Working with BIG DATA

Tim Smith CERN/IT

Open Access and Research Data Session

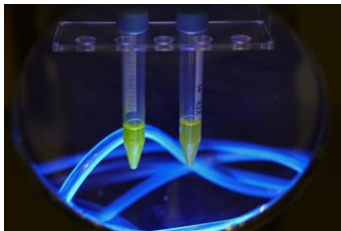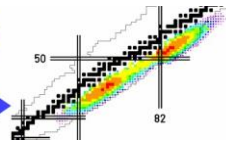# Research Big Data
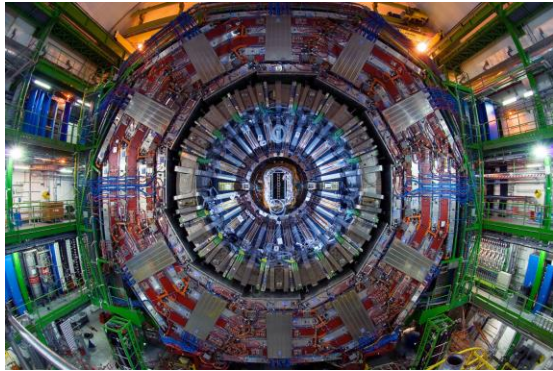
How big is BIG ?

40M collisions /sec

150M readout channels

100 PB

100 PB

10 PB

1 PB

100 TB

10 TB

1 TB

100 GB

10 GB

1 GB

10 PB/s

Trigger

10 TB/s

Filter

6 GB/s

3 years
LHC Data

