

# Big Data in High Energy Physics

Randall Sobie  
Institute of Particle Physics  
University of Victoria

## High Energy Physics (Particle Physics)

The area of physics that studies the fundamental particles of nature and their interactions.

### Accelerators



SLAC Linear Accelerator

### Underground labs



Sudbury Solar Neutrino Detector

### Orbiting labs



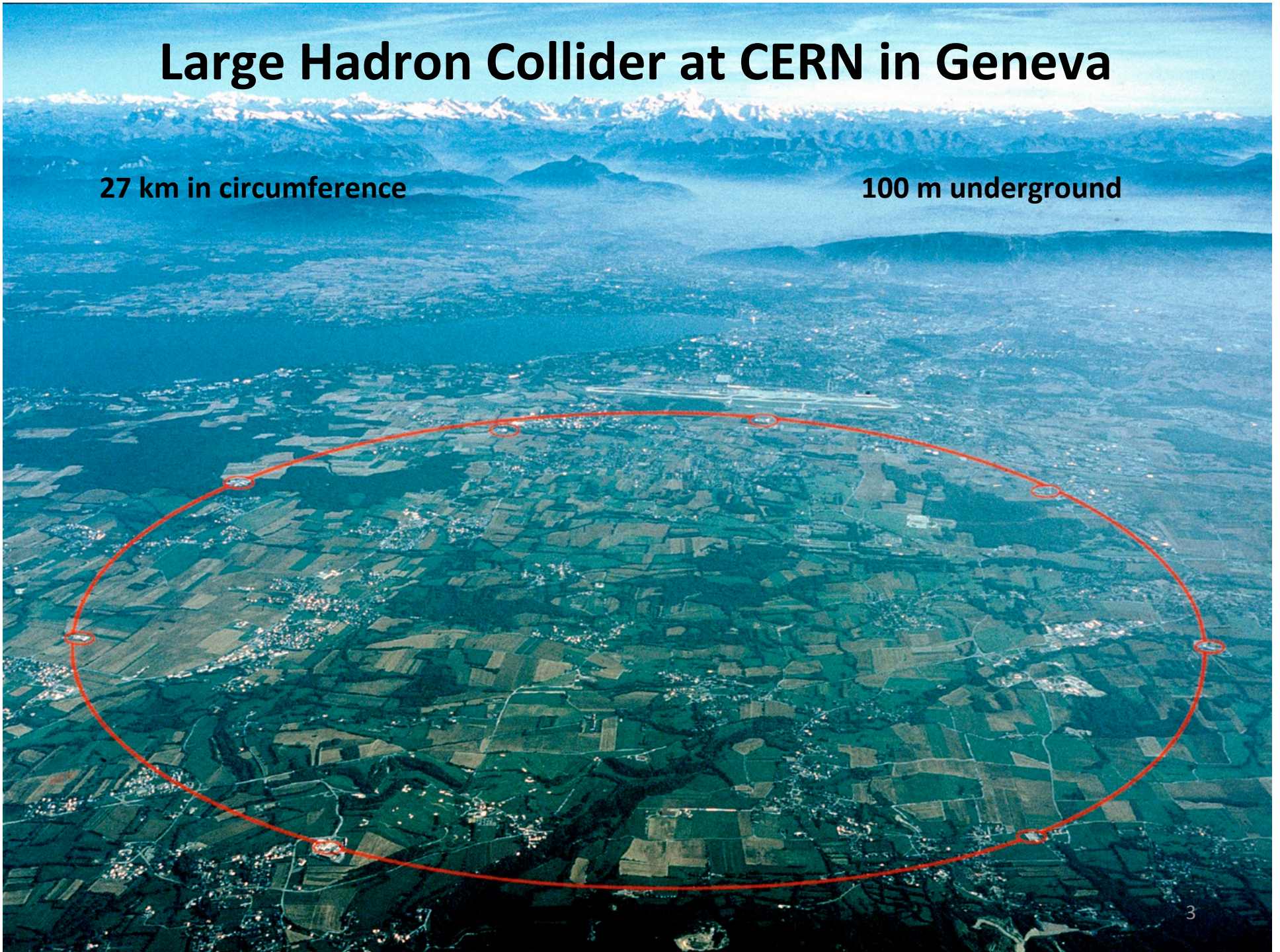
Alpha Magnetic Spectrometer  
Space Station



