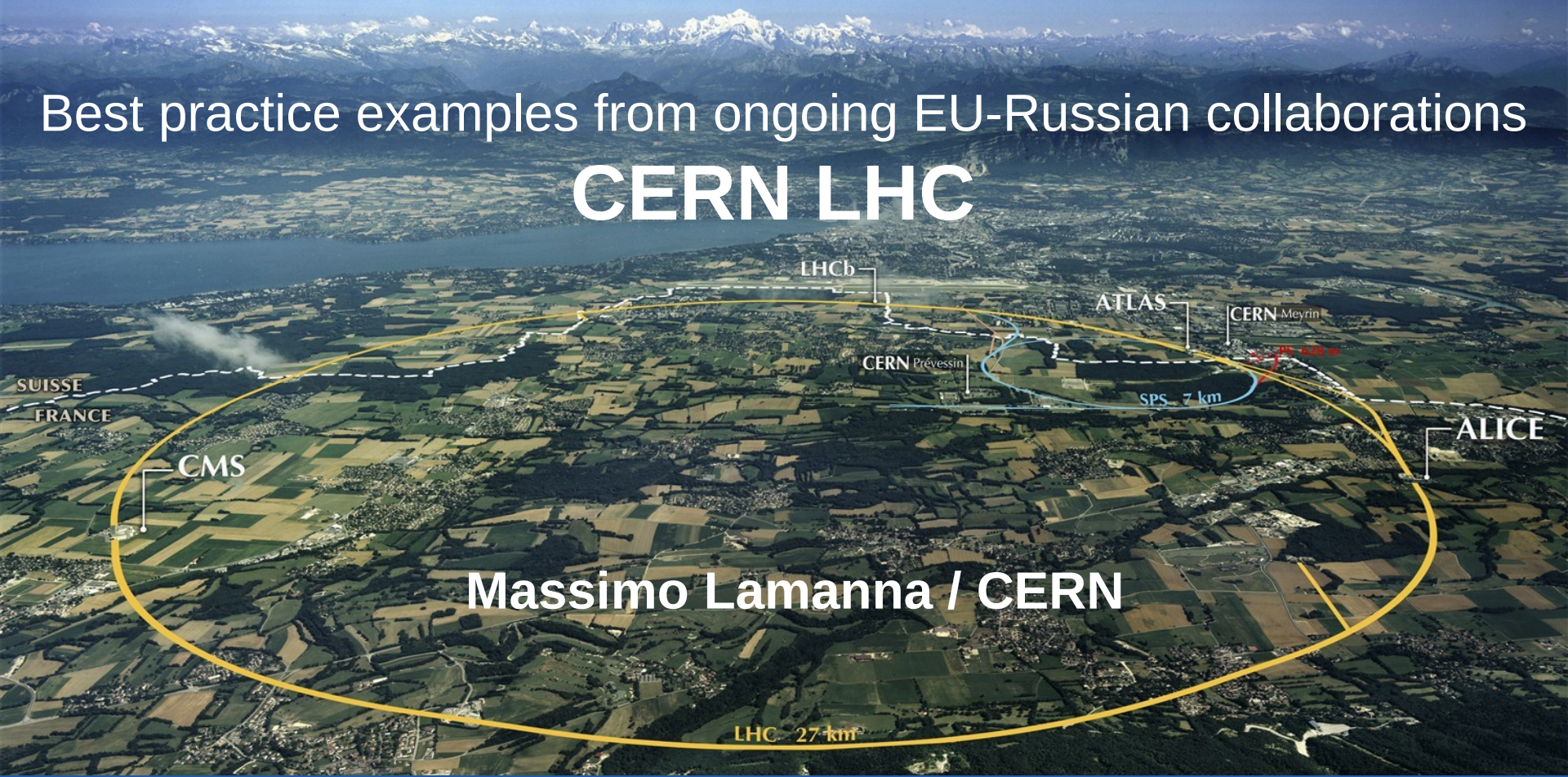

CERN

Best practice examples from ongoing EU-Russian collaborations

CERN LHC



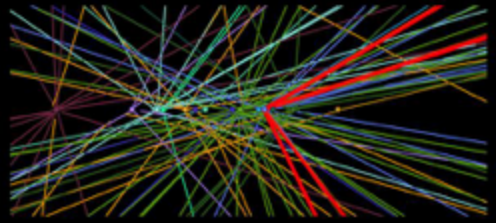
Massimo Lamanna / CERN

Table of Content

- **Introduction**
  - Big Data: LHC computing
- **Collaborations**
  - Established ones
    - **EOS**
  - Starting…
    - **CERNBox**
  - … and future perspectives
    - **SWAN**

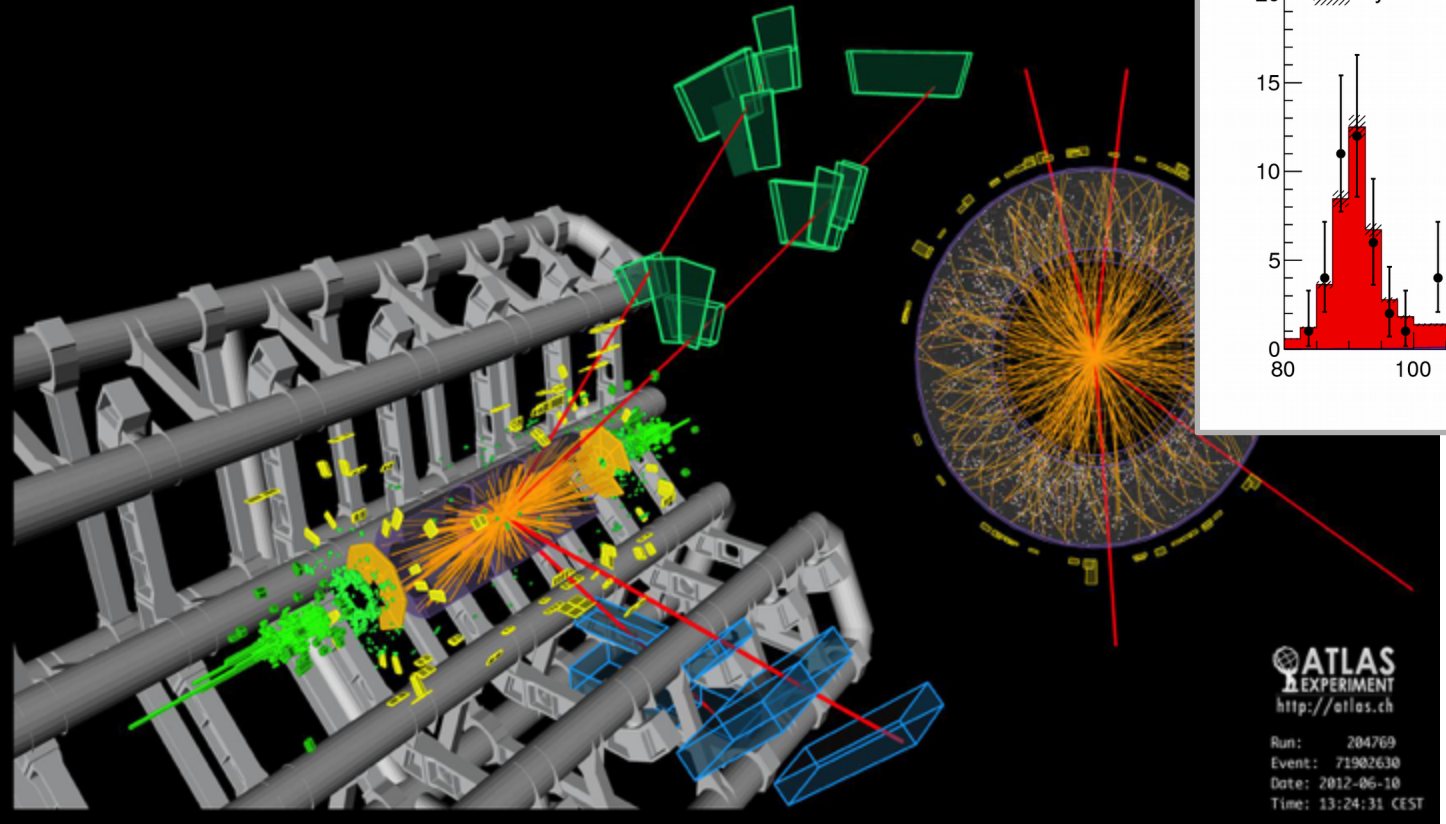
Focus on Big Data only

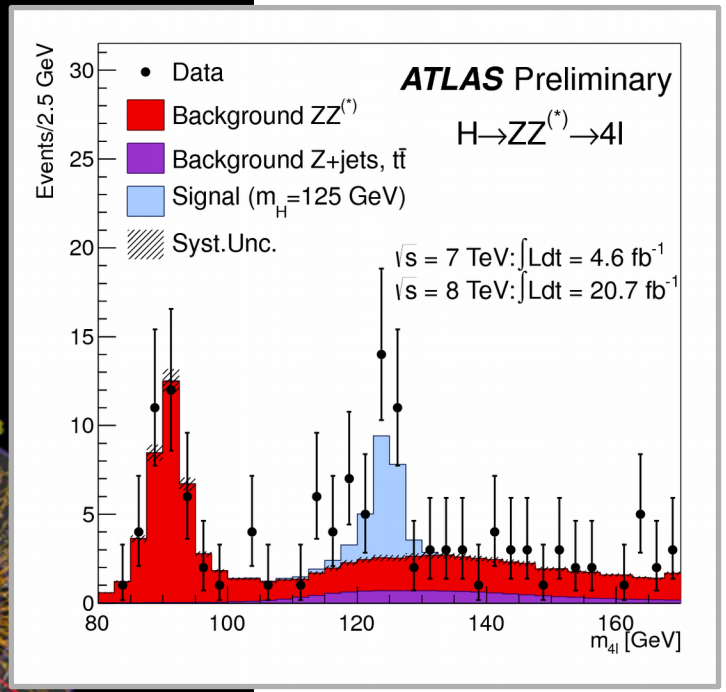
Credits: slides from the EOS, CERNBox and SWAN teams



