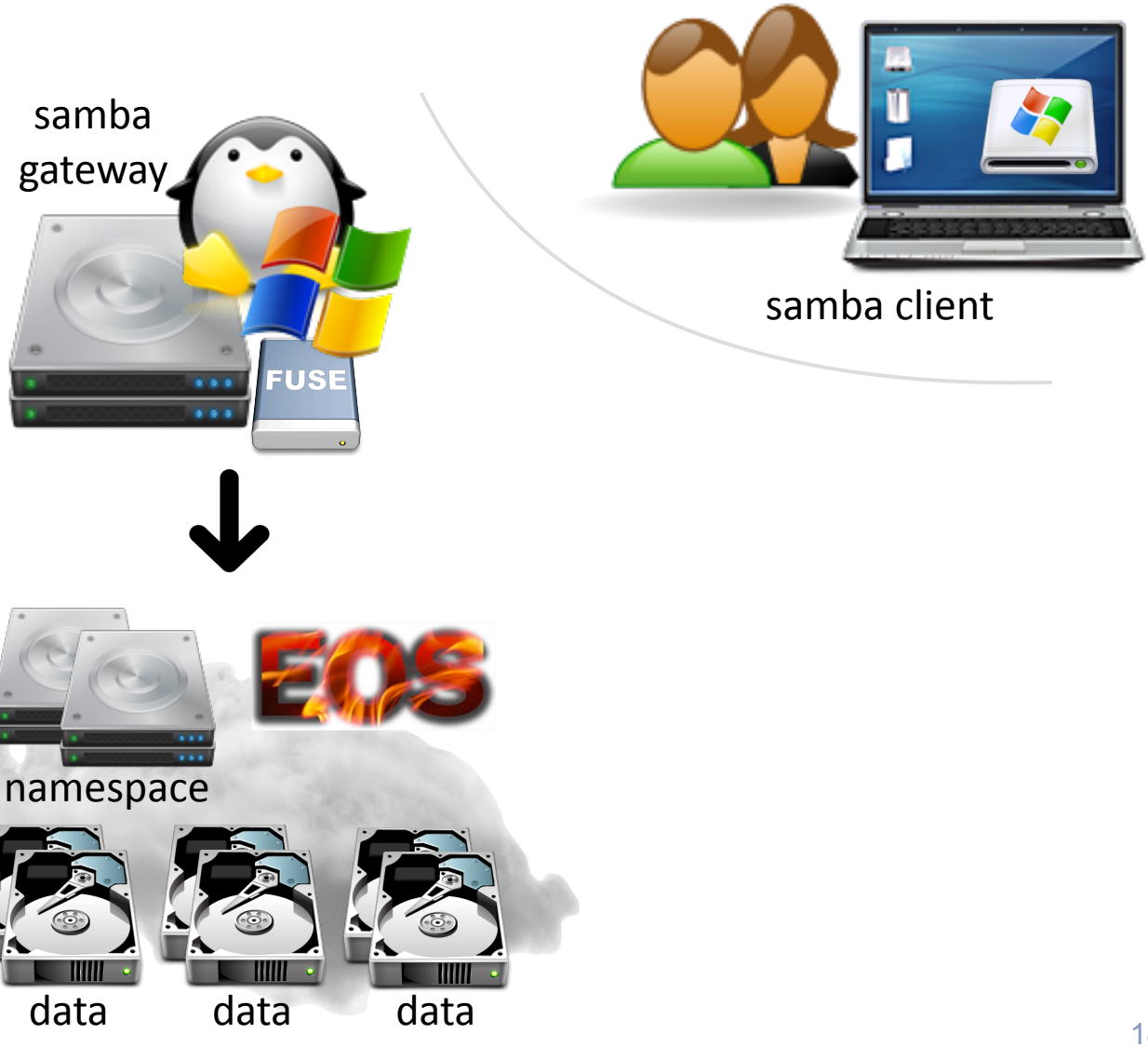


Access Methods: WebDAV





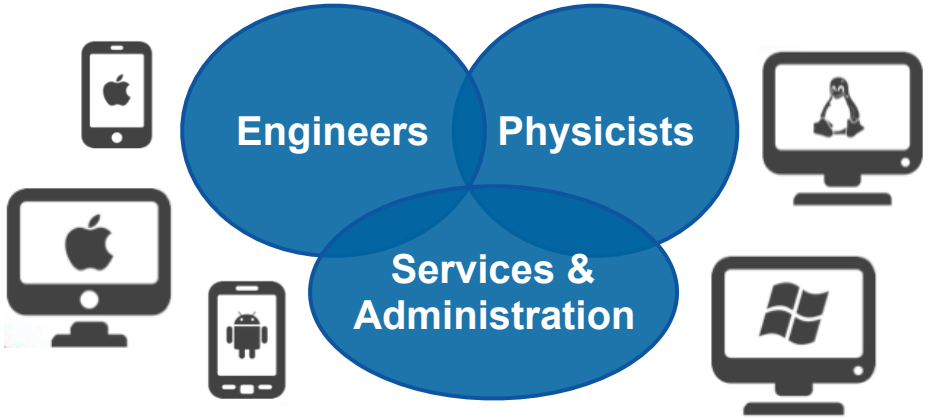
Access Methods: SAMBA

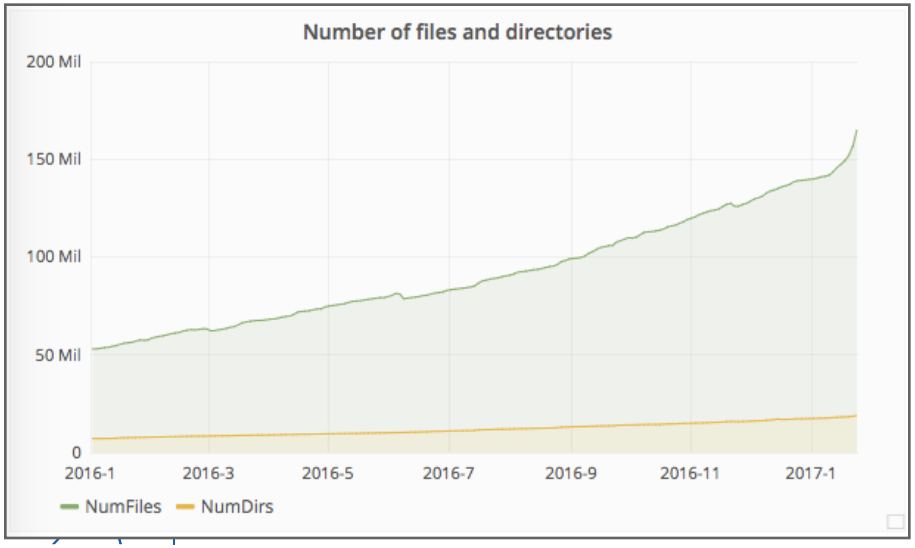


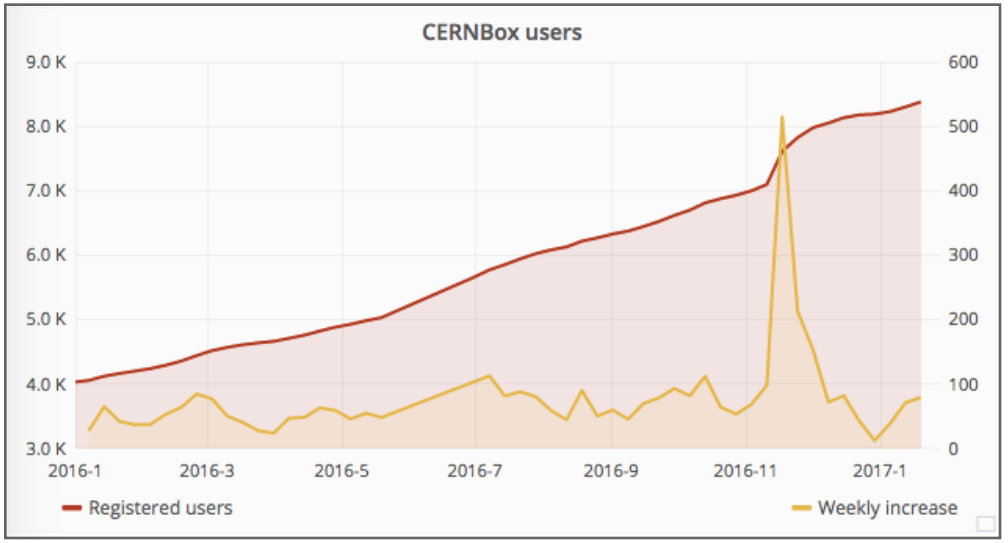
User Community and Service Numbers

CERNBox Service Numbers

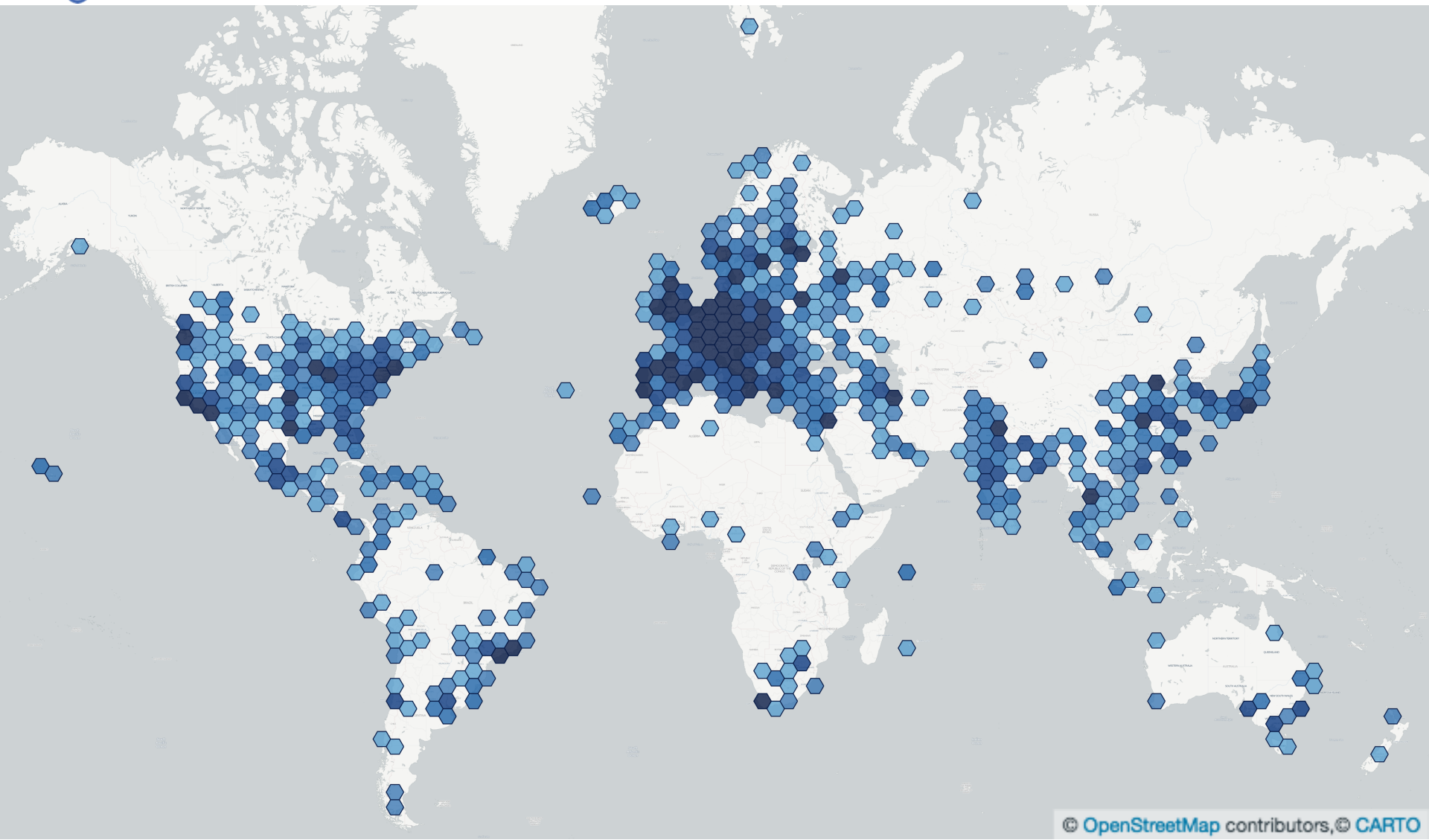
	Jan 2016	Jan 2017
Users	4074	8411
# files	55 Million	176 Million
# dirs	7.2 Million	19 Million
Used Raw Space	208 TB	806 TB
Deployed Raw Space	1.3 PB	3.2 PB

