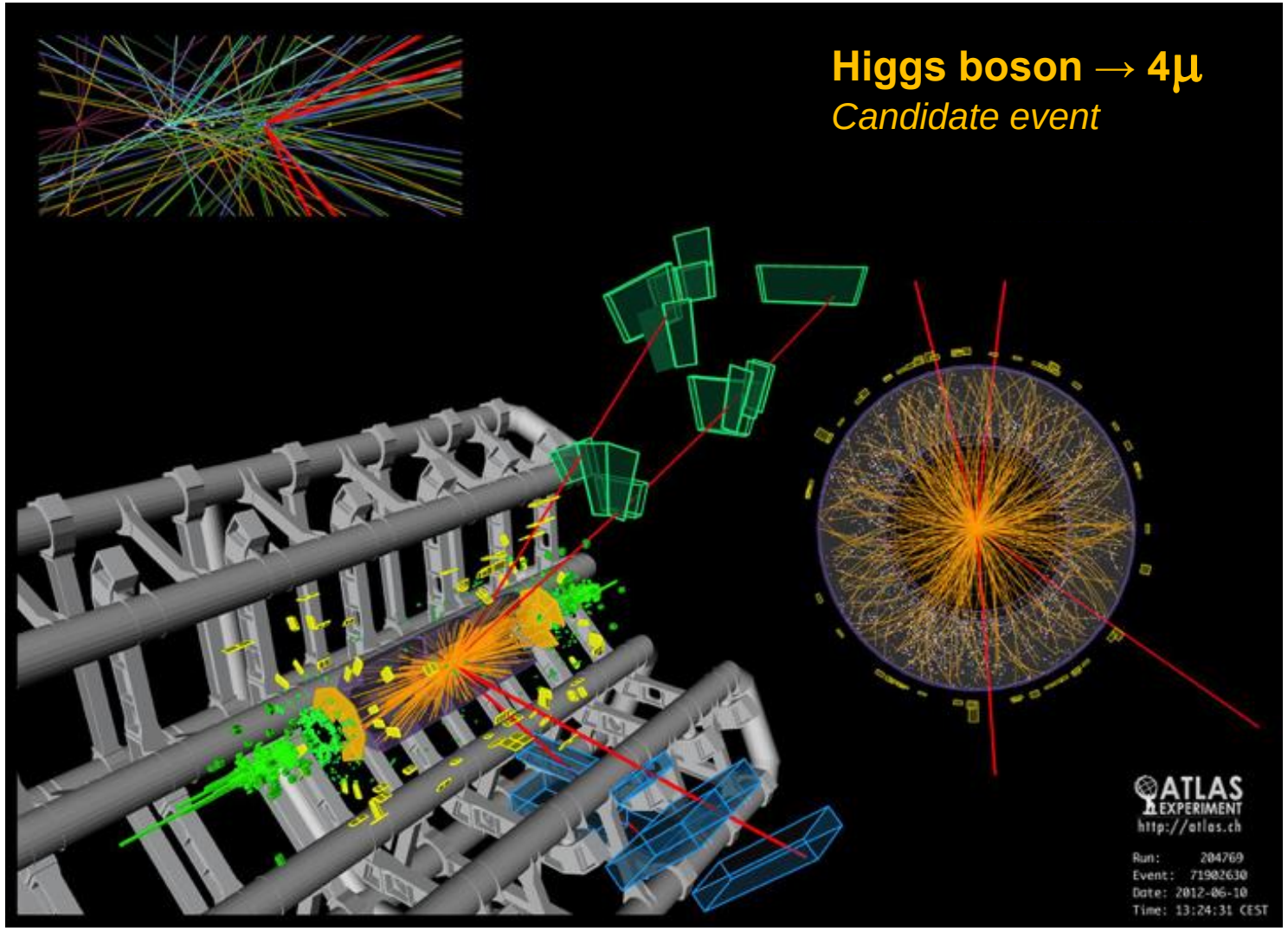