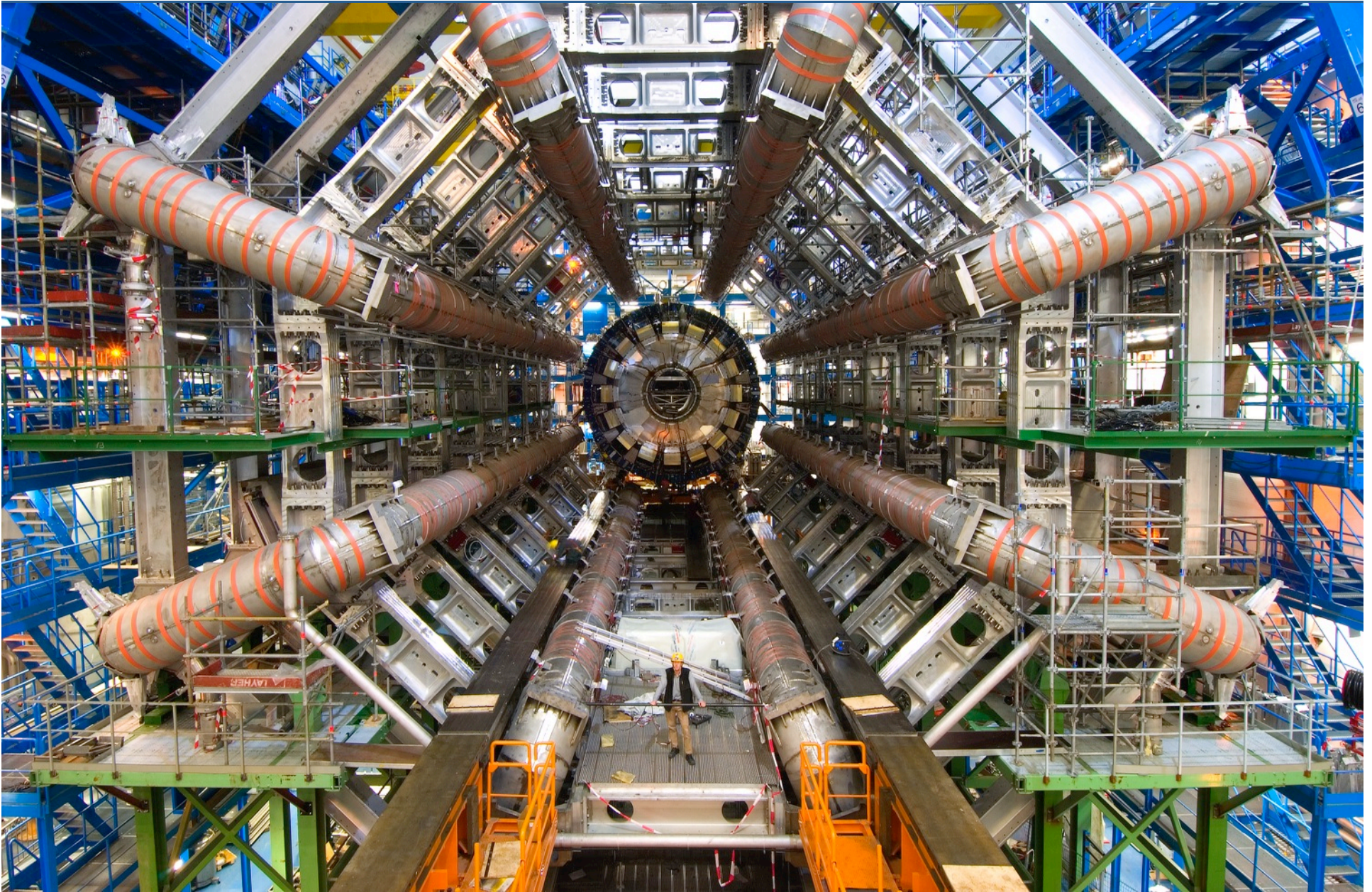
# Large Hadron Collider at CERN in Geneva