Higgs boson → 4μ

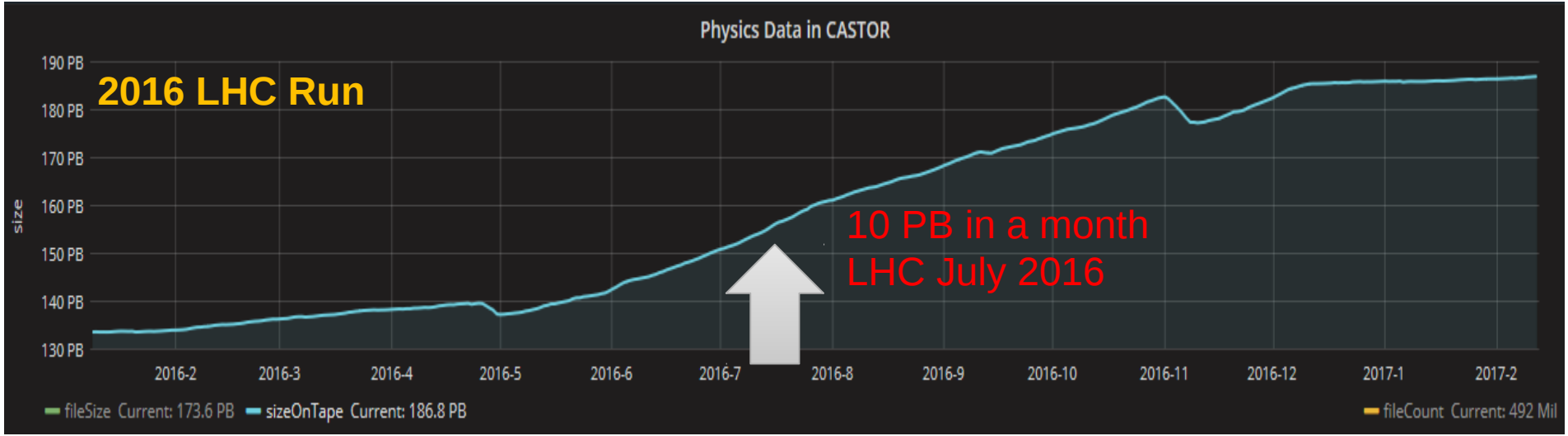
Candidate event

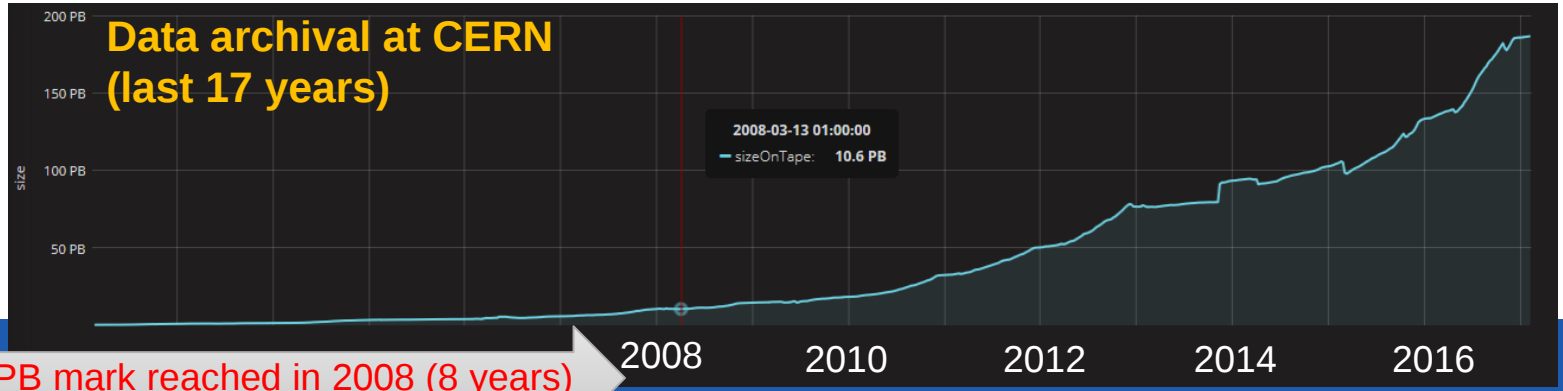




ATLAS data in Run1 2011 and 2012

The plot correspond to o(10) PB

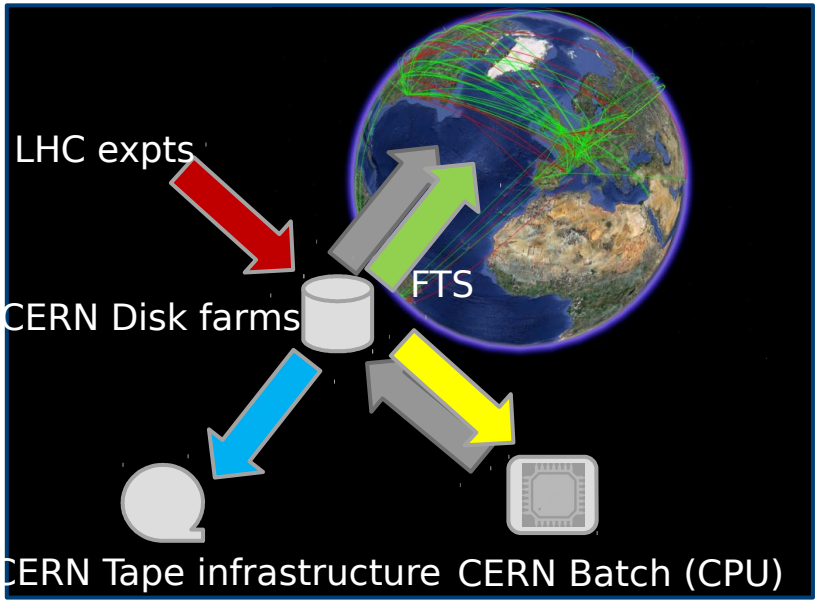




10 PB mark reached in 2008 (8 years)

WLCG

Worldwide LHC Computing Grid



Make LHC computing possible

Worldwide infrastructure (collaboration) open to all LHC physicists

Computing/storage resources at CERN: ~ 20%; 80% across about 200 sites worldwide

Data Reconstruction

Goals: data quality and immediate access for analysis

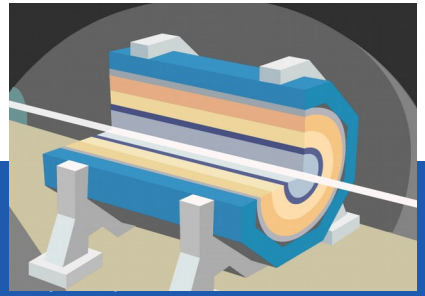
Organised activity dominated by heavy processing and replication (each expt: 1-8 GByte/s)

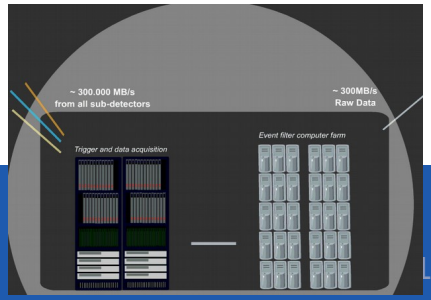
Data Analysis

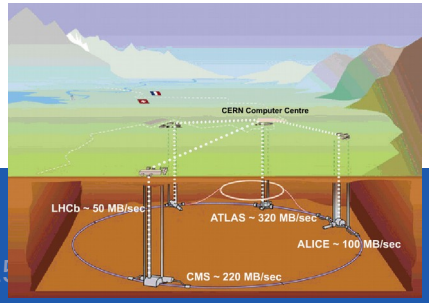
Goals: extract physics quantities (discovery)

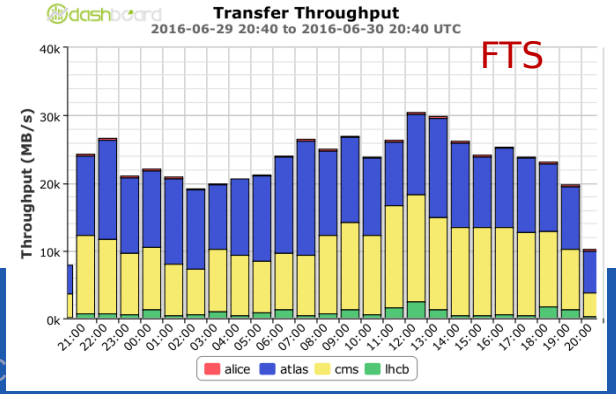
Individual activities dominated by event selection and sharing (thousands of physicists)

