Presentation on

WLCG: Worldwide Large Hadron Collider (LHC) Computing Grid

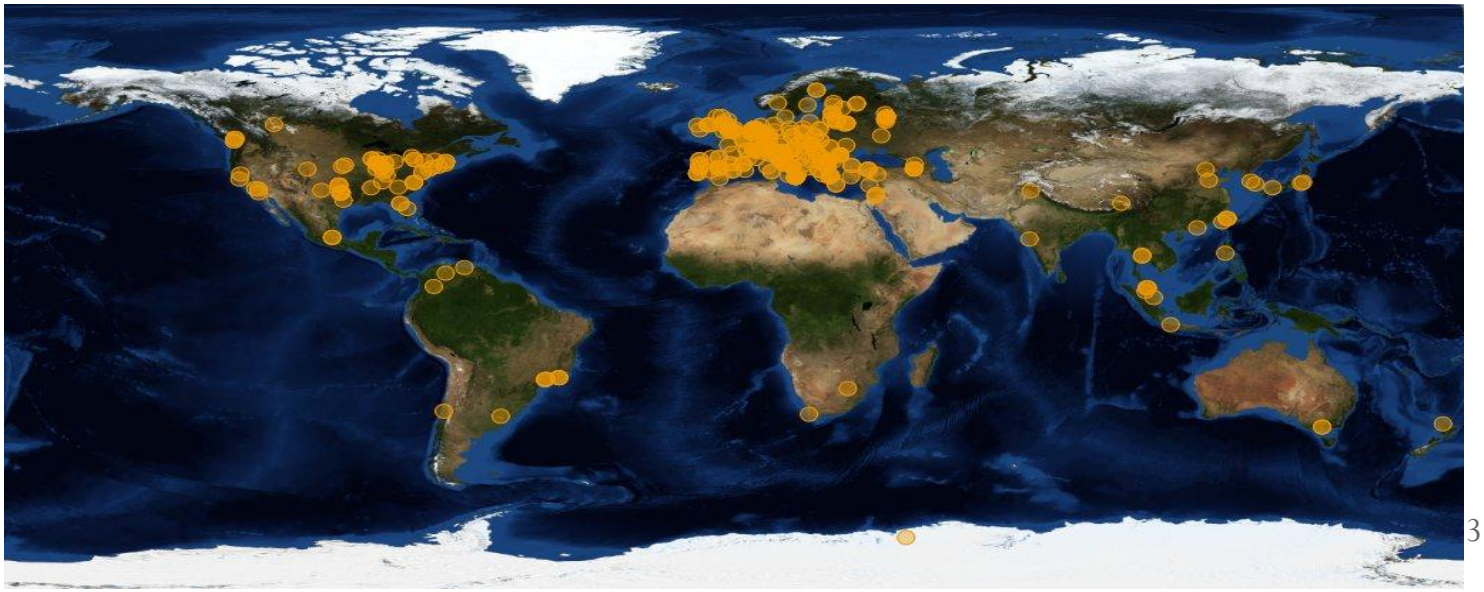
Md Anowarul Abedin
Master Student, Erasmus Mundus PERCCOM
ITMO Seminar

OUTLINE

- What is WLCG
- Why WLCG
- WLCG in Detail
- Grid for Big Data
- Grid Architecture
- Using the Grid
- Data Security of the Grid
- What Now?
- WLCG for Bigger Data
- References

WHAT IS WLCG

- International collaborative project
- Largest computing grid of the world
- The most sophisticated data-taking & analysis system [3]
- 170 computing centers from 42 countries [2]
- Consists of a grid-based computer network infrastructure
- Objective is to handle the data produced by Large Hadron Collider (LHC) experiments