Primary Storage

6 GB/s
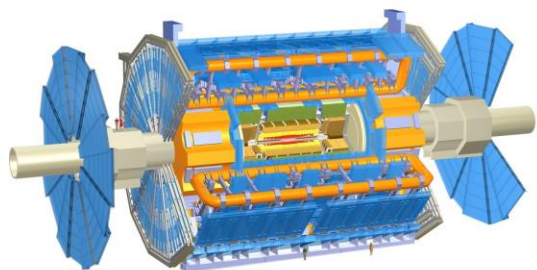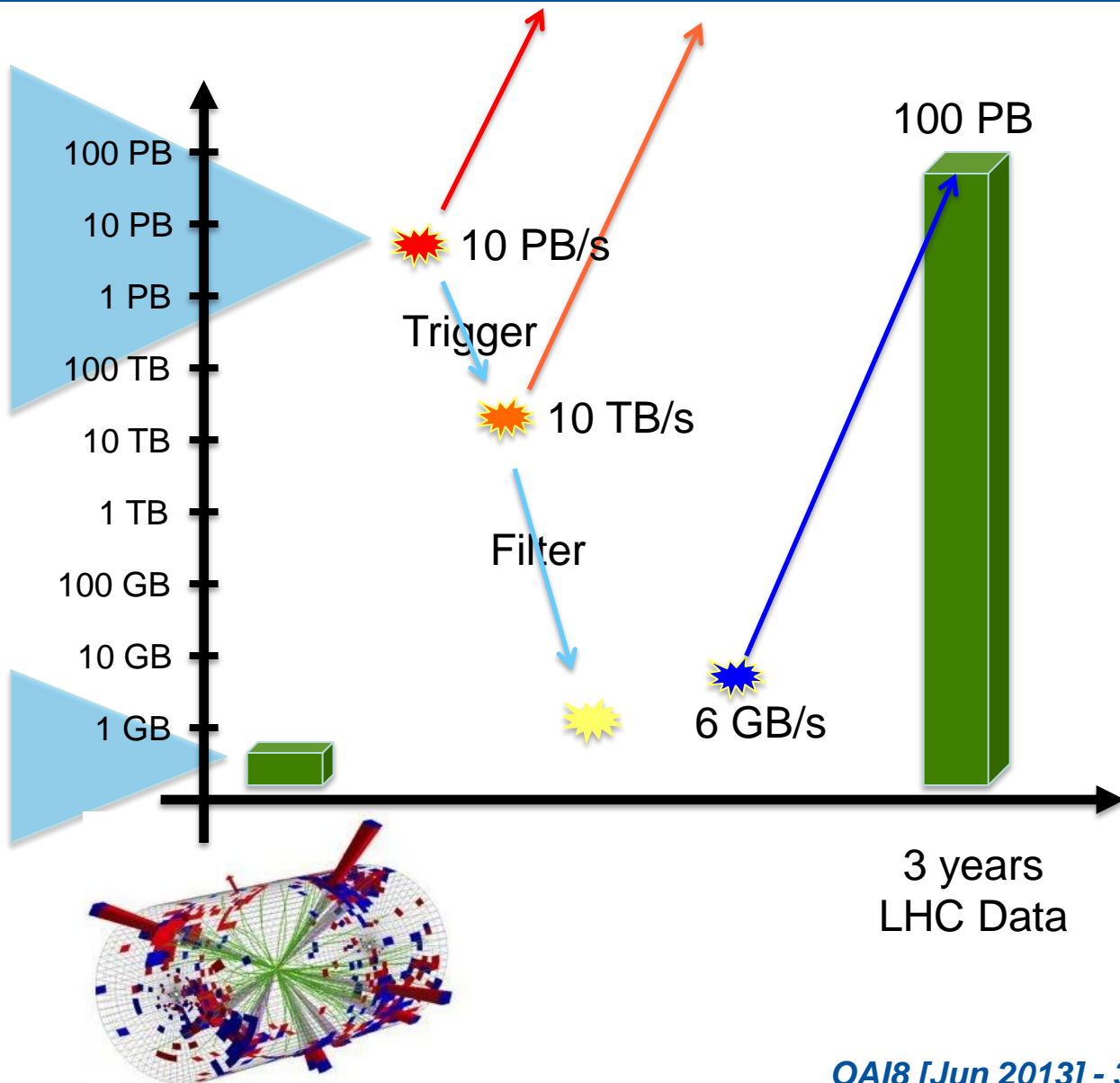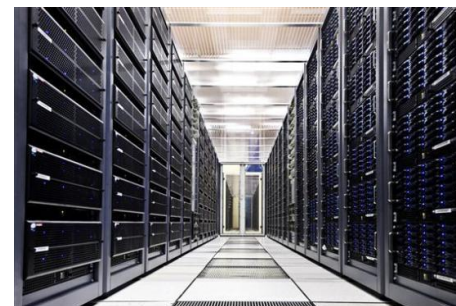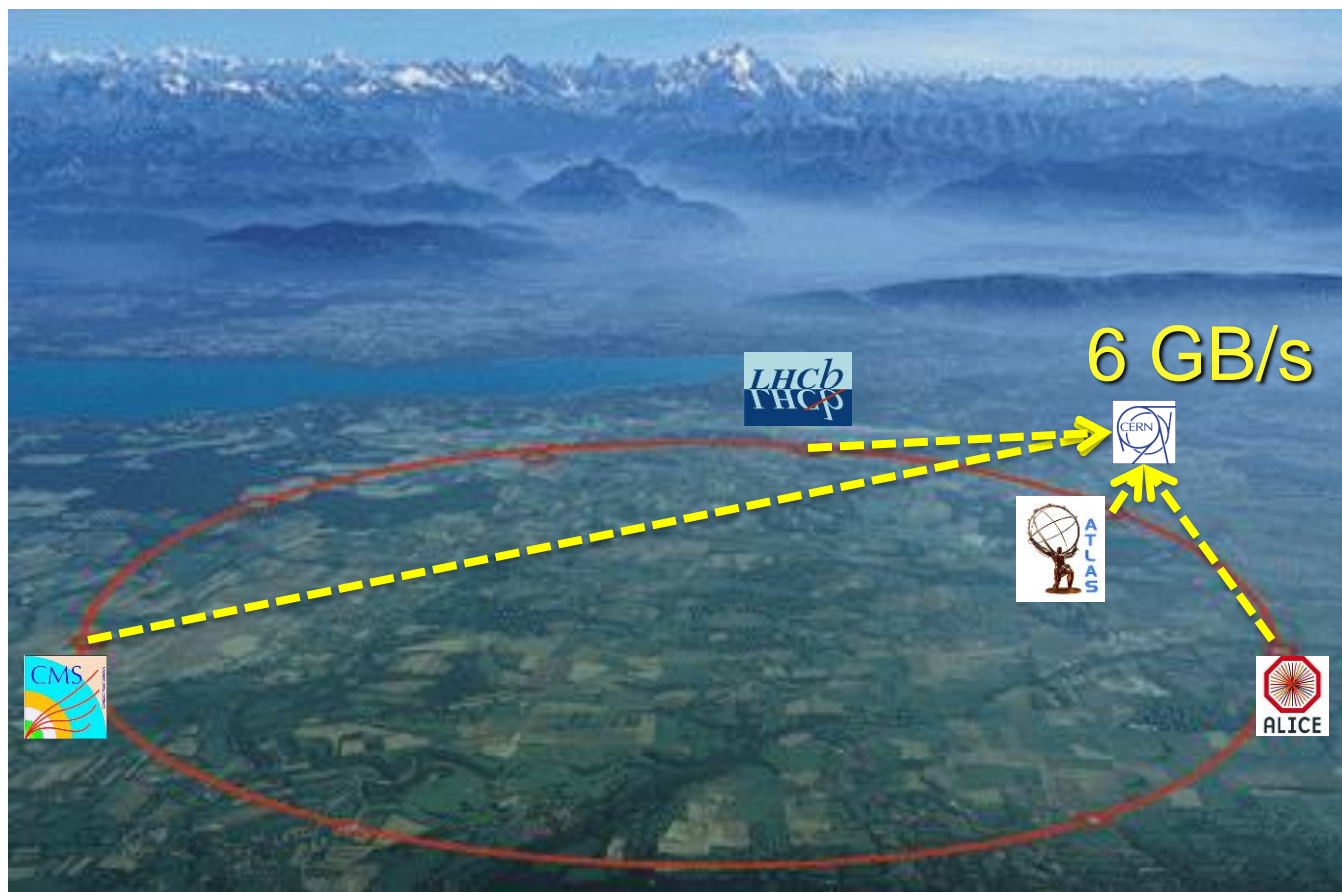
80,000 Disks
88,000 CPU Cores