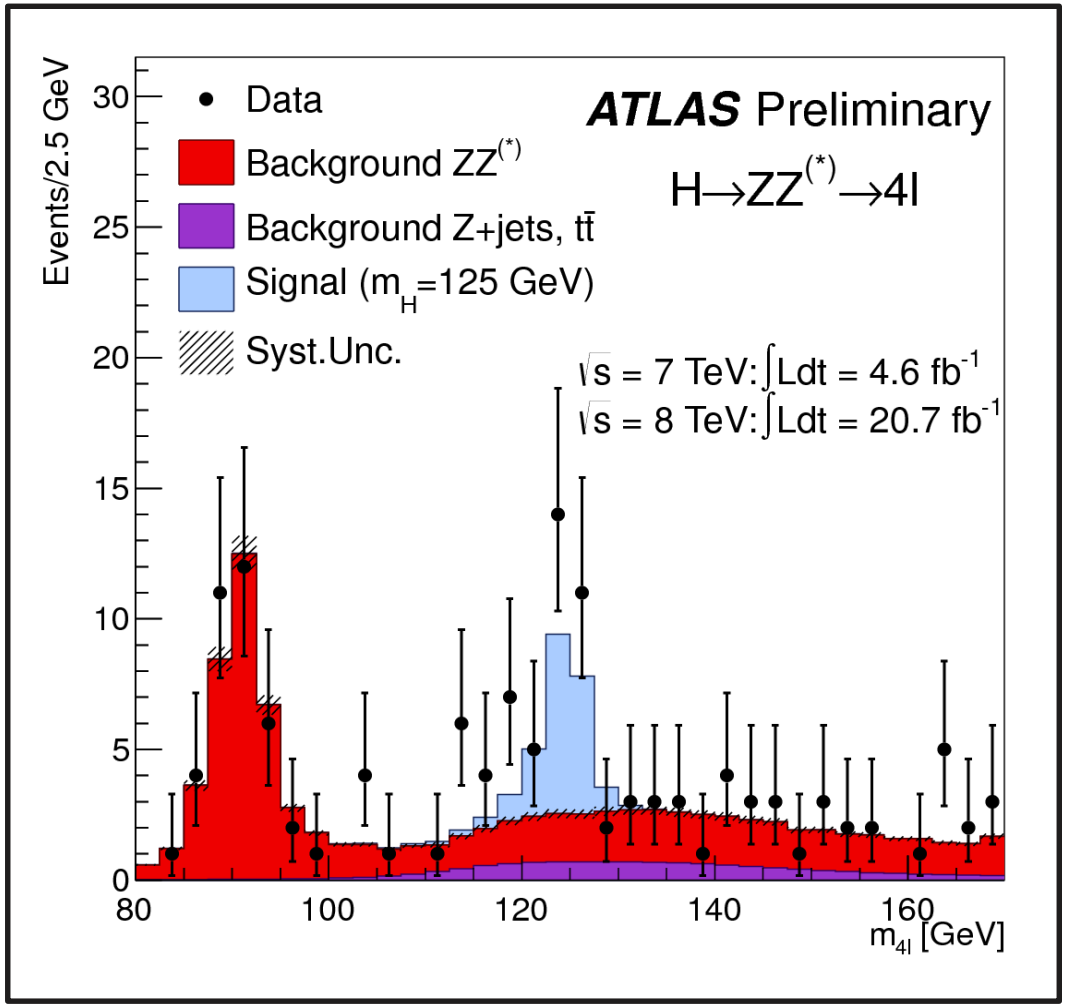
CERN - JPMorgan

