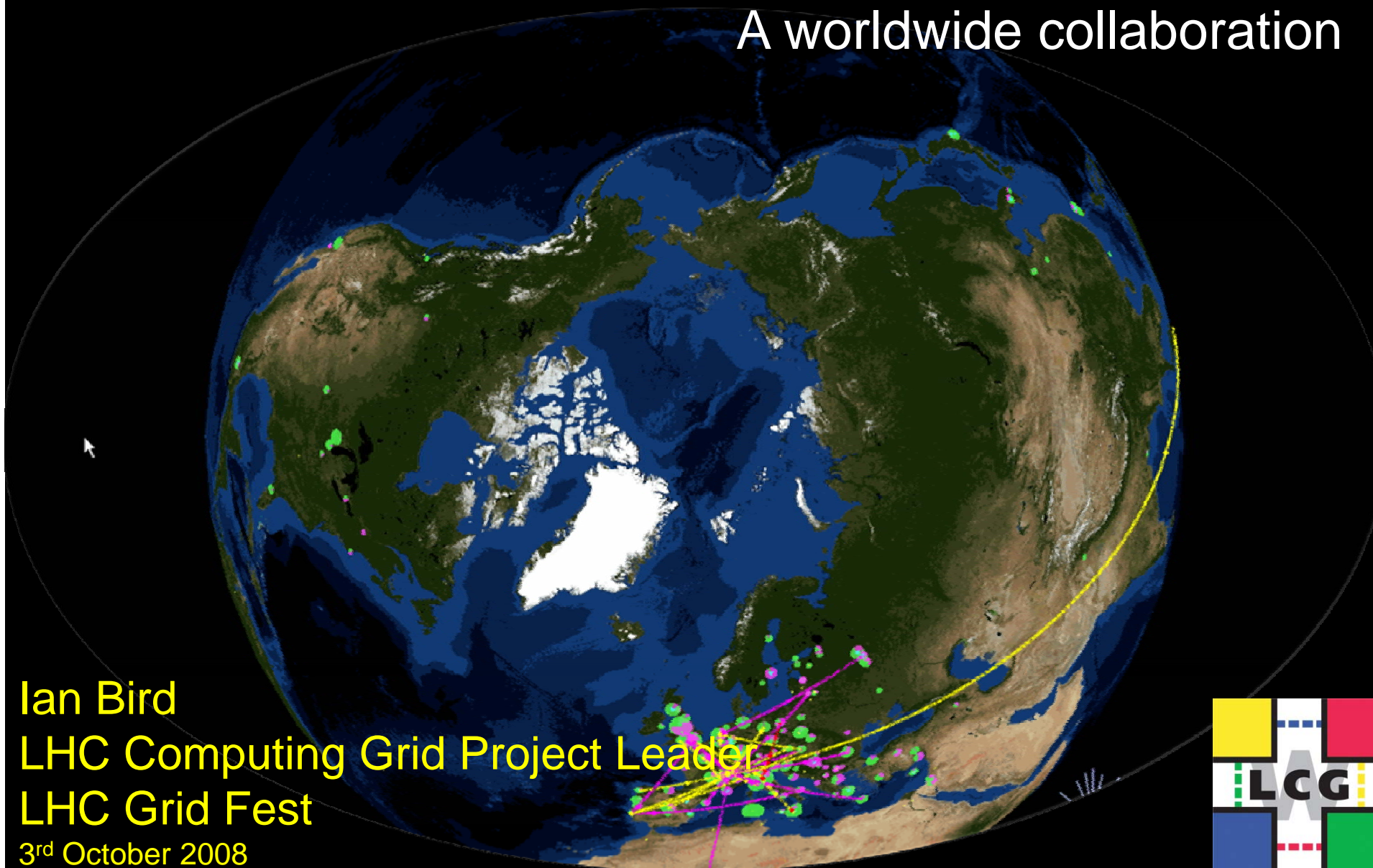




# THE LHC GRID SERVICE

A worldwide collaboration



Ian Bird  
LHC Computing Grid Project Leader  
LHC Grid Fest  
3<sup>rd</sup> October 2008