(Detector) simulation







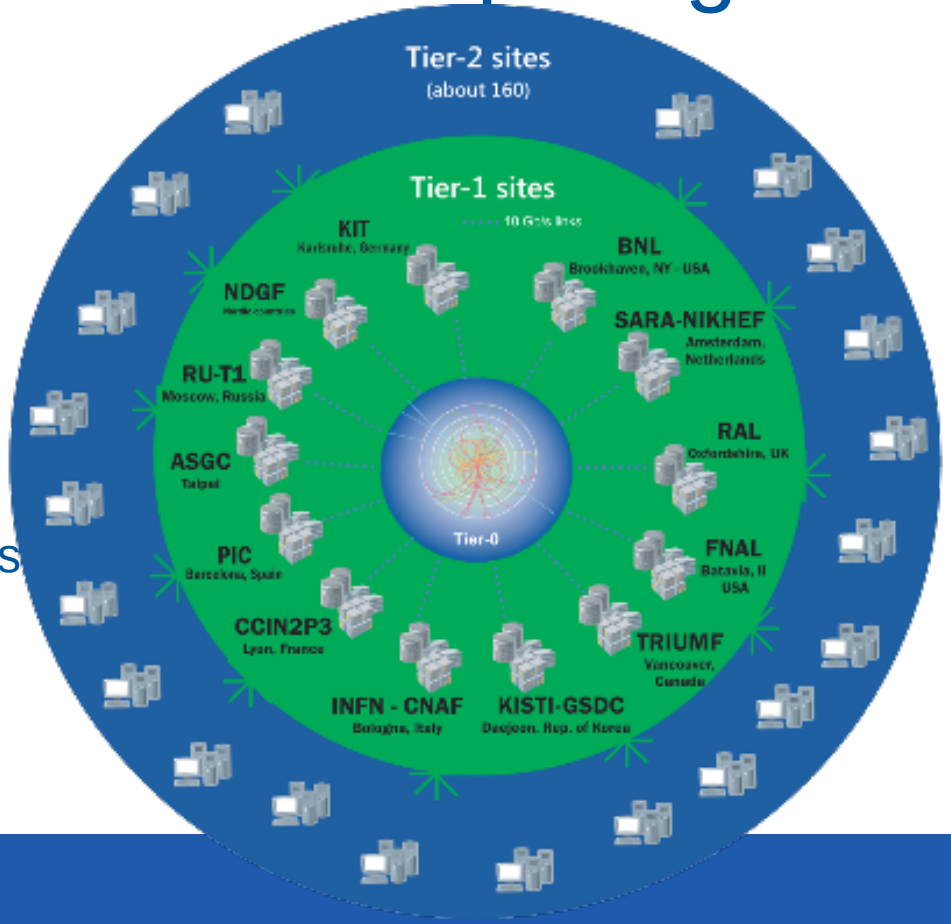


Worldwide LHC Computing Grid

Tier-0:
data recording, reconstruction and distribution

Tier-1: permanent storage, re-processing, analysis

Tier-2: Simulation, user analysis



~170 sites, 40 countries

~500k CPU cores

500 PB of storage

2+ million jobs/day

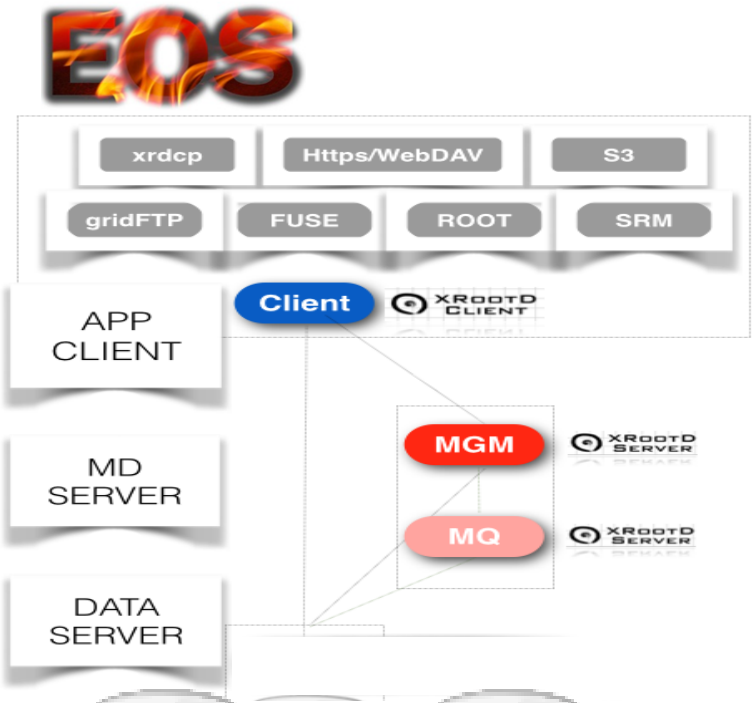
Multiple 10-100 Gb links

LCG:
Initial description: 2001
Tech. Design Report: 2005

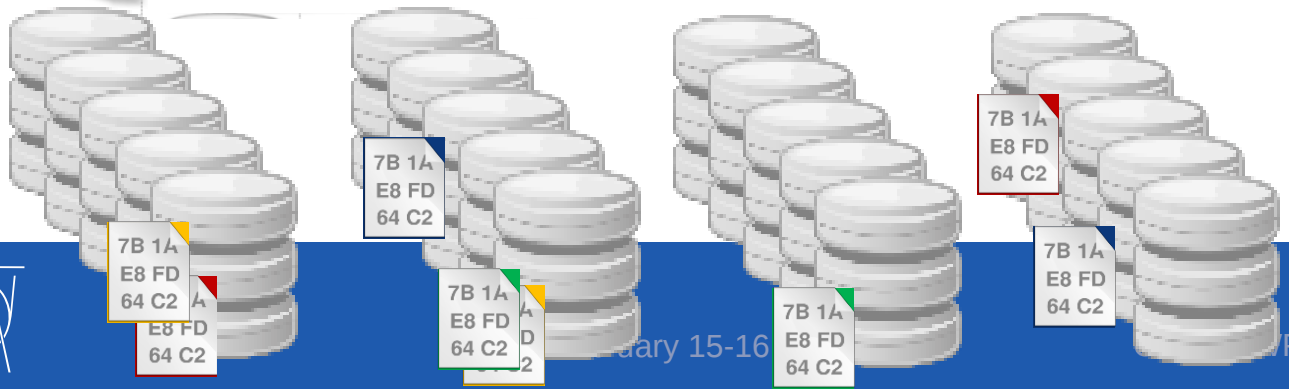
Table of Content

- **Introduction**
  - Big data: LHC computing
- **Collaborations**
  - Established ones
    - **EOS**
  - Starting…
    - **CERNBox**
  - … and future perspectives
    - **SWAN**

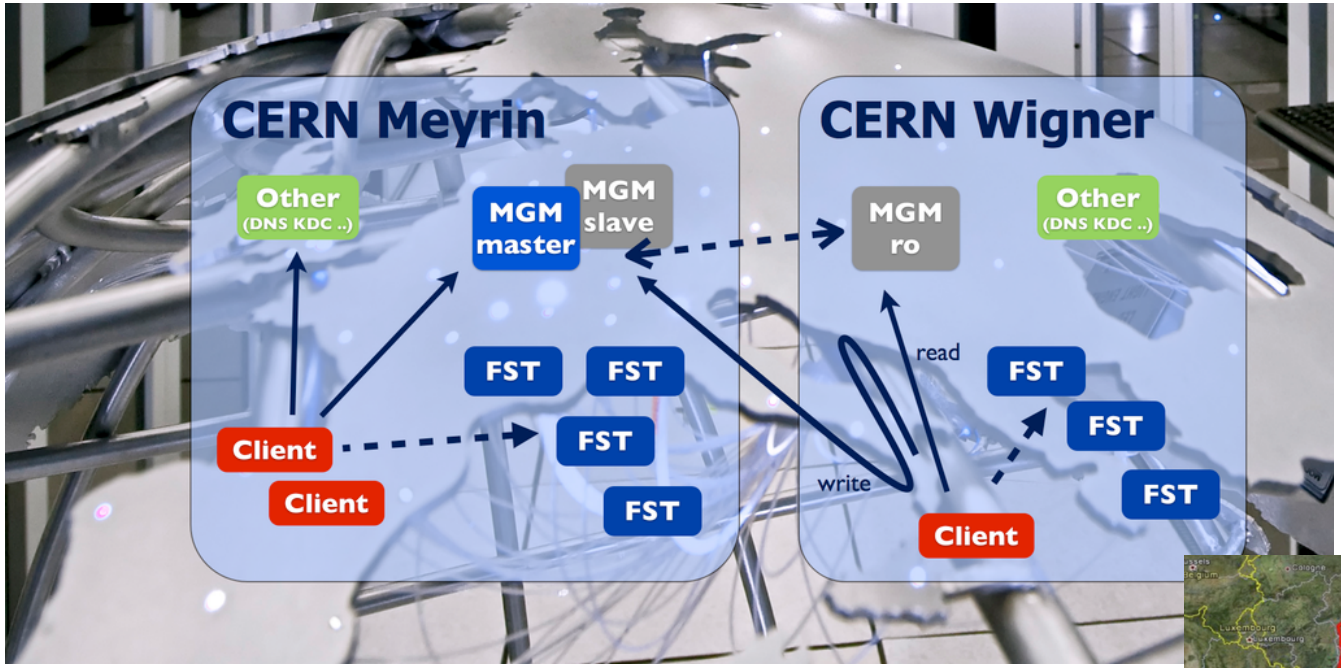

EOS
CERN
Open Source Storage



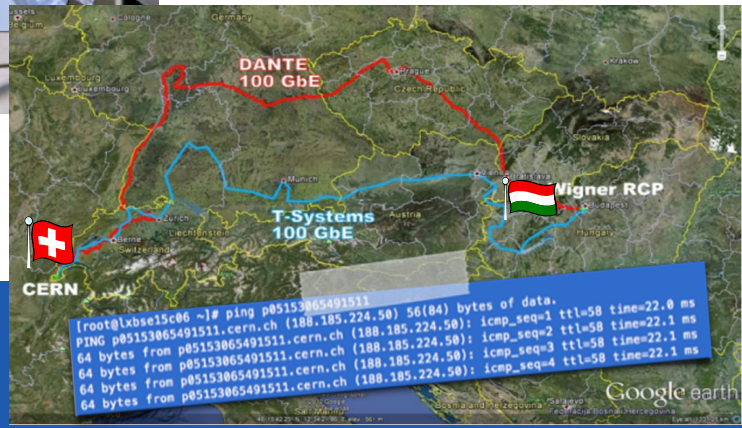
- **EOS: Large disk farms for physics and beyond**
  - Developed at CERN
  - LHC: PBs for 100s/1000s independent scientists
  - 200 PB JBOD installed (CERN installations)
- **Strategic points**
  - Distill **20+ years** of experience data management
  - Ultra-fast name space
  - Arbitrary level of data durability: cross-node file replication or RAIN on **commodity** hardware
  - Large **protocol choice**