Massimo LAMANNA
CERN dept IT

JPMorgan 7-MAY-2018

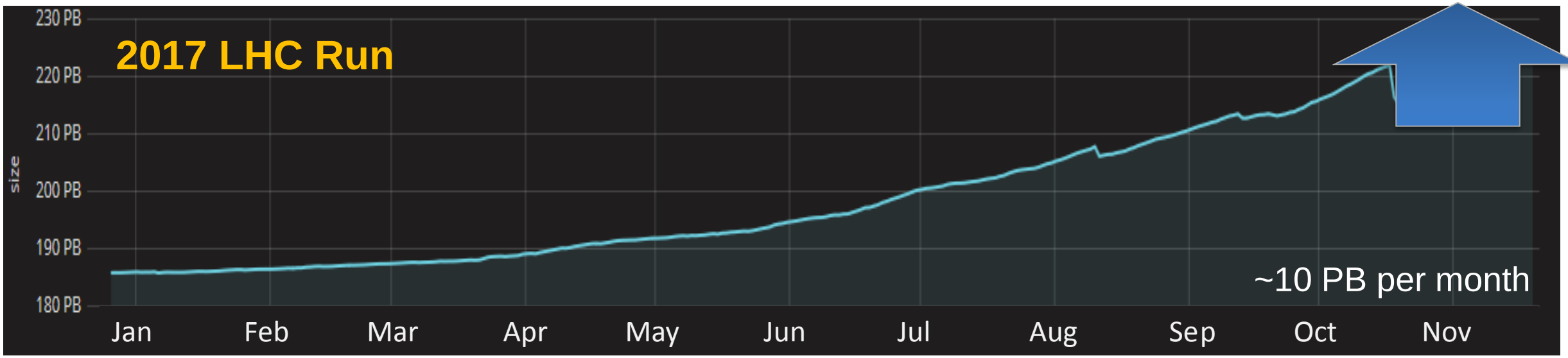
High-Energy Physics **is** a Data Science *par excellence*