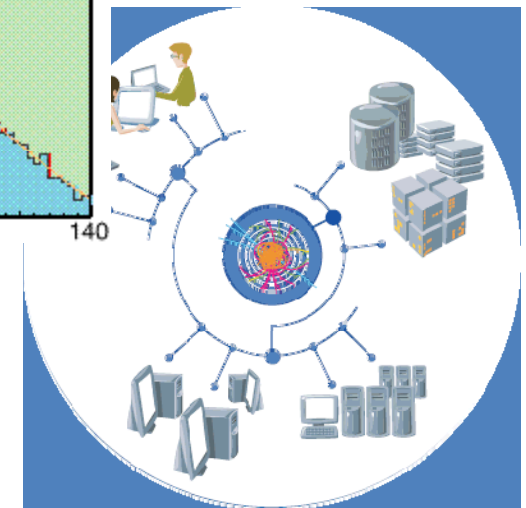
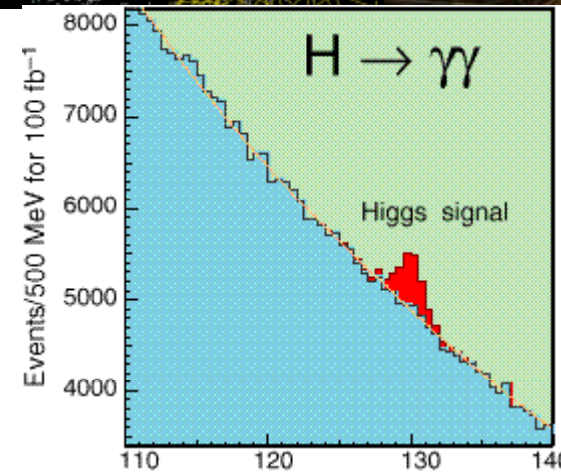
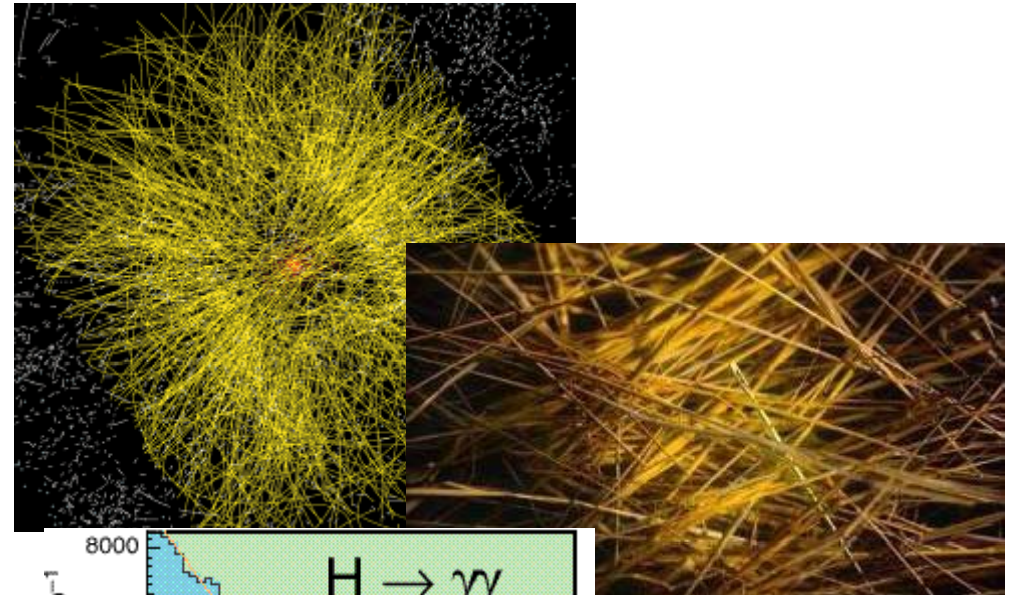
# Introduction

- The LHC Grid Service is a worldwide collaboration between:
  - 4 LHC experiments and
  - ~140 computer centres that contribute resources
  - International grid projects providing software and services
- The collaboration is brought together by a MoU that:
  - Commits resources for the coming years
  - Agrees a certain level of service availability and reliability
- As of today 33 countries have signed the MoU:
  - CERN (Tier 0) + 11 large Tier 1 sites
  - 130 Tier 2 sites in 60 “federations”
    - Other sites are expected to participate but without formal commitment