Our “20-ms-large” computer centre

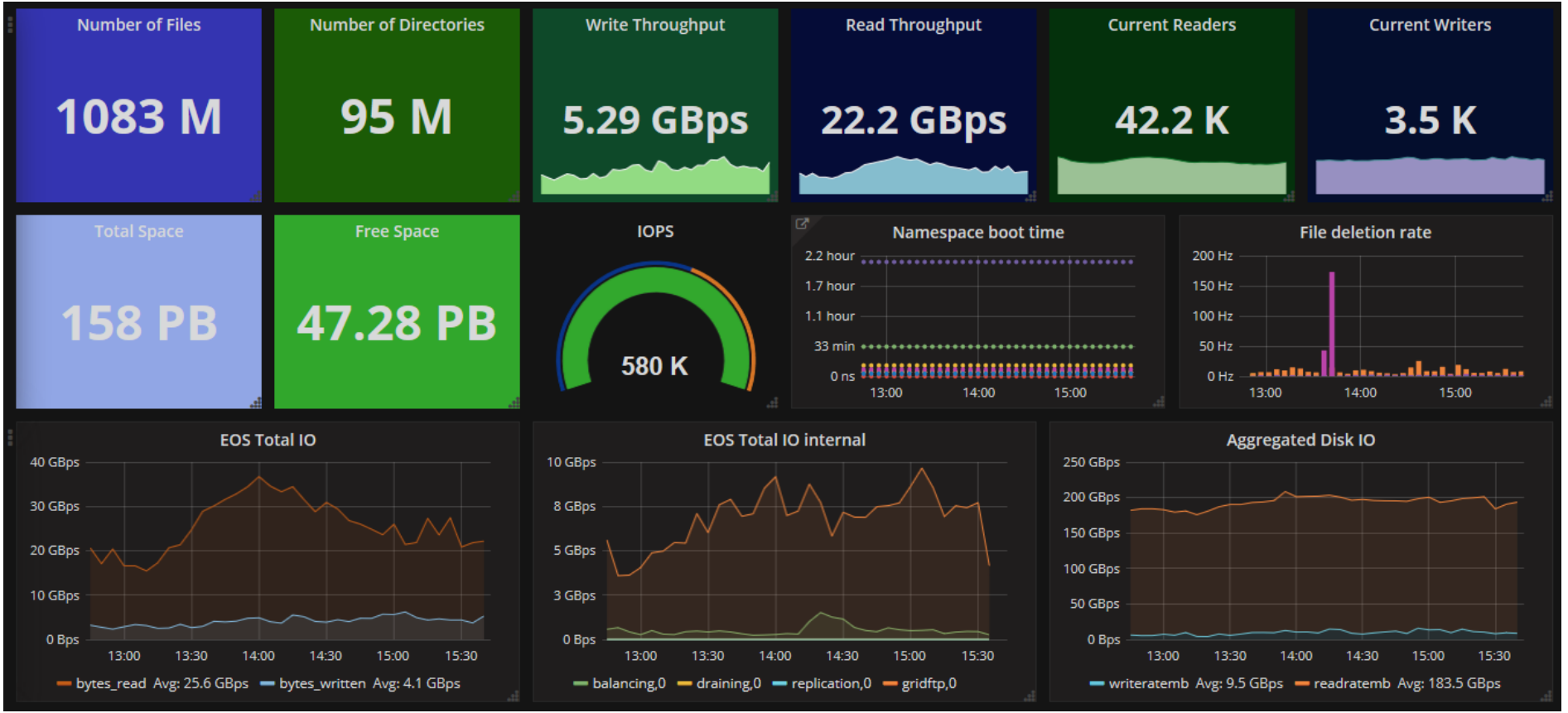


Autonomic,
Locality,
Disaster recovery/
business continuity



MGM = NameSpace/Metadata

FST = Disk servers


Number of Files
1083 M
Number of Directories
95 M
Write Throughput
5.29 GBps
Read Throughput
22.2 GBps
Current Readers
42.2 K
Current Writers
3.5 K
Total Space
158 PB
Free Space
47.28 PB
IOPS
580 K
Namespace boot time
2.2 hour
1.7 hour
1.1 hour
33 min
0 ns
13:00
14:00
15:00
File deletion rate
200 Hz
150 Hz
100 Hz
50 Hz
0 Hz
13:00
14:00
15:00
EOS Total IO
40 GBps
30 GBps
20 GBps
10 GBps
0 Bps
13:00
13:30
14:00
14:30
15:00
15:30
bytes_read Avg: 25.6 GBps
bytes_written Avg: 4.1 GBps
EOS Total IO internal
10 GBps
8 GBps
5 GBps
3 GBps
0 Bps
13:00
13:30
14:00
14:30
15:00
15:30
balancing,0
draining,0
replication,0
gridftp,0
Aggregated Disk IO
250 GBps
200 GBps
150 GBps
100 GBps
50 GBps
0 Bps
13:00
13:30
14:00
14:30
15:00
15:30
writeratemb Avg: 9.5 GBps
readratemb Avg: 183.5 GBps

1st EOS workshop (February 2-3 2017)

Integration and support

Disk technology

National Research Centre (NRC)
"Kurchatov Institute"

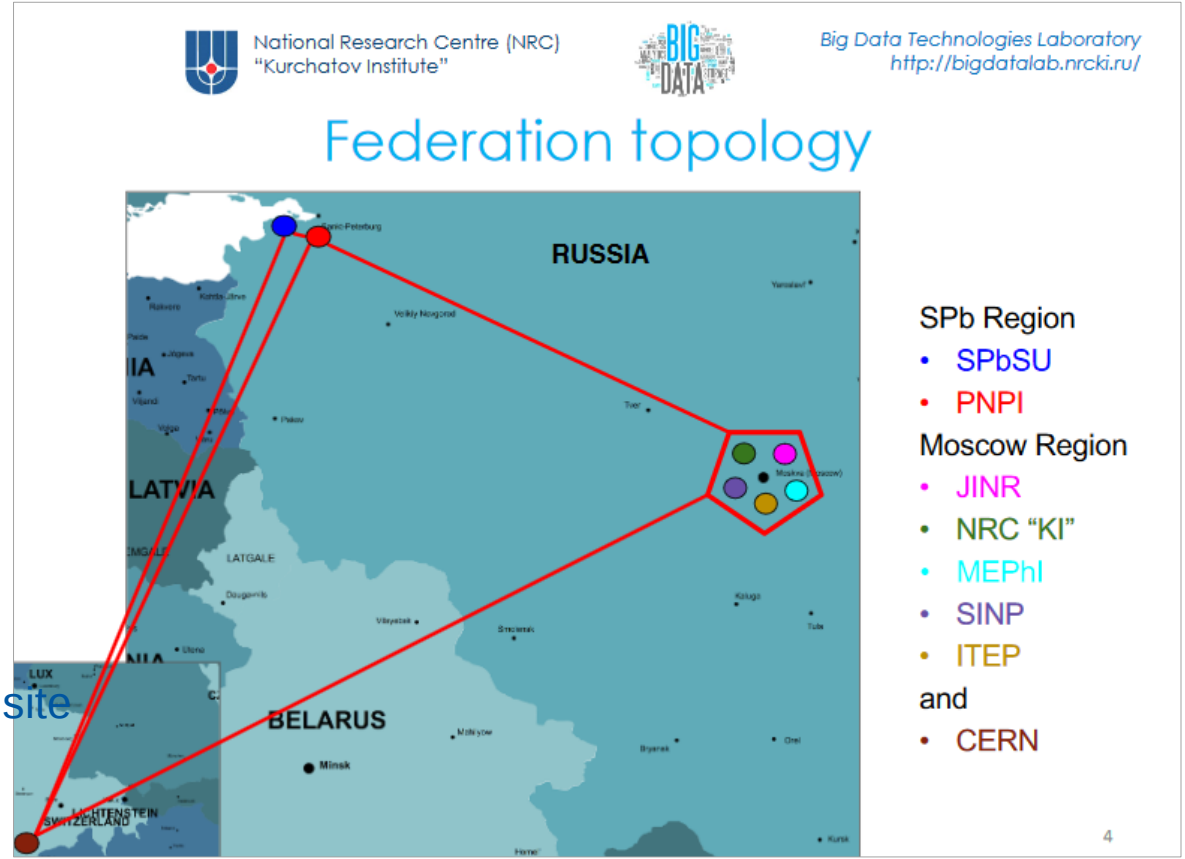
Big Data Technologies Laboratory
http://bigdatalab.nrcki.ru/

Russian Federated Data Storage System Prototype

Andrey Kiryanov, Alexei Klimentov, Andrey Zarochentsev

on behalf of BigData lab @ NRC "KI" and
Russian Federated Data Storage Project

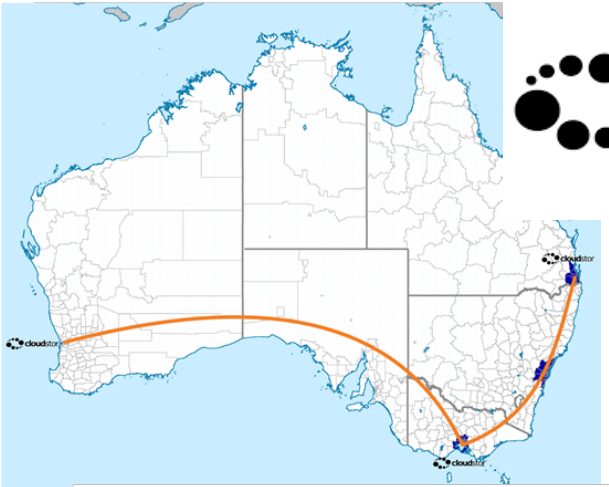
- HEP communities
  - Collaboration
  - Complementarity
- Federation
  - Moscow area
  - St Petersburg area
  - (CERN)
  - Sites from Russian Data Intensive Grid And WLCG site
- EOS workshop
  - A.Kyrianov et al.



EOS AT 6,500 KILOMETRES WIDE

An Australian Experience

David Jericho – Solutions Architect, AARNet



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!

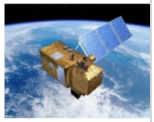
Joint Research Centre (JRC)

Science Service of the European Commission

"Earth Observation & Social Sensing Big Data Pilot Project"

- The EU ***Copernicus*** Programme with the ***Sentinel*** fleet of satellites acts as a game changer by bringing EO in the Big Data era:
  - *expected 10TB/day of **free and open** data*
  - *Requires new approaches for data management and processing*
- Pilot project launched in January 2015
- Major goal: set up a central infrastructure for storing and processing of Earth Observation and Social Sensing data at JRC


Sentinel-1 (Credits: ESA/P. Carr


Sentinel-2 (Credits: ESA/P. Carr

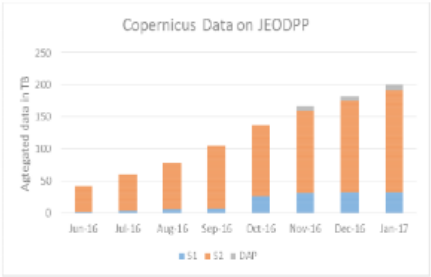

Sentinel-3 (Credits: ESA/J. Huar

Joint Research Centre

European Commission

EOS set-up at JRC

- Installation and configuration at JRC with strong support from CERN storage team
- Current set-up:
  - *1.4 PB gross capacity*
  - *10 FST nodes, each with one JBOD of 24x6 TB disks*
  - *Using replica 2*
- Further extension planned
  - *2017: extend to ~6 PB gross*



Joint Research Centre

European Commission

A. Burger and P.Soille (JRC)

Table of Content

- **Introduction**
  - Big data: LHC computing
- **Collaborations**
  - Established ones
    - **EOS**
  - Starting…
    - **CERNBox**
  - … and future perspectives
    - **SWAN**

CERNBox

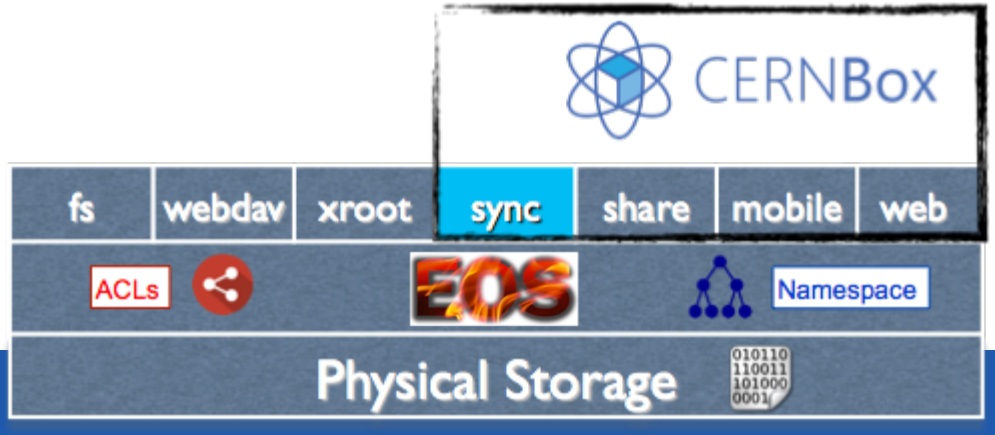
CERNBox

- Starting point: Dropbox-like service
  - Cloud synchronisation service
  - Just the starting point!