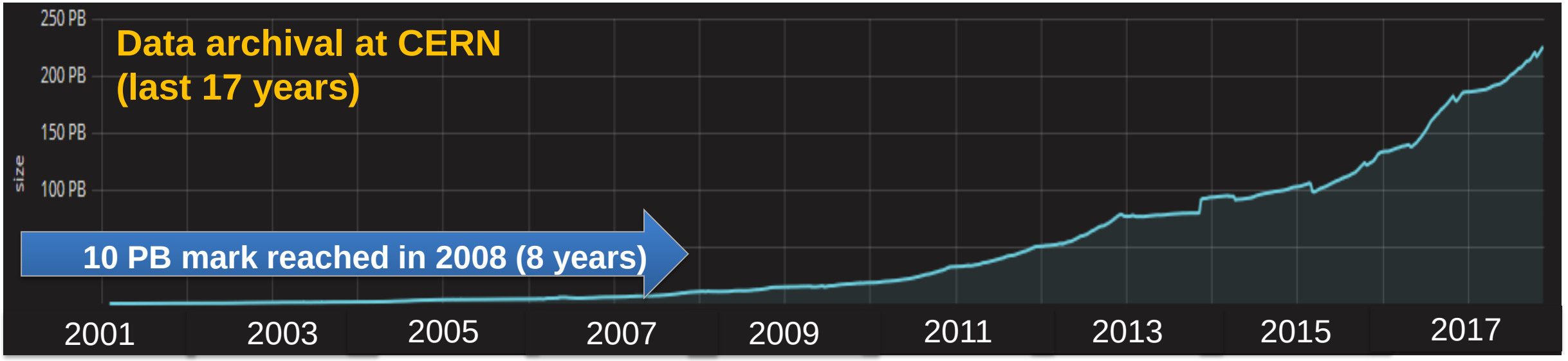




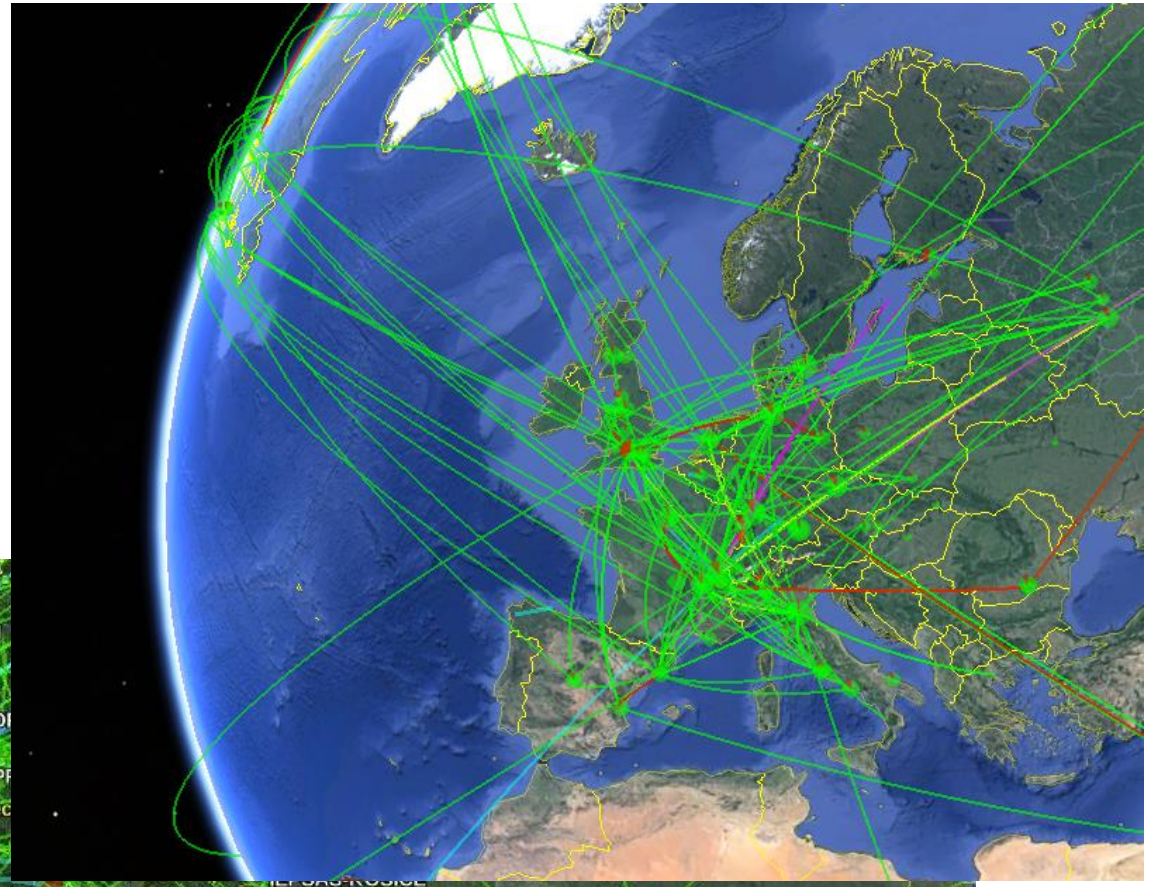
ATLAS data from 2011 and 2012
This plot correspond to o(10) PB. The analysis require CERN + ~200 site (WLCG)