27 km in circumference

100 m underground

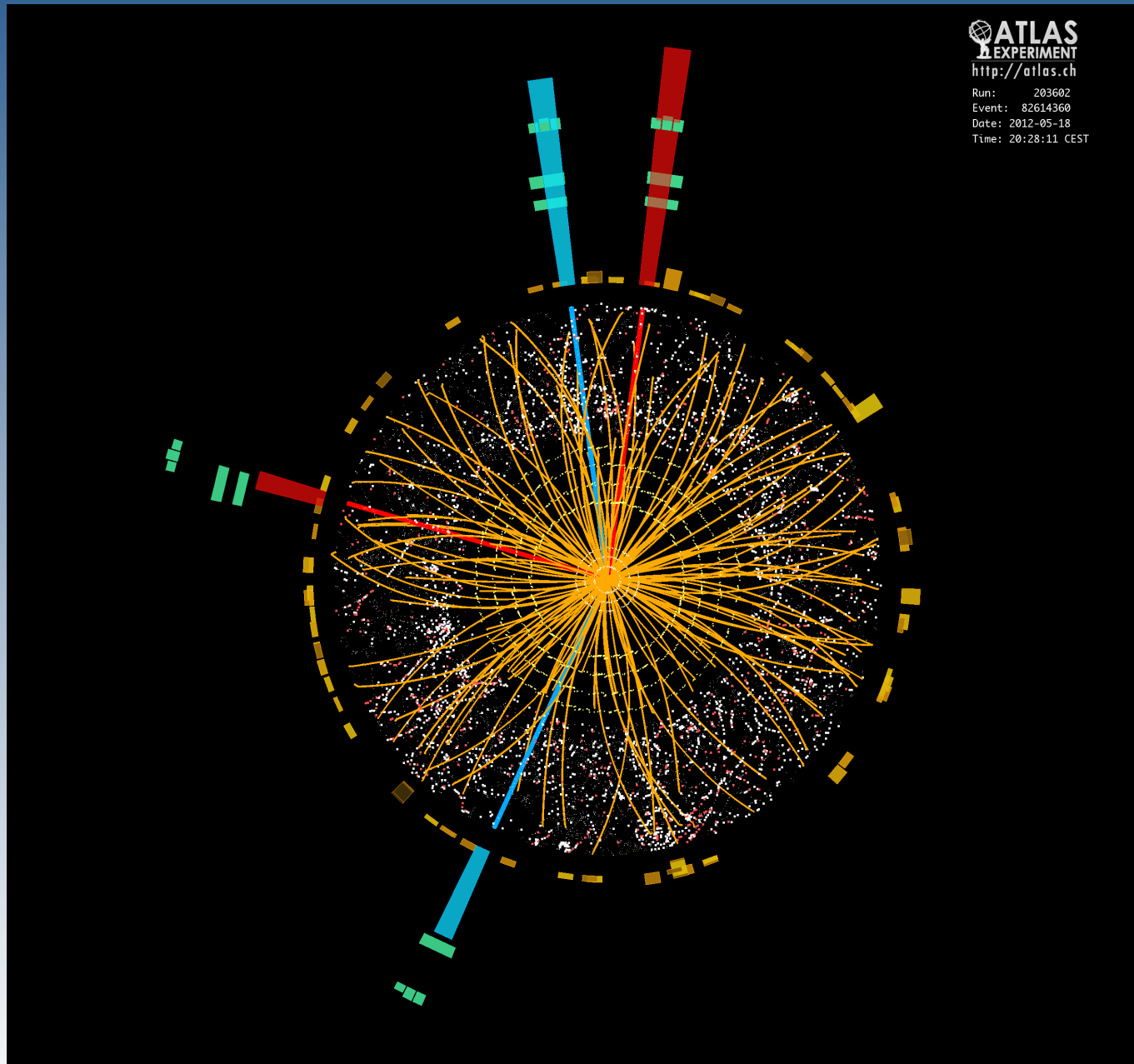


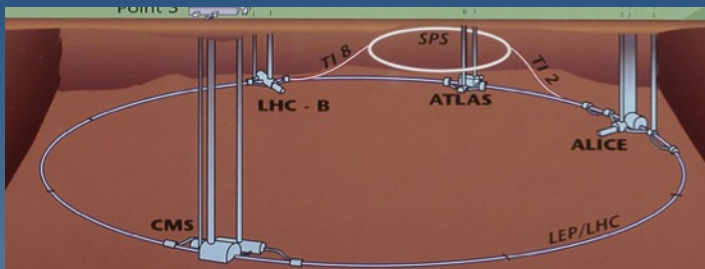




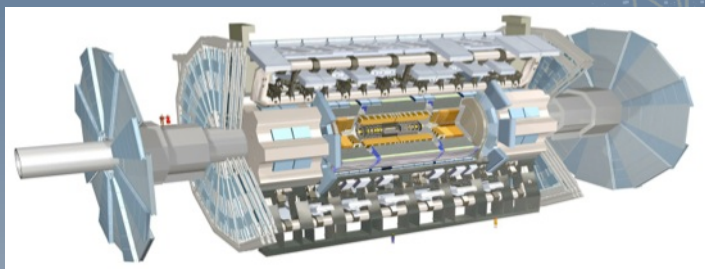
ATLAS detector during construction in 2005







