Karlsruhe Institute of Technology

LHC Computing Grid today

Did it work ?

Sept. 9th 2011, Günter Quast

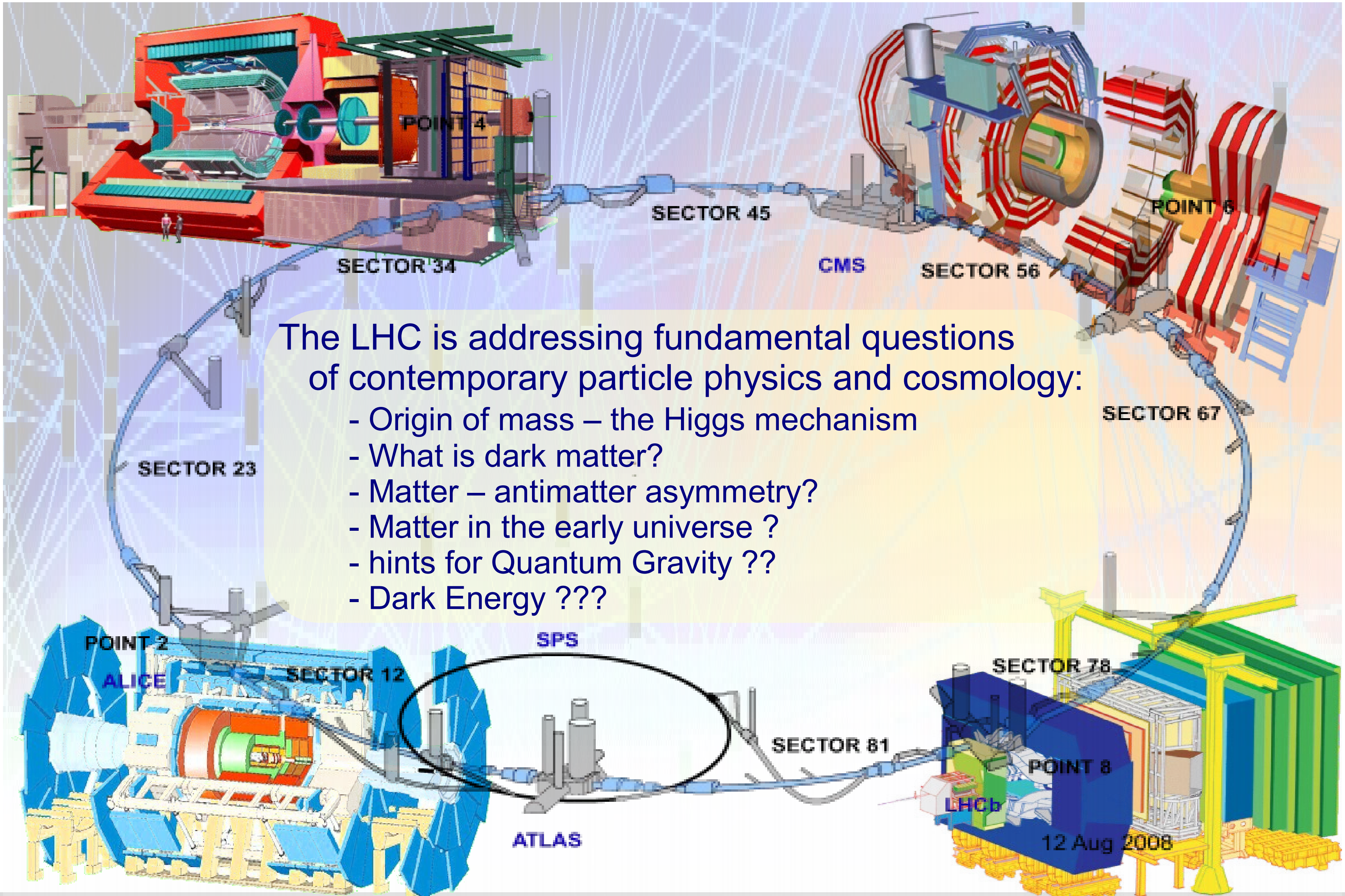
Institut für Experimentelle Kernphysik

GridKa School

9th International

GridKa School 2011

Large Hadron Collider and Experiments



The LHC is addressing fundamental questions of contemporary particle physics and cosmology:

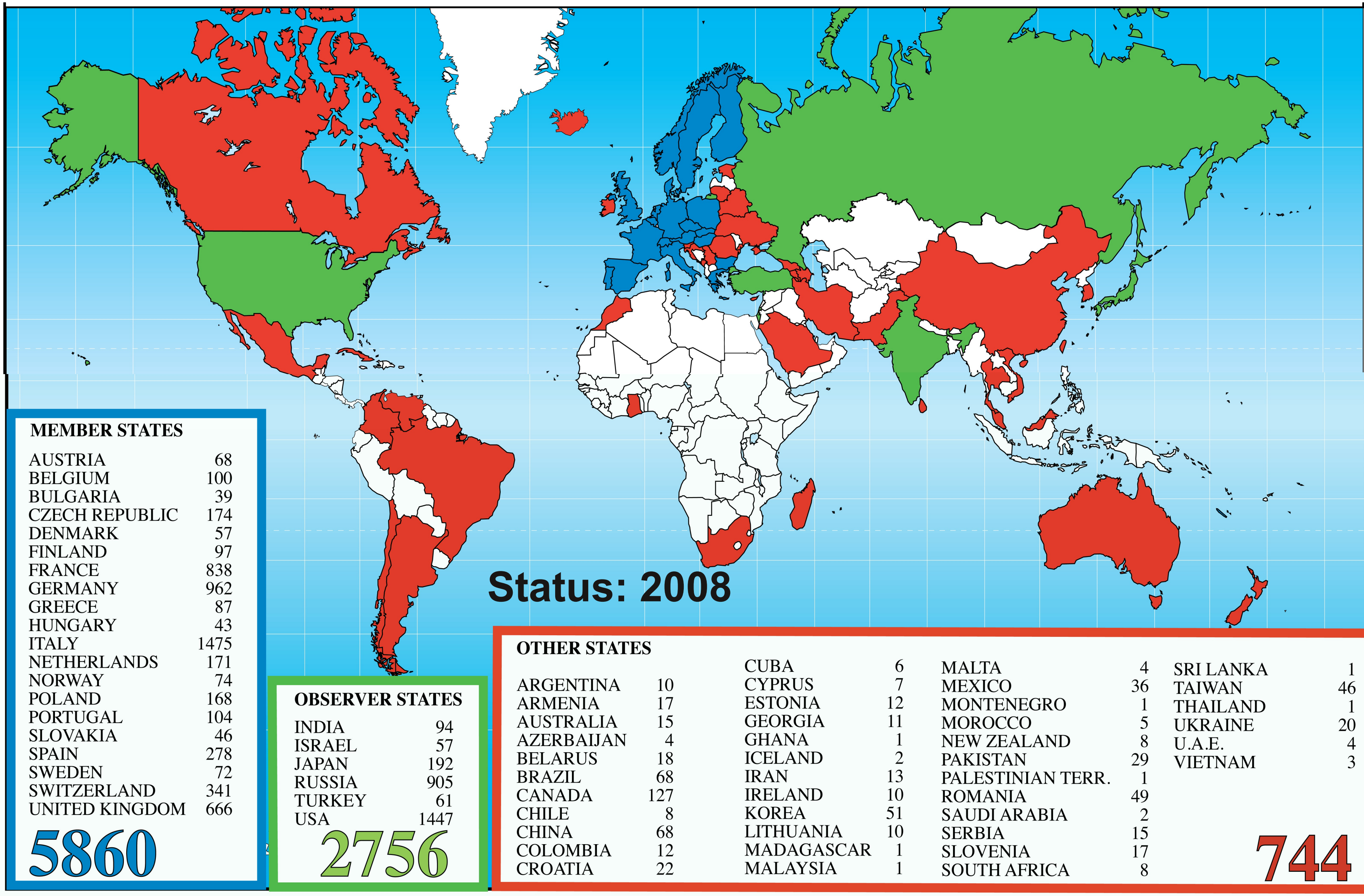
- Origin of mass – the Higgs mechanism
- What is dark matter?
- Matter – antimatter asymmetry?
- Matter in the early universe ?
- hints for Quantum Gravity ??
- Dark Energy ???

Particle Physics is international teamwork

Working @Cern: 290 institutes fromEurope, ~ 6349 Users
318 institutes elsewhere, ~ 3766 Users

Particle Physics international

Karlsruhe Institute of Technology



MEMBER STATES

AUSTRIA	68
BELGIUM	100
BULGARIA	39
CZECH REPUBLIC	174
DENMARK	57
FINLAND	97
FRANCE	838
GERMANY	962
GREECE	87
HUNGARY	43
ITALY	1475
NETHERLANDS	171
NORWAY	74
POLAND	168
PORTUGAL	104
SLOVAKIA	46
SPAIN	278
SWEDEN	72
SWITZERLAND	341
UNITED KINGDOM	666

5860

OBSERVER STATES

INDIA	94
ISRAEL	57
JAPAN	192
RUSSIA	905
TURKEY	61
USA	1447

2756

OTHER STATES

ARGENTINA	10
ARMENIA	17
AUSTRALIA	15
AZERBAIJAN	4
BELARUS	18
BRAZIL	68
CANADA	127
CHILE	8
CHINA	68
COLOMBIA	12
CROATIA	22
CUBA	6
CYPRUS	7
ESTONIA	12
GEORGIA	11
GHANA	1
ICELAND	2
IRAN	13
IRELAND	10
KOREA	51
LITHUANIA	10
MADAGASCAR	1
MALAYSIA	1
MALTA	4
MEXICO	36
MONTENEGRO	1
MOROCCO	5
NEW ZEALAND	8
PAKISTAN	29
PALESTINIAN TERR.	1
ROMANIA	49
SAUDI ARABIA	2
SERBIA	15
SLOVENIA	17
SOUTH AFRICA	8
SRI LANKA	1
TAIWAN	46
THAILAND	1
UKRAINE	20
U.A.E.	4
VIETNAM	3

744

Grid for LHC

- Given the international and collaborative nature of HEP
  Computing must be distributed
  → harvest intellectual contributions from all partners,
  also funding issues

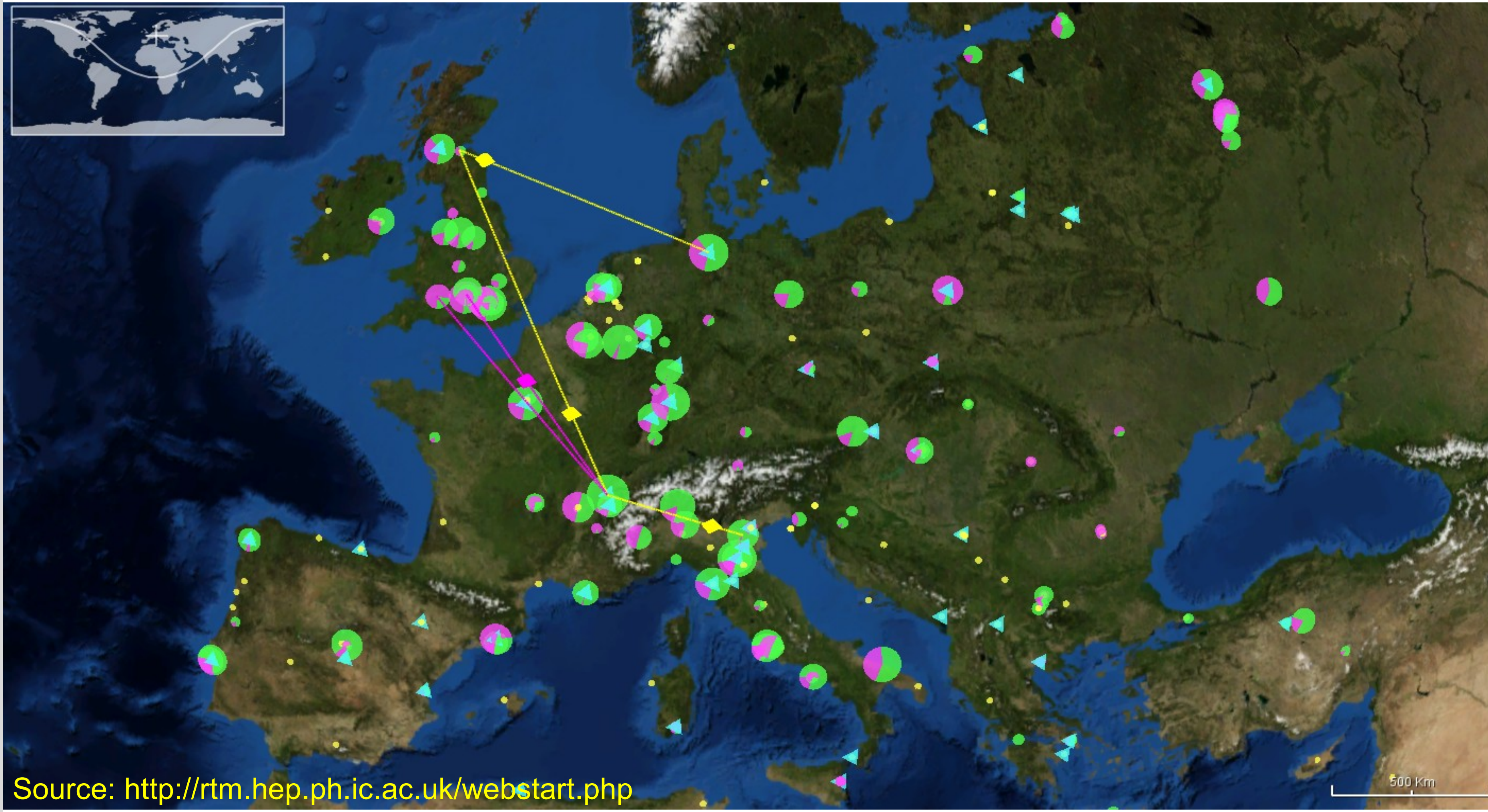
- Early studies in 1999 (Monarc Study group) suggested
  a hierarcical approach, following the typical data reduction schemes ususally adopted in data analysis in high enery physics

- Grid paradigm came at the right time and was adopted by LHC physicists as the base line for distibuted computing