Big-data as seen from the CERN data archive

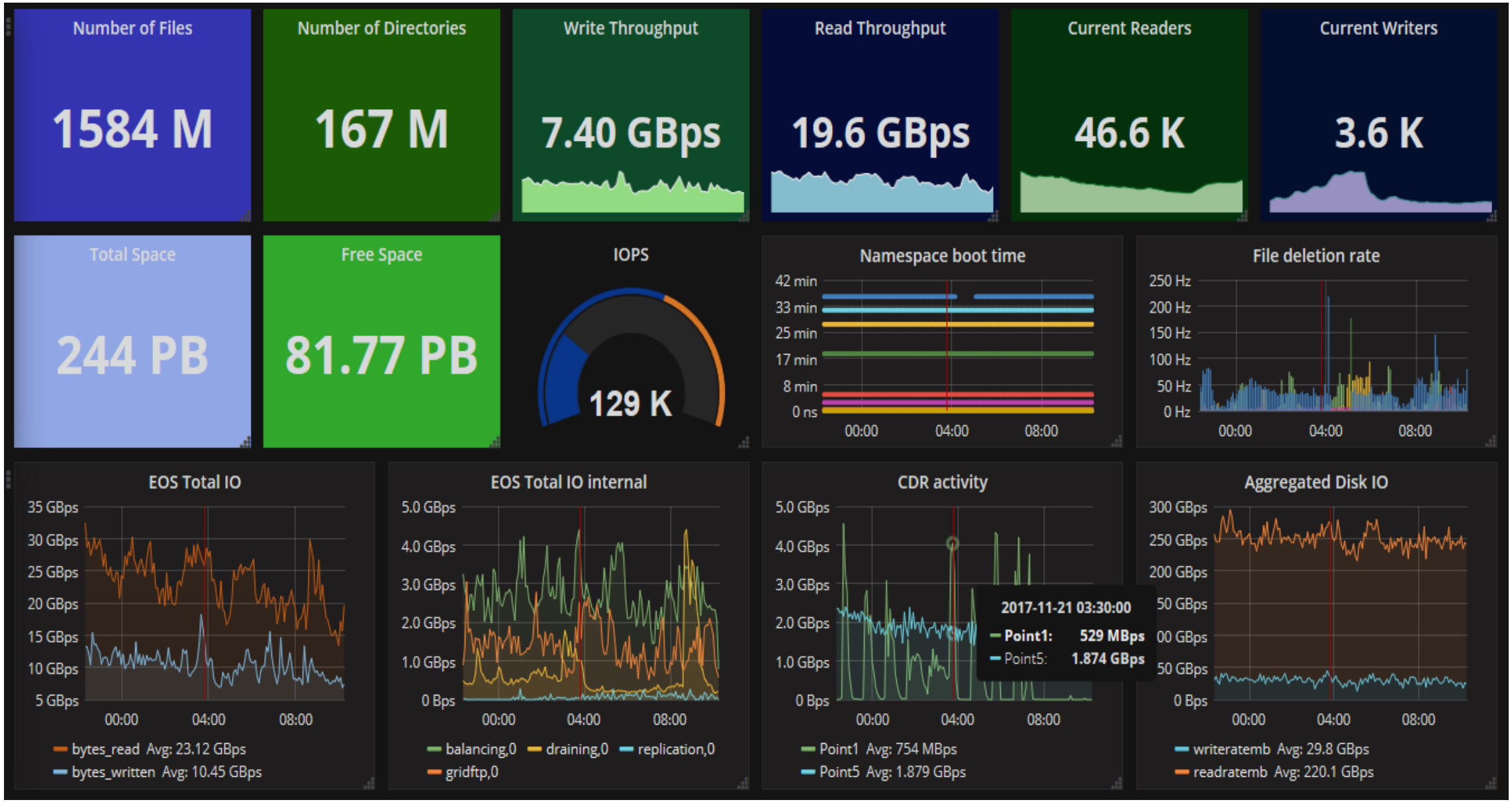




Worldwide LHC Computing Grid



CERN disk farm



A multiscience data infrastructure based on EOS

cloudstor

EOS

SOLUTIONS WE HAVE TRIED

- Hadoop
  - MapR, Hortonworks, Apache official
- XtreemFS
- Ceph
- GlusterFS
- pNFS
- OrangeFS

... and others

SUCCESSES WE'VE HAD

- IT WORKS!
  - Stable, server issues have been almost exclusively container related
  - Fast
- Obvious write latency penalty
  - Users don't notice
- Hello all, I know it's Monday...
  - CERN have been very responsive, THANKYOU!

JRC-CERN collaboration

Joint Research Centre / EU commission

Earth Observation in the Big Data Era

JEODPP
JRC Earth Observation Data and Processing Platform

- The EU ***Copernicus*** Programme with the ***Sentinel*** satellites acts as a game changer by bringing EO into the Big Data era:
  - *expected 10TB/day of **free and open** data*
  - *various spatial, spectral, and temporal resolutions*
- Set-up of a collaborative platform at JRC for storing, analyzing, and sharing Earth Observation data

Sentinel-1 (Credits: ESA/P. Carril)

Sentinel-2 (Credits: ESA/P. Carril)

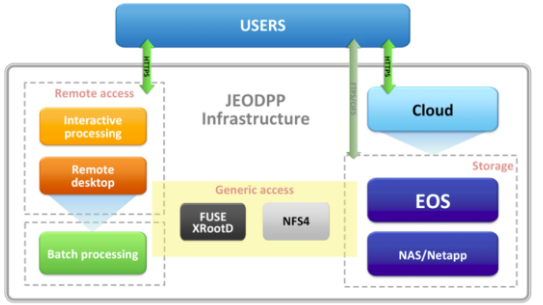
Sentinel-3 (Credits: ESA/J. Huart)