40 million collisions per second



100,000 collisions selected



200 events per second

## WLCG Computing Grid

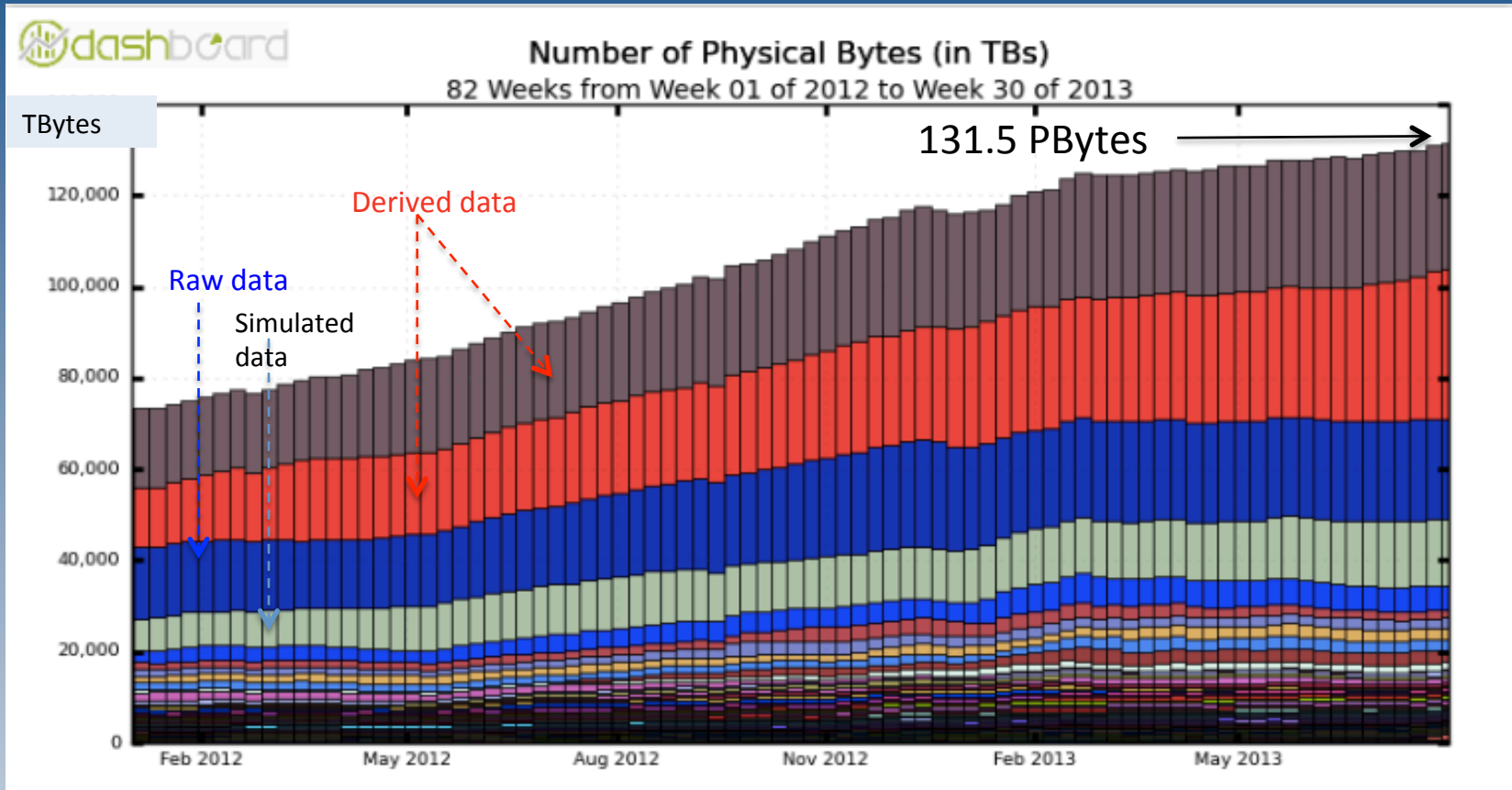
CERN Tier 0

10 Tier1 sites

60+ Tier2 sites

140 PB data





**Raw data** is distributed to 10 *Tier 1* centres around the world

TRIUMF Laboratory in Vancouver is the Canadian ATLAS Tier 1 centre

Other samples distributed to ~60 *Tier-2* centre