18,000 1GB NICs
2,700 10GB NICs

Problems Size Brings

- Few places can store it
  - Lots of Copies ?
  - 10% at 10 centres
- Hard to transport
- Aggregation easier for processing than storage

- **M**odels **o**f **N**etworked **A**nalysis at **R**egional **C**entres
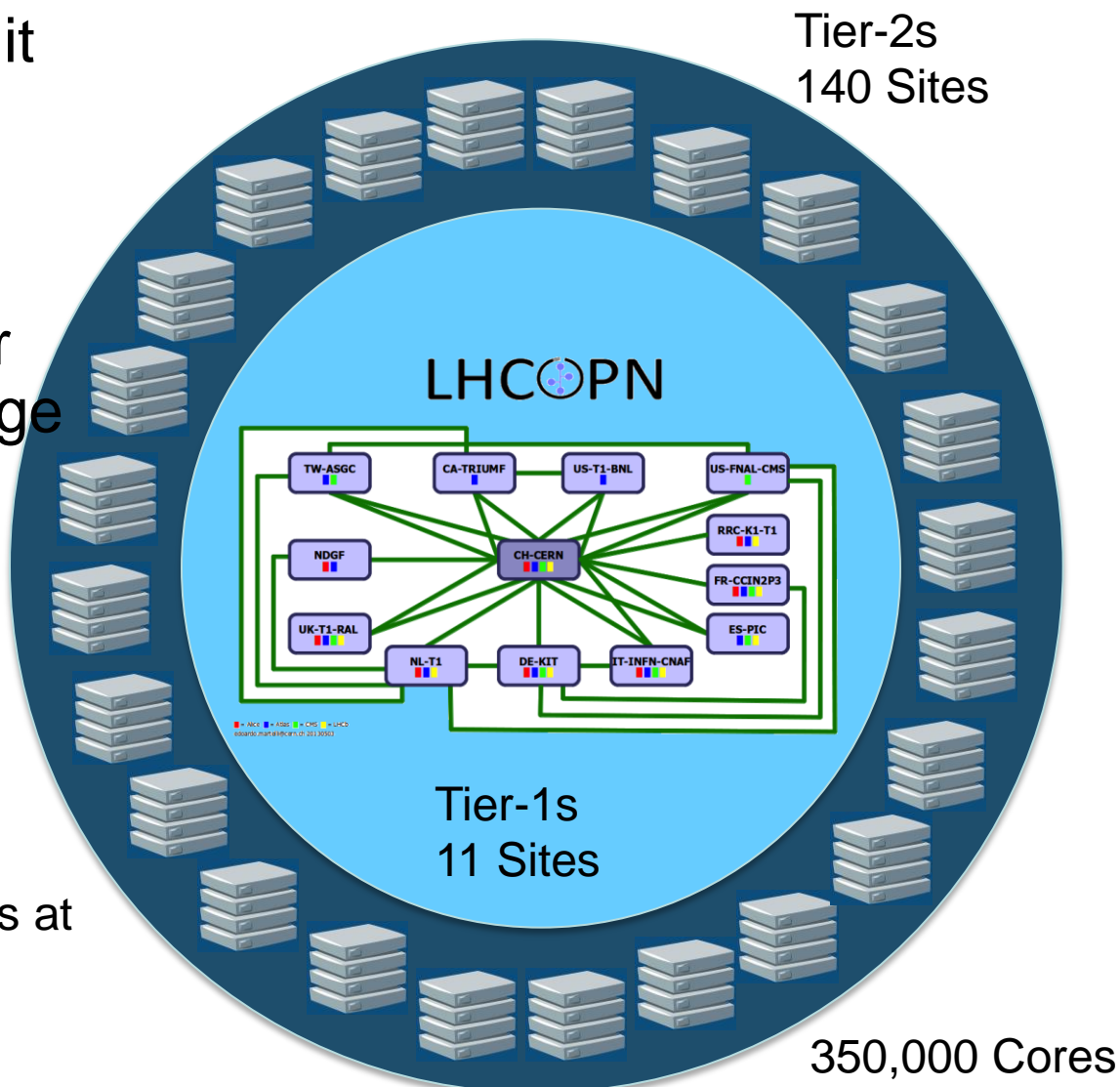
x 2 locations @ CERN

Problems Size Brings

- Few places can store it
  - Lots of Copies ?
  - 10% at 10 centres
- Hard to transport
- Aggregation easier for processing than storage

- Models of Networked Analysis at Regional Centres

Tier-2s
140 Sites

LHC OPN

Tier-1s
11 Sites

350,000 Cores

Worldwide LHC Computing Grid

- **Distributed Data Management**
  - Limited Network resources
  - Optimize / minimize movement
  - File placement logic
  - Deterministic / Static
- **Site Data Management**
  - HSMs
  - Transparent file access and movement
    - Disk-Tape migration/recall

Running jobs: 125377.0
Transfer rate: 6.02 GiB/sec

© 2010 Google
© 2010 Europa Technologies
US Dept of State Geographer
© 2010 Tele Atlas