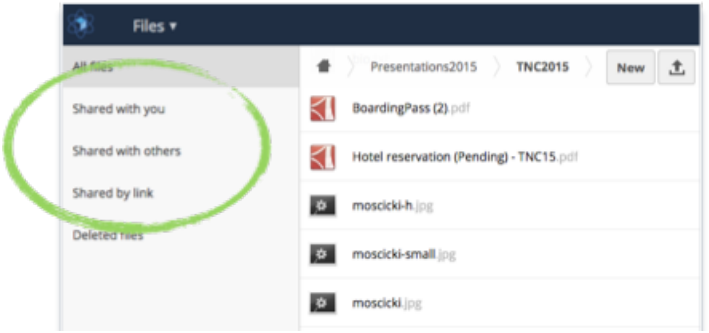
- Innovative way to offer storage
  - Sync and share from ownCloud GmbH
  - EOS as a back-end (all LHC data!)
  - New way to interact with your data

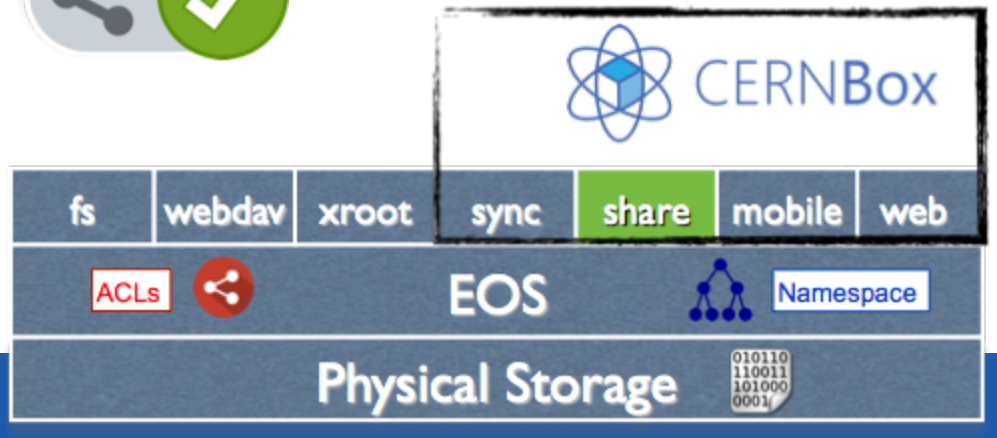
- Strong interest
  - In HEP: from V.Ilin and his group (Moscow)
  - Broader scientific/university community

Access Methods: Sync



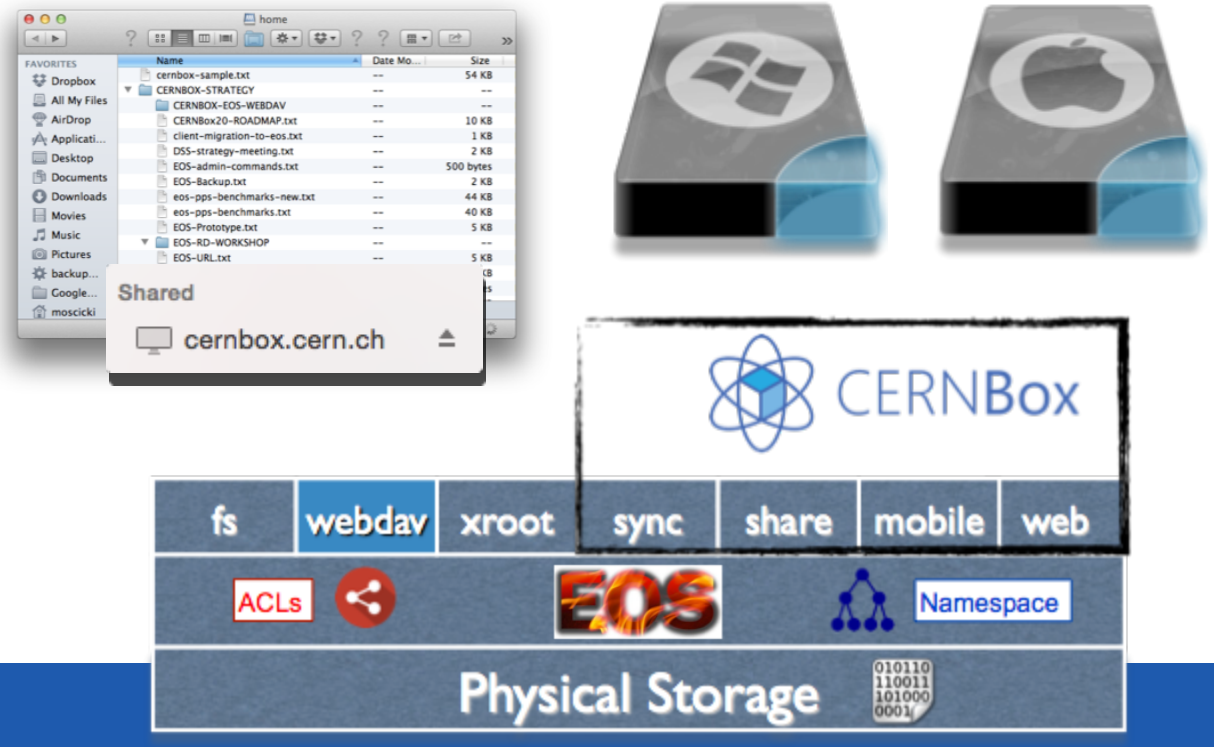
Access Methods: Sharing





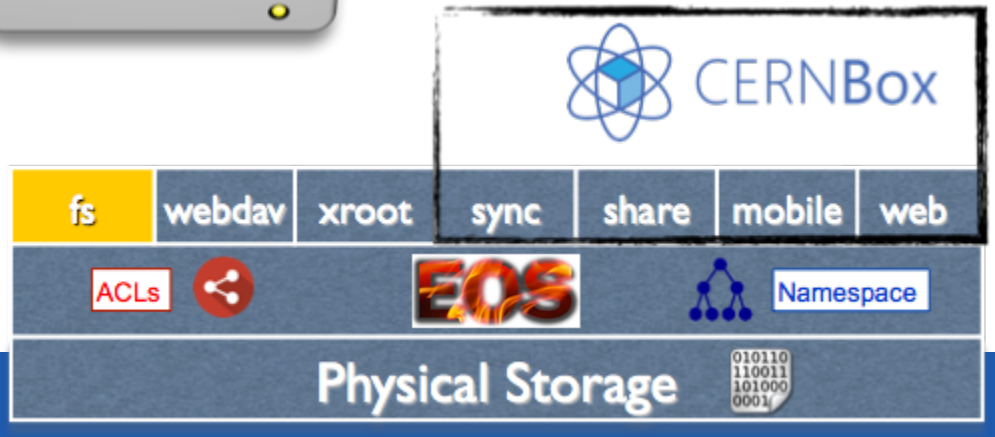
Access Methods: Mobile & Web

Access Methods: WebDAV

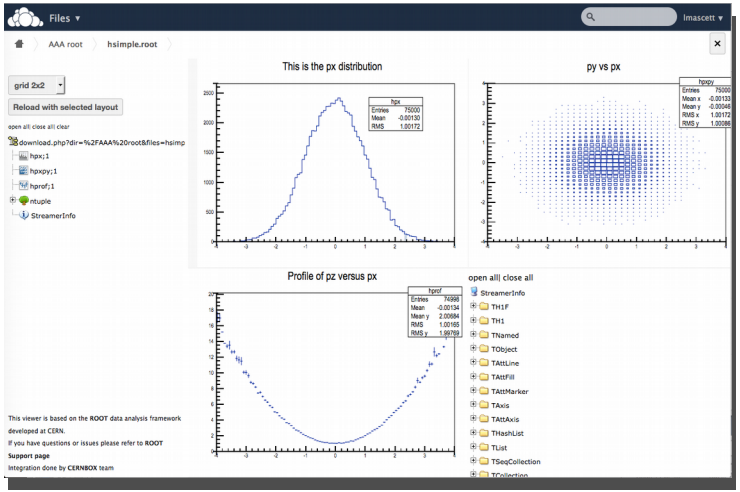


Access Methods: FUSE

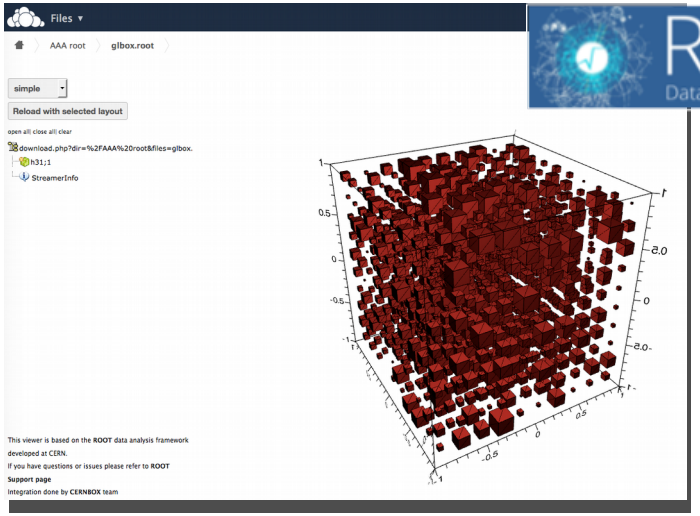

```
[lmascett@lxplus2015 ~]#
[lmascett@lxplus2015 ~]# df -H -t fuse
Filesystem      Size  Used Avail Use% Mounted on
eosuser         506T   70T  437T  14% /eos/user
eosatlas         36P   17P   20P  45% /eos/atlas
eosalice         20P   11P  8.5P  57% /eos/alice
eoscms           28P   14P   15P  49% /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
```

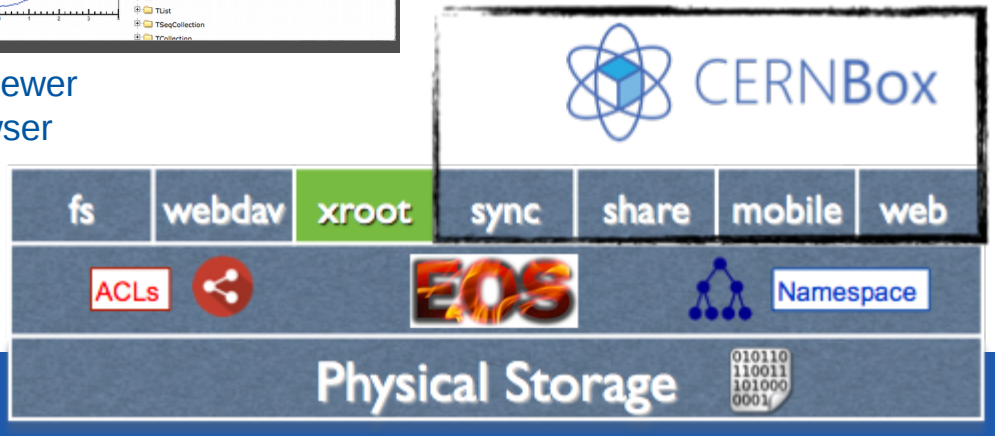


Optimised access



Embedded ROOT viewer
in CERNBox browser





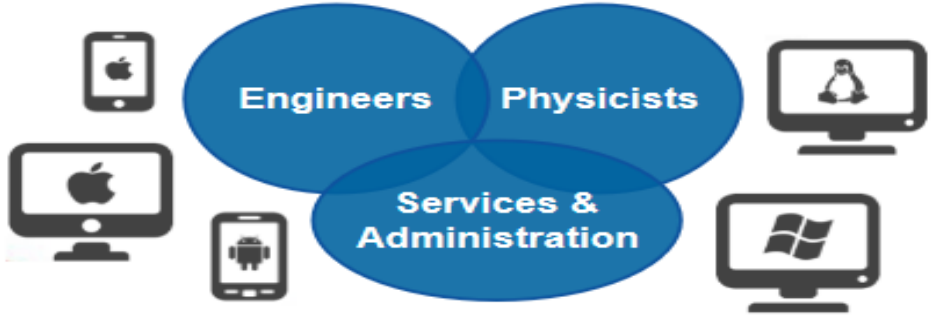
3rd Cloud Services for Synchronisation and Sharing (CS3)

