

# Scientific HPC Infrastructures

## Introduction

Boris Pará<sup>1,2</sup>    Tomáš Rebok<sup>1,2</sup>    Zdeněk Šustr<sup>2</sup>

<sup>1</sup>Masaryk University

<sup>2</sup>CESNET

February 20, 2017



- ▶ Get a different perspective (academic/non-commercial)
- ▶ Introduce scientific computing environments
- ▶ Get an idea of what is available for researchers
- ▶ Briefly mention existing HPC infrastructures

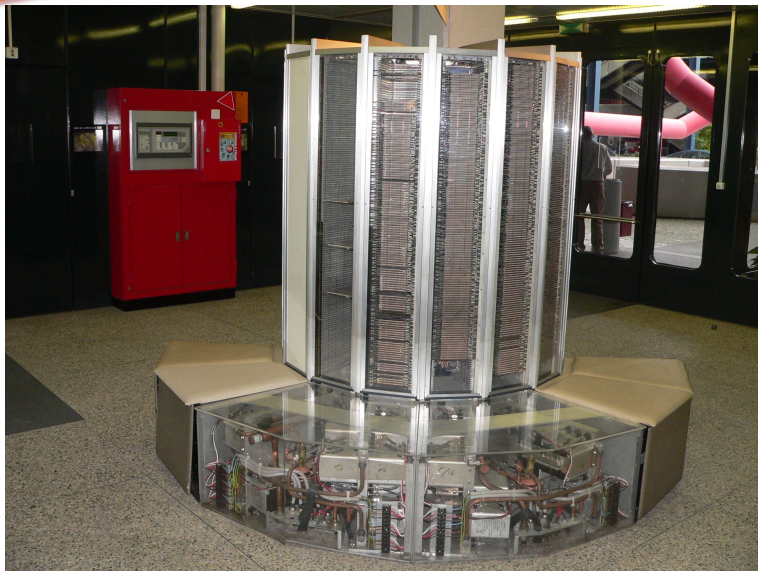
- ▶ For anyone doing any kind of research
- ▶ Widely ranging in scale and requirements
- ▶ Computing-intensive tasks
- ▶ Data-intensive tasks (*"Big Data"*)
- ▶ Reliably sharing open data & protecting sensitive data

- ▶ Computing with a particular set of requirements
- ▶ Working on demanding problems
  - ▶ performance (CPU, IO, bandwidth)
  - ▶ storage capacity (short-term, long-term)
  - ▶ scale (1k+ machines, distributed logically and geographically)
- ▶ Raw computing power and storage capacity, no “tricks”

- ▶ High-energy physics (Alice, Atlas, CMS, LHCb)
- ▶ Biology (genomics, protein analysis)
- ▶ Medicine (patient data analysis, drug development)
- ▶ Digital data preservation (truly long-term)
- ▶ Environmental sciences & Earth exploration
- ▶ Astronomy, . . .

# – HPC Infrastructures –

- ▶ Supercomputers (Cray, Deep Blue, ...)
- ▶ Local computing clusters (small scale)
- ▶ Grid (→ HEP) (large scale)
- ▶ *Cloud* (generic, structurally opaque)

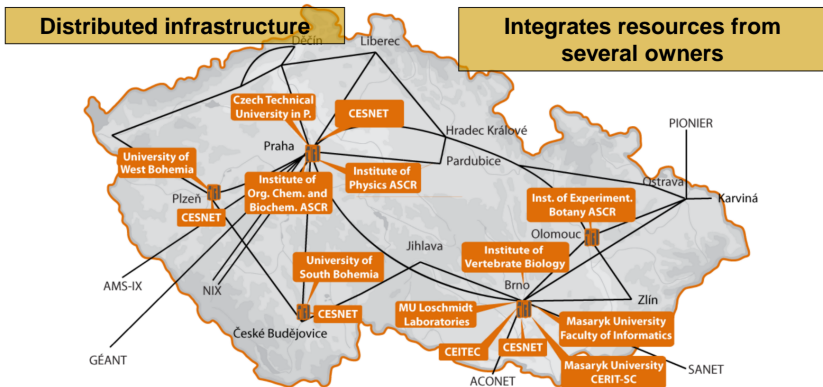








- ▶ CESNET
  - ▶ MetaCentrum (→ NGL\_CZ)
  - ▶ data, networking, identity management, research
  - ▶ see <http://www.metacentrum.cz/en/>
- ▶ CERIT-SC
  - ▶ transformation of Supercomputing Center Brno (SCB)
  - ▶ see <http://www.cerit-sc.cz/en/>
- ▶ IT4Innovations
  - ▶ National Supercomputing Center (academia + industry)
  - ▶ see <http://www.it4i.cz/?lang=en>