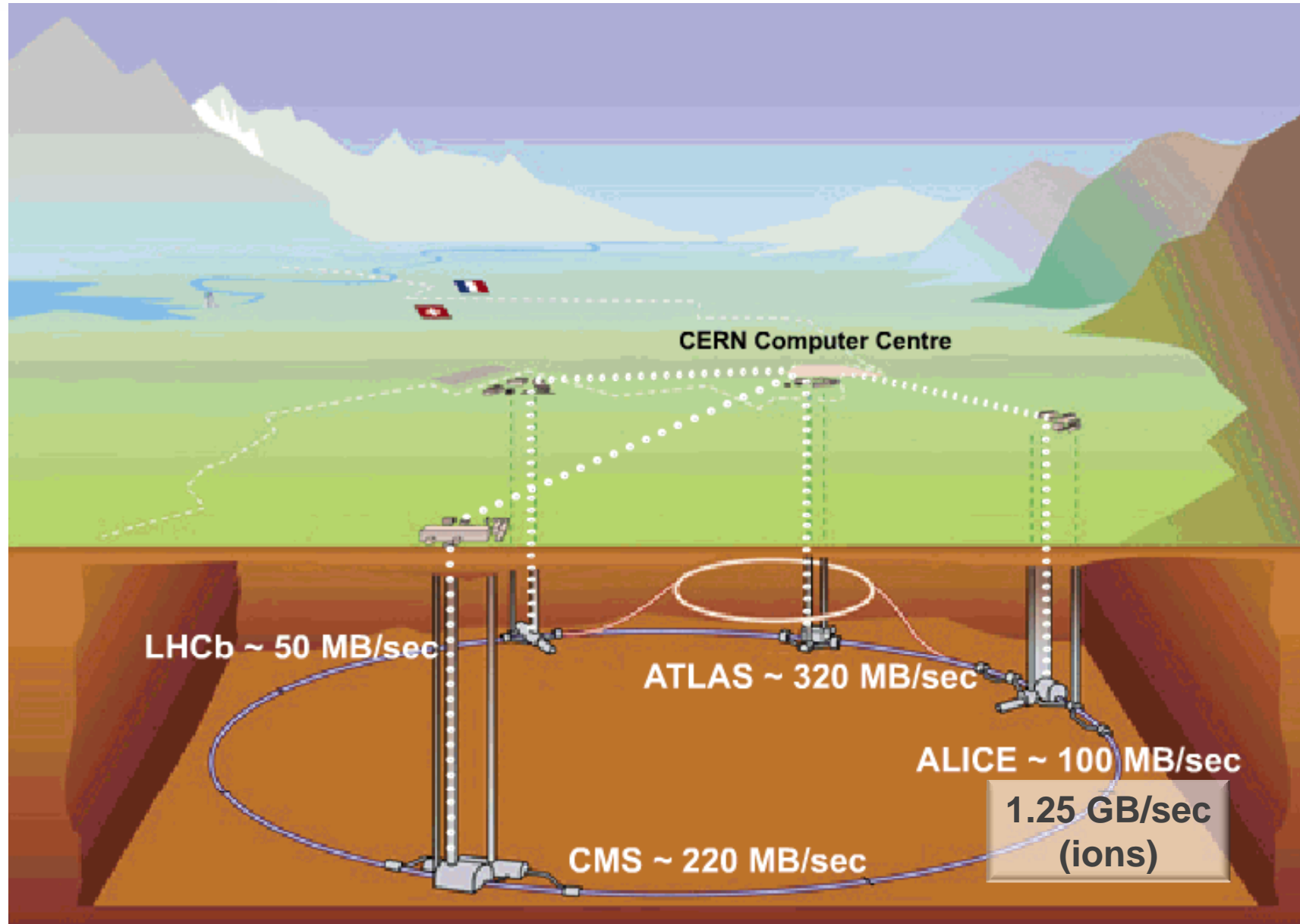
# The LHC Computing Challenge

- Signal/Noise:  $10^{-9}$
- Data volume
  - High rate \* large number of channels \* 4 experiments
  - **15 PetaBytes of new data each year**
- Compute power
  - Event complexity \* Nb. events \* thousands users
  - **100 k of (today's) fastest CPUs**
  - **45 PB of disk storage**
- Worldwide analysis & funding
  - Computing funding locally in major regions & countries
  - Efficient analysis everywhere
  - **GRID technology**