- Major contributions by physicists to developments in Grid computing

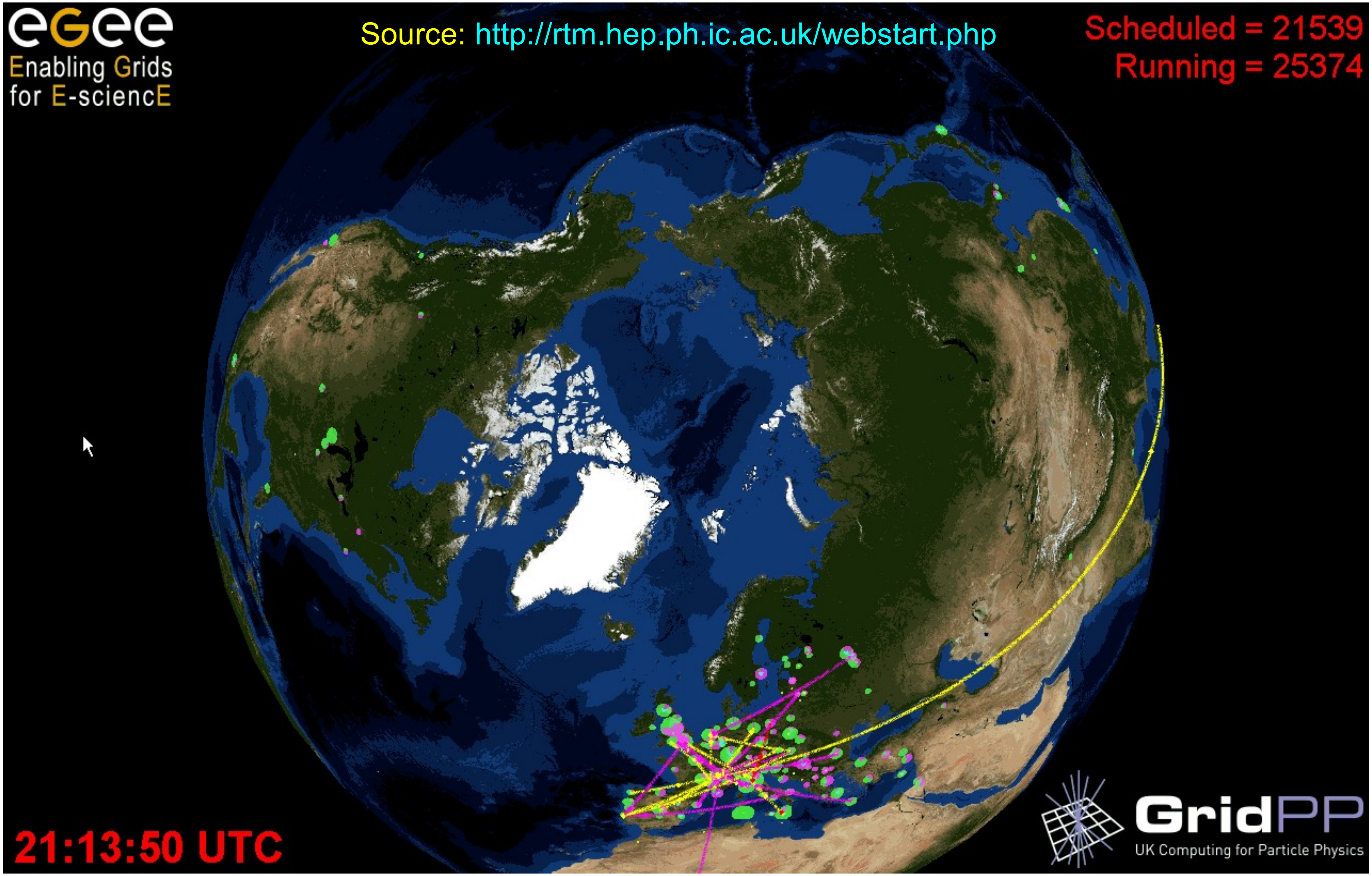
- Other HEP communities also benefit and contributed

WLHC Computing Grid in action



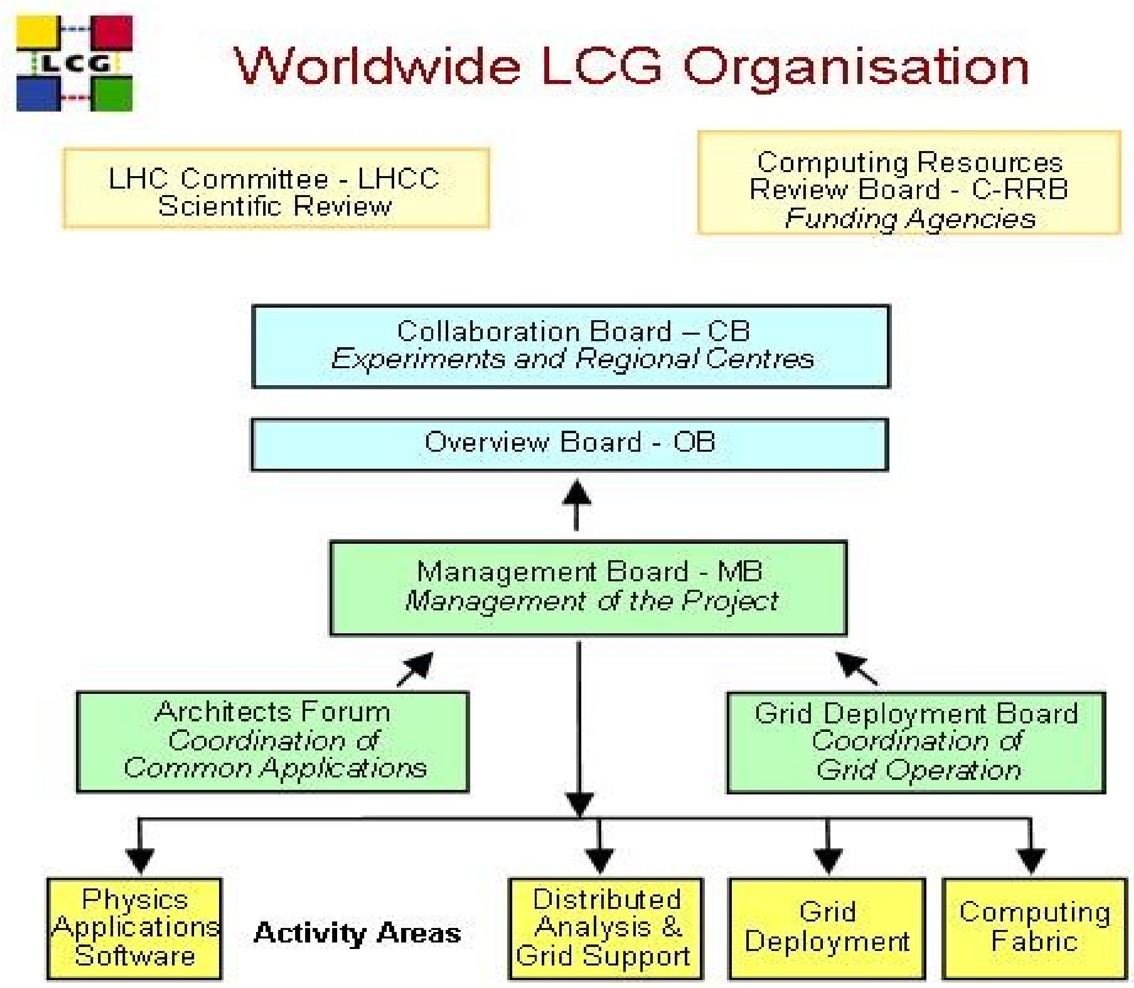
A truly international, world-spanning Grid for LHC data processing, simulation and analysis

WLHC Computing Grid in action



A truly international, world-spanning Grid for LHC data processing, simulation and analysis

Organisation of the World-wide LHC computing Grid



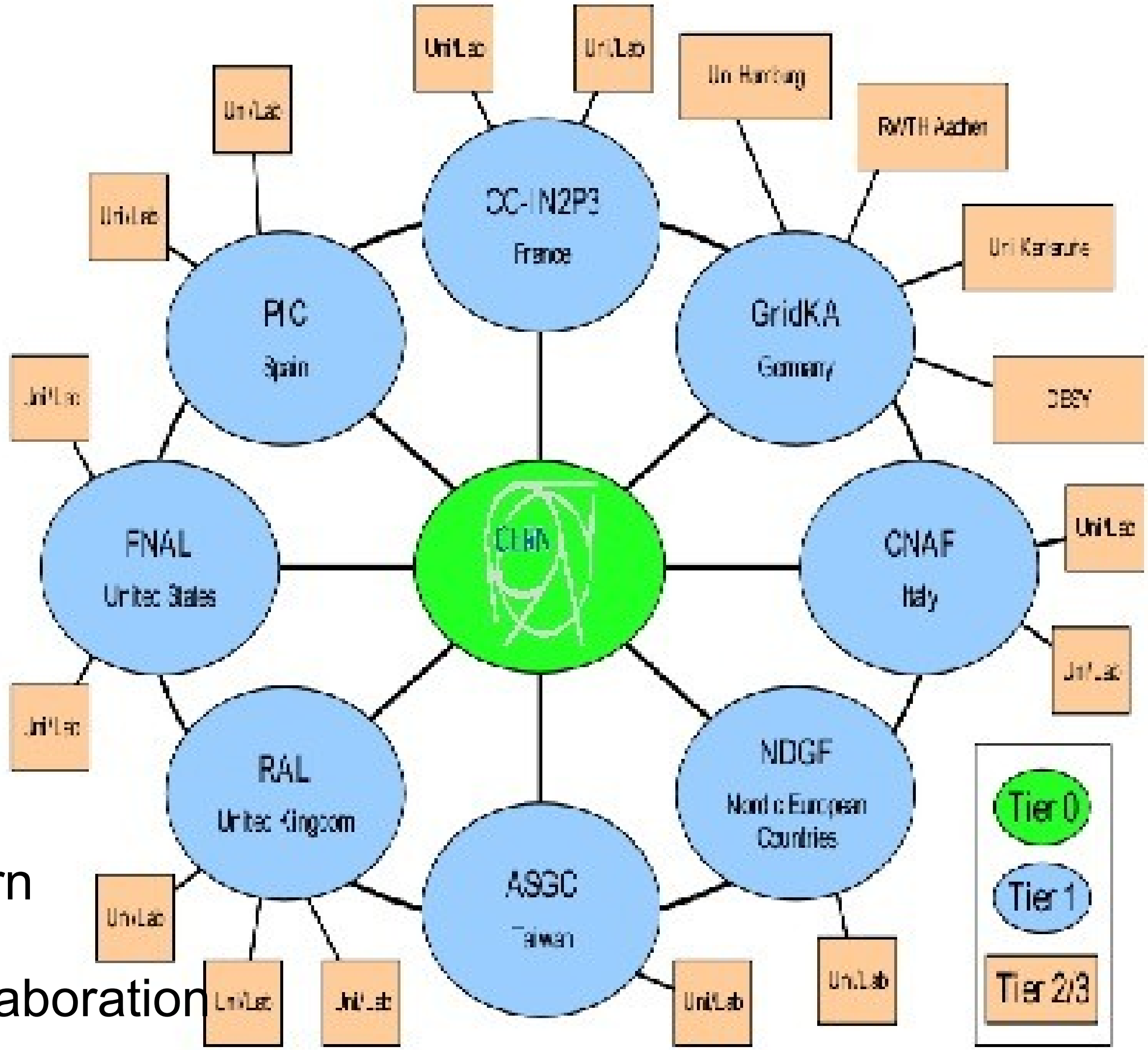
Grids can`t work without an organisational structure representing all parties involved

Structure of the LHC Grid

A grid with hierarchies
and different tasks
at different levels

In addition, it is a **“Grid of Grids”** with interoperability between different middlewares:

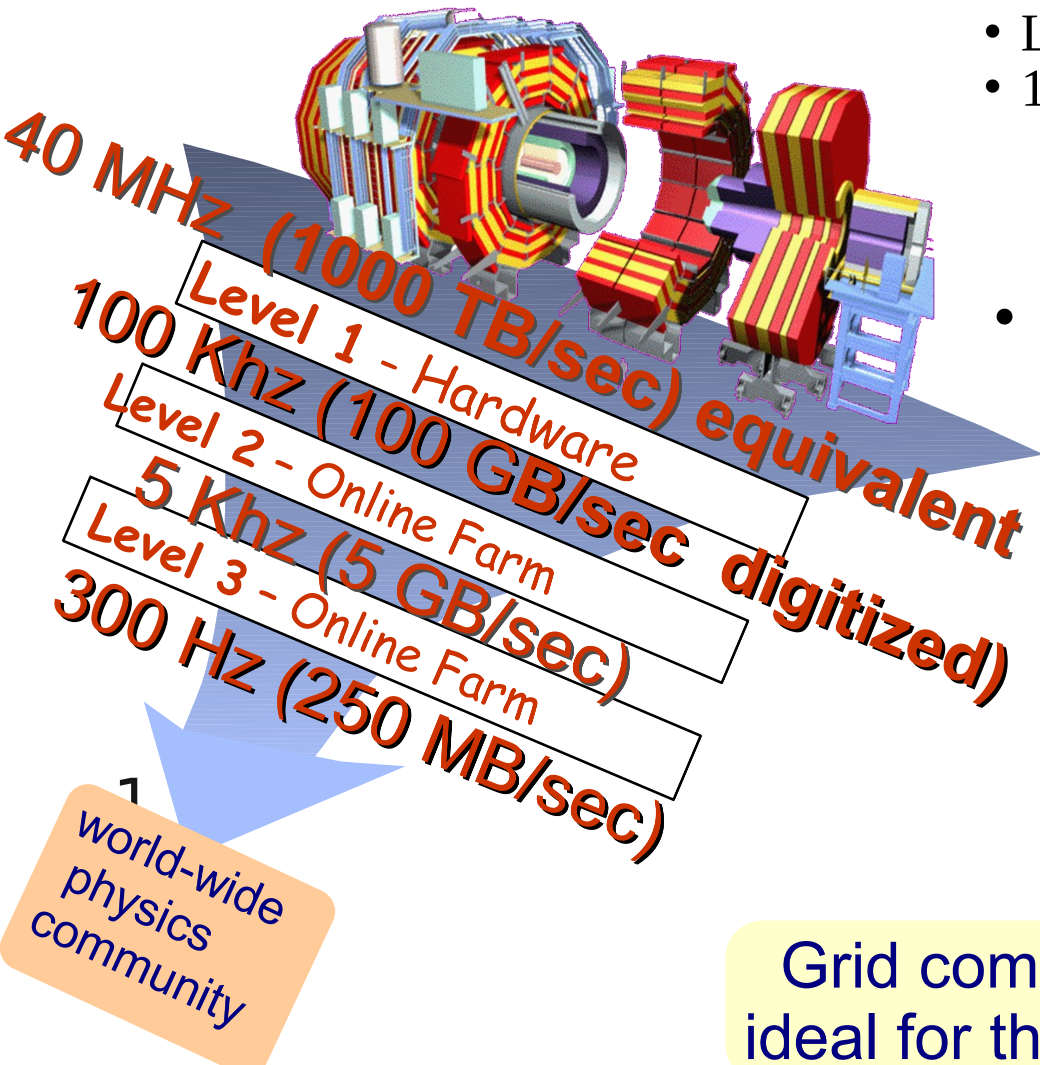
- EGEE middleware in most of Europe
- Open Science Grid in USA
- NorduGrid in Northern Europe
- Alien by the Alice collaboration