WHY WLCG

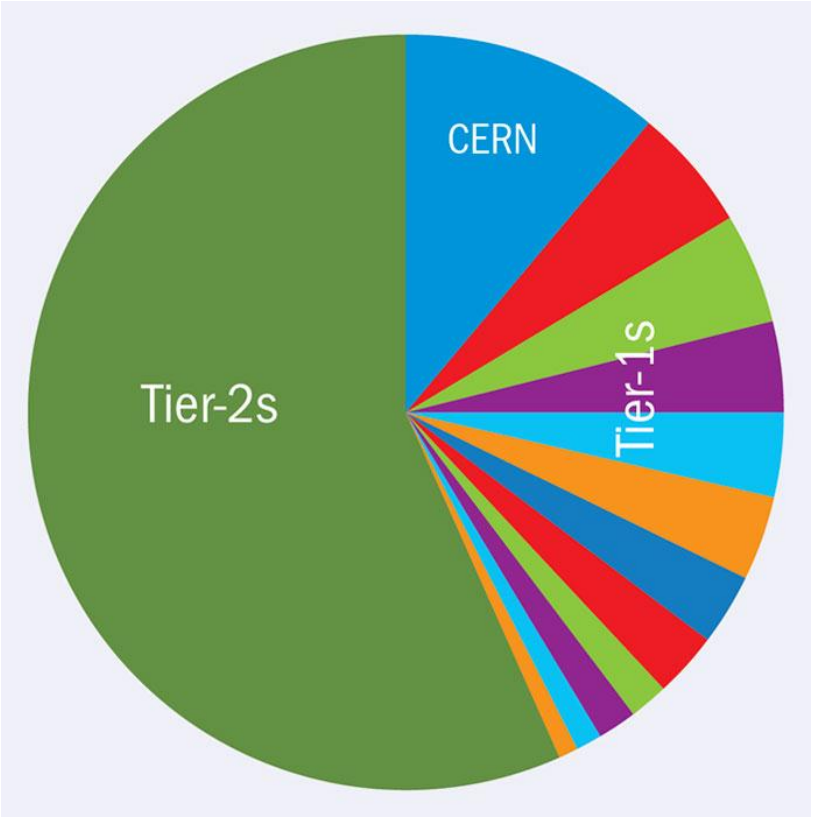
- To prove or disprove the existence of the Higgs-Boson particle
- Higgs-Boson particle is sought for over 40 years
- Higgs-Bosons might not be seen in lower energy experiments
- A very powerful particle accelerator is used for a large number of collisions and huge energy
- Such a collider would also produce huge data requiring analysis
- Therefore advanced computing facilities are needed to process the data
- No center could provide such computing requirements alone
- Financial and technical supports are distributed throughout the world

WLCG IN DETAIL

- Started in 2002
- Designed and Co-ordinated by CERN- European Organization for Nuclear Research [1]
- Technology initially proposed by Ian Foster and Carl Kesselman in 1999 [3]
- Incorporates both private fiber optic cable links and high-speed public Internet
- Data from 4 main LHC Experiments: ALICE, ATLAS, CMS and LHCb
- The primary configuration in the grid is based on Scientific Linux