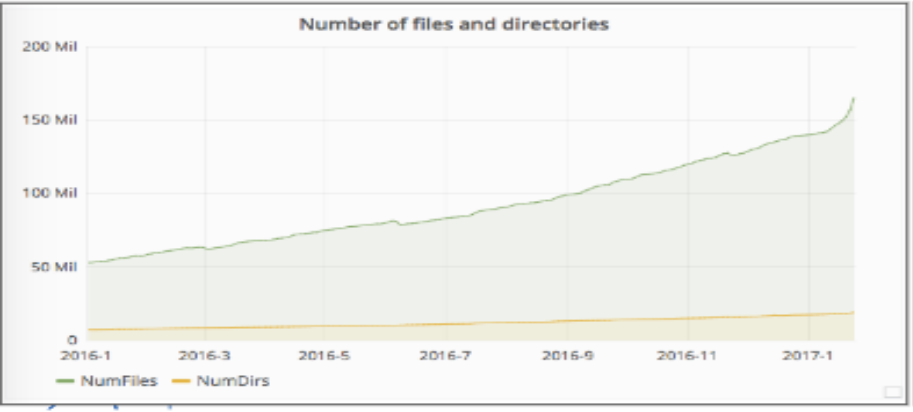
Novel applications, cloud storage technology, collaborations

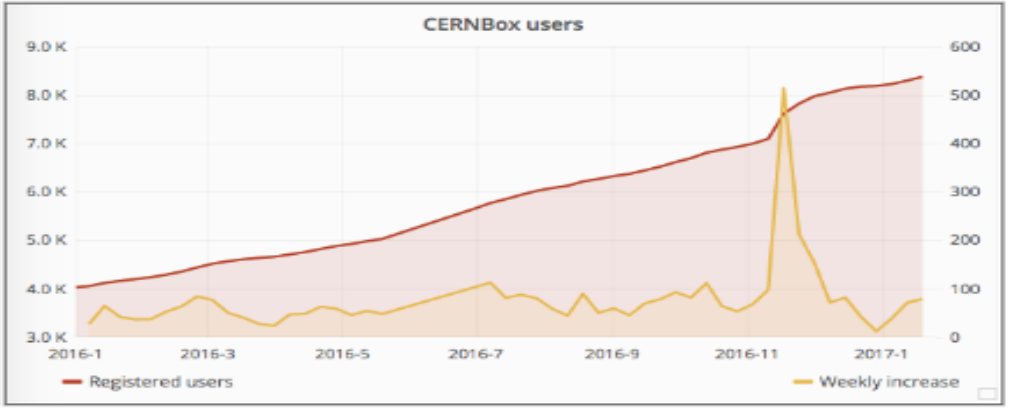
Amsterdam January 2017 *** 120+ participants *** 15+ companies

CERNBox Service Numbers

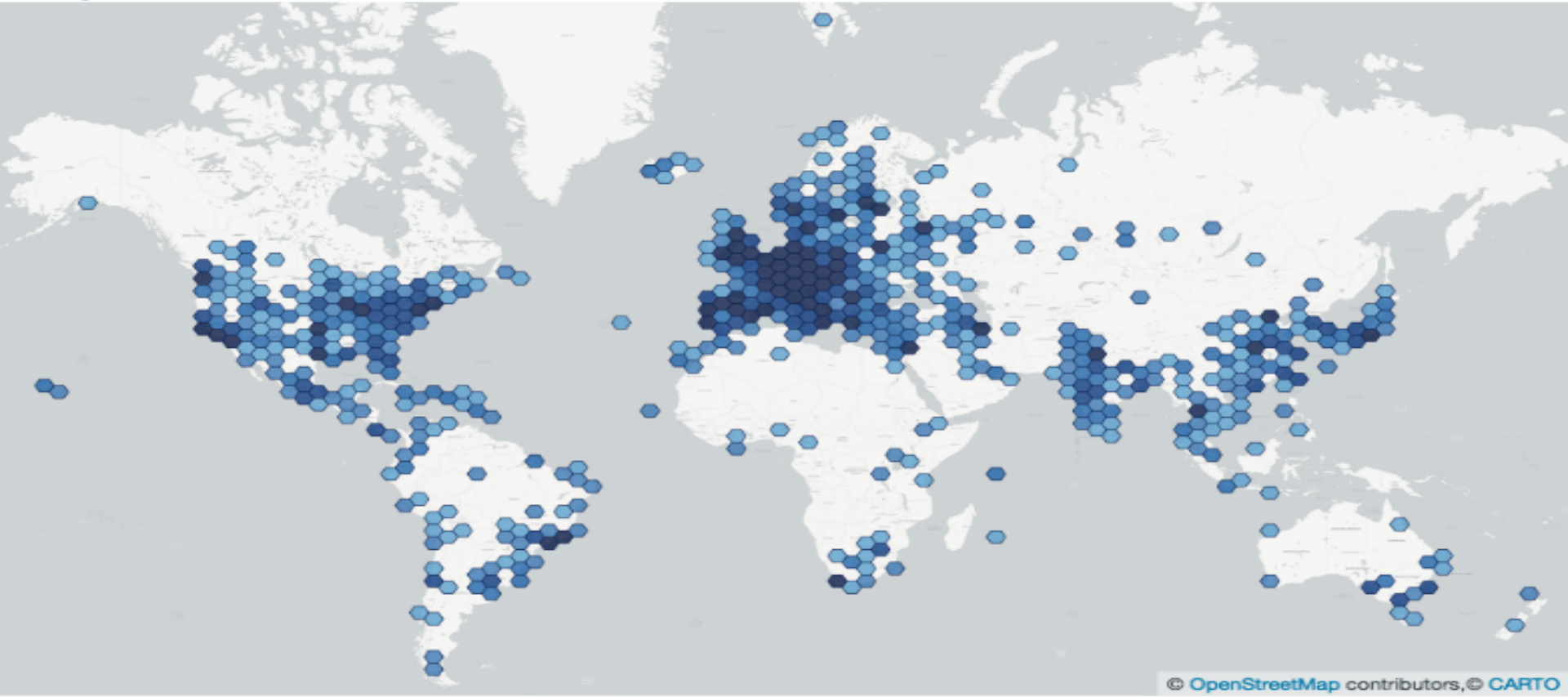
	Jan 2016	Jan 2017
Users	4074	8411
# files	55 Million	176 Million
# dirs	7.2 Million	19 Million
Used Raw Space	208 TB	806 TB
Deployed Raw Space	1.3 PB	3.2 PB

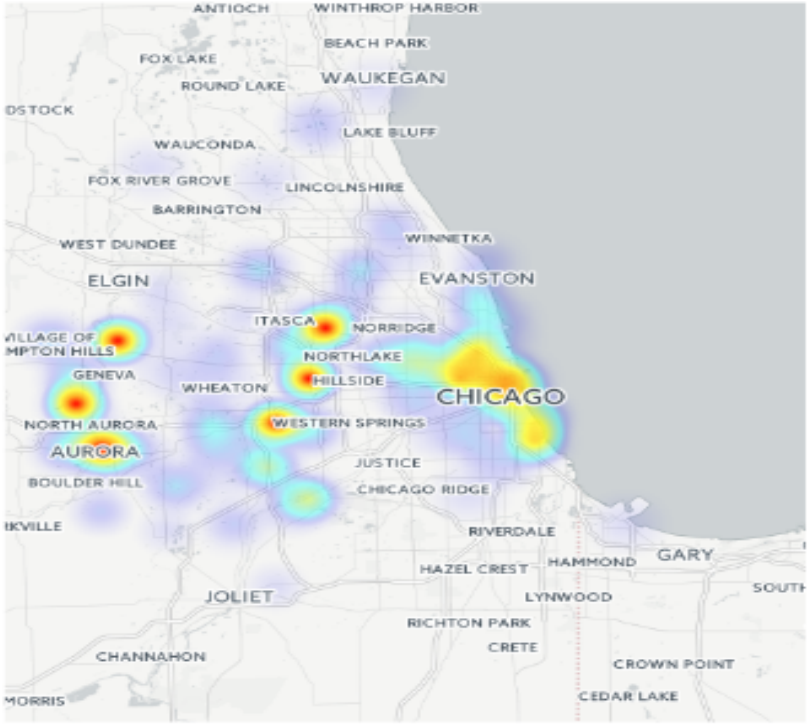






CERNBox Clients




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

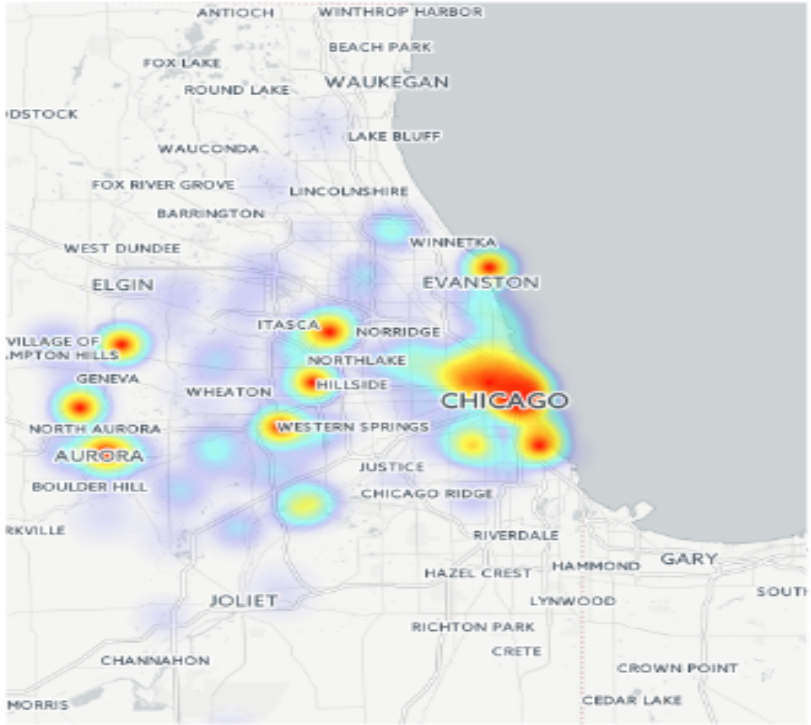
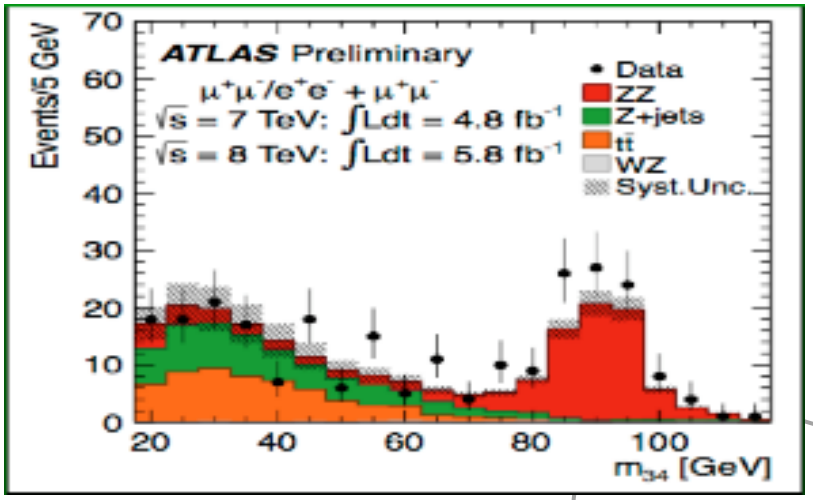