LHC Experiments are huge data sources

- ~ 100 Millionen detector cells
- LHC collision rate: 40 MHz
- 10-12 bit/cell

→ ~1000 Tbyte/s raw data

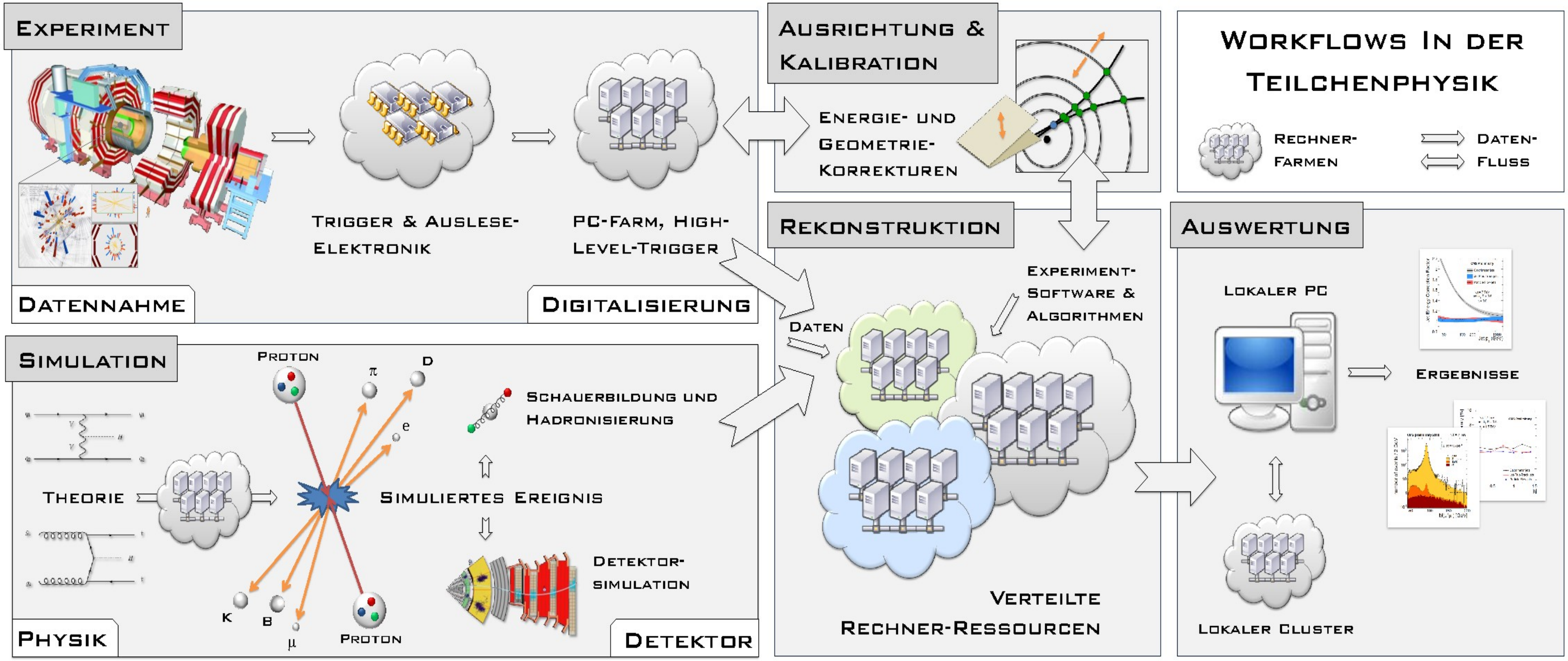


- Zero-suppression and Trigger reduce this to „only" some 100 Mbyte/s

i.e. 1  /sec

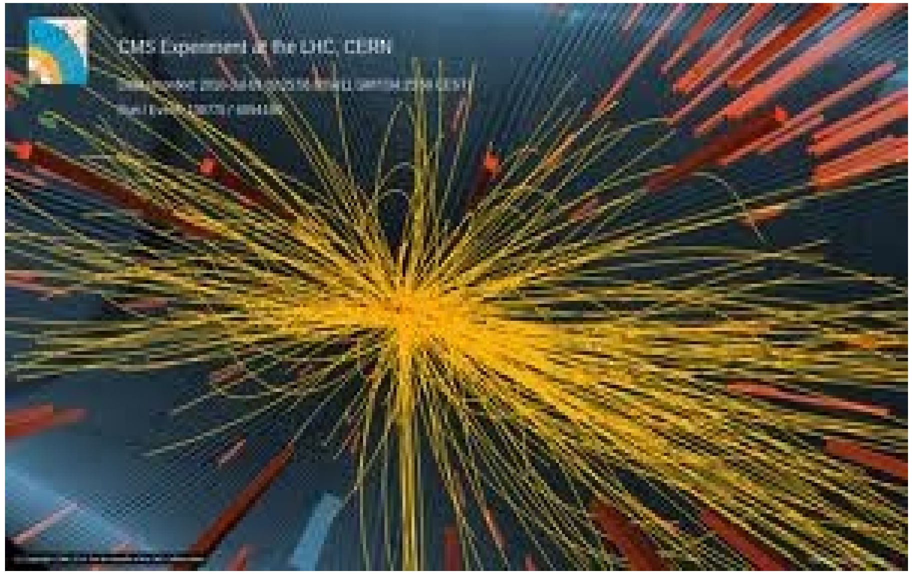
Grid computing ideal for the job !

What Particle Physicists do on the Grid



- CPU-intense simulation of particle physics reactions and detector response
- processing (=reconstrucion) of large data volumes
- I/O-intense filtering and distribution of data
- transfer to local clusters and workstations for final physics interpretation

Why Grid is well suited for HEP



Experimental HEP codes - key characteristics:

- modest memory requirement (~2GB) & modest floating point
  → **perform well on PCs**
- independent events
  → **easy parallelism**
- large data collections (TB → PB)
- shared by very large collaborations

WLCG: Grid with hierarcical structure

Tier-0 the accelerator centre

- Data acquisition & initial processing
- Long-term data curation
- Distribution of data to T1/T2

Tier-1

11 Tier-1 Centres

- "online" to the data acquisition process → high availability
- Managed Mass Storage → grid-enabled data service
- Data-intensive analysis
- National, regional support

Tier-2 150 Centres in 60 Federations in 35 countries

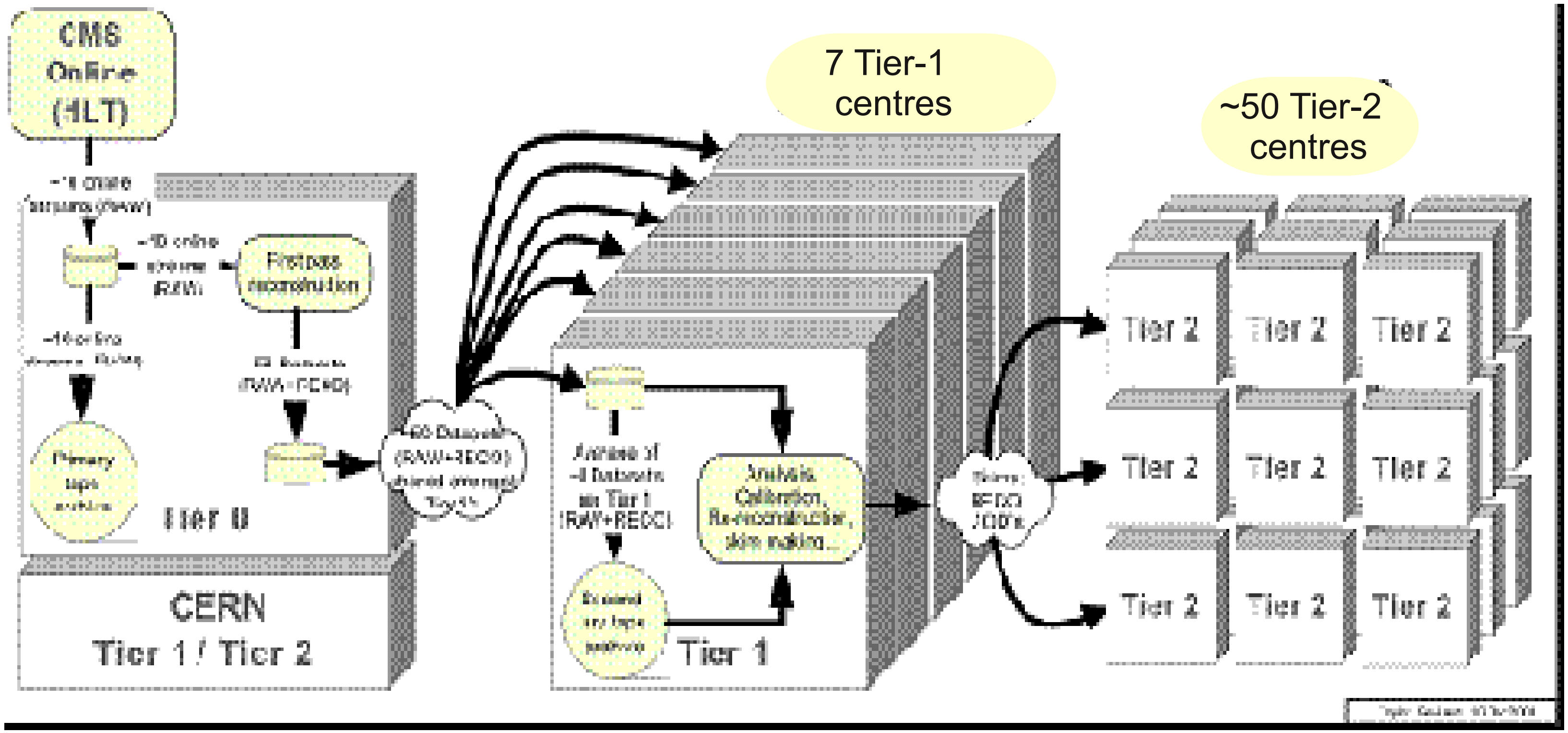
- **End-user (physicist, research group) analysis** & collaboration with T3 (= institute recources) **– where the discoveries are made**
- Monte Carlo Simulation

Tier-3

several 100 grid-enabled PC clusters @ institutes

Grid Structure

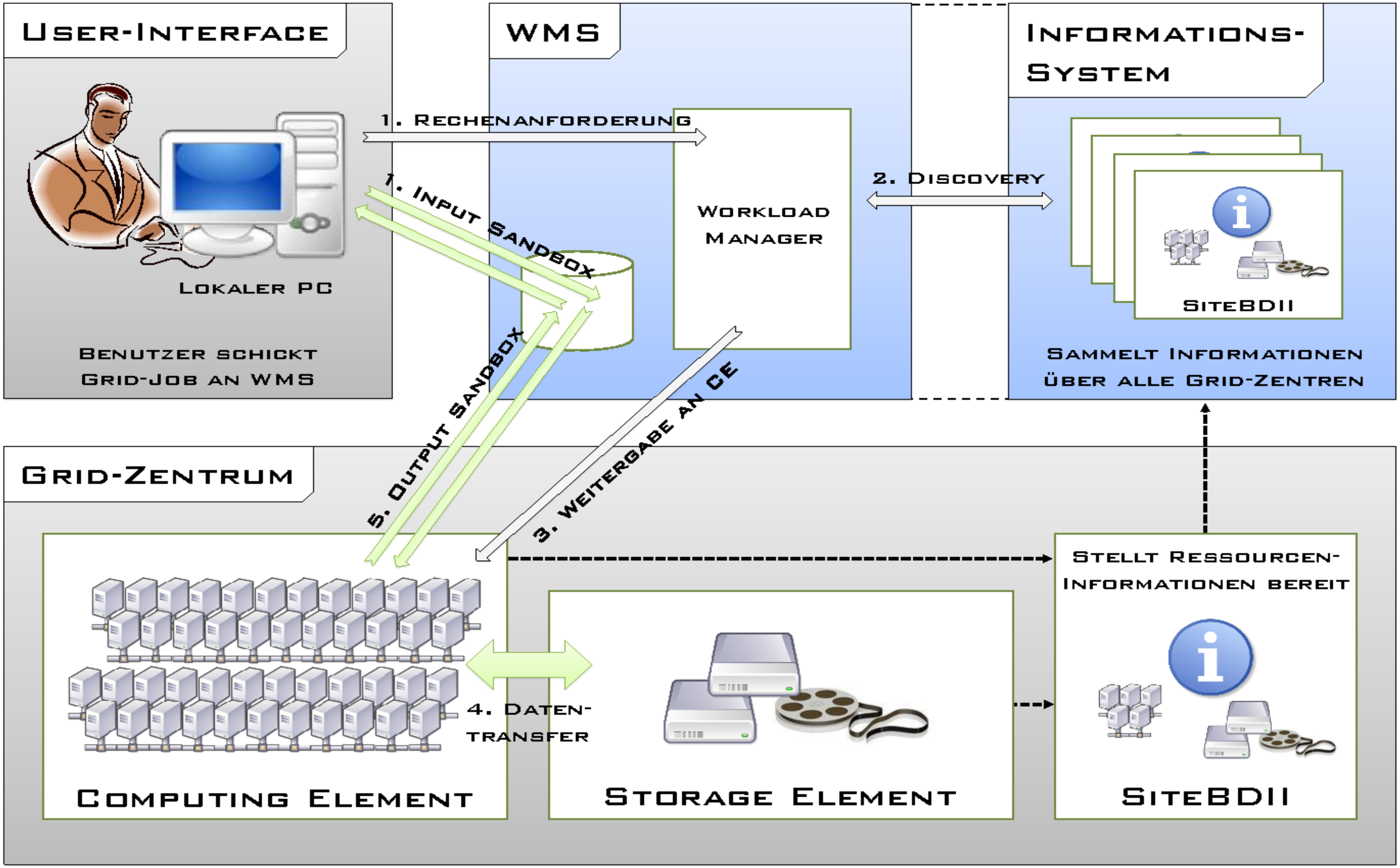
Example: CMS computing model



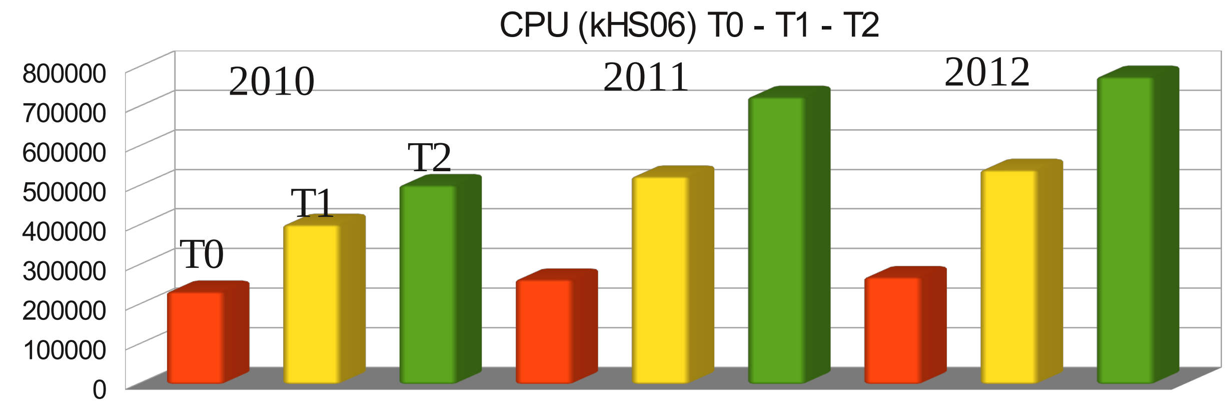
LHC-Experiments

- typically share big Tier-1s, take responsibility for experiment-specific services
- have a large number of Tier2s, usually supporting only one experiment
- have an even larger number of Tier-3s without any obligations towards WLCG