Copernicus

Joint Research Centre

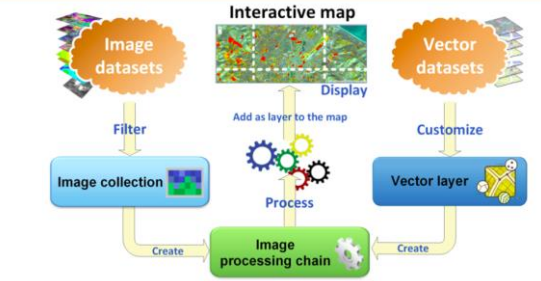
European Commission

Connecting storage and processing via cloud sharing services



From data to interactive display

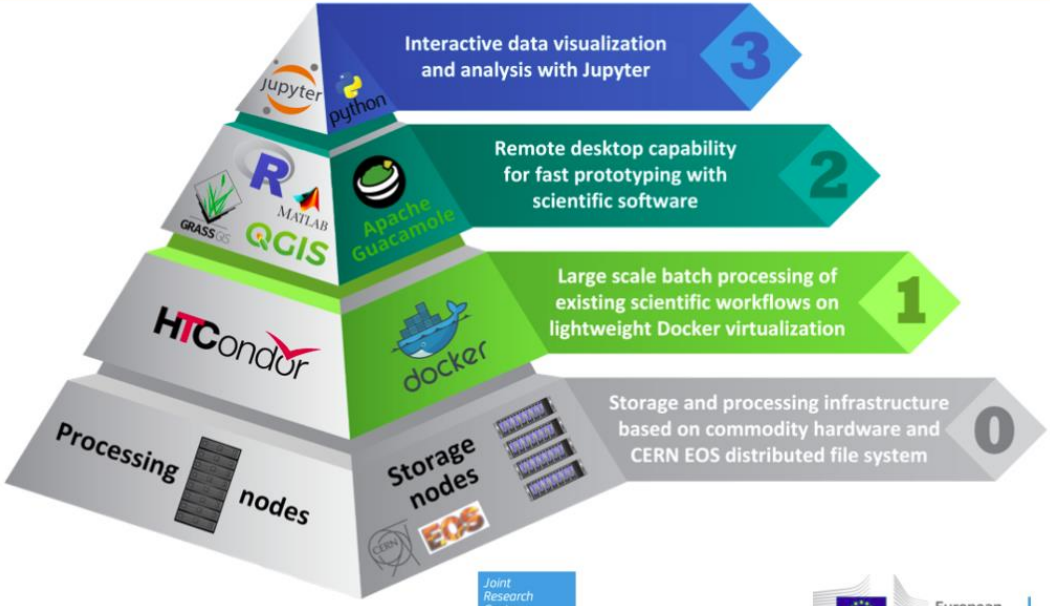
JEODPP



Source: Soille et al., FGCS, 2017, DOI: 10.1016/j.future.2017.11.007

JRC Earth Observation Data and Processing Platform

JEODPP
JRC Earth Observation Data and Processing Platform