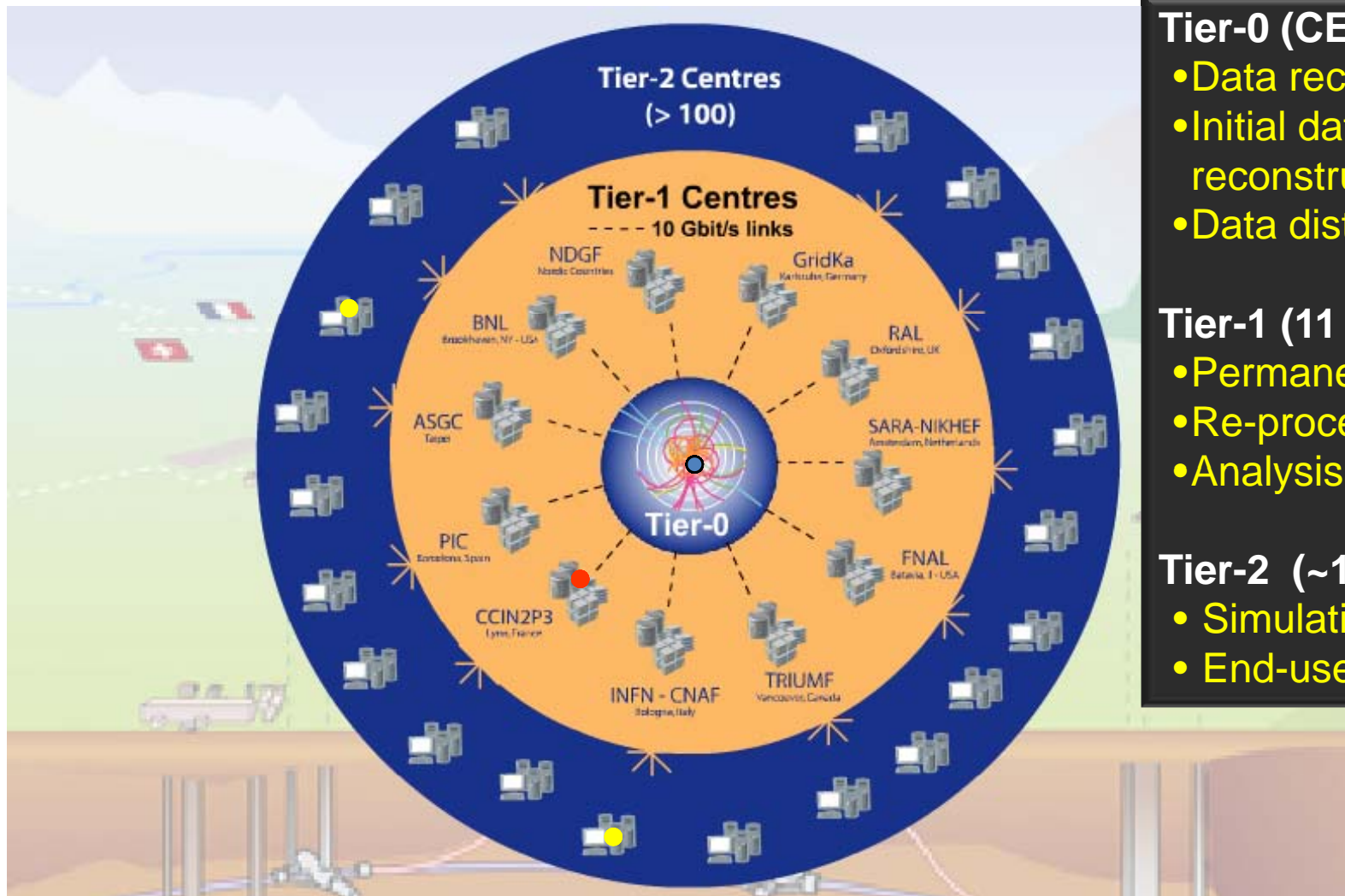




# Tier 0 at CERN: Acquisition, First pass processing Storage & Distribution



# Tier 0 – Tier 1 – Tier 2



## Tier-0 (CERN):

- Data recording
- Initial data reconstruction
- Data distribution

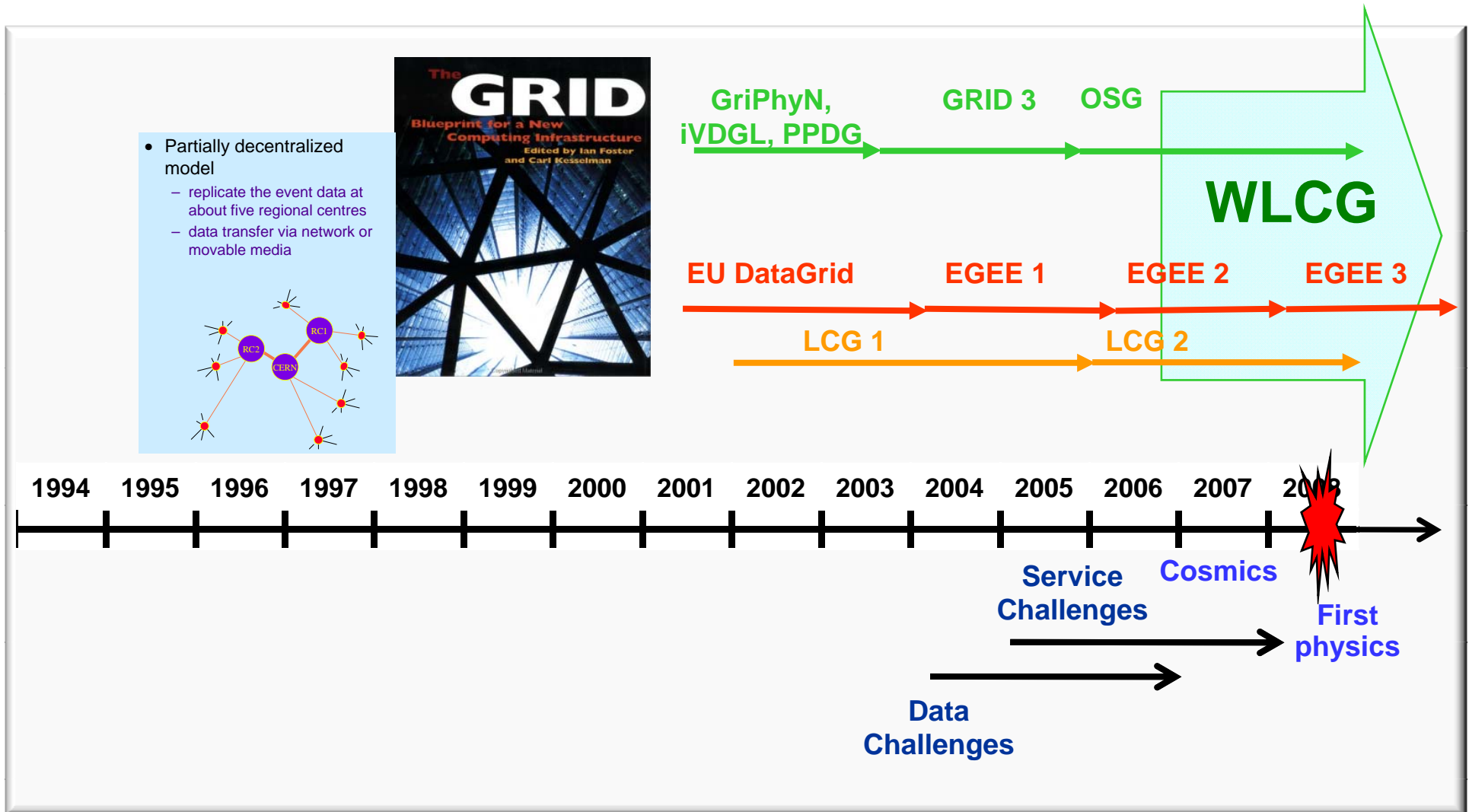
## Tier-1 (11 centres):