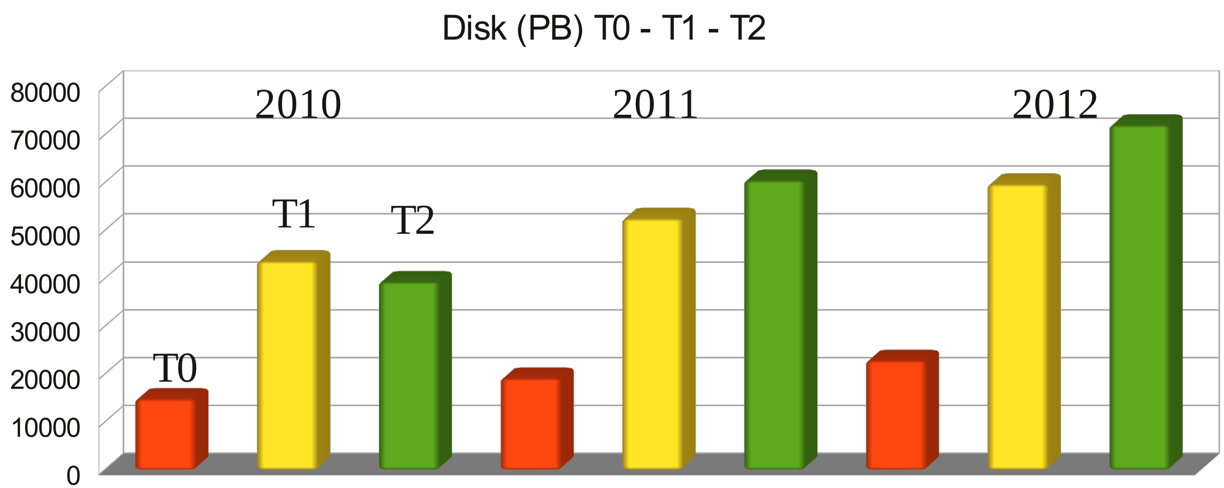
Typical workflows on the Grid



How big is WLCG today ?



Total CPU 2011:
1500 kHS06
approx. equiv.
150'000 CPU cores



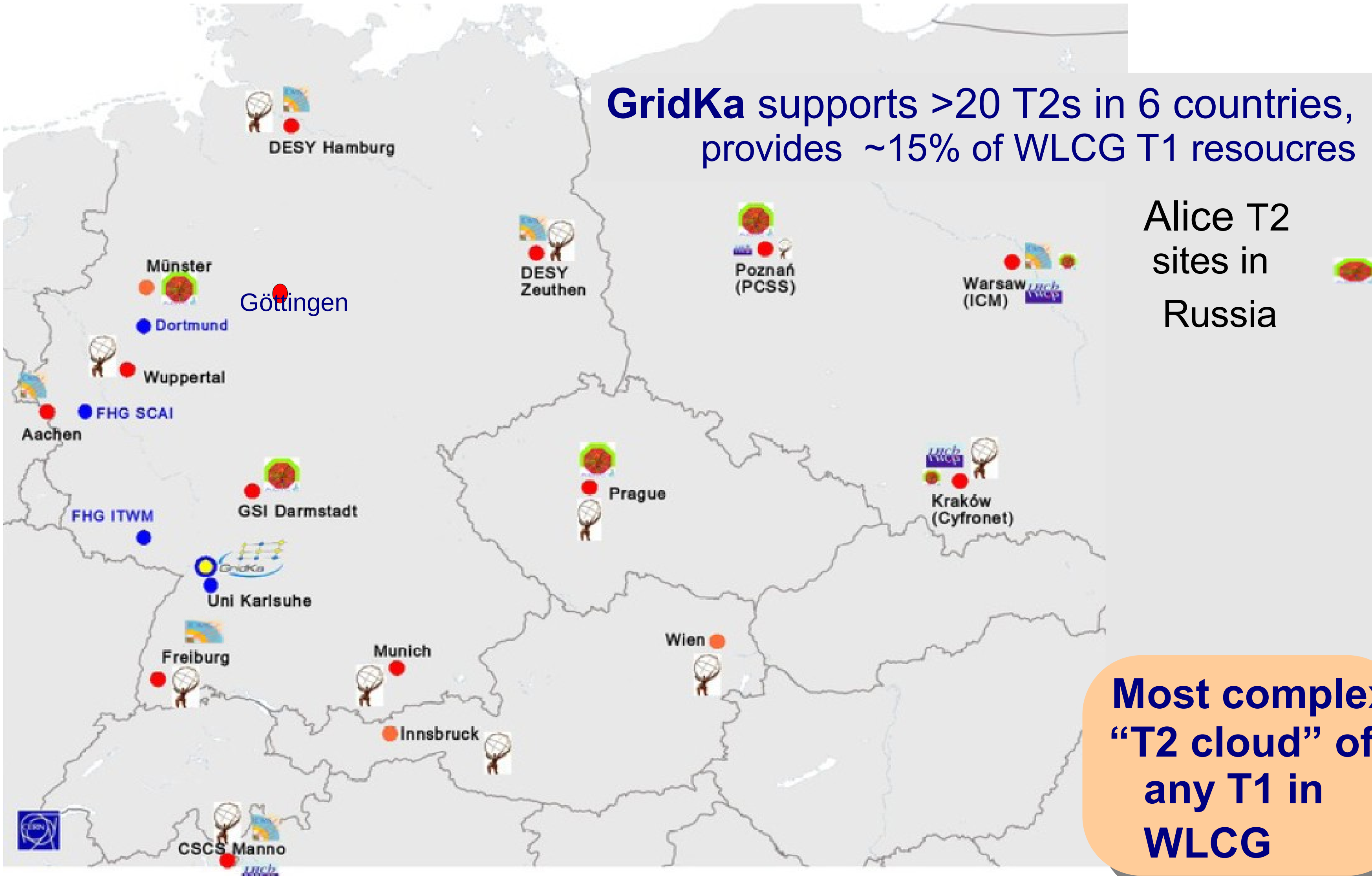
Total Disk 2011:
130 PB,
same amount as
Tape Storage 2011:
130 PB
(40 PB at CERN,
none @ T2s)

2012 numbers still being negotiated !

The largest Science Grid in the World

A closer look to the surroundings of GridKa

GridKa supports >20 T2s in 6 countries, provides ~15% of WLCG T1 resoucres



Most complex "T2 cloud" of any T1 in WLCG

The World-wide LHC Computing Grid

After almost 2 years of experience with LHC operation:

How well did it work ?

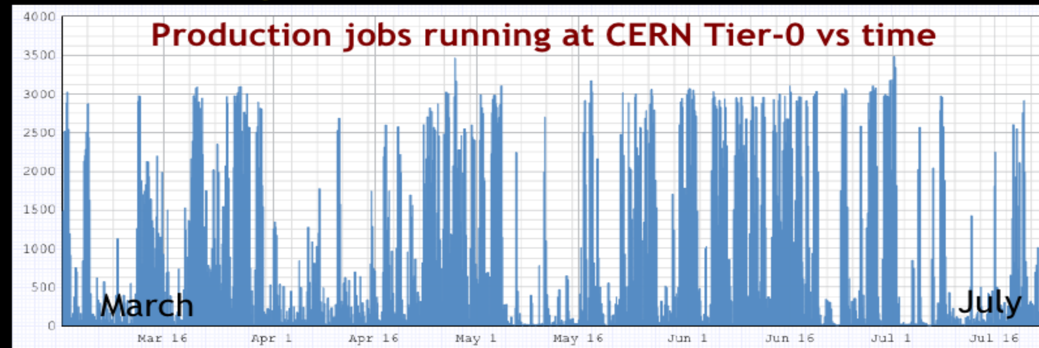
Almost 2 years of experience - Did it work ?

Up to the users to give feedback:

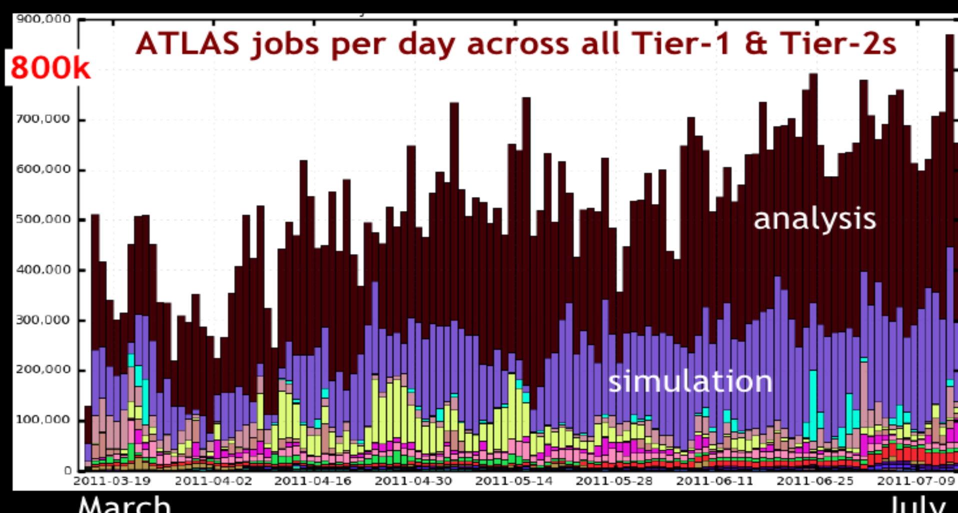
D. Charlton,ATLAS, EPS HEP 2011

Computing Grid Delivers Physics

Data preparation:
- First-pass reco. at Tier-0 within ~2 days
- Calibration/DQ good for physics analysis
- Data analysable on Grid within ~1 week



Tier-1 and Tier-2's process ~⅔ M jobs per day
- simulation
- re-reconstruction (campaigns)
- group production (ntuples…)
- physics analysis



The high quality computing system allows us to show results on data taken until the end of June

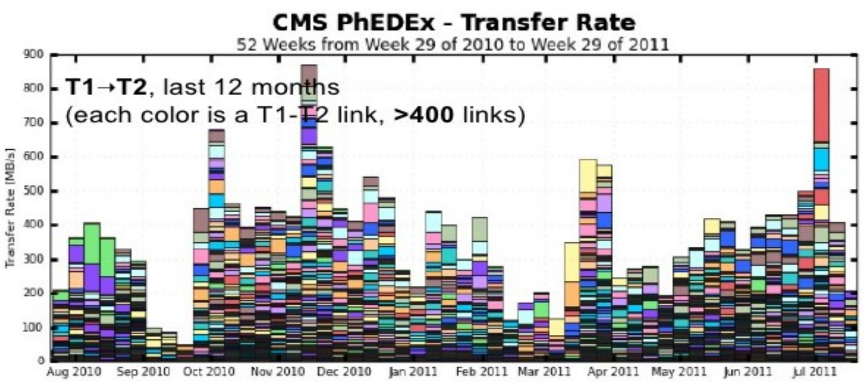
Payback for the years of investment and hard work

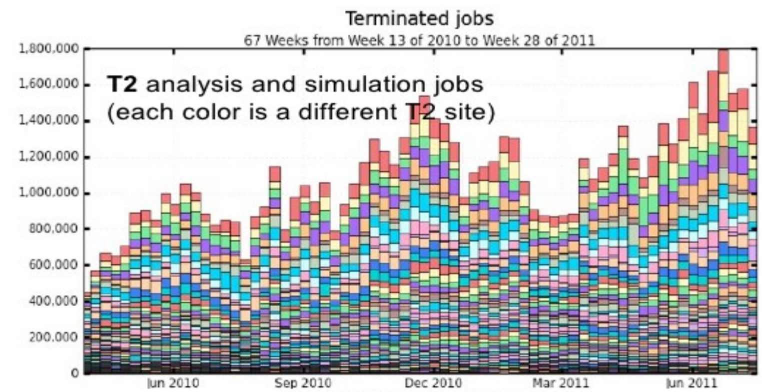
Did it work (2)

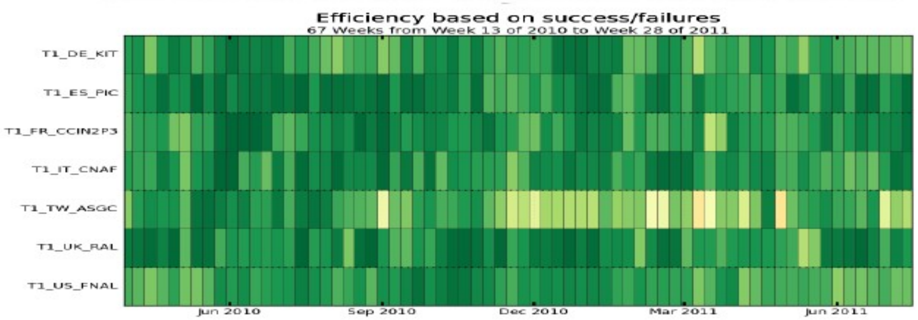
G. Tonelli, CMS, EPS HEP 2011

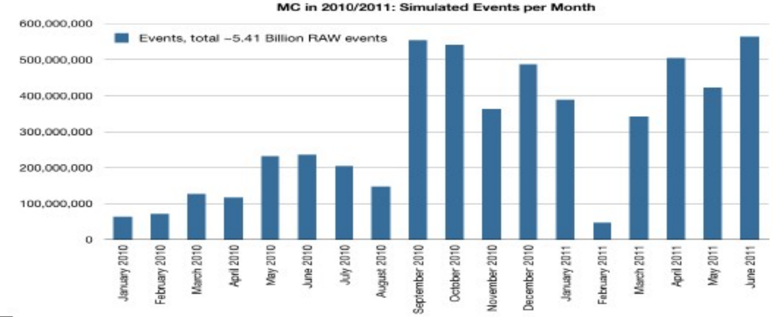
Offline and Computing running smoothly

•Smooth **Tier-0** operation, keeping up with the data taking. Increase in **Tier-1** utilization, for reprocessing and skimming jobs; High usage of **Tier-2** for analysis, **>400 (800)** individual users per week (month). More than 5.4 Billions MC events.









G. Tonelli, CERN/INFN/UNIPI *HEP_2011_GRENOBLE* *July 25 2011* *36*

Did it work ?