©2010 Google

Research Data Infrastructure of today

- ## Distributed Data Management
  - Network: a resource to schedule
  - Dynamic data placement
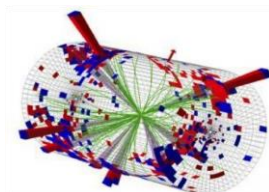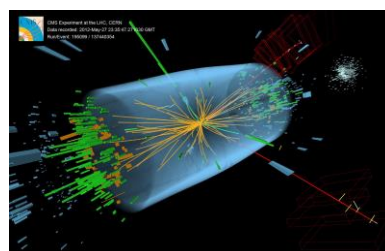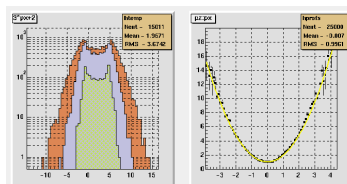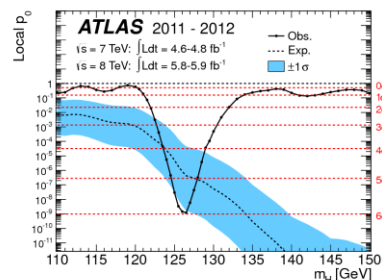 - Data transfer services
  - Expt replica management rules
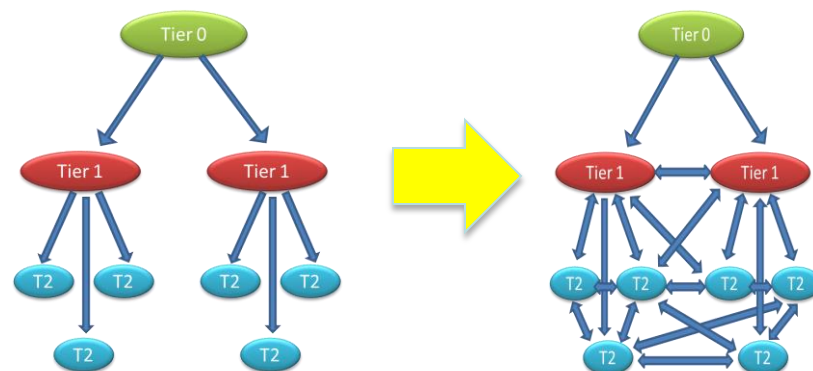
- ## Site Data Management
  - Indep. technology choices
  - Decoupled tiers
  - Disk caches
    - Managed by owners
  - Bulk 3rd party migration to tertiary by owners
- AAA: any data, any time, any where ☺
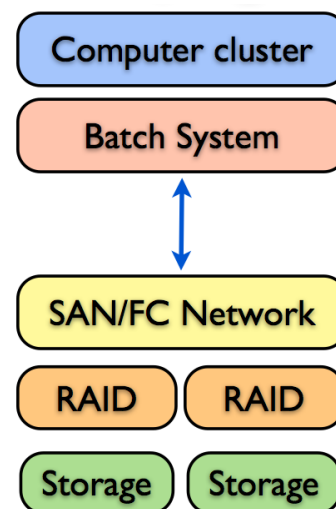
Data Reduction / Analysis