A GRID FOR BIG DATA

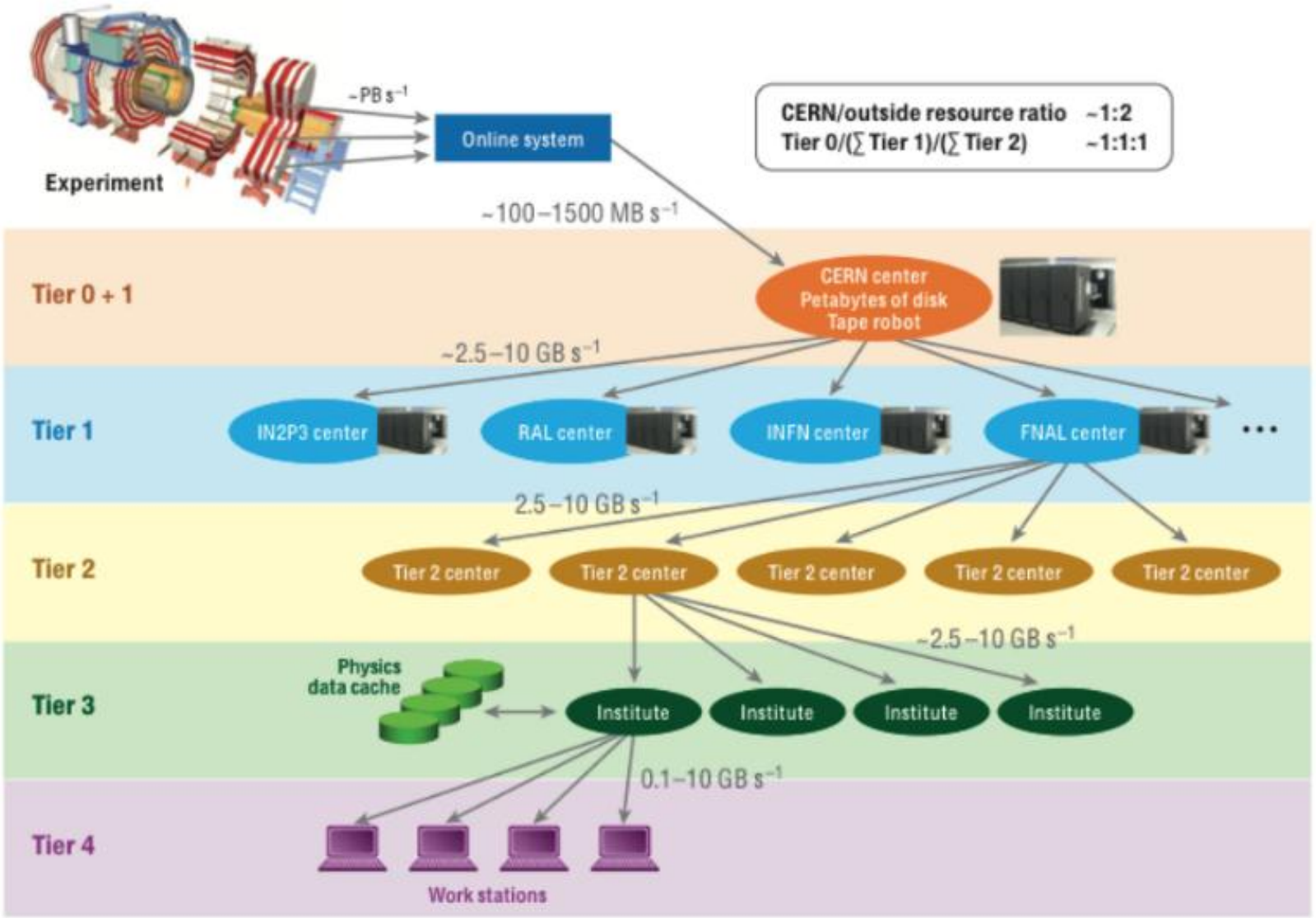
- 99% data is filtered out, still around 30 petabytes per year [3]
- Equivalent to nearly 9 million HD movies
- 1 billion collisions every second
- Peak of 6 gigabytes-per-second storage capacity
- Storage and Computing capability
- Already stored more than 100 petabytes from previous experiments



GRID ARCHITECTURE – 5 Tiers

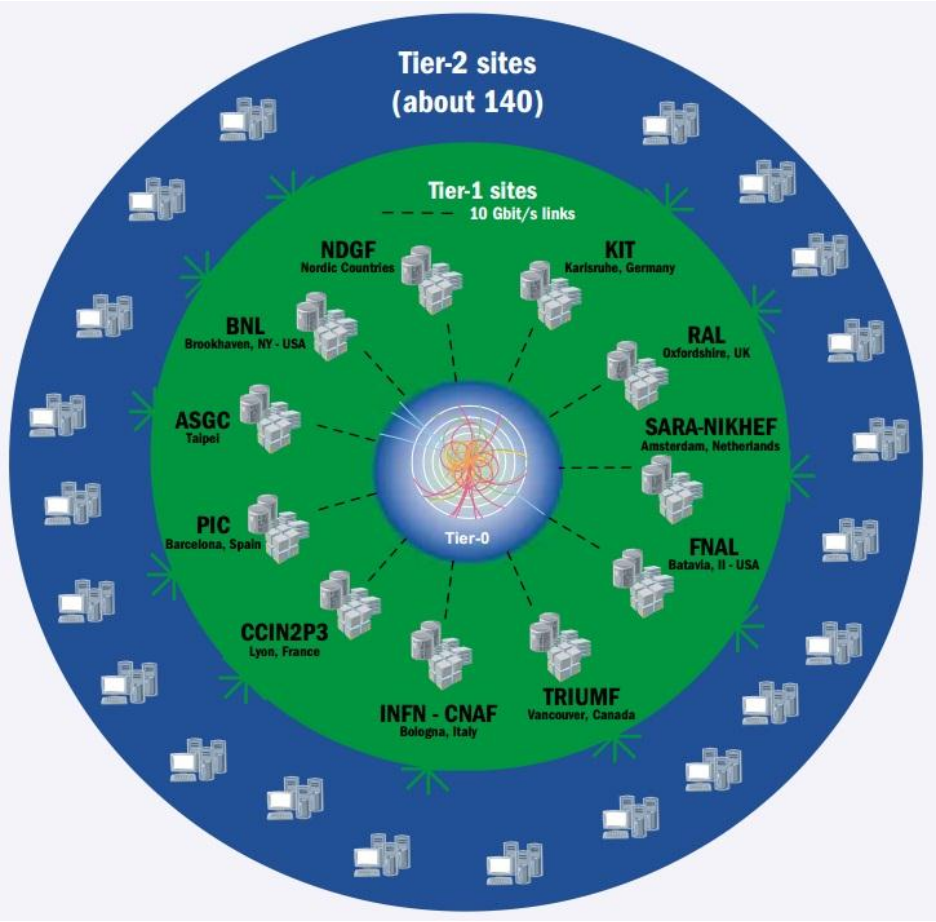
- Tier-0: CERN Datacenter in Geneva
  - Data Recording
  - Reconstruction
- Tier-1: 11 major Datacenters around the world, connected directly with the Datacenter; 10 Gbps communication
  - Data Storage
  - Reprocessing
- Tier -2: Connected with Tier-1, normally in the same geographical location of connecting Tier-1 site
  - Similation
  - Analysis
- Tier-3: Different institutes and Physics Labs
- Tier-4: End work stations