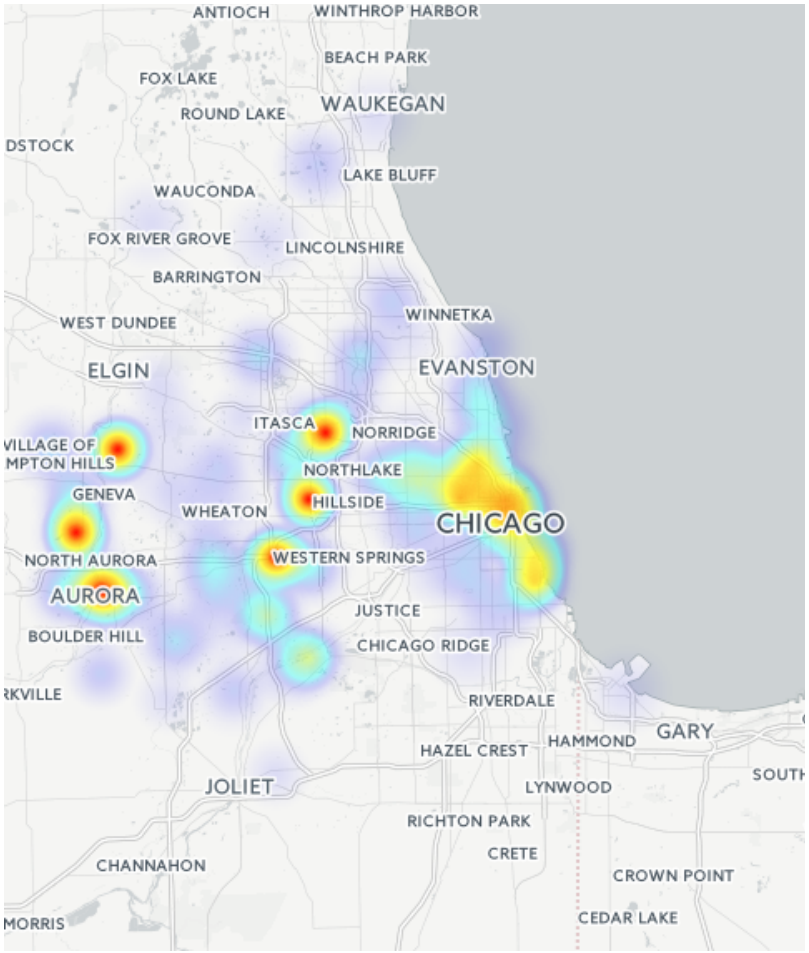


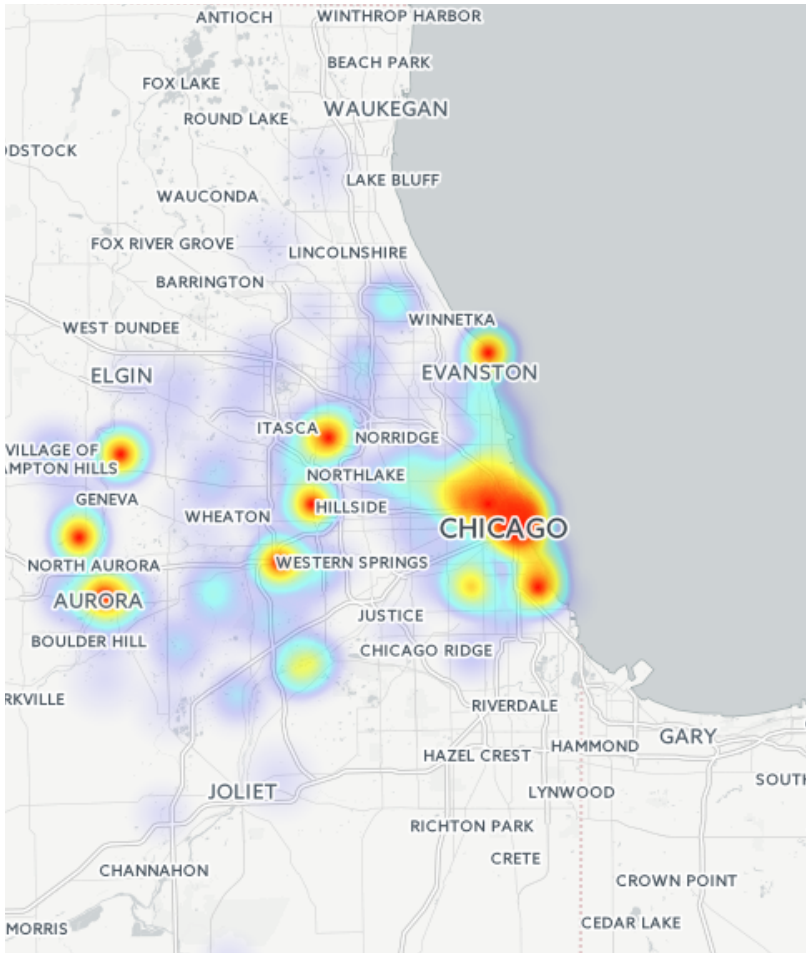




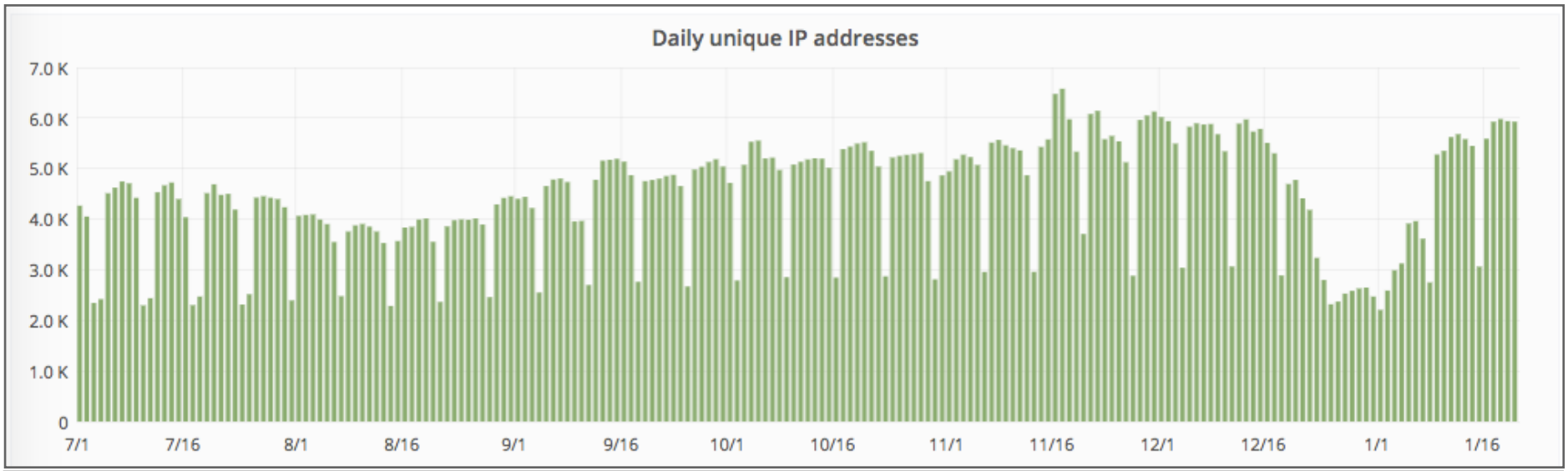
CERNBox Clients

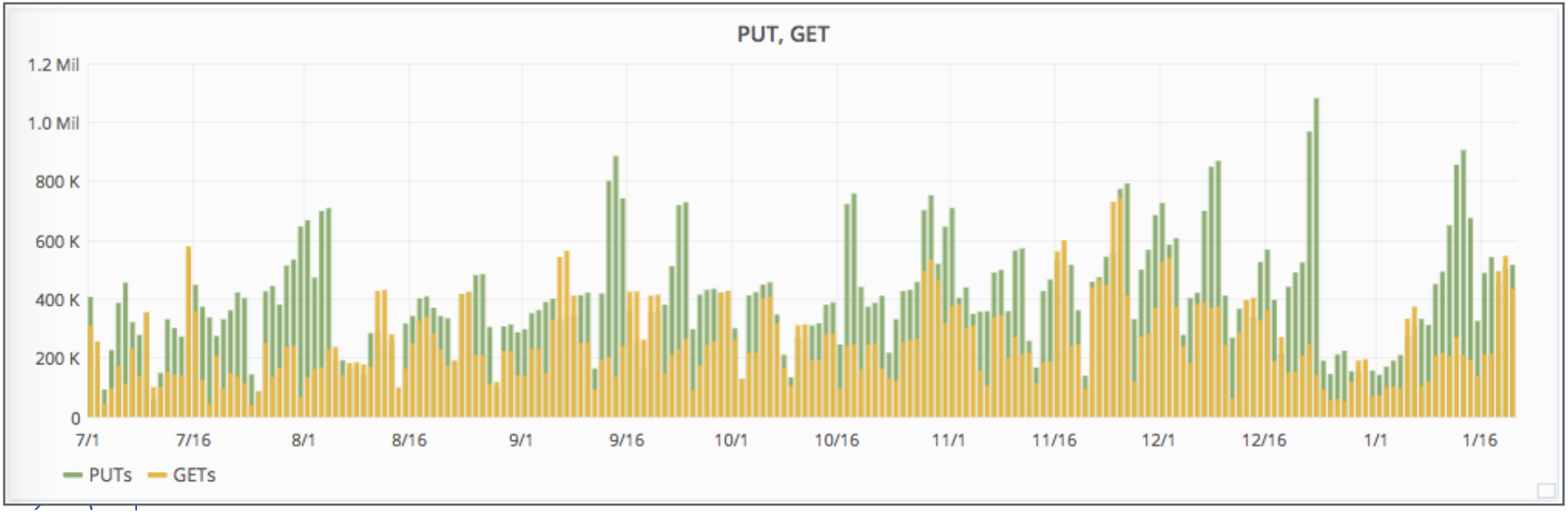



ANTIOCH
WINTHROP HARBOR
BEACH PARK
FOX LAKE
ROUND LAKE
WAUKEGAN
LAKE BLUFF
WAUCONDA
FOX RIVER GROVE
LINCOLNSHIRE
BARRINGTON
WEST DUNDEE
WINNETKA
ELGIN
EVANSTON
ITASCA
NORRIDGE
NORTHLAKE
GENEVA
WHEATON
HILLSIDE
CHICAGO
NORTH AURORA
WESTERN SPRINGS
AURORA
JUSTICE
BOULDER HILL
CHICAGO RIDGE
RIVERDALE
HAZEL CREST
HAMMOND
GARY
JOLIET
LYNWOOD
RICHTON PARK
CRETE
CHANNAHON
CROWN POINT
CEDAR LAKE
MORRIS