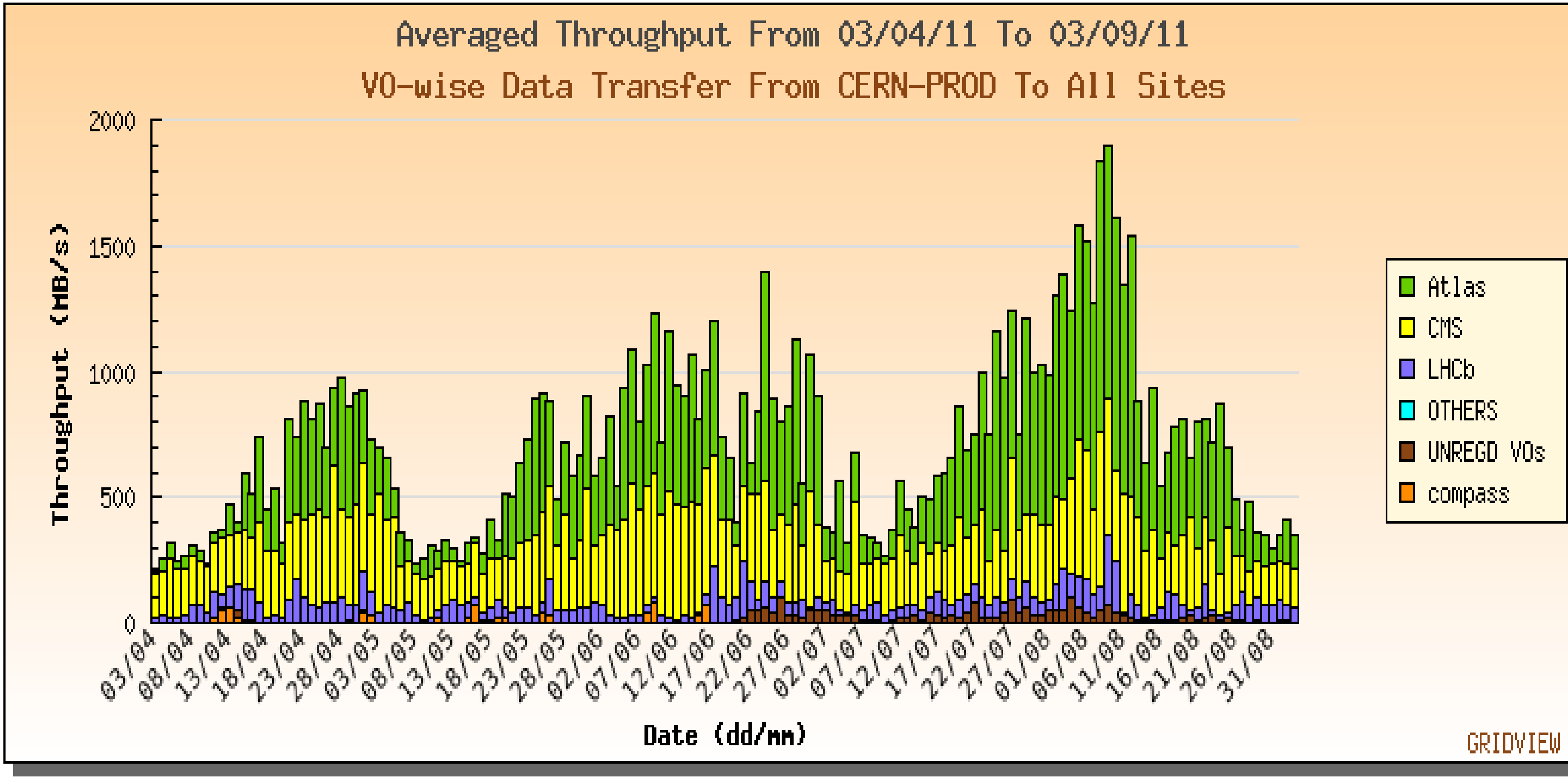
Obviously it did!

- Grid infrastructure for the LHC performed extremely well
- physics results from freshly recorded data
- but: effort for running computing infrastructure is high!

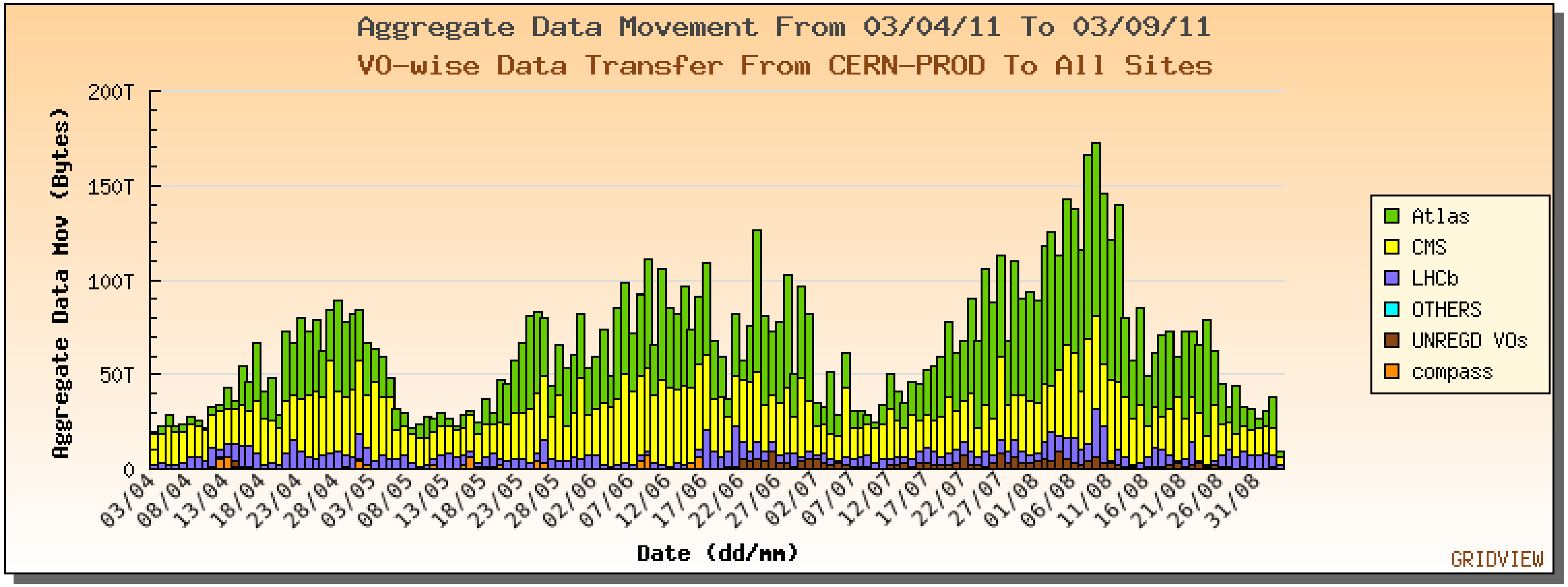
There are challenges ahead!

Let`s have a look in detail ...