GRID ARCHITECTURE – 5 Tiers (2)



USING THE GRID

- With more than 8000 LHC physicists across the world
- 4 main experiments – ALICE, ATLAS, CMS and LHCb
- 3 new experiments - LHCf, MoEDAL and TOTEM from 2015
- Near real-time access and analysis
- Control Station and Monitoring
- Tier-0:
  - Data Recording from LHC
  - Reconstruction of the data
- Tier-1:
  - Data Storage facility
  - Reprocessing of the data
- Tier -2:
  - Simulation from the gathered data
  - Analysis of the data



DATA SECURITY OF THE GRID

- An AAA System with Public Key infrastructure
- WLCG participants are bound by a set of security policies, that are approved by the Management Board
- Top-level Grid Security Policy:
  - Grid Security Policy (Version 5.7a)
- For all Users:
  - Grid Acceptable Use Policy (Version 4.2a)
- For all Sites:
  - Grid Site Operations Policy (Version 1.4a)
  - Site Registration Security Policy (Version 3.2a)
- For all Virtual Organizations:
  - VO Security Policy

WHAT NOW?

- WLCG enabled the discovery of the Higgs-Boson July 2012 [2]
- First "Long Shutdown" (LS1):14 February 2013, restarts 05 April 2015
- LS1 was used for maintenance, repair and upgrade works
- Season-2 plans to double the collision energy!

TODAY IS THE SCHEDULED DATE for Launching 13 Tera Electonvolts (TeV)!! [5] 10:40 am, we have achieved it!

LIVE ON: http://run2-13tev.web.cern.ch/

WLCG WITH BIGGER DATA

- More energy, more collision, Bigger Data! [4]
- Magnetic tape named CERN Advanced Storage system (CASTOR)
- EOS disk pool system – a system for fast analysis access
- Concurrent user support in EOS
- Introduced a data 'chunking'
- Up to 10 gigabytes-per-sec
- 140 petabytes of storage; 99.5% availability!

REFERENCES

1. Andrey Shevel, "Cluster, Grid, Cloud computing systems – Introduction Lecture" , Seminar in ITMO for PERCCOM, 25 May 2015, St. Petersburg, Russia
2. WLCG, available at: <http://wlcg.web.cern.ch/> , accessed at: <02 June 2015>
3. WLCG, available at: < http://wlcg-public.web.cern.ch/ > , accessed at: <03 June 2015>
4. LHC Season 2: CERN computing ready for data torrent, available at: <http://home.web.cern.ch/about/updates/2015/06/lhc-season-2-cern-computing-ready-data-torrent>, accessed at: <02 June 2015>
5. LHC Season 2: New Frontier in Physics, Available at: <http://run2-13tev.web.cern.ch/>, accessed at: <02 June 2015>

Thank You