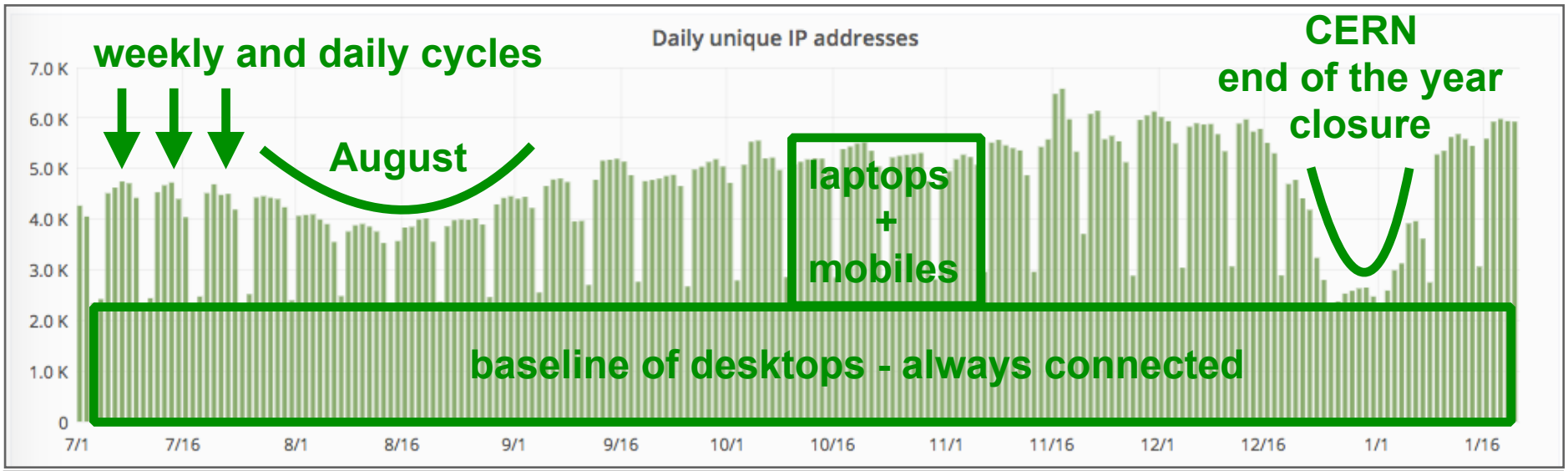

ANTIOCH
WINTHROP HARBOR
BEACH PARK
FOX LAKE
ROUND LAKE
WAUKEGAN
LAKE BLUFF
WAUCONDA
FOX RIVER GROVE
LINCOLNSHIRE
BARRINGTON
WEST DUNDEE
WINNETKA
ELGIN
EVANSTON
ITASCA
NORRIDGE
NORTHLAKE
GENEVA
WHEATON
HILLSIDE
CHICAGO
NORTH AURORA
WESTERN SPRINGS
AURORA
JUSTICE
BOULDER HILL
CHICAGO RIDGE
RIVERDALE
HAZEL CREST
HAMMOND
GARY
JOLIET
LYNWOOD
RICHTON PARK
CRETE
CHANNAHON
CROWN POINT
CEDAR LAKE
MORRIS