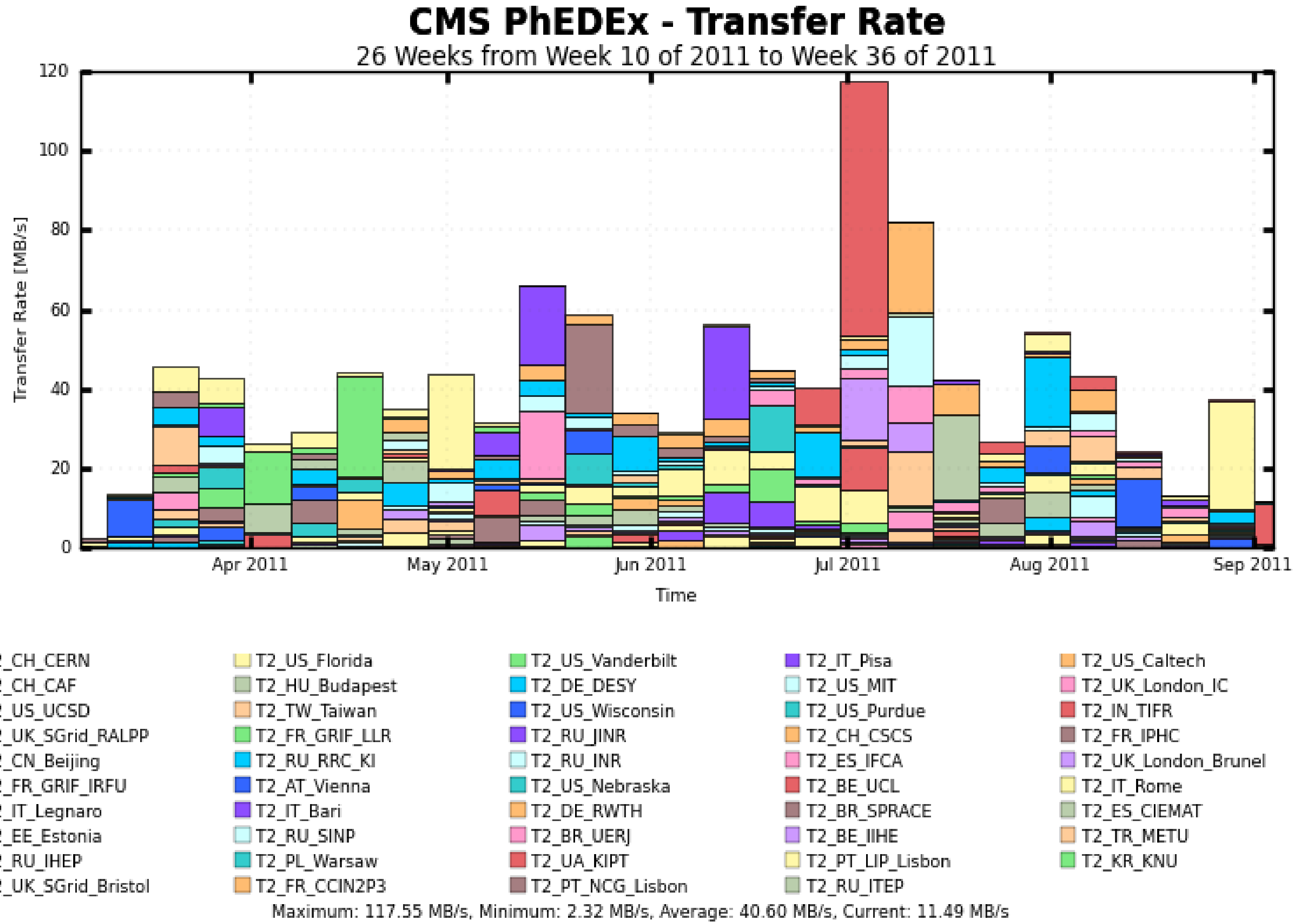
Data Export Rate from CERN



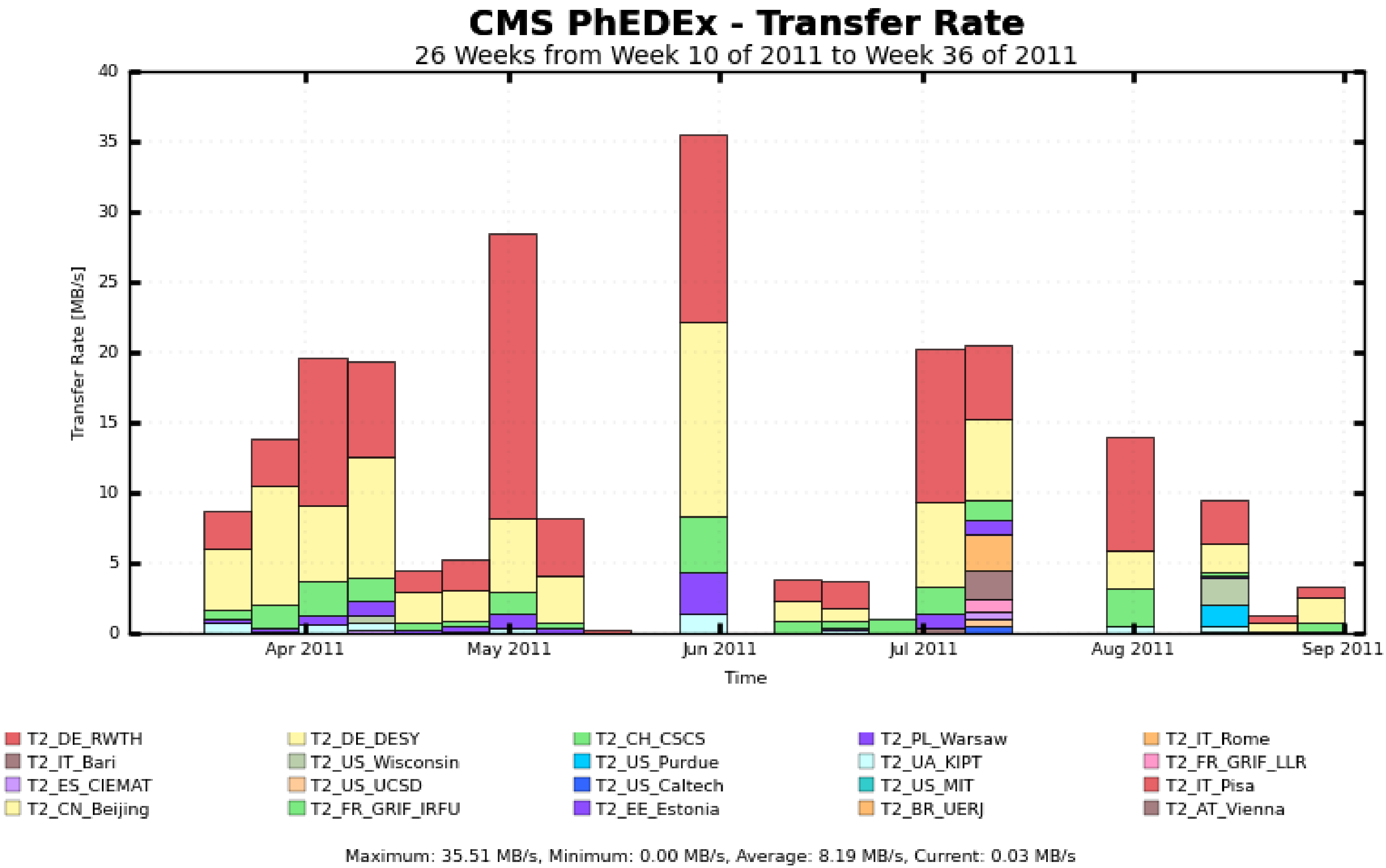
Daily data export from CERN



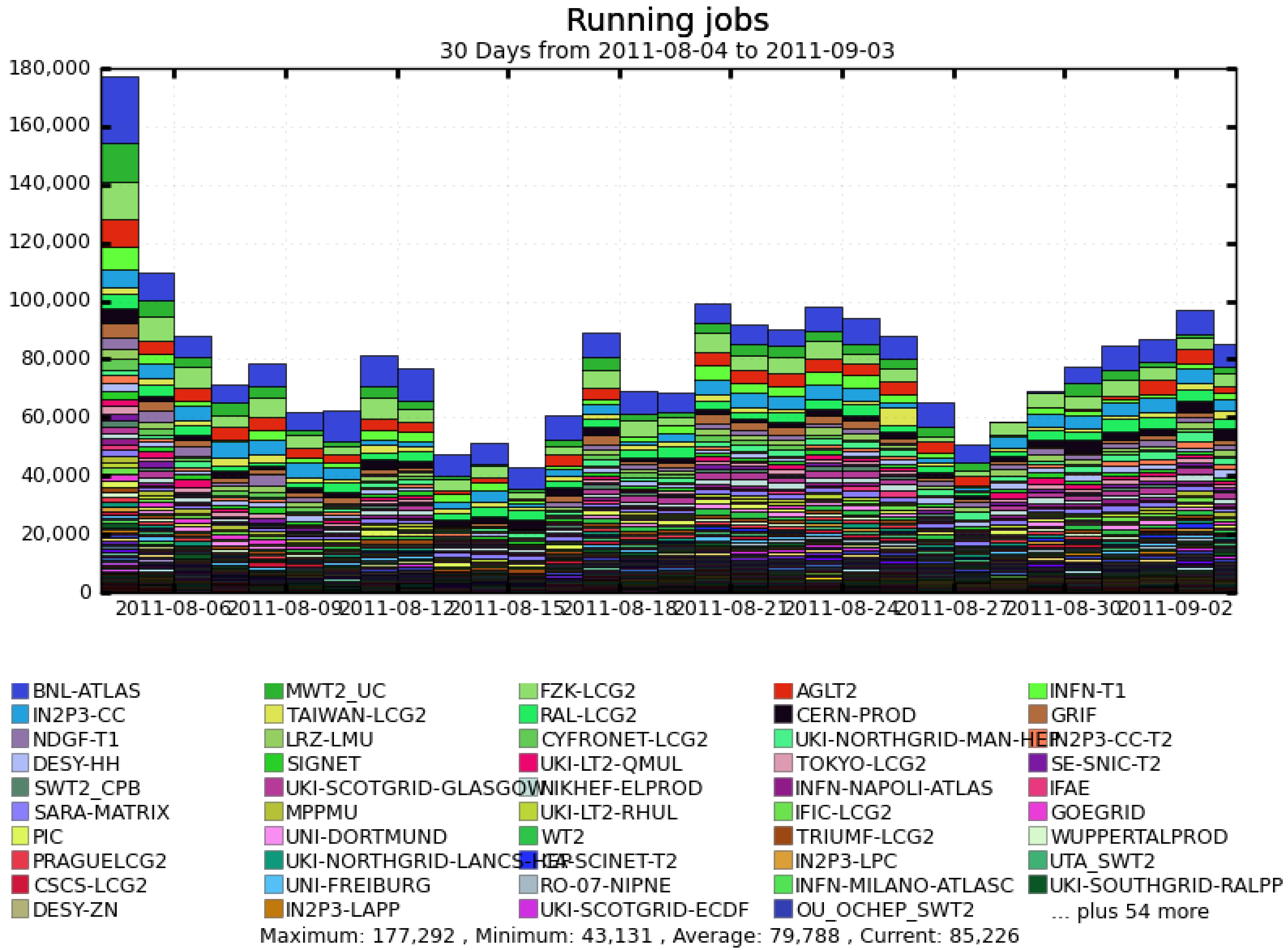
Data Export from T1: example: CMS @ KIT



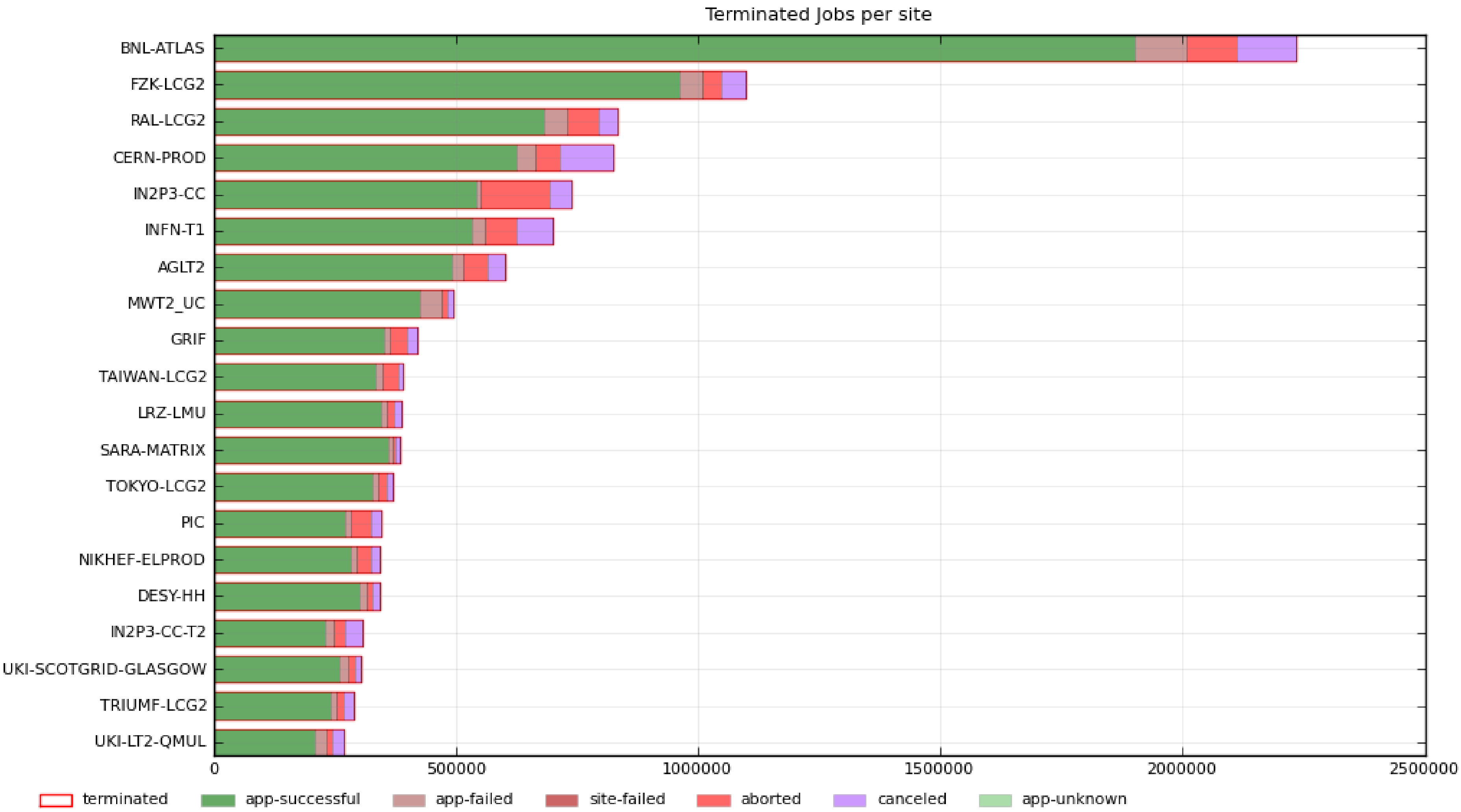
Data Import T2->T1 (MC): example CMS@KIT



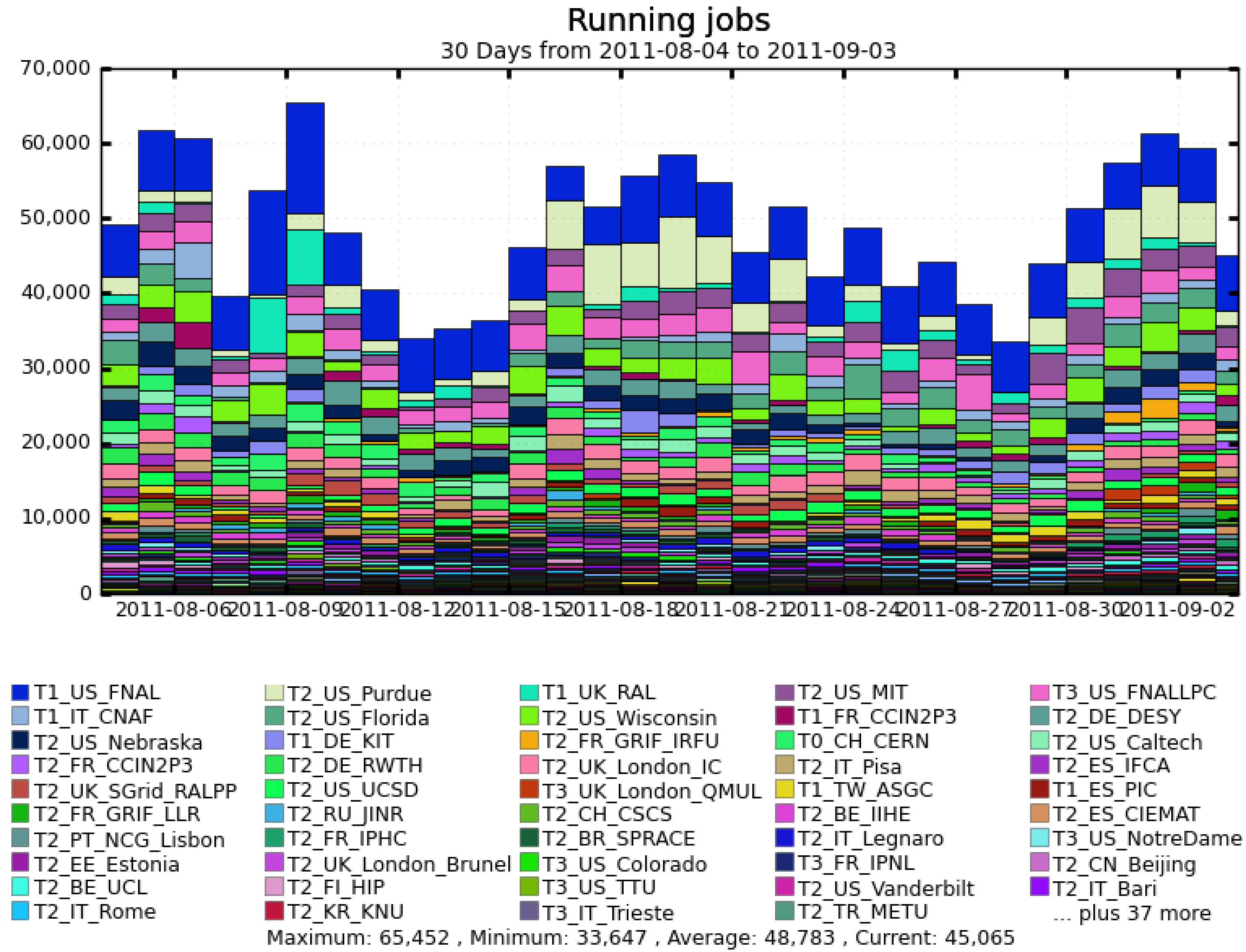
Processed Jobs ATLAS



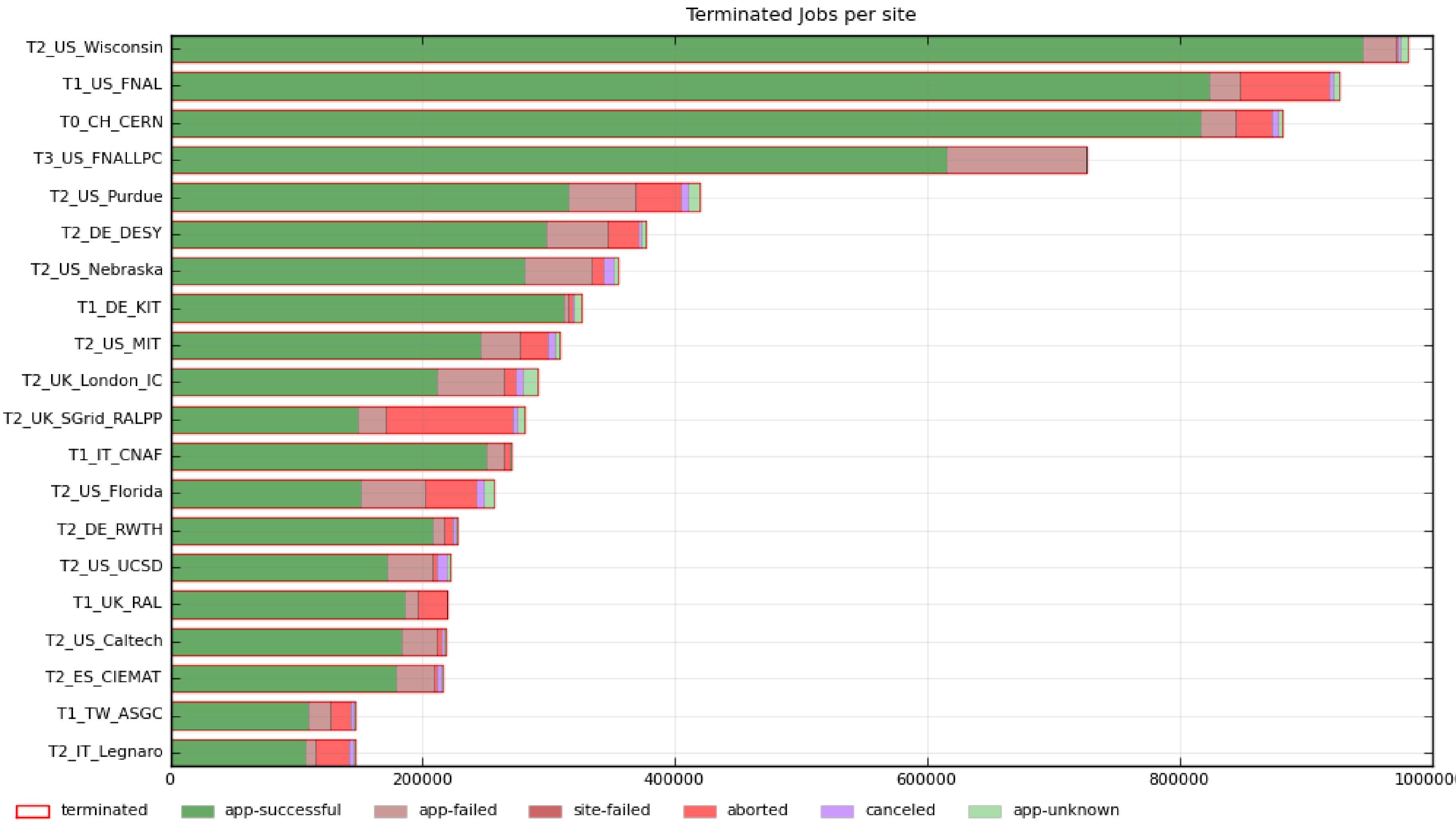
Pocessed jobs success rates (ATLAS)