- Permanent storage
- Re-processing
- Analysis

## Tier-2 (~130 centres):

- Simulation
- End-user analysis

# Evolution of Grids



# Preparation for accelerator start up

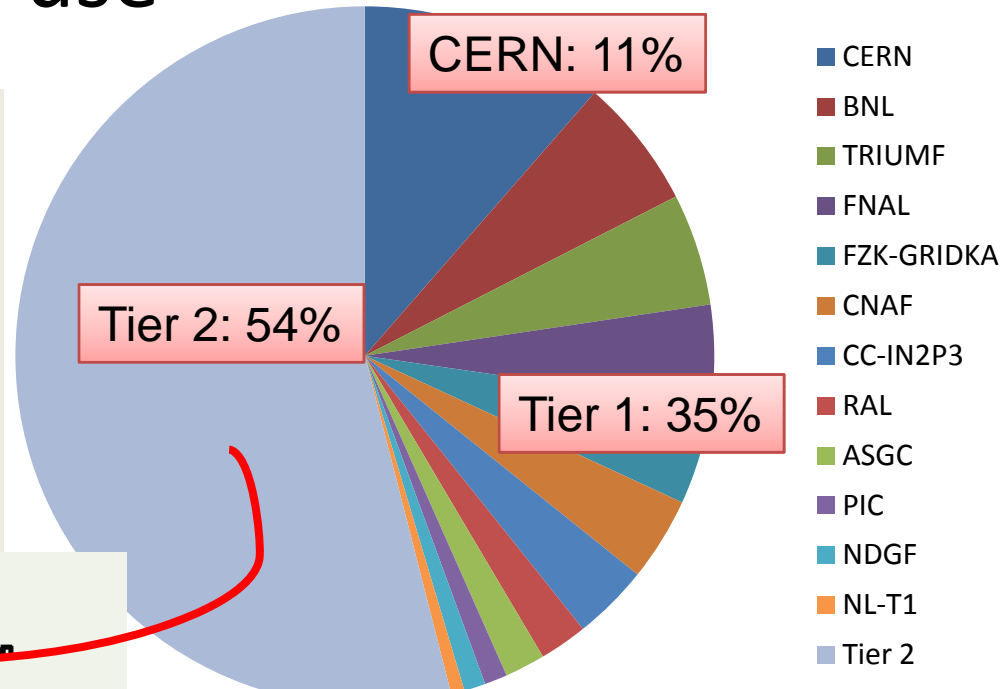
- Since 2004 WLCG has been running a series of challenges to demonstrate aspects of the system; with increasing targets for:
  - Data throughput
  - Workloads
  - Service availability and reliability
- Culminating in a 1 month challenge in May with
  - All 4 experiments running realistic work (simulating what will happen in data taking)
  - Demonstrated that we were ready for real data
- In essence the LHC Grid service has been running for several years



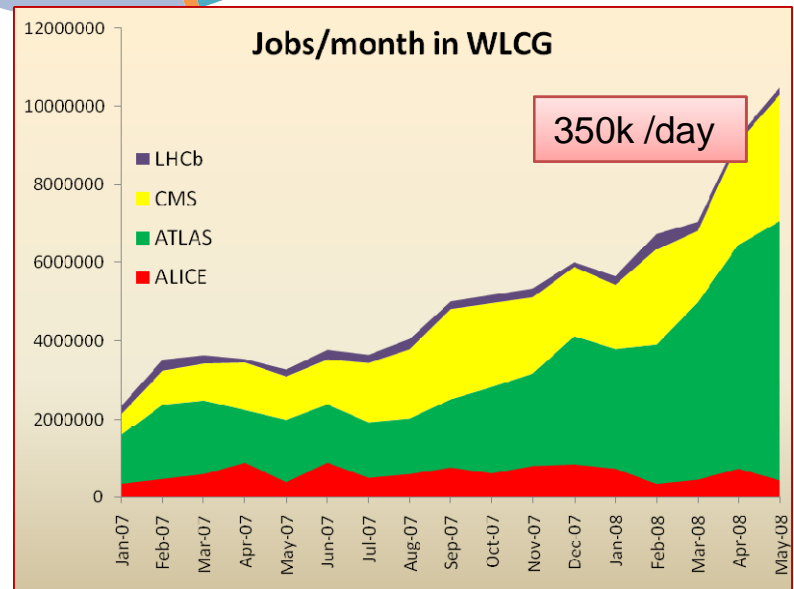
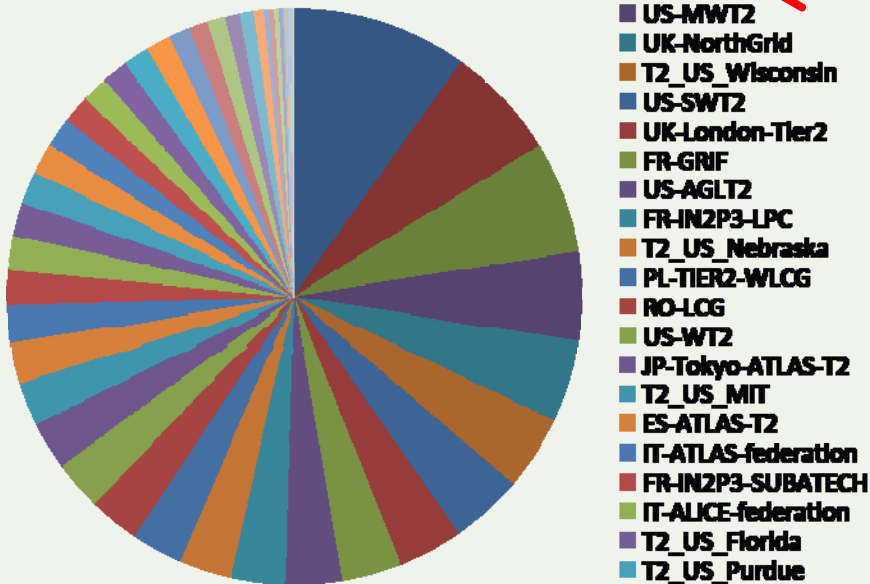
# Recent grid use