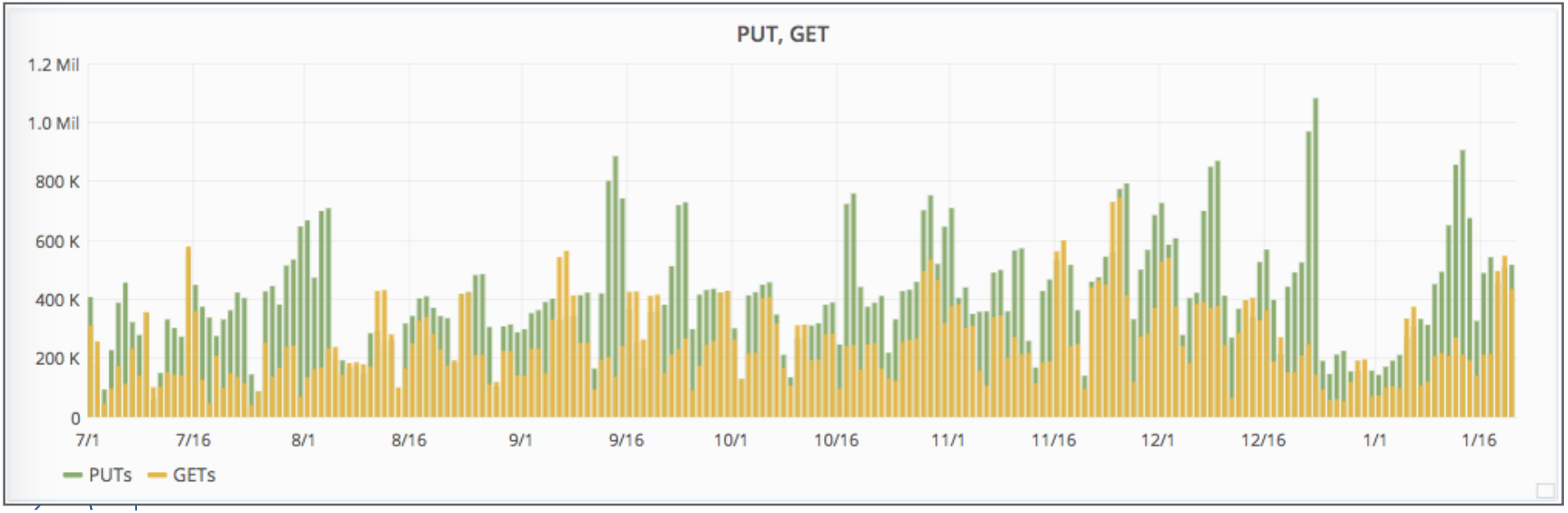
CERNBox daily clients



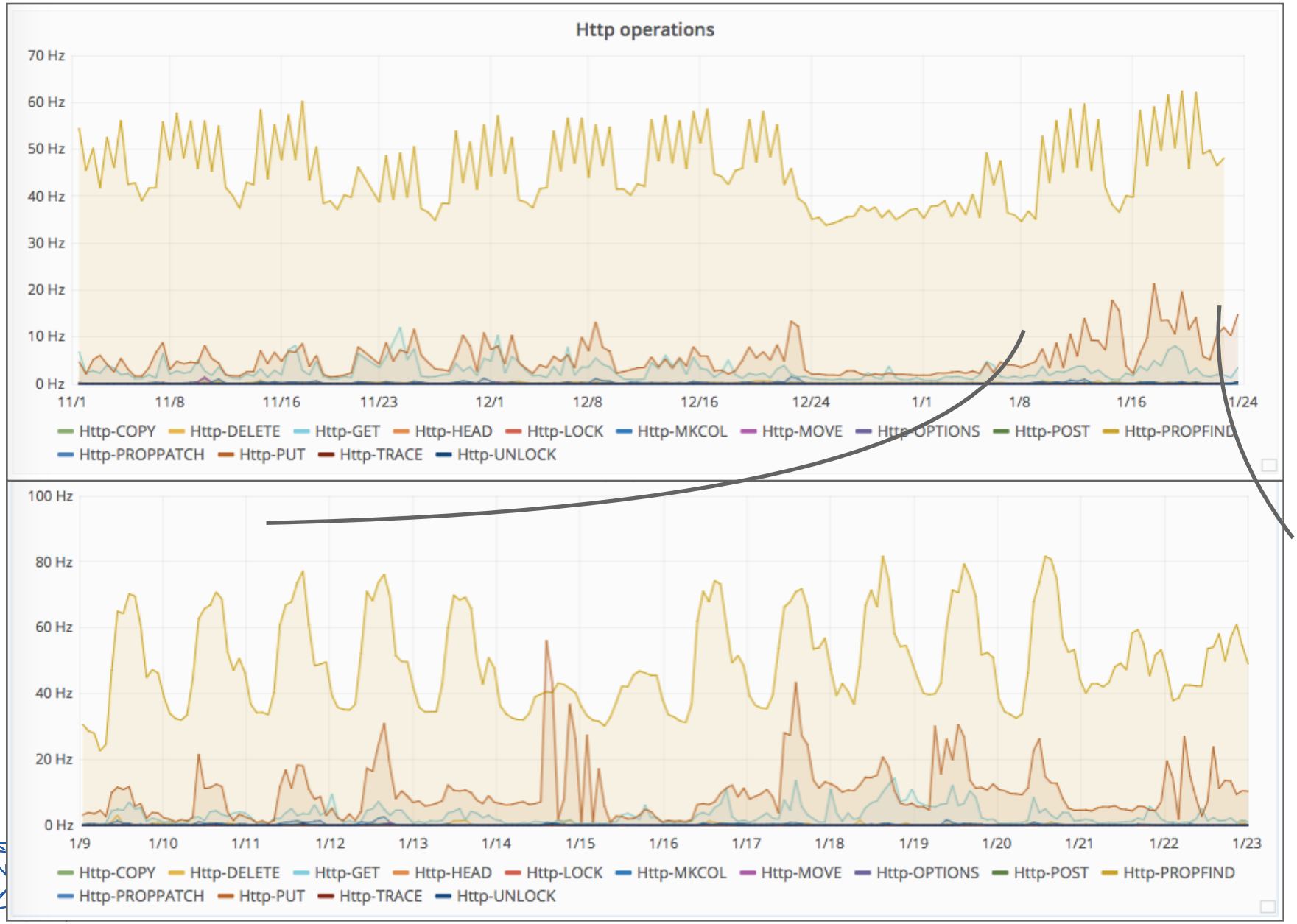


CERNBox daily clients



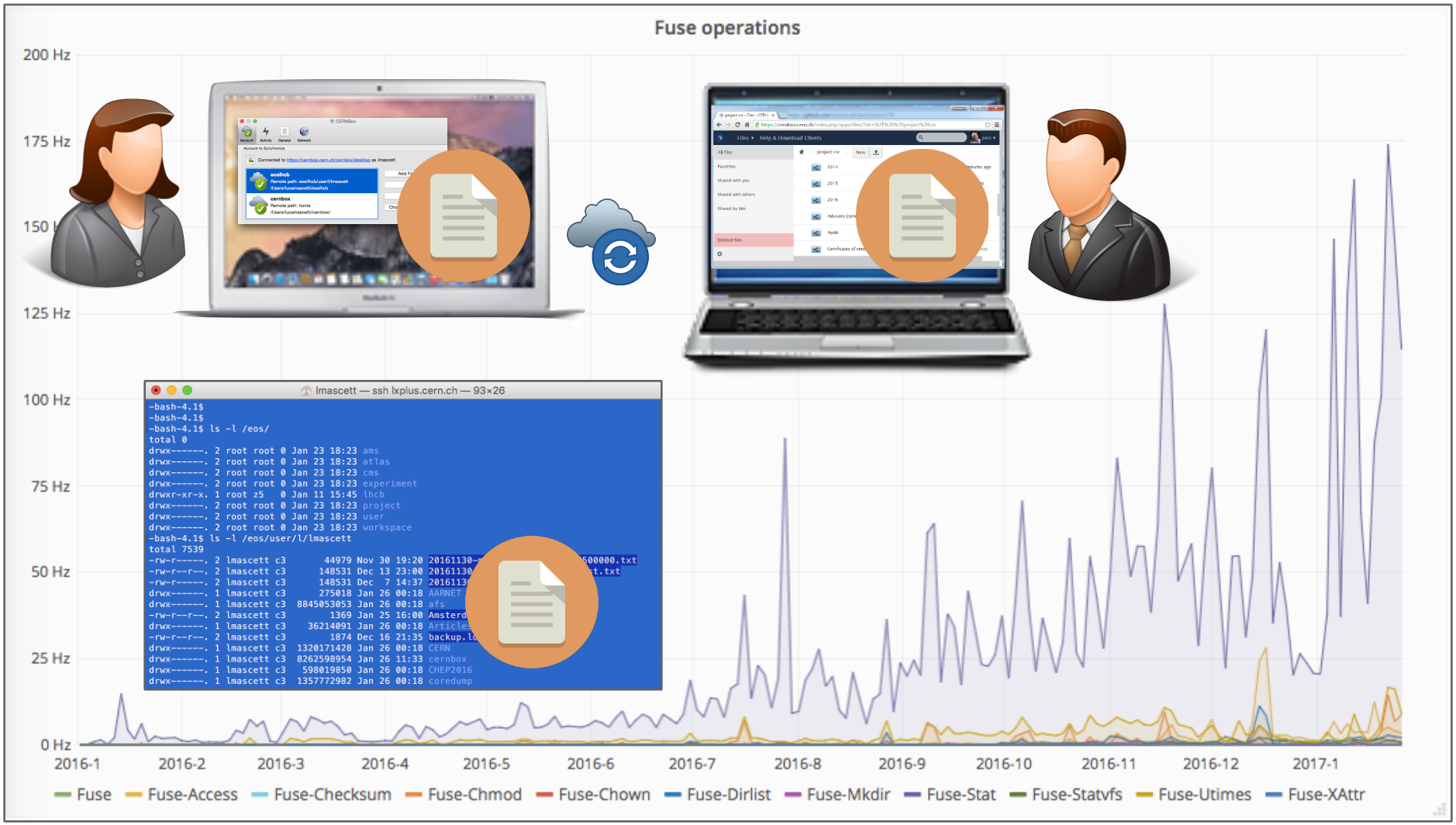