McGill, Toronto, SFU and Victoria host T2 centres

Analysis jobs are sent to T2 centre that hosts the data

# Big Data requires very high speed networks

## HEPnet Canada

Responsible for national and international networks in for  
Canadian Subatomic Physics community

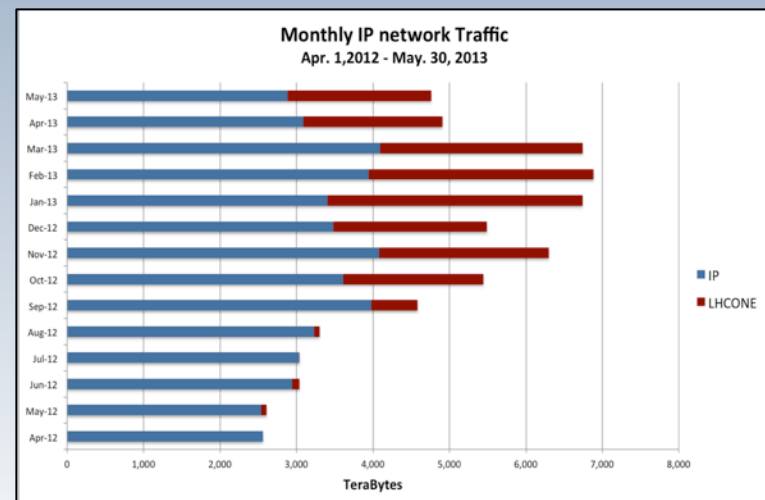
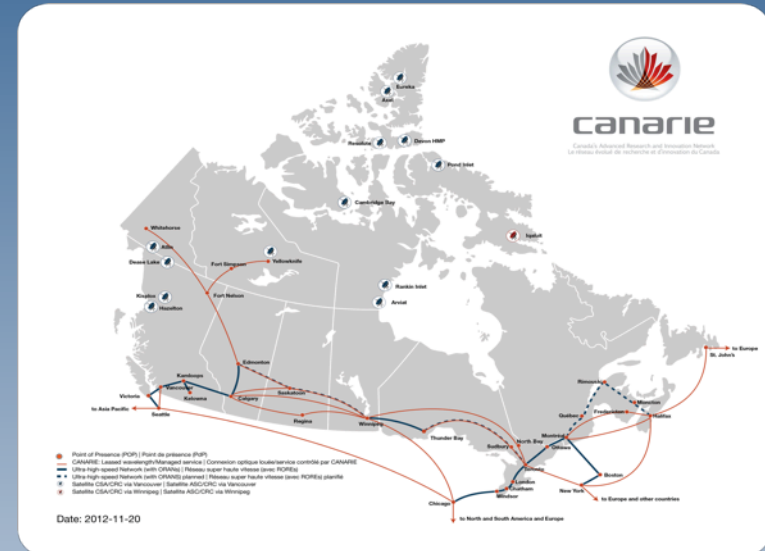
5G lightpath from Vancouver-Geneva