ANTIOCH
WINTHROP HARBOR
BEACH PARK
FOX LAKE
ROUND LAKE
WAUKEGAN
LAKE BLUFF
WAUCONDA
FOX RIVER GROVE
LINCOLNSHIRE
BARRINGTON
WEST DUNDEE
WINNETKA
ELGIN
EVANSTON
ITASCA
NORRIDGE
NORTHLAKE
GENEVA
WHEATON
HILLSIDE
CHICAGO
NORTH AURORA
WESTERN SPRINGS
AURORA
JUSTICE
BOULDER HILL
CHICAGO RIDGE
RIVERDALE
HAZEL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

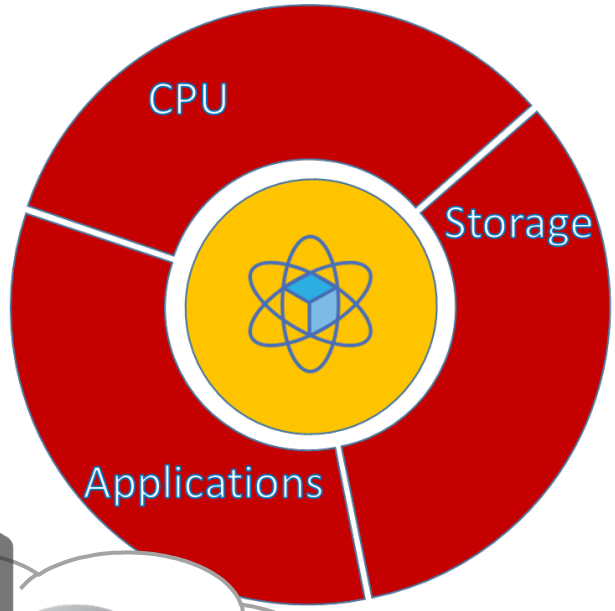
Table of Content

- **Introduction**
  - Big data: LHC computing
- **Collaborations**
  - Established ones
    - **EOS**
  - Starting…
    - **CERNBox**
  - … and future perspectives
    - **SWAN**

`Cloud analysis: **SWAN** project

with CERN Physics Department

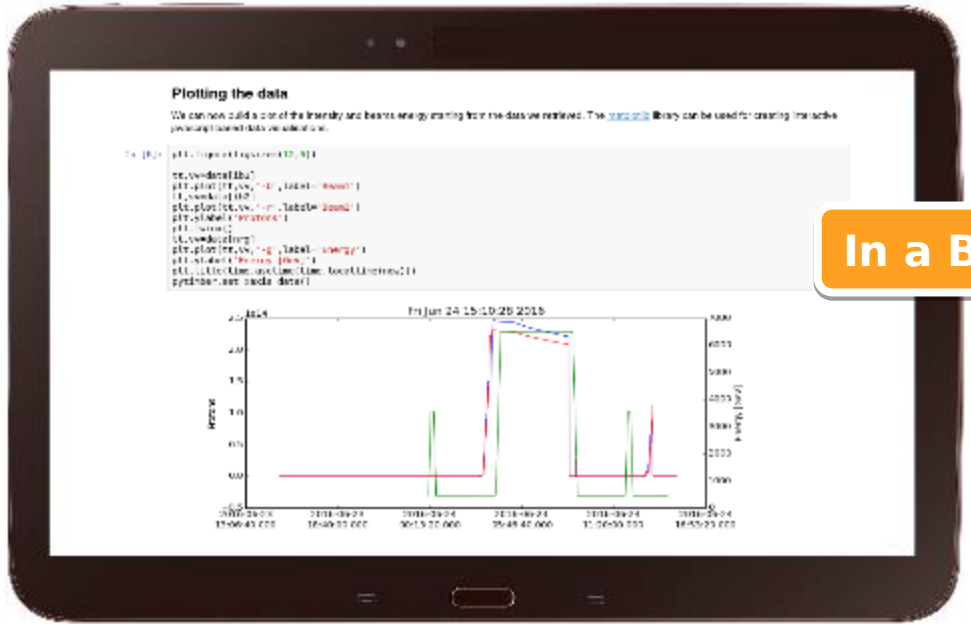




Lots of activity in previous projects with several Russian groups, notably with V. Korenkov (JINR Dubna)

ROOT is the CERN data analysis framework: http://root.cern.ch

Jupyter Notebook: A web-based interactive computing interface and platform that combines code, equations, text and visualisations



Interface: The Notebook

Text

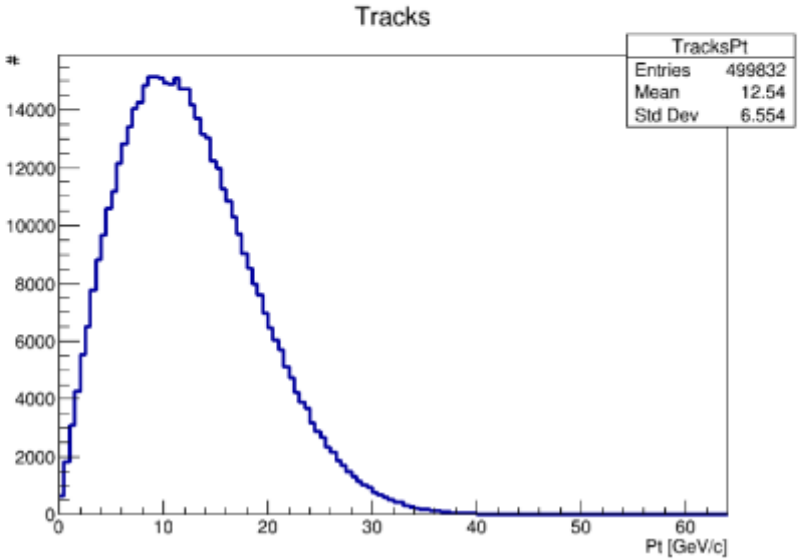
Code

Graphics

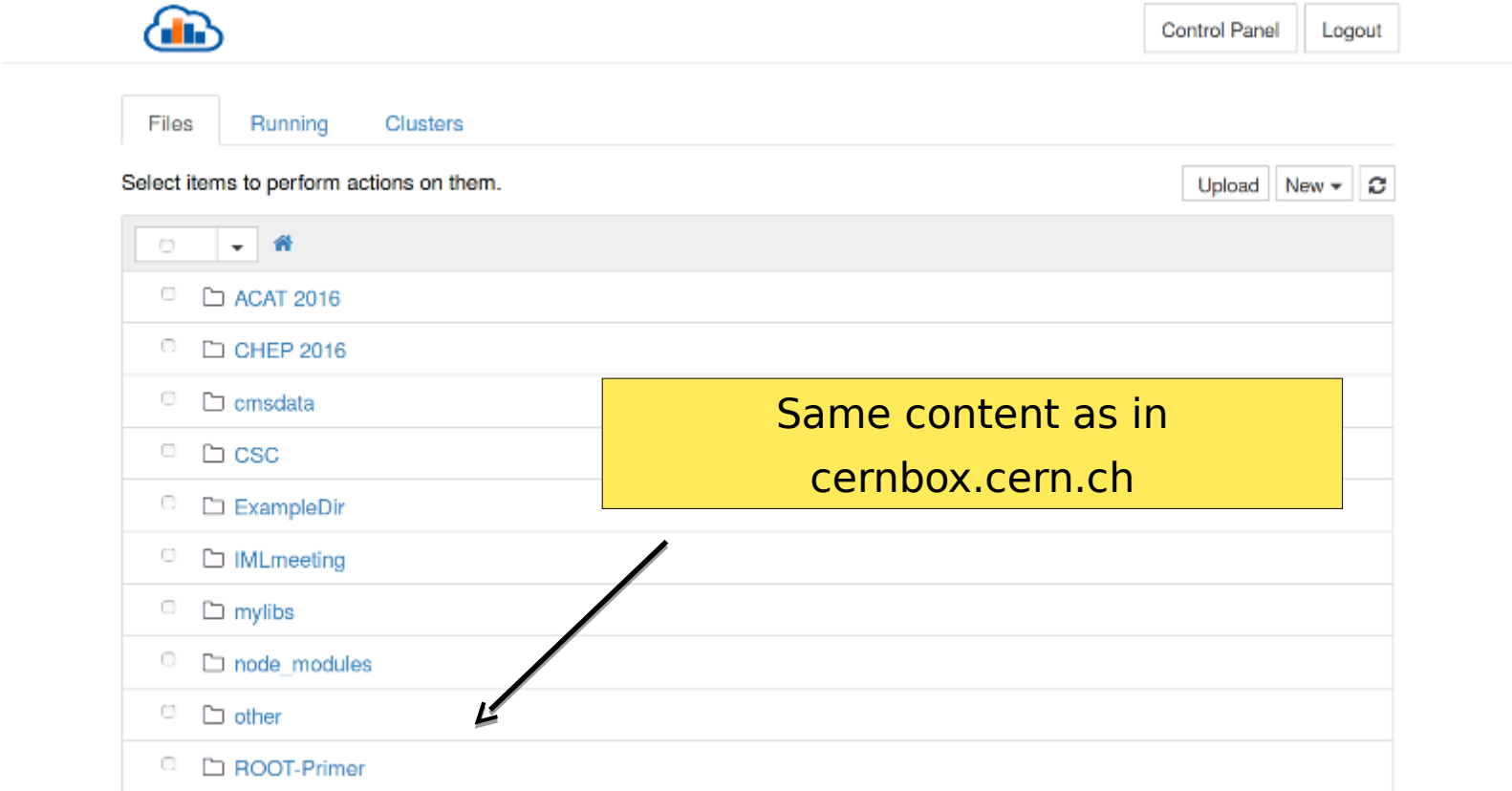
Access TTree in Python using PyROOT and fill a histogram

Loop over the TTree called "events" in a file located on the web. The tree is accessed with the dot operator. Same holds for the access to the branches: no need to set them up - they are just accessed by name, again with the dot operator.