Processed Jobs - CMS

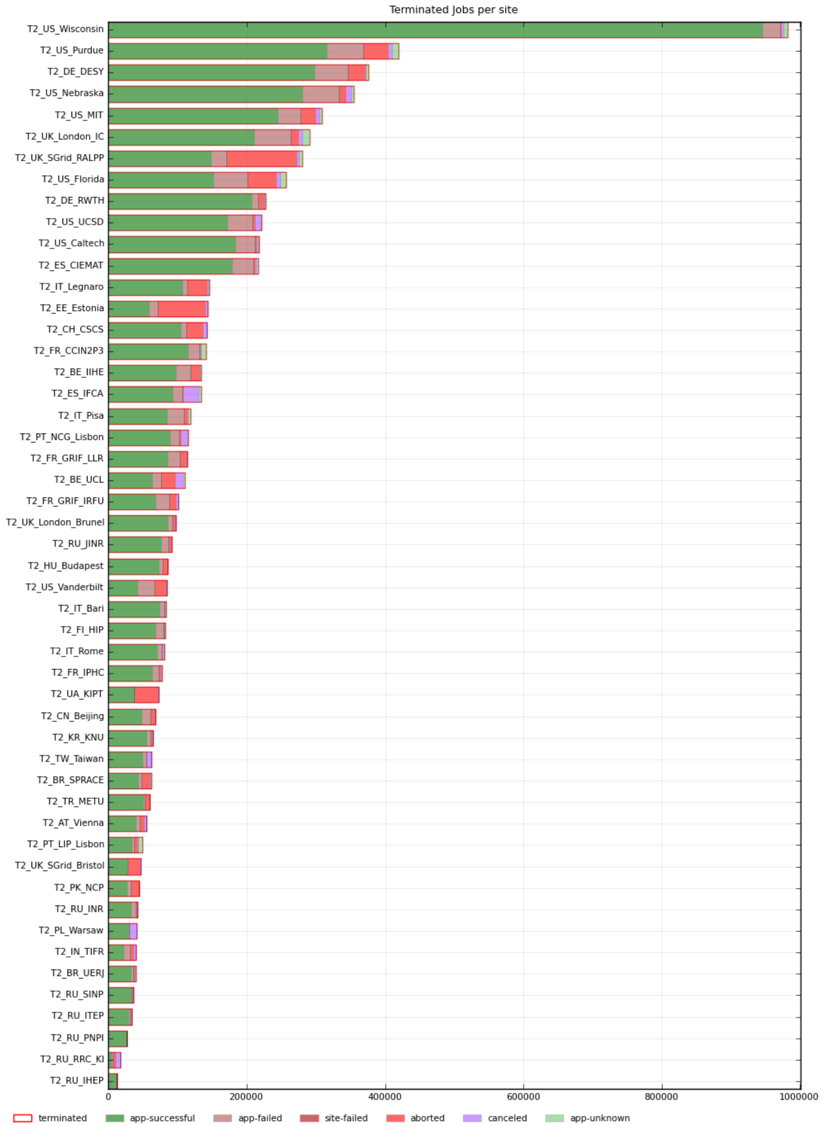


Processed Jobs success rates (CMS)



Job success rates @ T2s

(CMS T2s shown here ATLAS looks similar)



T2s see a mixture of
- MC production jobs
- user analyis and skimming jobs

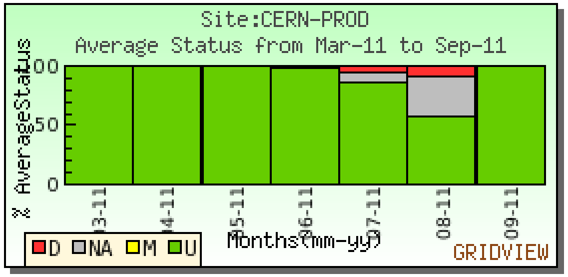
Job success rates not in all cases above 90%

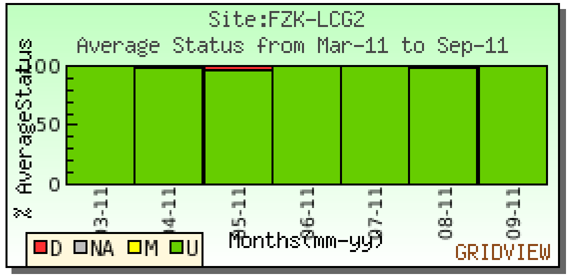
90% success rate would be considered very low for a classical computer centre

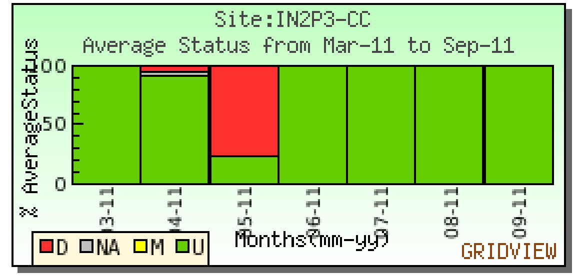
This must improve …

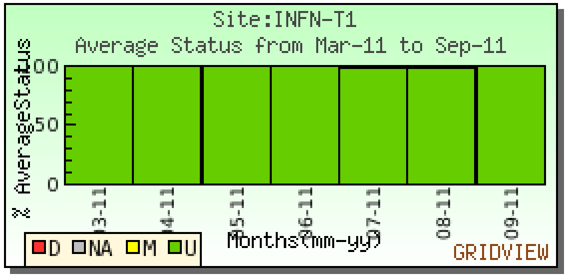
Not easy to disentangle failures of the system from "user errors"

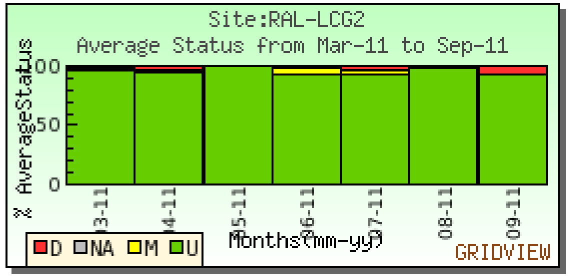
Site Reliability (examples)

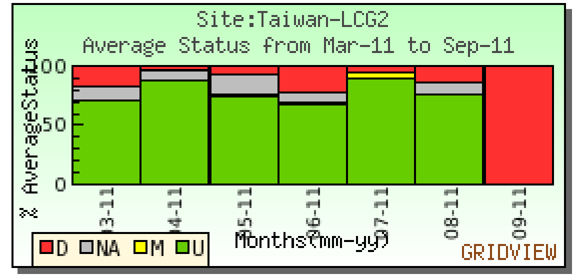


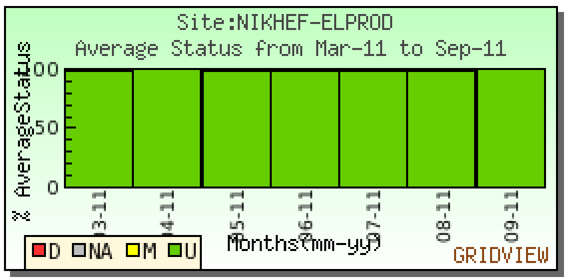


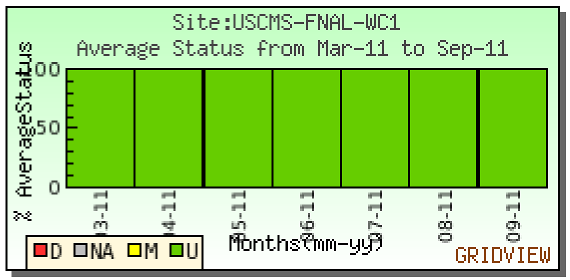


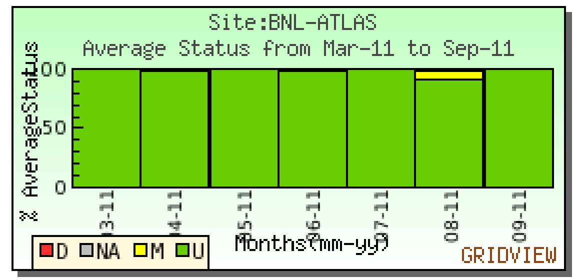


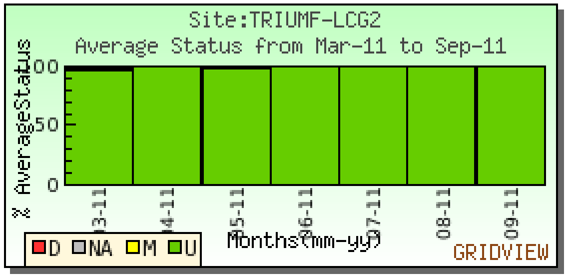


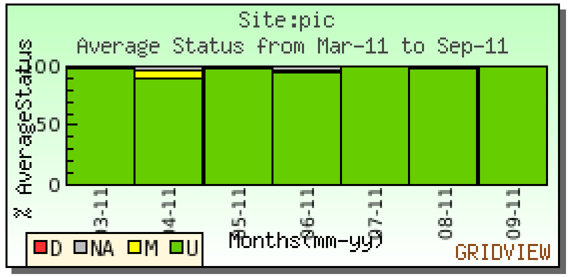


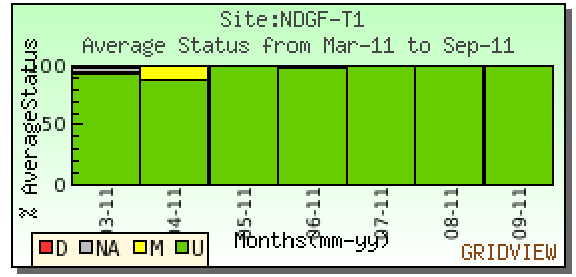










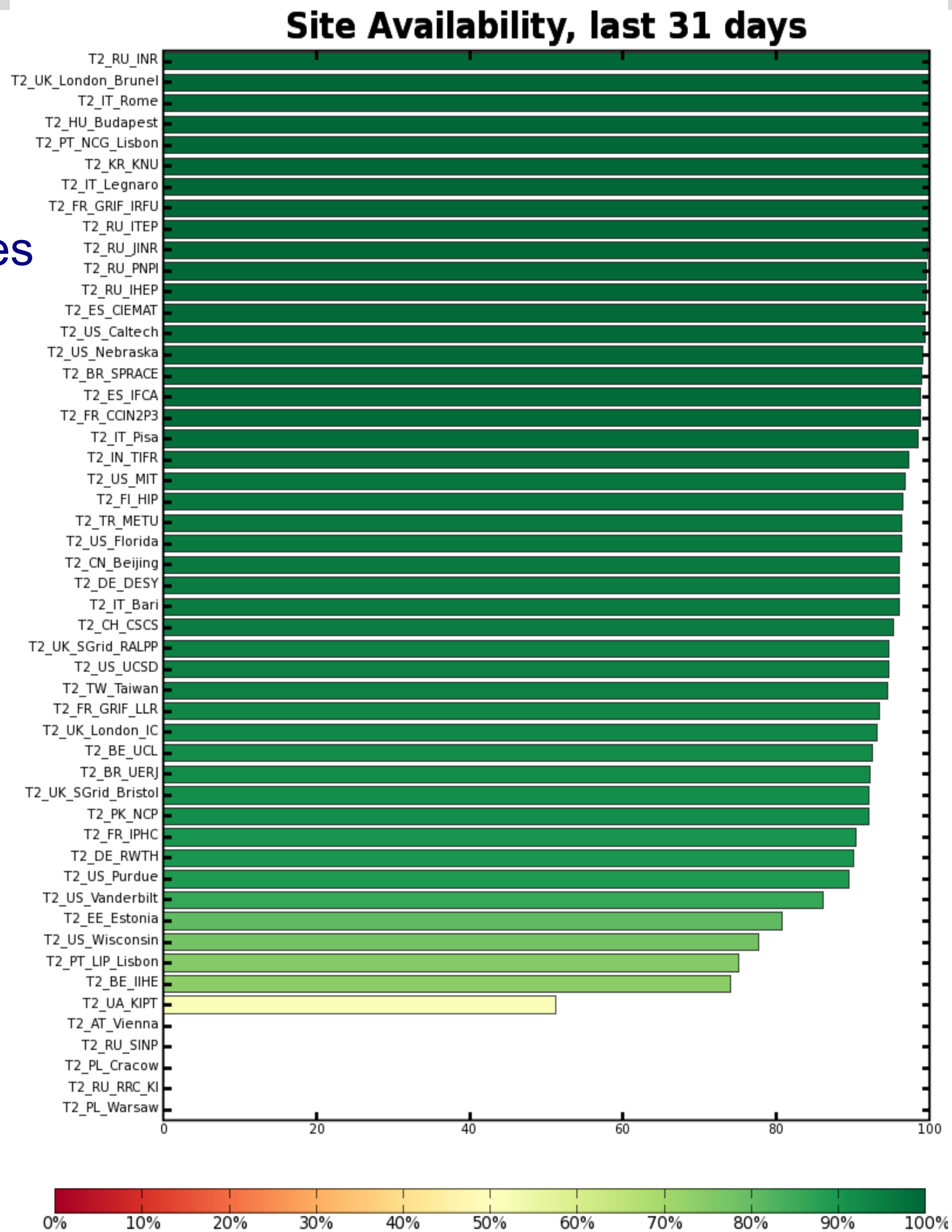


T2 performance

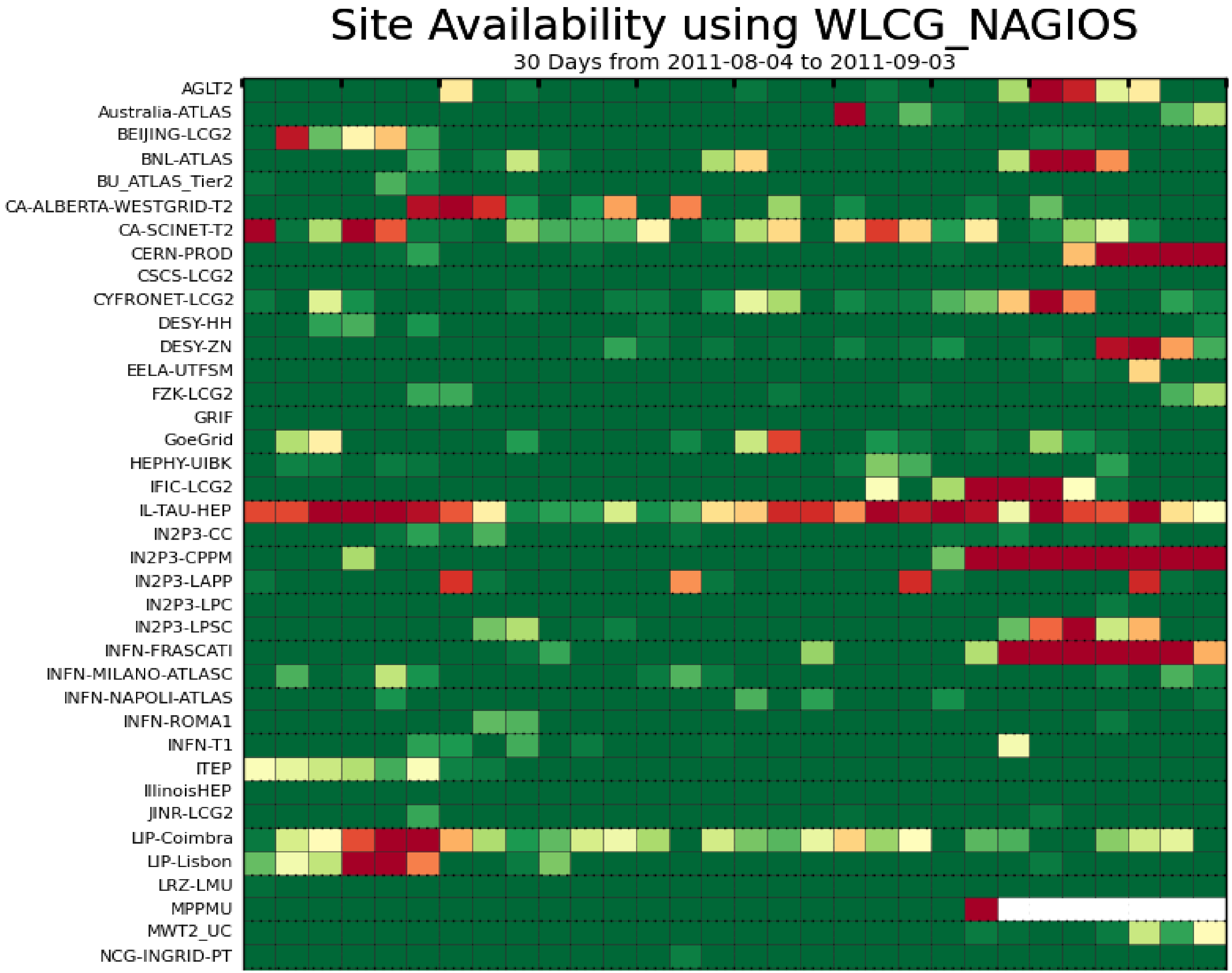
"Availability" of CMS T2 sites

There are sites with peroformance issues !