- Publication

- Reduced
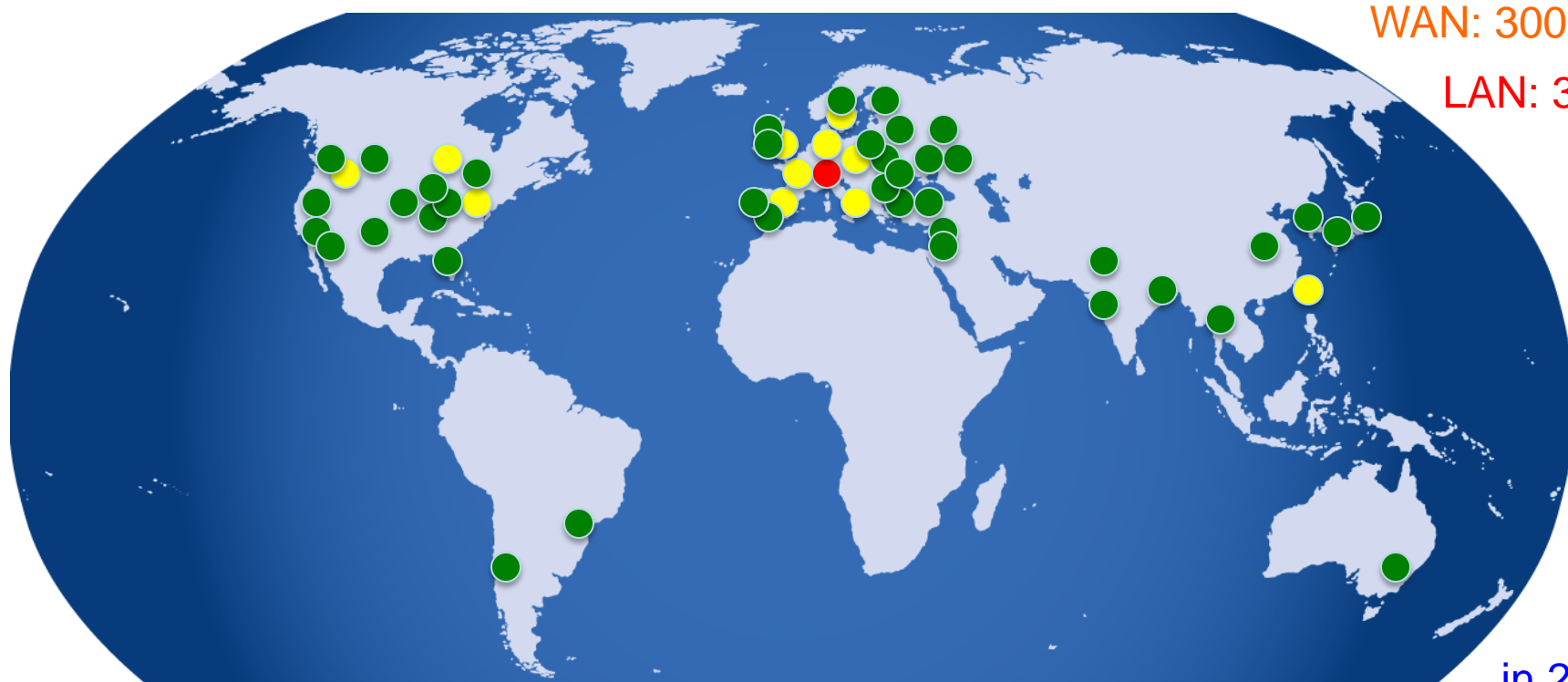
- Reconstructed

- Raw

File Size

Files

Researchers
T2s, T1s

Analysis Coordinators
T1s

Production Managers
T0, T1s

Data Inflation

- Static: Storing 100 PB was a good challenge ☺
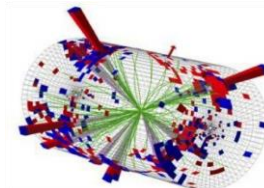- Dynamic: Analysing it means transformation, reduction, transport, replication, regeneration
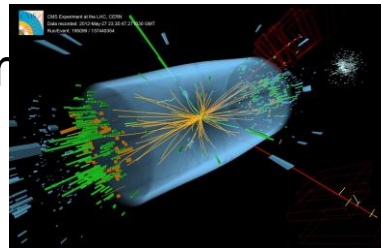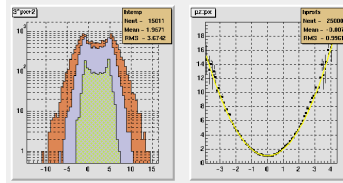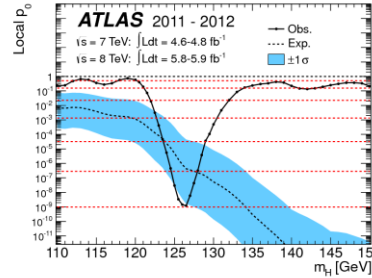
WAN: 300 PB

LAN: 3 EB

in 2012

More than Data

- Papers
- Tabular Data
- Correlation Matrices

- Internal Notes
- Wikis
- Presentations

- Quality monitoring data
- Filter / selection algorithm
- Formatters

- Calibration Data
- Conditions Data
- Log Books

Researchers
T2s, T1s

Analysis Coordinators
T1s

Production Managers
T0, T1s

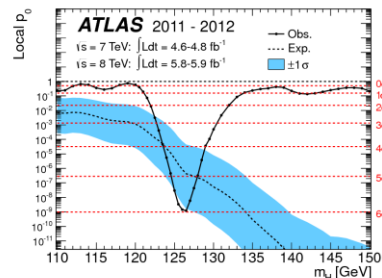SW: 10M LoC   Contextual metadata   Workflows