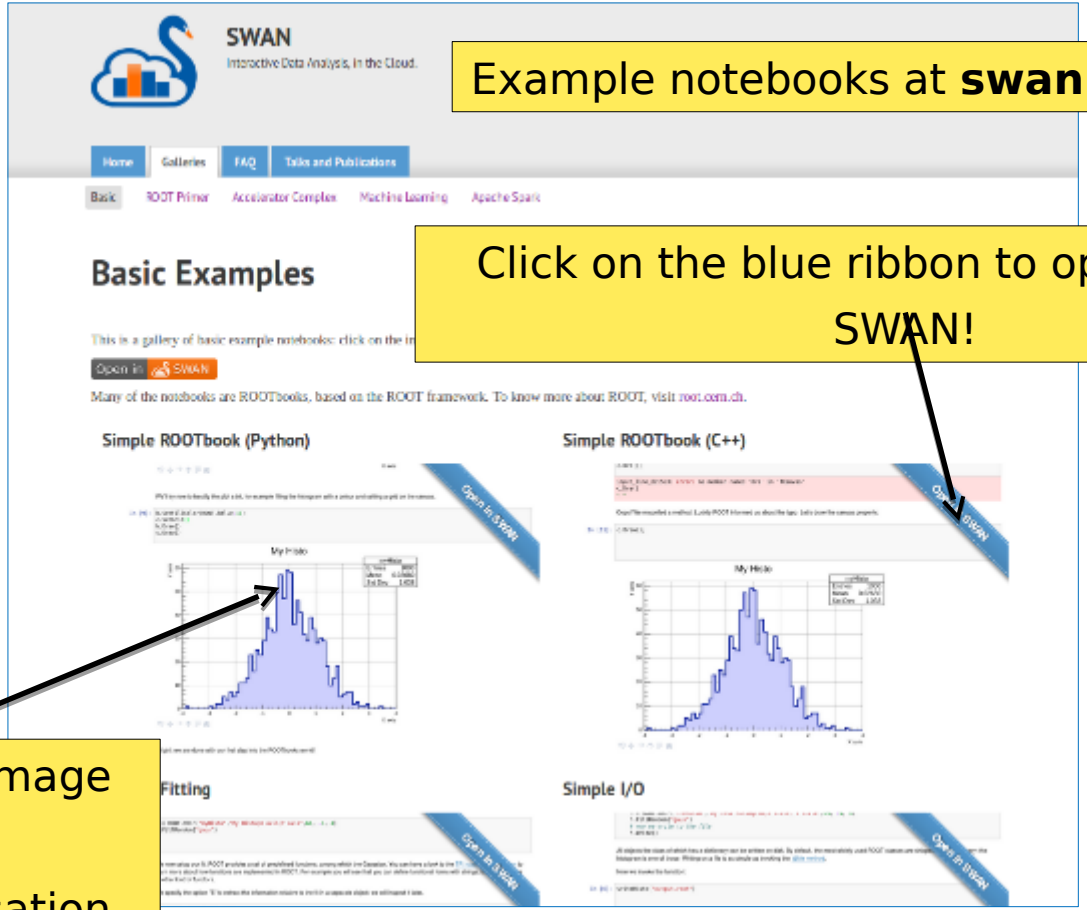
In [1]:
```
import ROOT

f = ROOT.TFile.Open("http://indico.cern.ch/event/395198/material/0/0.root");
h = ROOT.TH1F("TracksPt","Tracks;Pt [GeV/c];#",128,0,64)
for event in f.events:
    for track in event.tracks:
        h.Fill(track.Pt())
c = ROOT.TCanvas()
h.Draw()
c.Draw()
```



CERNBox as Home



Notebook Galleries



Example notebooks at **swan.web.cern.ch**

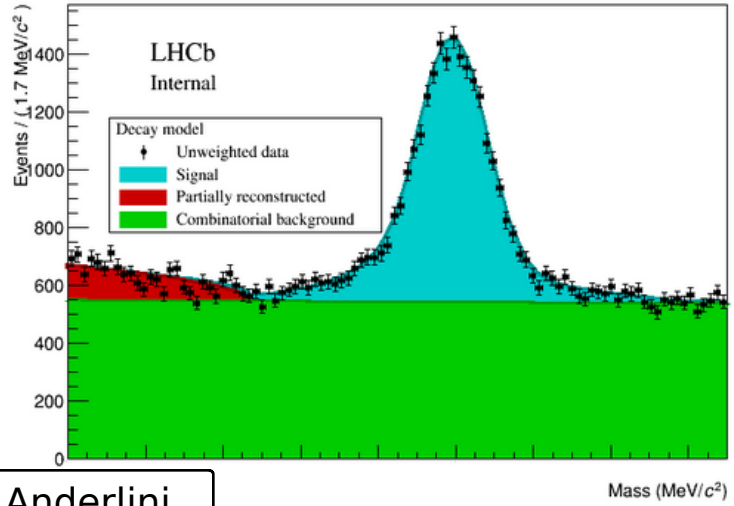
Click on the blue ribbon to open them in SWAN!

Click on the image for a static visualisation

SWAN Use Cases

```
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()
```

Results coming from real data! (published now)



L. Anderlini

Physics Analysis

Rare B meson decay in LHCb

- Read data from EOS
- Setup complex fit
- Document and inspect results

- SWAN as platform for outreach
  - Introductory course about experimental HEP for future high school teachers

Particle open data teaching (Hiukkasfysiikan avoin data opetuksessa)

Esim-pseudorapiditeetti-mittatarkkuus (autosaved)

Control Panel | Logout

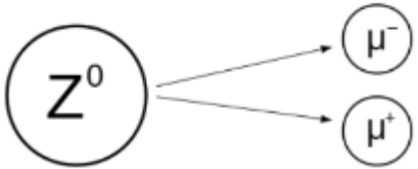
File Edit View Insert Cell Kernel Help | Python 2

Markdown | CellToolbar

Lähdetäänpä tutkimaan!

Lähdetään seuraavaksi tarkastelemaan, miten pseudorapiditeetin vaikutus mittatarkkuuteen voidaan havaita CMS-ilmaisimen keräämän oikean datan avulla. Käytetään CMS:n vuodelta 2011 kerättyä dataa [1], josta on valittu 10851 törmäystapahtumaa (events) tiedostoon "Zmumu_Run2011A_massoilla.csv". (Karsinta on suoritettu koodilla, joka on avoimesti saatavilla osoitteessa https://github.com/tpmccauley/dimuon-filter.)

Tiedostoon on valittu niitä törmäystapahtumia, joissa syntynyt Z-bosoni on hajonnut myoniksi μ^- ja antimyoniksi μ^+. Ilmaisin on havainnut nämä myonit ja mitannut niiden liikemäärät.

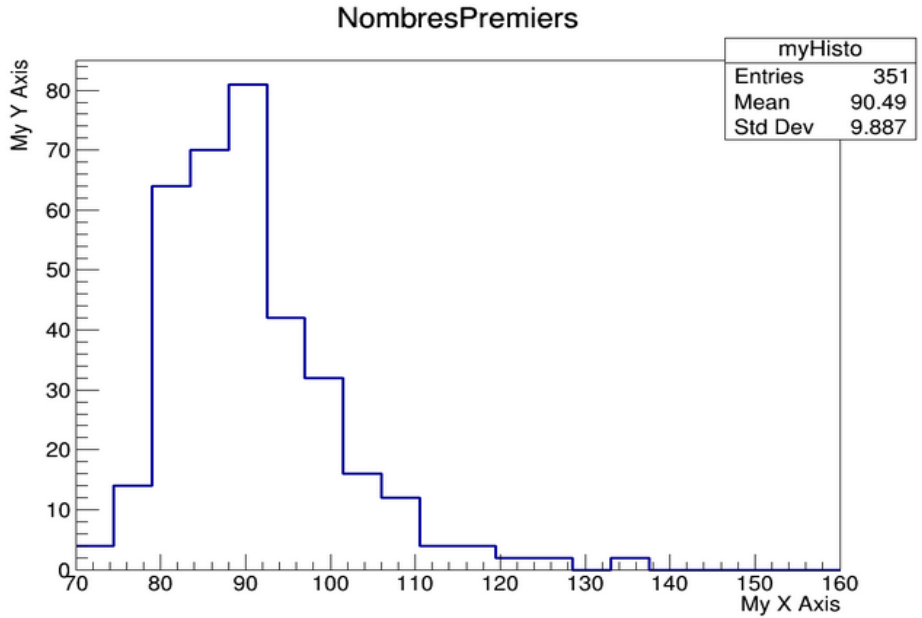


P. Rikkilä

Education

```
In [138]: import ROOT
htemp = ROOT.TH1F("myHisto","NombresPremiers;My X Axis;My Y Axis",20,70,160)
for i in range(len(data)):
    d = data[i][0]
    htemp.Fill(float(d))
c = ROOT.TCanvas("myCanvas","myCanvasTitle",1024,768)
htemp.Draw()
c.Draw()
```

```
TROOT::Append:0: RuntimeWarning: Replacing existing TH1: myHisto (Potential memory leak).
TCanvas::Constructor:0: RuntimeWarning: Deleting canvas with same name: myCanvas
```

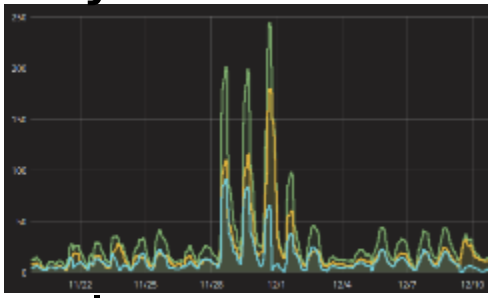


Mano S. (14 years old), K12 student

- Approaches programming for the first time
- Verifies numerically what he learned at school
- Shares results with his supervisor and classmates

- Practical Statistics for Particle Physics Analyses

https://indico.cern.ch/event/545212/



- CERN Summer Student Program: ROOT

https://indico.cern.ch/event/536772/

- CERN School of computing: Parallelization lectures

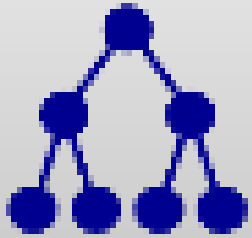
http://indico.cern.ch/event/502875/

- Data Science @ LHC Workshop, Multivariate analysis tutorial

http://indico.cern.ch/event/395374/

Summary

- **Solid foundations**
  - 200 PB LHC disk infrastructure
    - Steadily growing!
  - HEP collaborations
- **Strategic partnership**
  - HEP computing evolution
  - Cloud storage enables new use cases
    - and new ways to work and to collaborate








EOS


101
011

www.cern.ch