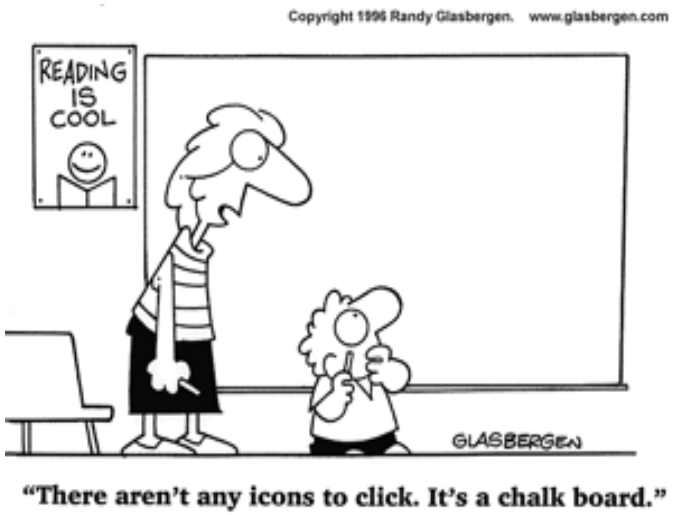


Joint Research Centre

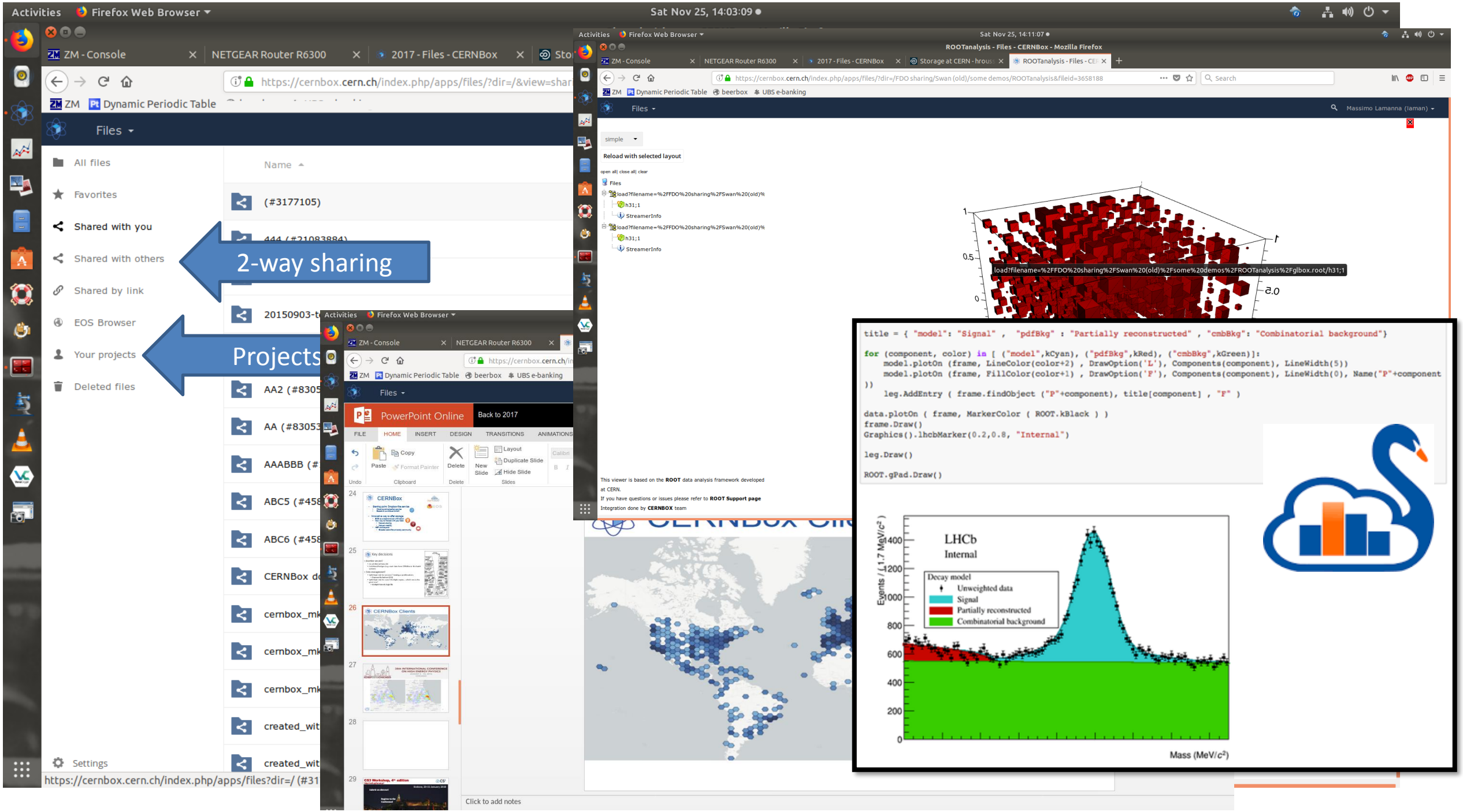
European Commission

CERNBox

- Starting point: Dropbox-like service
  - Users (implicit) request
  - Based on ownCloud GmbH



- Innovative way to offer storage
  - EOS as a back-end (all LHC data!)
  - New way to interact with your data
    - Natural sharing
    - Natural mobility
    - Data analysis
    - Web based tools (e.g. prepare this PowerPoint)