Data for Tomorrow
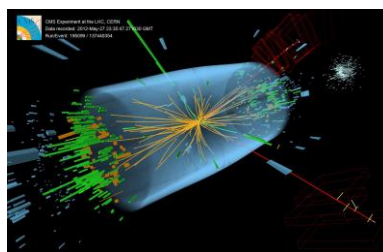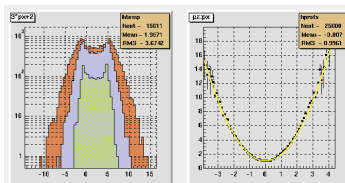
- ## Digital Libraries
  - DOIs, ORCID
  - DataCite, OAIS
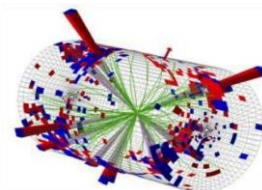
- ## Custom systems
  - DBs
  - HEP formats

- ## Mass Storage
  - GUIDs
  - Access
  - Bit Preservation

iNSPIRE

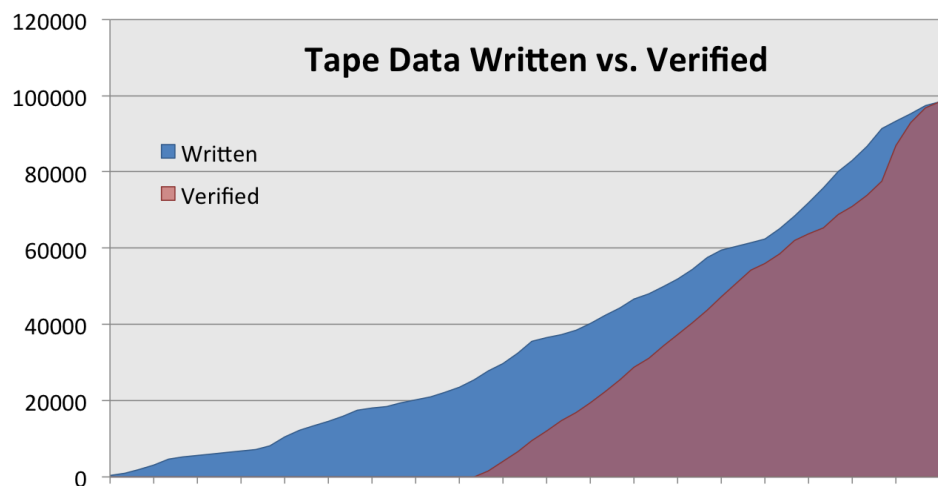INVENIO

30 TBs !

DPHEP

CASTOR
CERN Advanced STORage manager

Bit Access / Bit Preservation

- Media Verification
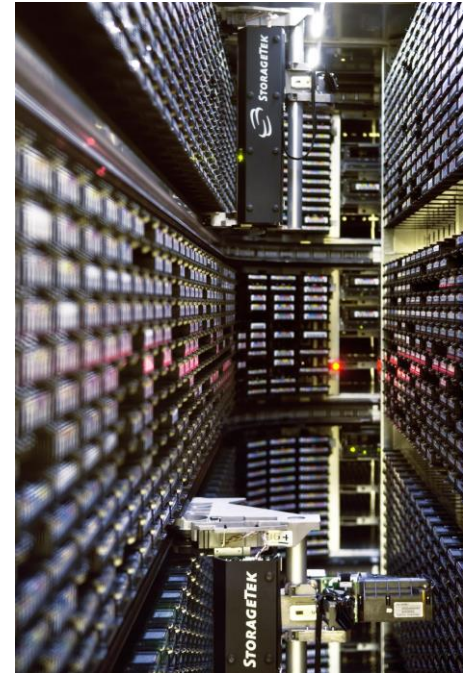 - Hot / Cold Data
  - Catching and correcting errors while you still can
  - 10% of production drive capacity for 2.6 years