CERNBox HTTP Operations

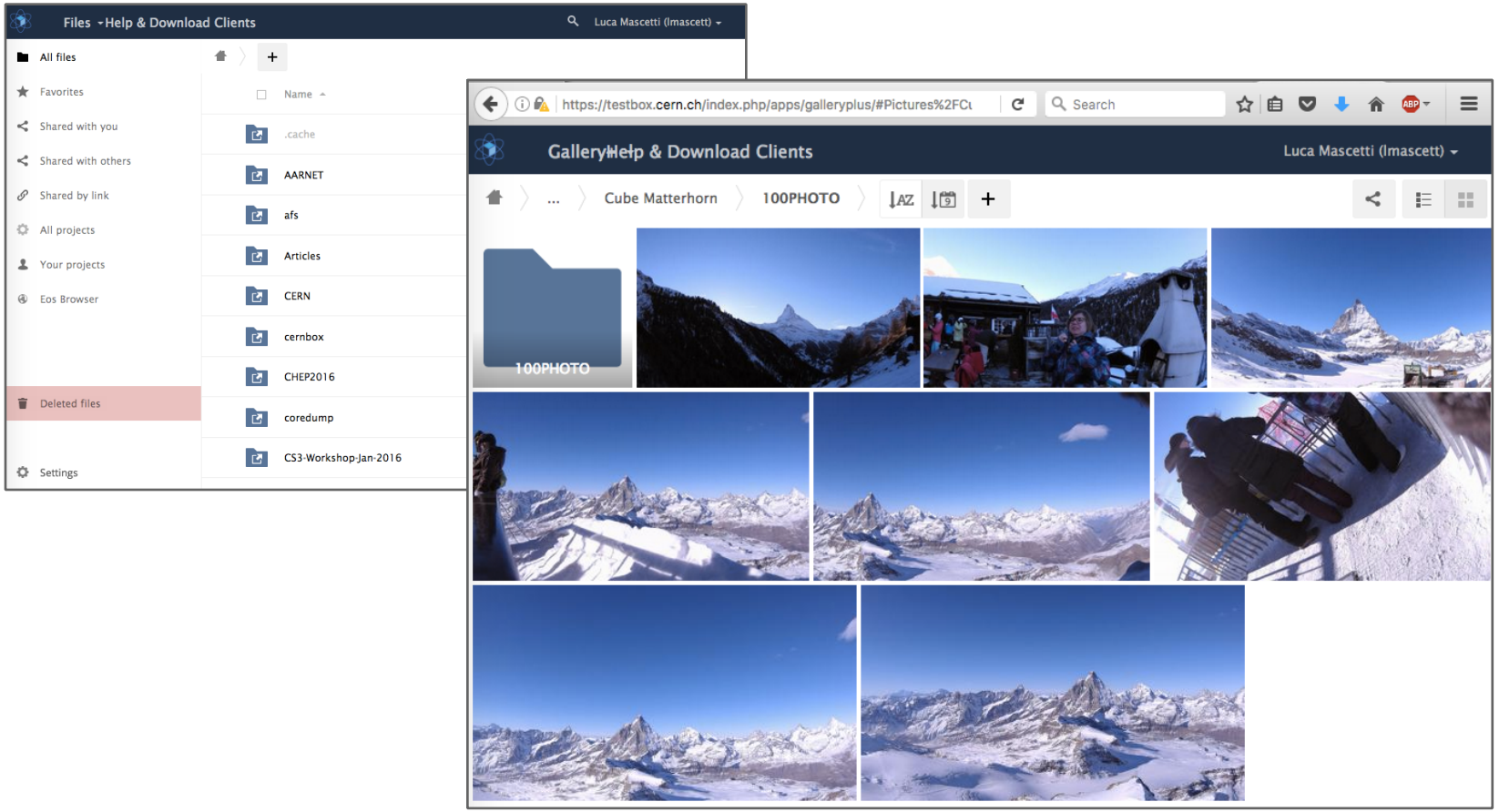


New features and ongoing work

Filesystem Access (FUSE)

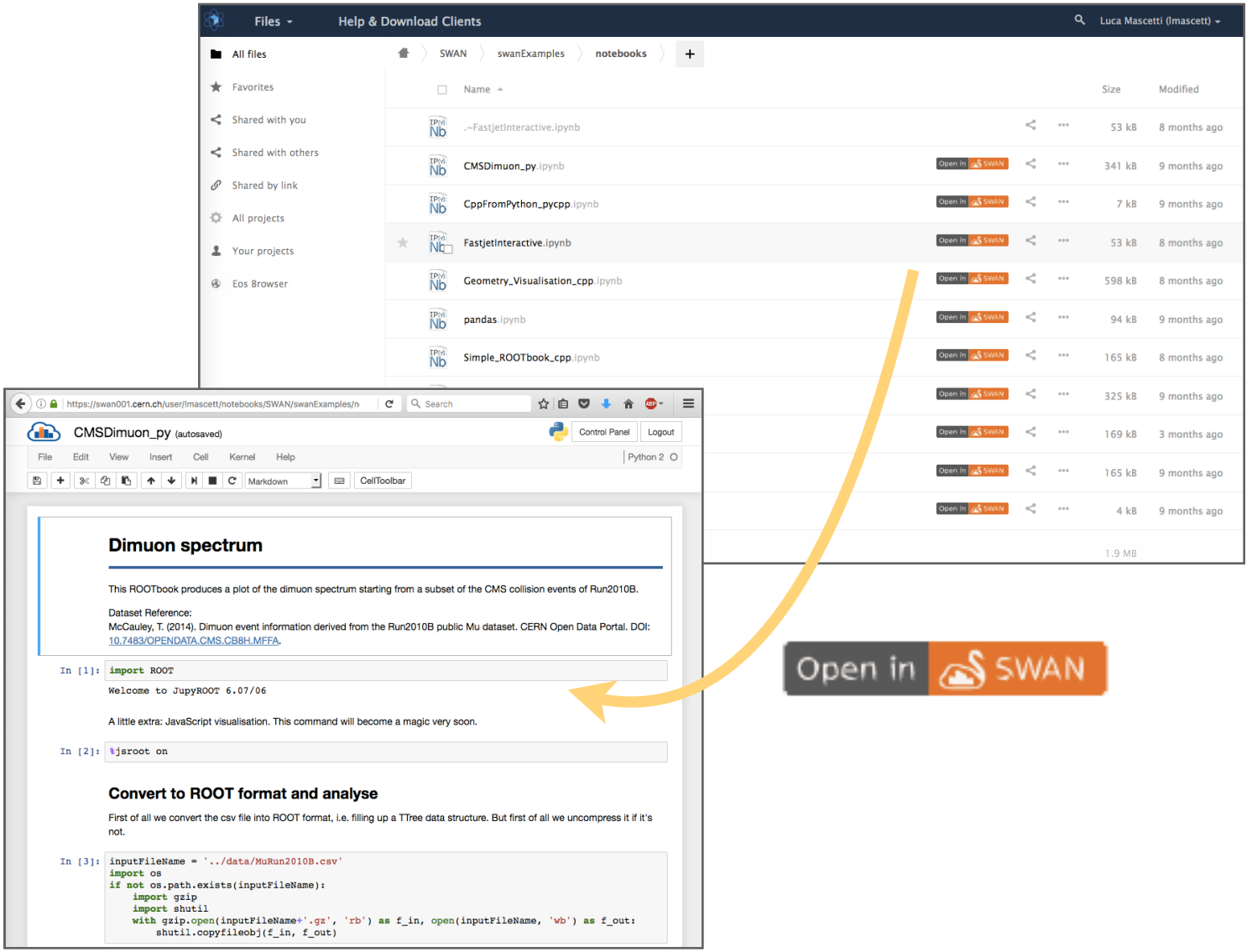