- The grid concept really works – all contributions – large & small contribute to the overall effort!

CPU Usage Early 2008

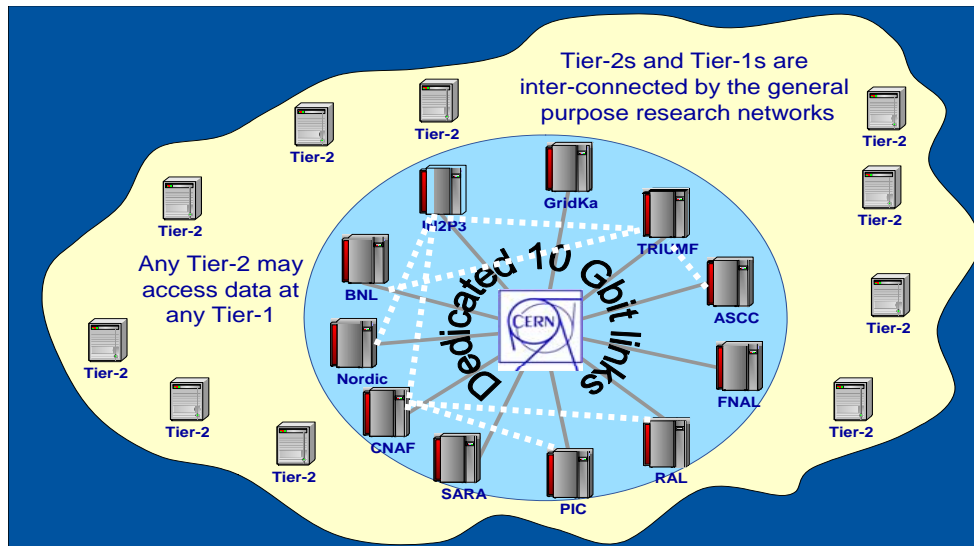


Tier 2 federation use





# Data transfer out of Tier 0



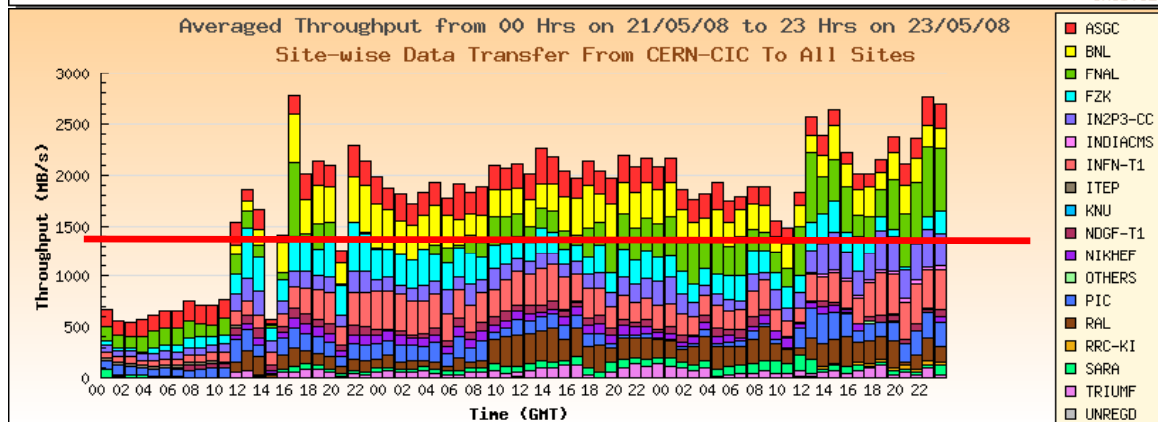
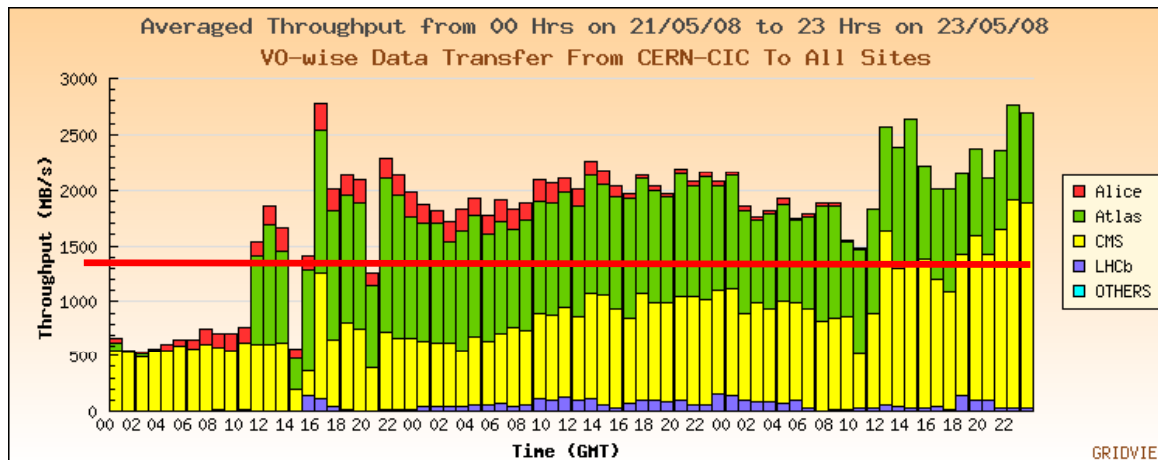
- Full experiment rate needed is 650 MB/s