Bit Rot

Tape Data Written vs. Verified
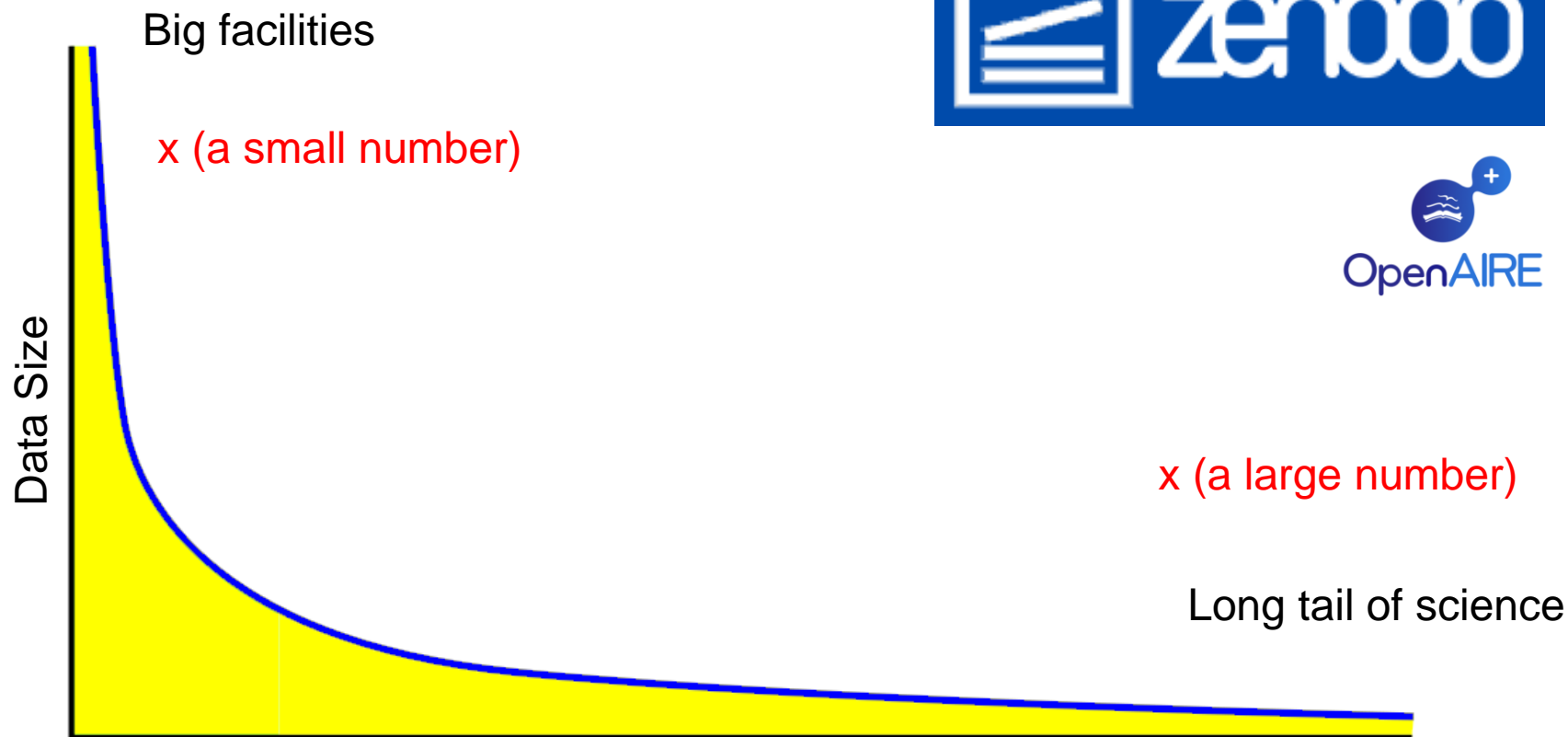- Written
- Verified

  - (0.000065% data loss)

Bit Access / Bit Preservation

- ## Media Migration
  - – Drive and Media obsolescence
  - – 50% of current drive capacity for 2 years

Little Big Data

Big facilities

x (a small number)

Data Size

x (a large number)

Long tail of science

Concluding Remarks

- Data Management best done by the data owner
  - Make data management services available
  - Empower the users!
- Don't assume Researchers behavioral patterns
  - Decouple the layers
  - Offer integratable building blocks, not an integrated solution
- Static Data is one thing, planning for Active Data quite another
  - (The number you know) x (a large factor) ☺
- Housekeeping (media verification and migration) takes non-negligible resources
- Preservation means combining expertise
  - Storage managers, SW engineers, Librarians and Researchers

www.cern.ch