10G VPN for global LHC traffic  
*HEP is 50% of the Canadian research traffic*

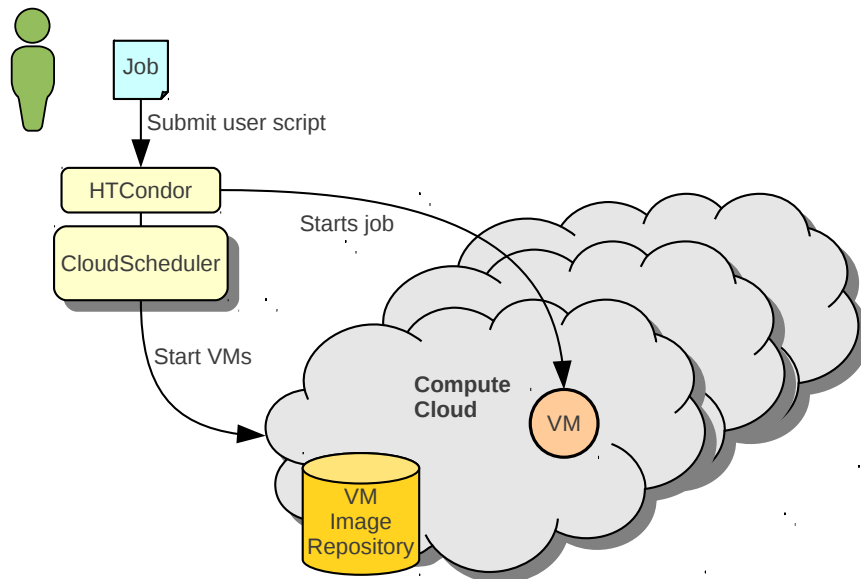
Testing 100G transatlantic link

High-speed network demo projects  
*Set WAN world records*

Exploring Software Define Networks (SDN)  
*Application controls the network*







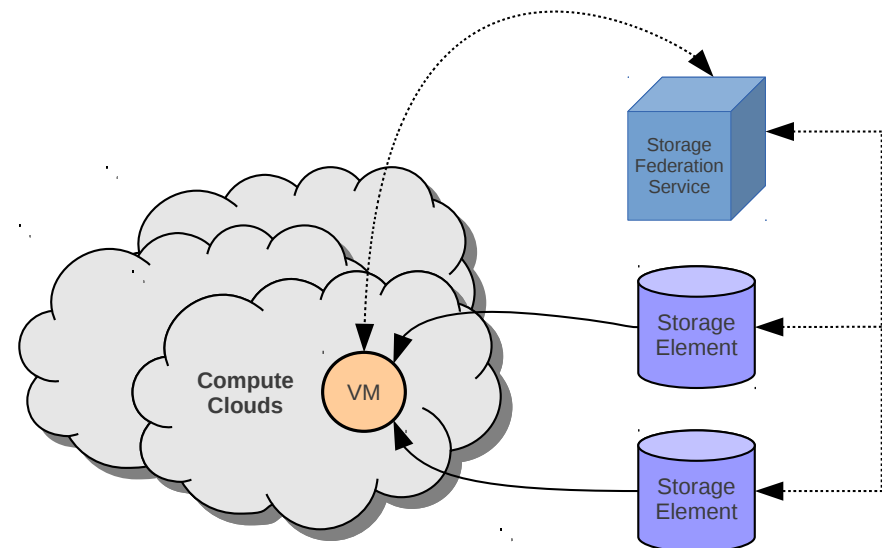
## Batch cloud computing

Using 30 clouds in 3 continents

Workloads in excess of 2000 jobs for ATLAS and BelleII experiments

## Storage federations

VM queries SF service to get reference to nearest storage element with a copy of the file



Big Data help us understand our universe

This requires world-class networks and fast storage facilities

*Multi-100G networks is changing our computing models*

Migrating to an on-demand model using computing clouds and federated storage