- Desire capability to sustain twice that to allow for Tier 1 sites to shutdown and recover

- Have demonstrated far in excess of that

- All experiments exceeded required rates for extended periods, & simultaneously

- All Tier 1s achieved (or exceeded) their target acceptance rates





# Production Grids

- WLCG relies on a *production quality* infrastructure
  - Requires standards of:
    - Availability/reliability
    - Performance
    - Manageability
  - Will be used 365 days a year ... (has been for several years!)
  - Tier 1s must store the data for at least the lifetime of the LHC - ~20 years
    - Not passive – requires active migration to newer media
- Vital that we build a fault-tolerant and reliable system
  - That can deal with individual sites being down and recover





WLCG depends on two major science grid infrastructures ....

**EGEE** - Enabling Grids for E-Science

**OSG** - US Open Science Grid

... as well as many national grid projects



*A map of the worldwide LCG infrastructure operated by EGEE and OSG.*

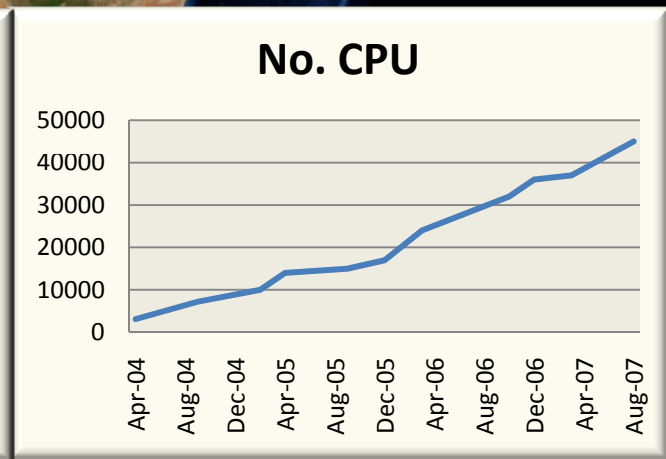
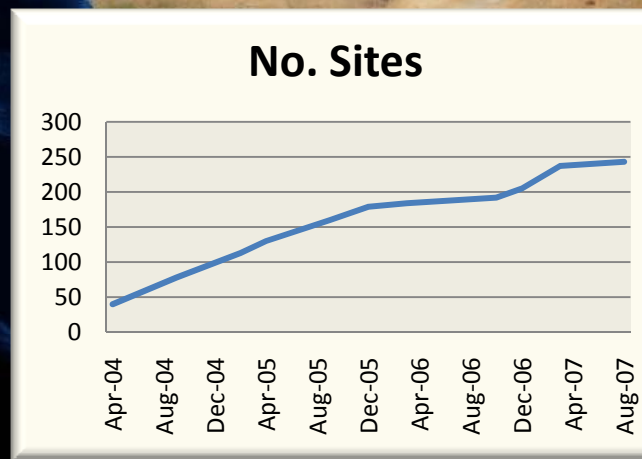
**Grid infrastructure project co-funded by the European Commission -  
now in 2<sup>nd</sup> phase with 91 partners in 32 countries**



**240 sites**  
**45 countries**  
**45,000 CPUs**  
**12 PetaBytes**  
**> 5000 users**  
**> 100 VOs**  
**> 100,000 jobs/day**

- Archeology**
- Astronomy**
- Astrophysics**
- Civil Protection**
- Comp. Chemistry**
- Earth Sciences**
- Finance**
- Fusion**
- Geophysics**
- High Energy Physics**
- Life Sciences**
- Multimedia**
- Material Sciences**

...