### **Computing services:**

- ▶ grid computing
- ▶ cloud computing
- ▶ Hadoop
- ▶ specialized environments

### **Storage services:**

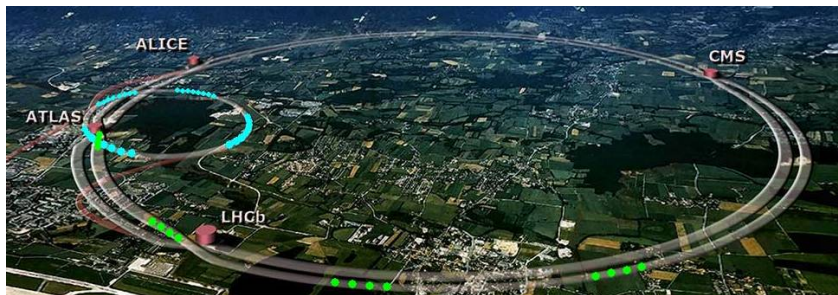
- ▶ NGI storage systems
- ▶ CESNET DU

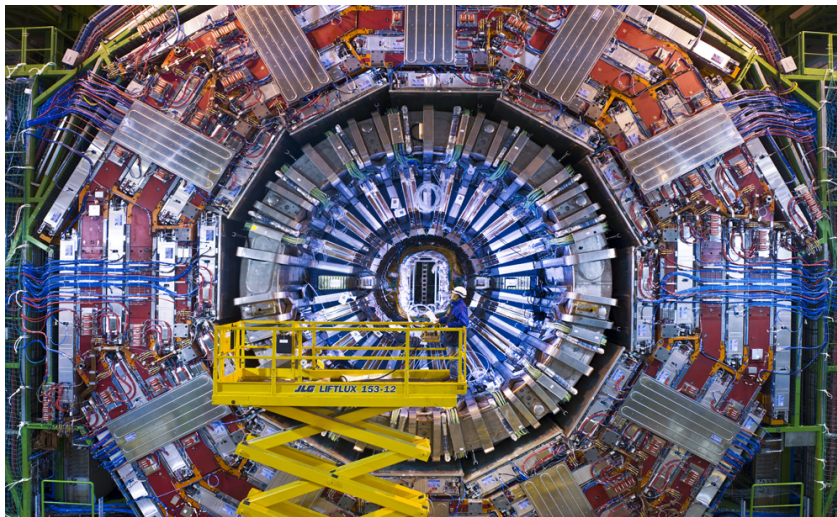
### **Collaboration & Support services:**

- ▶ interdisciplinary and infrastructure research (CERIT-SC)
- ▶ user support activities

- ▶ EGI (formerly known as *European Grid Infrastructure*)
- ▶ WLCG (*The Worldwide LHC Computing Grid*)
  - ▶ High-energy physics community (*HEP*)
  - ▶ CERN (*The European Organization for Nuclear Research*)
- ▶ PRACE Research Infrastructure









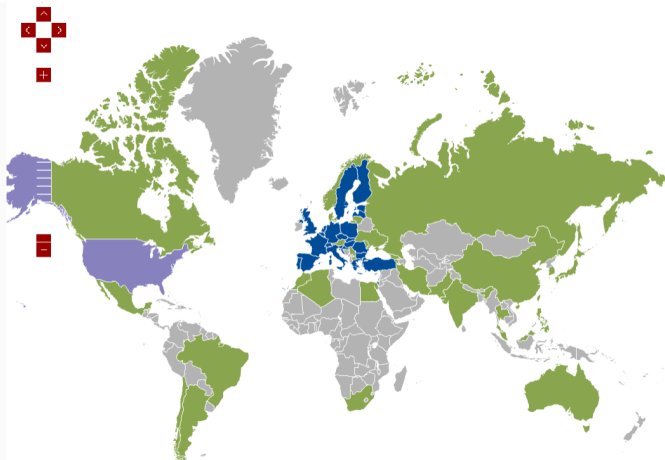


Data production rate (GB/s)	25
T0 Power capacity (MW)	3.5
T0 Logical CPUs (cores)	100,000
WLCG size (sites)	170+
WLCG Logical CPUs (cores)	627,964
WLCG Storage (GB)	333,002,607

EGI is a publicly-funded federation of over 300 data centers across the globe with:

- ▶ over 45,000 users from a wide range of fields
- ▶ access to over 650,000 logical CPUs and 500 PB of storage
- ▶ a wide range of services for compute, storage, and support

The federation is coordinated by the **EGI Foundation** created to govern on behalf of the participants of the **EGI Council**.



Caption (as of July 2015)  
 blue: EGI council members  
 green: integrated resource infrastructure providers  
 purple: peer resource infrastructure providers

- ▶ No longer just for high-energy physics science
- ▶ Bio-informatics, ecology, chemistry, geology, . . .
- ▶ So-called “The Long Tail of Science” user communities
- ▶ Scientific gateways (DIRAC, Chipster, COMPSs, Catania Science Gateway)

# Quick Quiz

– That's All Folks! –

...

Do you have any questions?

- ▶ ask **NOW!**
- ▶ ask directly at [parak@cesnet.cz](mailto:parak@cesnet.cz) or [sustr4@cesnet.cz](mailto:sustr4@cesnet.cz)
- ▶ send your questions to [cloud@metacentrum.cz](mailto:cloud@metacentrum.cz)