Gallery View



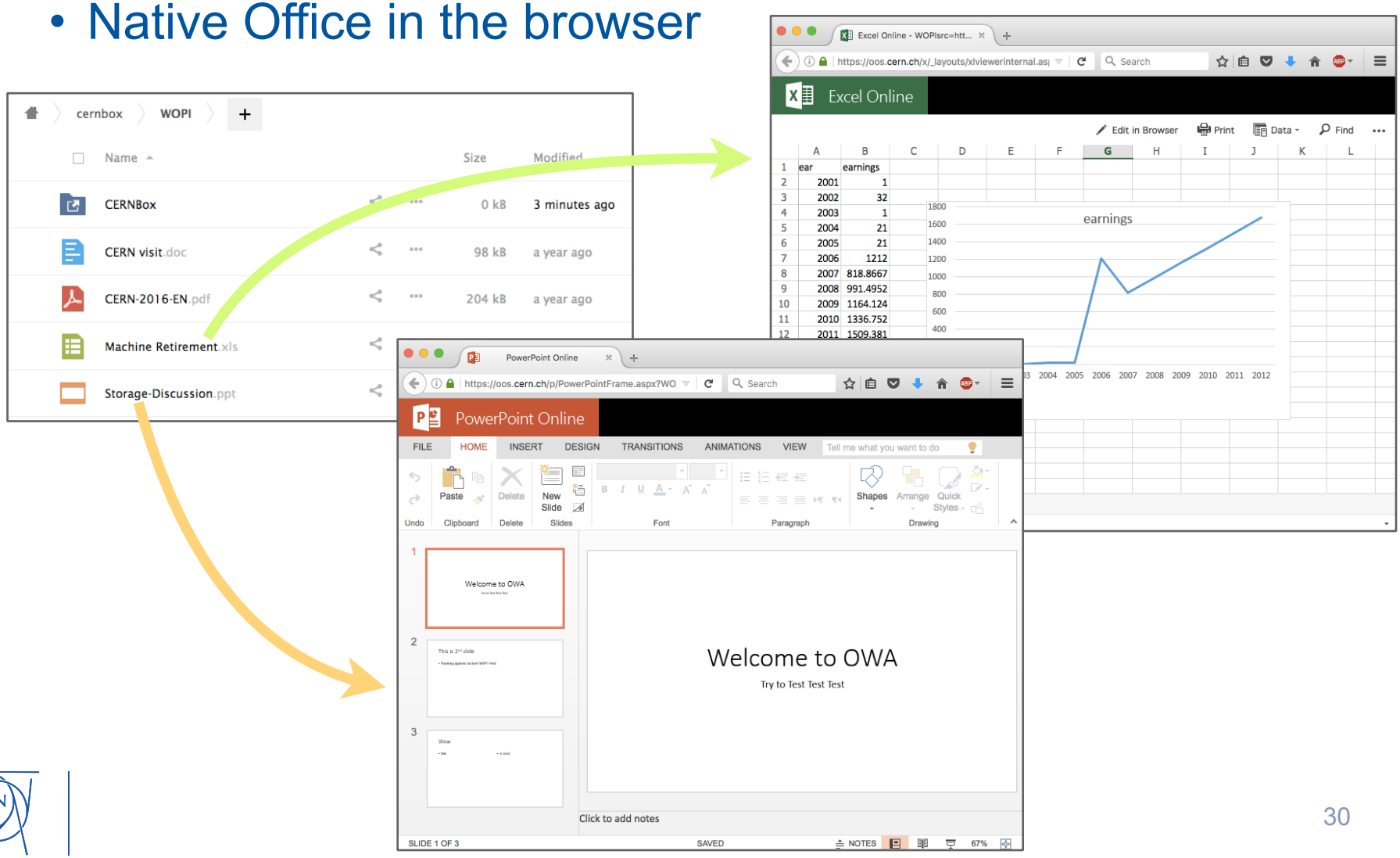
- Working on the enable the Gallery View
  - Integration between the Gallery App and EOS