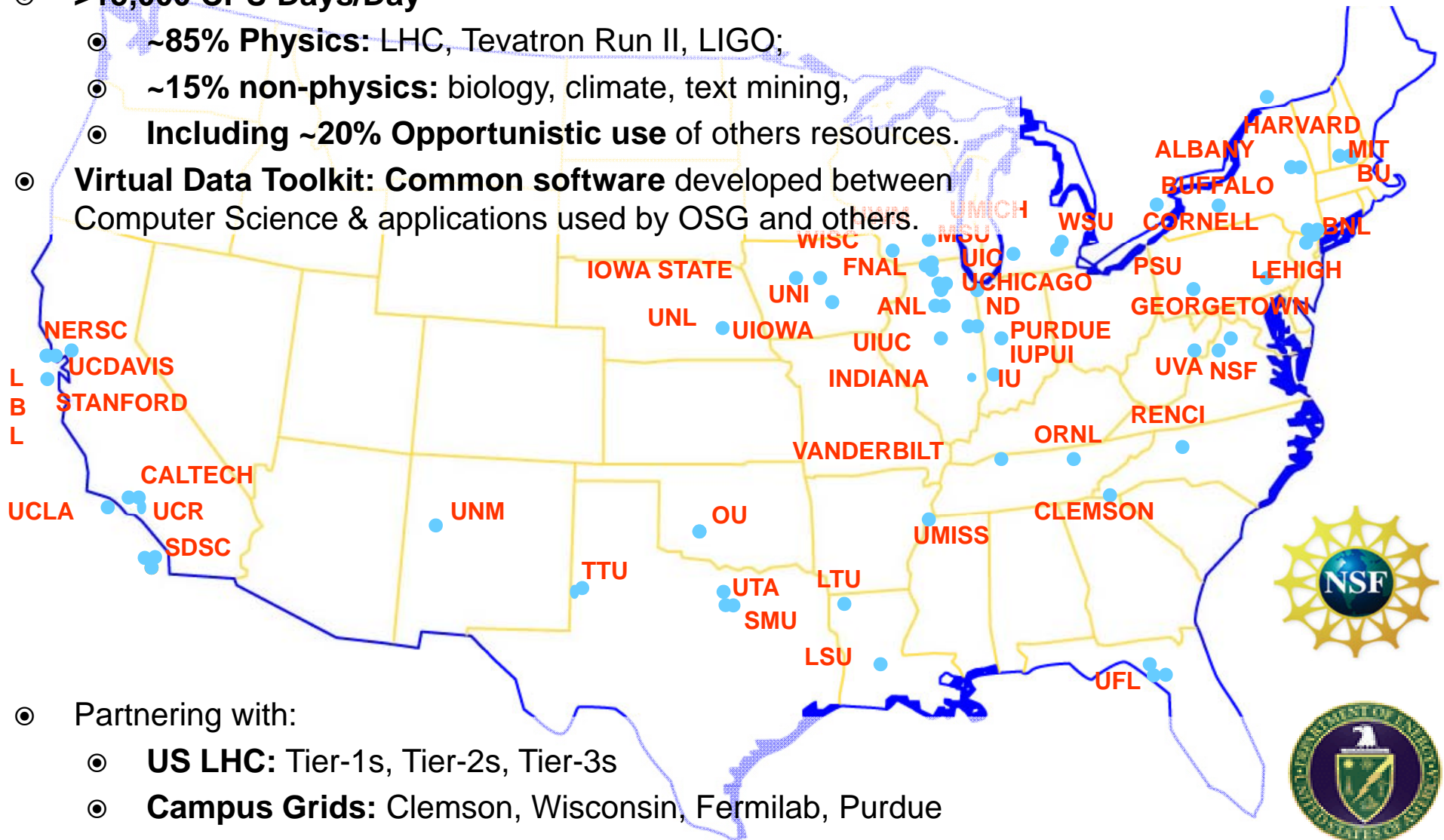


## OSG Project :

Supported by the Department of Energy & the National Science Foundation



- ◎ Access to 45,000 Cores, 6 Petabytes Disk, 15 Petabytes Tape
- ◎ >15,000 CPU Days/Day
  - ◎ ~85% Physics: LHC, Tevatron Run II, LIGO;
  - ◎ ~15% non-physics: biology, climate, text mining,
  - ◎ Including ~20% Opportunistic use of others resources.
- ◎ Virtual Data Toolkit: Common software developed between Computer Science & applications used by OSG and others.



- ◎ Partnering with:
  - ◎ US LHC: Tier-1s, Tier-2s, Tier-3s
  - ◎ Campus Grids: Clemson, Wisconsin, Fermilab, Purdue
  - ◎ Regional & National Grids: TeraGrid, New York State Grid, EGEE, UK NGS
  - ◎ International Collaboration: South America, Central America, Taiwan, Korea, UK.



# THE LHC GRID SERVICE

A worldwide collaboration

**Has been in production for several years**

**Is now being used for real data**

**Is ready to face the computing challenges  
as LHC gets up to full speed**