Typically, less well performing sites are very small !



T2 Perfromance – ATLAS example



T2 Performance Atlas (cont'd)

This is an example of a time-resolved measurement of site avaiability

Message similar as previously.

This kind of graphs helps the site responsibles to monitor their sites and act on problems.

Provided centrally by WLCG and expriments !

Again: Does it work ?

YES !

- Routinely runing ~150'000 jobs simultaneously on the Grid
- Shipping over 100 TB/day to T1 centres
- data distribution to T2 works well
- some T2s have performance issues
- very little is known about T3 usage and success rates - responsibility of the institutes
- plenty of resources available at LHC start-up, now approaching “resource limited operation”
- Users have adapted to the “GridWorld” - Grid is routinely used as a huge batch system, output is transferred home

but ...

Does it work?

Message:

it worked better than expeced by many,

but running such a complex computing infrastructure as the WLCG is tedious (and expensive!)

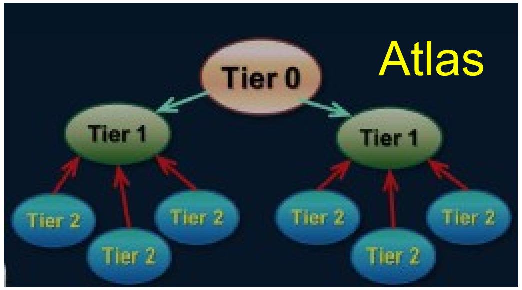
Reliability and cost of operation can be improved by

- simpified and more robust middleware
- redundancy of services and sites,
  requires dynamic placement of data and investment in network bandwidh
- automated monitoring and triggering of actions
- use of commercially supported approaches to distributed computing:
  - private clouds are particularly important for shared resources at universities
  - eventually off-load simple tasks (simulation, statistics calculations) to commercial clouds

Many of the new developments were addressed at this School

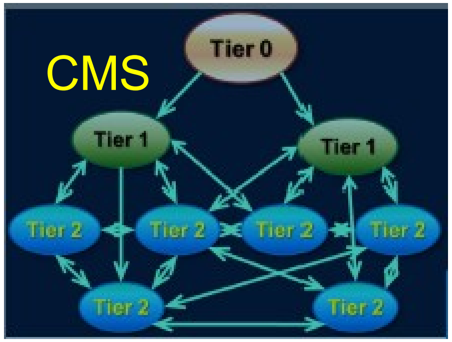
Let`s have a look at some future developments ...

LHCONE



ATLAS and CMS
computing models differ slightly

CMS already more "distributed"



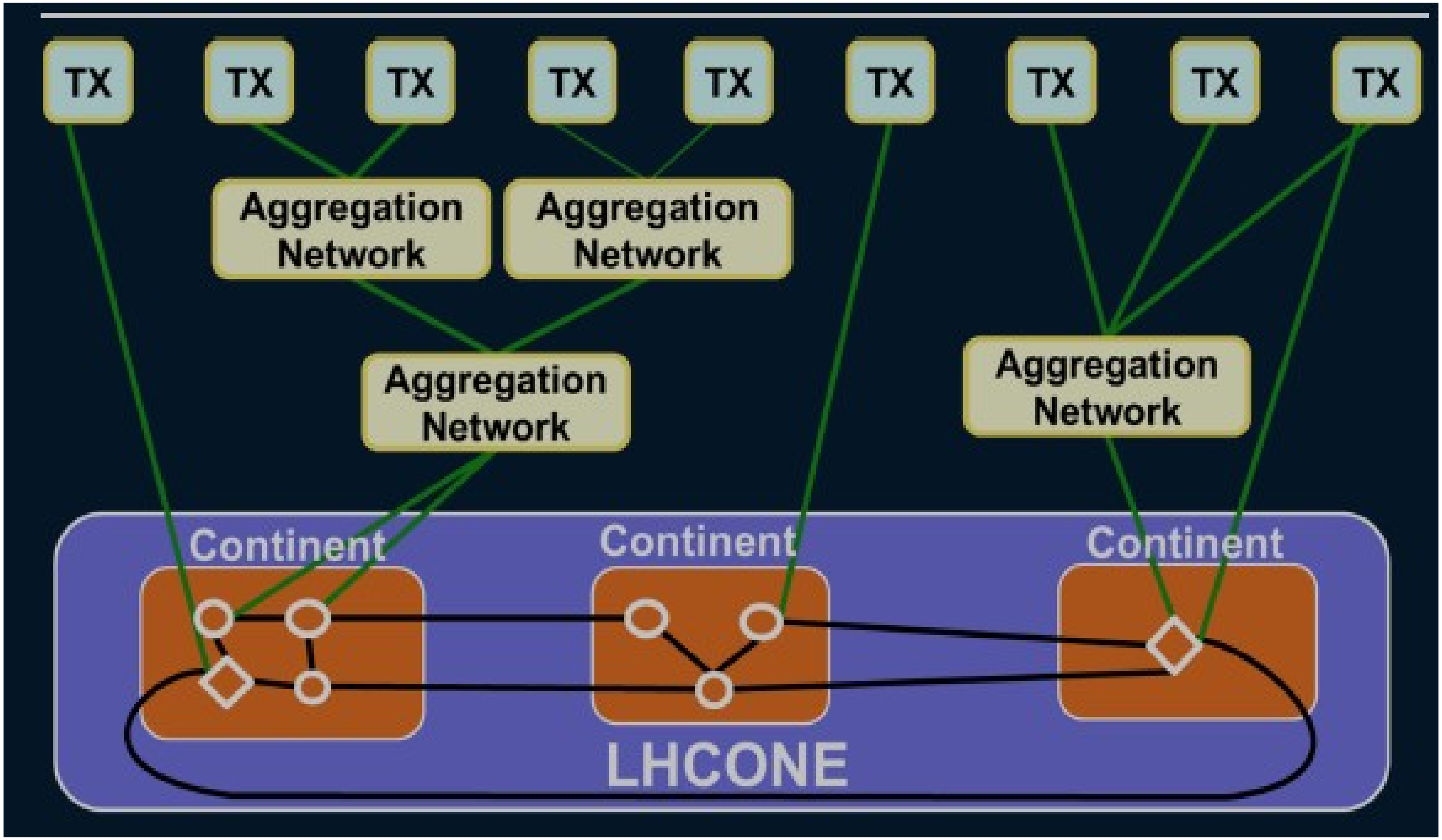
Aim of LHCONE project is

better trans-regional networking for data analyis,
complementary to **LHCOPN** network connecting LHC T1s

- **flat(er) hierarchy:** any site has access to any other site's data
- **dynamic data caching:** pull data "on demand"
- **remote data access:** jobs may use data remotely

by interconnecting open exchange points between regional networks

LHCONE (2)



Schamatic layout of LHCONE network infrastrucure

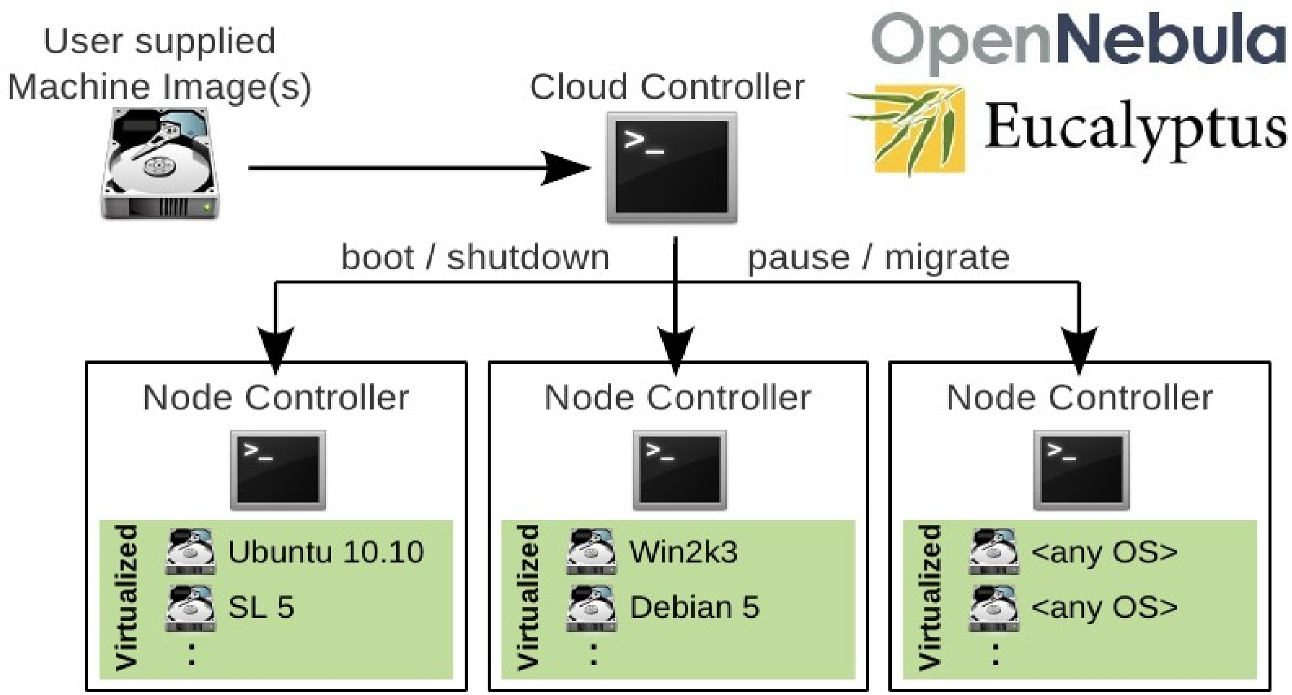
A dedicated HEP network infrastrucure – what is the cost ?

Virtualisation

You have heard a lot about clouds and virtualisation at this school

in a nutshell:

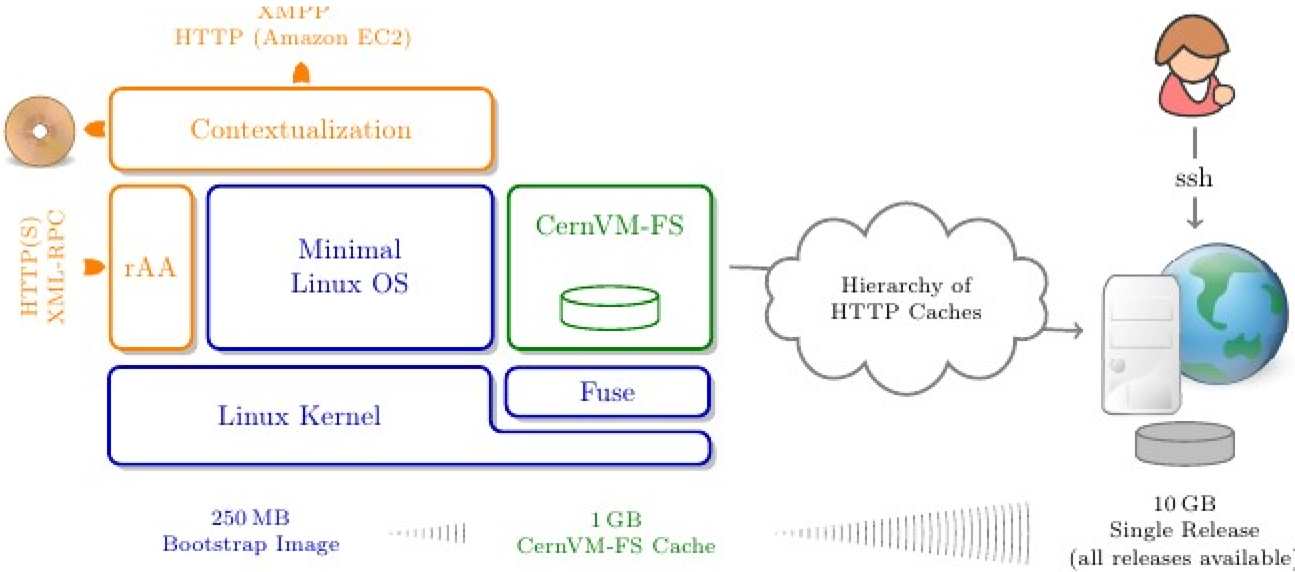
- Clouds offer "Infrastructure as a Service"
  - easy provision of resources "on demand"
    even by including (private) clould resources as a classical batch queue
    (e.g. ROCED project developed at EKP, KIT)
- independent of local hardware and operating system
  (Scientific Linux 5 for Grid middleware and experiment software)



CernVM & CernVM-FS

CernVM is a virtual machine (“Virtual Software Appliance”) based on Scientific Linux with CERN software environment, runs on

CernVM-FS is a client-server file system based on http and implemented as as user-space file system optimized for read-only access to software repositories with a performant caching mechanism. Allows a CernVM instance to efficiently access software installed remotely.



A recap of this School

This week, you have heard about many of the new developments:

- J. Templon, *Grid and Cloud*
  "Grids need Clouds to prosper, Clouds need Grids to scale"
- O. Synge, *Virtualisation*
- P. Millar, *Data Storage*
- C. Witzig, *European Grid Projects*
- T. Beckers, Storage Architectures for Petaflops Computing
- S. Reißer, *Grid User Support*
- N. Abdennadher, *Combining Grid, Cloud and Volunteer Computing*
- U. Schwickerath, *Cloud Computing*
- S. Maffioletti, *ARC for developers*
- T. Metsch, B. Schott, *Sustainable DCI Operations*
- A. Aeschlimann, *Grid and Cloud Security*

and a number of Hands-on workshops in parallel sessions.

**HEP Grid runs fine in its initial version,
but virtualisation and "clouds" offer new possibilities
for resource increase, efficiency, cost effectiveness , operation and reliability.**

It's up the you, the participants of this school, to shape the future !

Thanks to all speakers, session teams and organizers and also from my side.