Python Notebook Integration (SWAN)



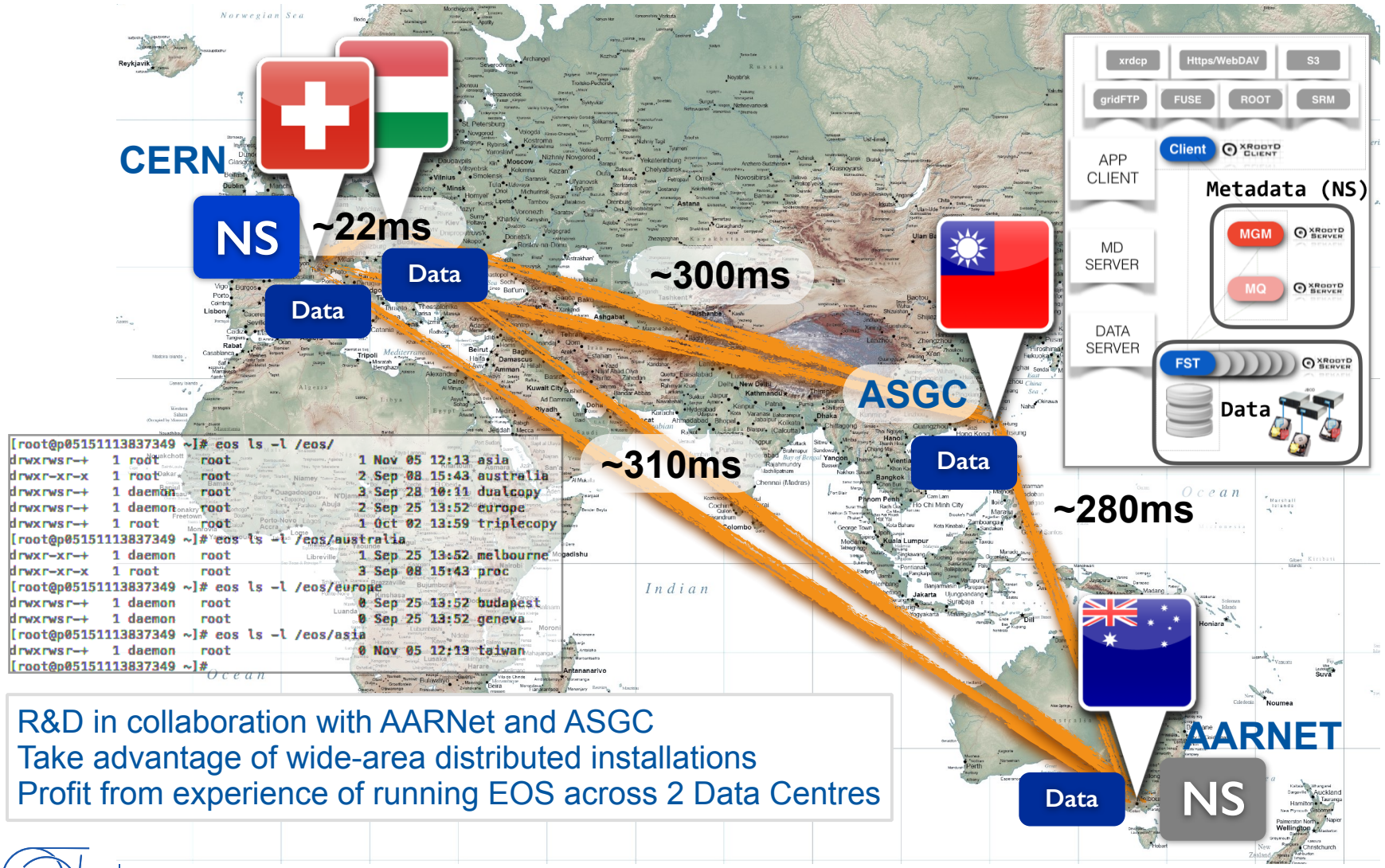
Office Online Integration (WOPI)

- Web Application Open Platform Interface (WOPI) protocol
  - Native Office in the browser