Activities
Firefox Web Browser
Sat Nov 25, 14:03:09
ZM - Console
NETGEAR Router R6300
2017 - Files - CERNBox
https://cernbox.cern.ch/index.php/apps/files/?dir=/&view=shar
ZM
Dynamic Periodic Table
Files
All files
Favorites
Shared with you
Shared with others
Shared by link
EOS Browser
Your projects
Deleted files
Settings
Name
(#3177105)
2-way sharing
Projects
AA2 (#8305
AA (#83053
AAABBB (#
ABC5 (#458
ABC6 (#458
CERNBox d
cernbox_mk
cernbox_mk
cernbox_mk
created_wit
created_wit
https://cernbox.cern.ch/index.php/apps/files?dir=/ (#31
PowerPoint Online
Back to 2017
FILE
HOME
INSERT
DESIGN
TRANSITIONS
ANIMATIONS
Click to add notes
Sat Nov 25, 14:11:07
ROOTanalysis - Files - CERNBox - Mozilla Firefox
Storage at CERN - hrouss
ROOTanalysis - Files - CER
https://cernbox.cern.ch/index.php/apps/files?dir=/FDO sharing/Swan (old)/some demos/ROOTanalysis&fileid=3658188
Search
beerbox
UBS e-banking
Massimo Lamanna (lamanna)
simple
Reload with selected layout
open all| close all| clear
load?filename=%2FFDO%20sharing%2FSwan%20(old)%
h31;1
StreamerInfo
This viewer is based on the ROOT data analysis framework developed at CERN.
If you have questions or issues please refer to ROOT Support page
Integration done by CERNBOX team
load?filename=%2FFDO%20sharing%2FSwan%20(old)%2Fsome%20demos%2FROOTanalysis%2Fglbox.root/h31;1
title = { "model": "Signal" , "pdfBkg" : "Partially reconstructed" , "cmbBkg": "Combinatorial background"}
for (component, color) in [("model",kCyan), ("pdfBkg",kRed), ("cmbBkg",kGreen)]:
model.plotOn (frame, LineColor(color+2) , DrawOption('L'), Components(component), LineWidth(5))
model.plotOn (frame, FillColor(color+1) , DrawOption('F'), Components(component), LineWidth(0), Name("P"+component))
leg.AddEntry (frame.findObject ("P"+component), title[component] , "F")
data.plotOn (frame, MarkerColor (ROOT.kBlack))
frame.Draw()
Graphics().lhcbMarker(0.2,0.8, "Internal")
leg.Draw()
ROOT.gPad.Draw()
LHCb
Internal
Decay model
Unweighted data
Signal
Partially reconstructed
Combinatorial background
Events / (1.7 MeV/c²)
Mass (MeV/c²)

CS3

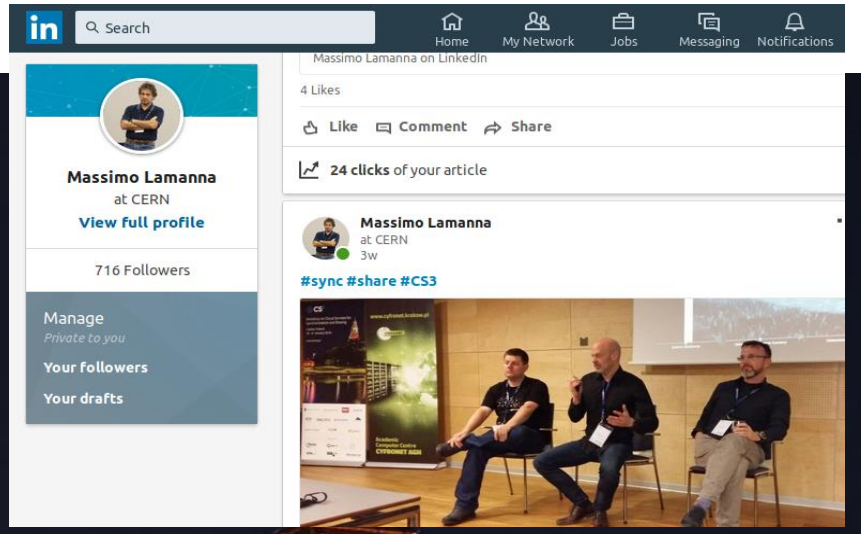
HOME BLOG ABOUT THIS WORKSHOP PROGRAMME CALL FOR ABSTRACTS PREVIOUS WORKSHOPS

Workshop on Cloud Services for Synchronisation and Sharing

29 - 31 January 2018

REGISTER PROGRAMME

Collabora Online · cynny space · DDN STORAGE · iRODS · Nextcloud · ONEDATA · ONLYOFFICE · ownCloud · PowerFolder · pydio · Seafile




CS3
Home
Programme
Register
Abstract
Info
Committees
To be announced next week
Cloud Services for Synchronisation and Sharing
28 - 30 January 2019, Roma
Previous Workshops
Krakow 2018 - Amsterdam 2017 - Zurich 2016 - Geneva 2014
CERN
INFN
Istituto Nazionale di Fisica Nucleare