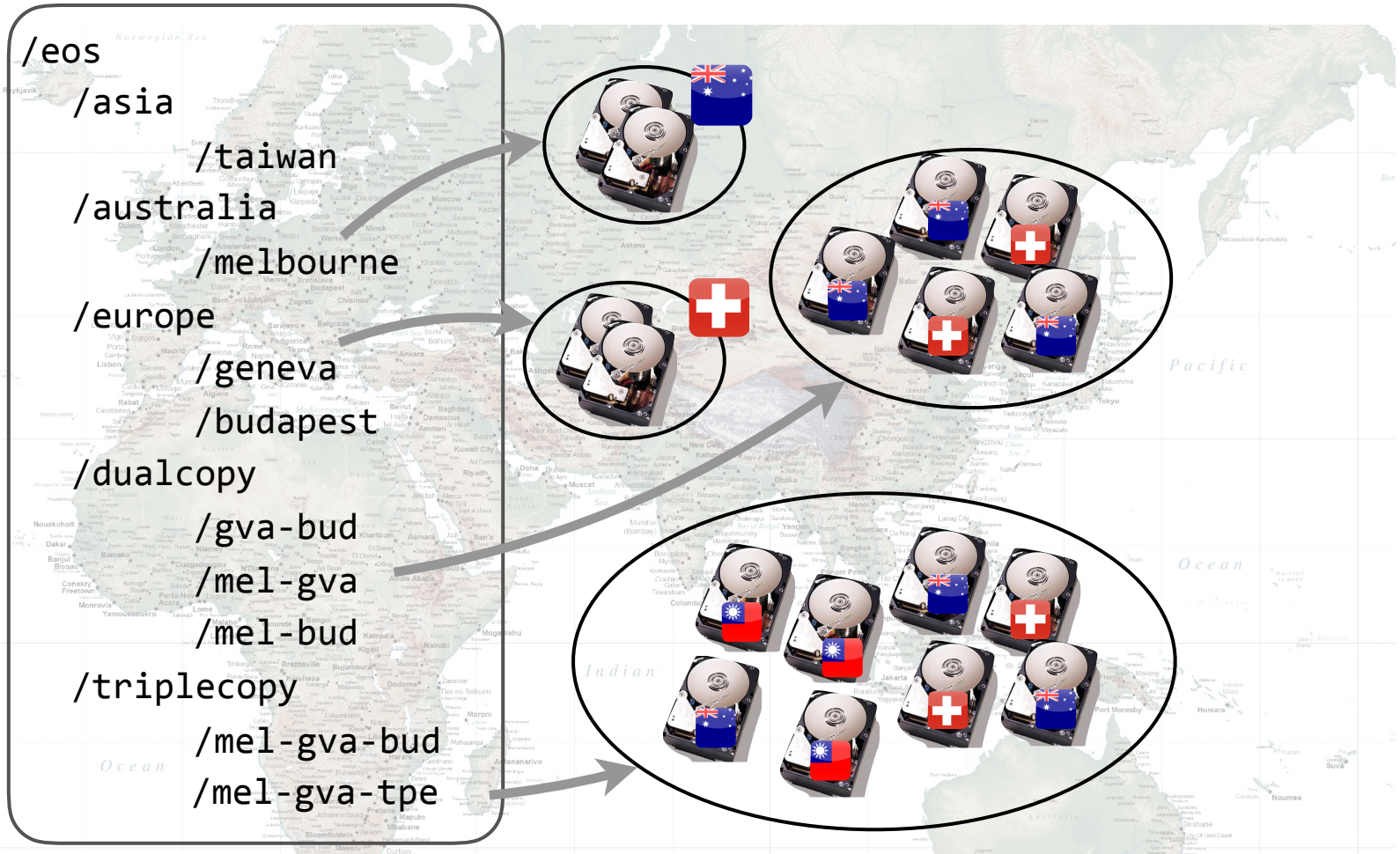
R&D Global EOS

Global EOS: World-Wide Deployment



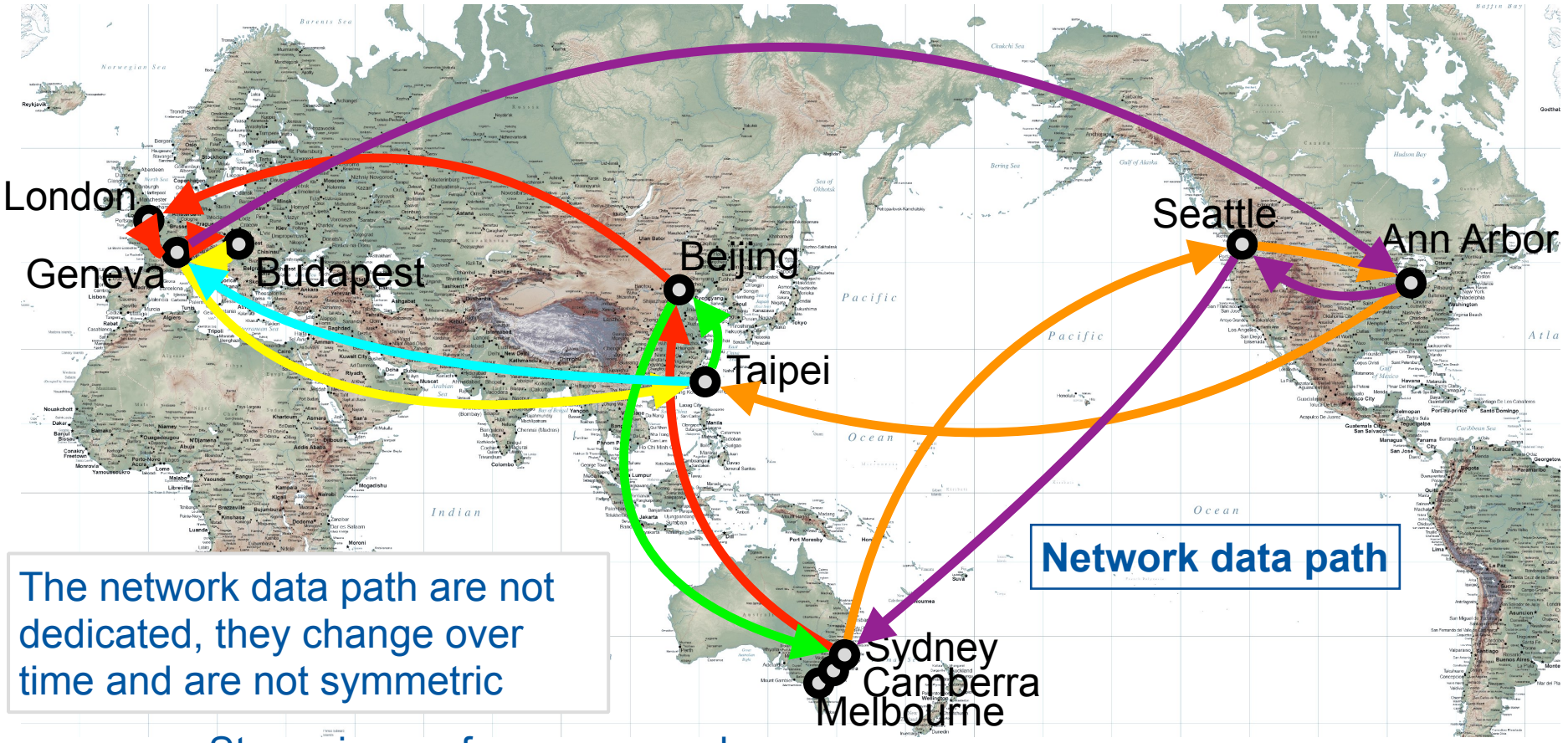
R&D in collaboration with AARNet and ASGC
Take advantage of wide-area distributed installations
Profit from experience of running EOS across 2 Data Centres

Data replication in storage pools



Storage pools were created with filesystems from all four sites. Files were replicated according to the different configured policy (e.g. 3 replicas: MEL-GVA-TPE).

Network Topology

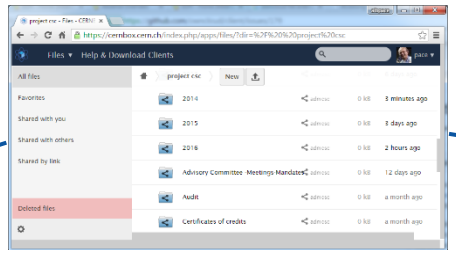


The network data path are not dedicated, they change over time and are not symmetric

- Streaming performance good
  - possible problems in case of packet drops (tcp window)
  - tcp settings could be optimised
- Latency in read hidden by the read-only NS
- Latency in write to contact the read-write NS

Overview and Summary

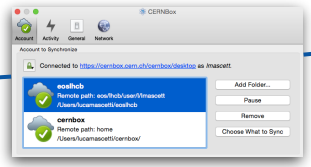
Future Home Directory ($HOME)



Browser and Mobile

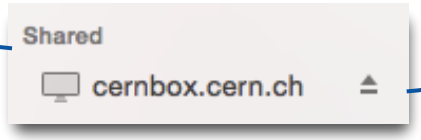
web access

offline access



Sync Client

online access



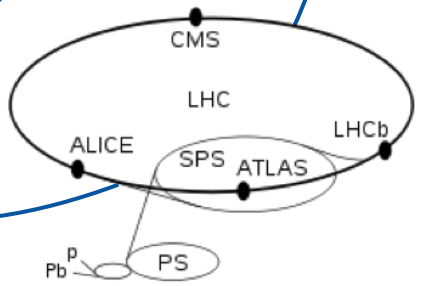
WebDAV & Fuse

direct access

http, webdav, xroot

Laptops, PCs & Mobile devices

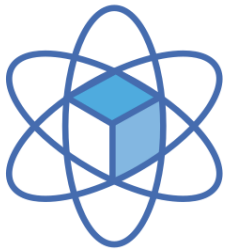
Batch Service



CERNBox + EOS scenarios

- end-user files and sharing
- project spaces
- physics analysis
- future home directory

Summary



CERNBox

- EOS/CERNBox innovative platform for scientific computing
  - scalable architecture
- Full integration with the entire LHC storage (100+ PB)
- Bring data closer to our users
  - New ways to interact with the data
- CERNBox as an extension of the Desktop
  - ongoing integration with multiple services

CERN

www.cern.ch