WHAT IS A T1?

Federico Carminati

The Grid of the Americas

February 8-11, 2011

THE CERN LARGE HADRON COLLIDER – LHC

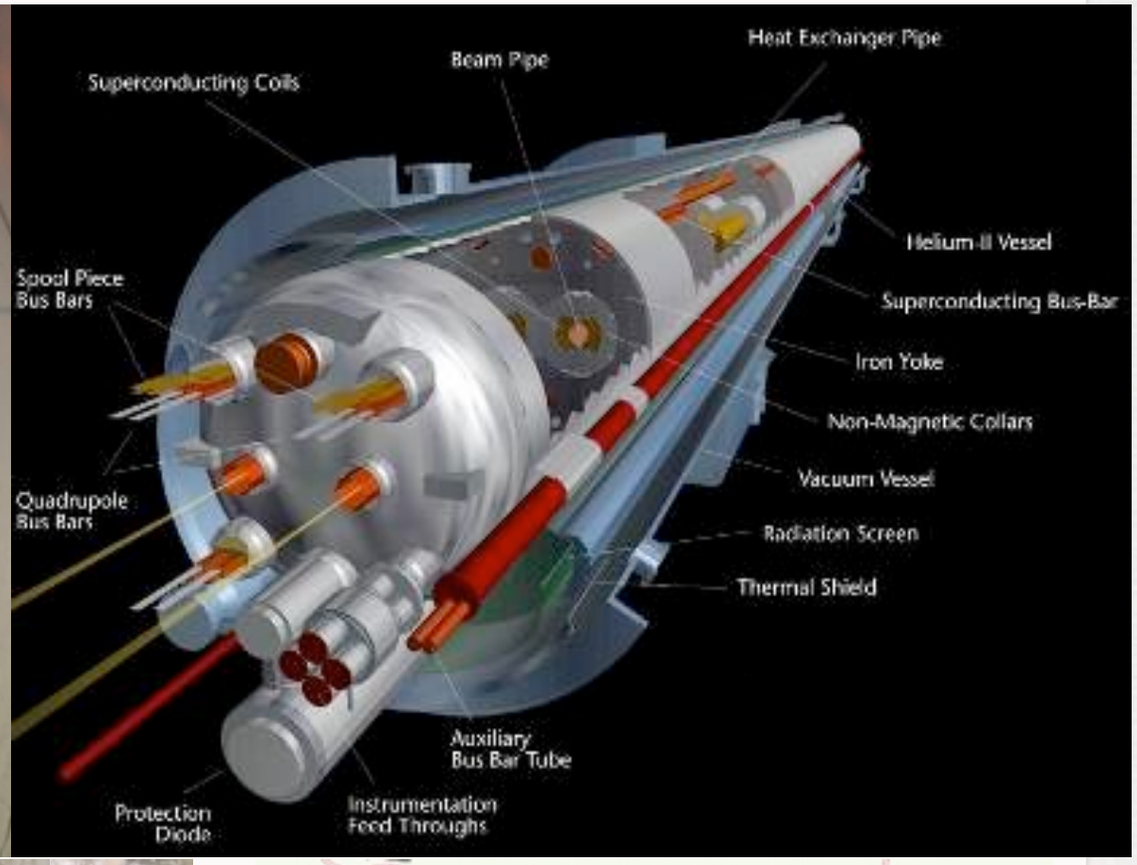
- LHC is a proton-proton and heavy ion collider
- Proton-proton center-of-mass energy $\sqrt{s_{pp}} = 14$ TeV
- Started Sept 2009

THE CERN LARGE HADRON COLLIDER – LHC

- LHC is a proton-proton and heavy ion collider
- Proton-proton center-of-mass energy $\sqrt{s_{pp}} = 14$ TeV
- Started Sept 2009

THE CERN LARGE HADRON COLLIDER – LHC

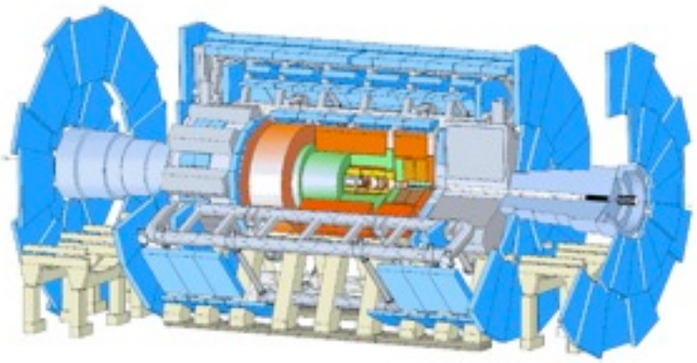
- LHC is a proton-proton and heavy ion collider
- Proton-proton center-of-mass energy $\sqrt{s_{pp}} = 14$ TeV
- Started Sept 2009



THE CERN LARGE HADRON COLLIDER – LHC

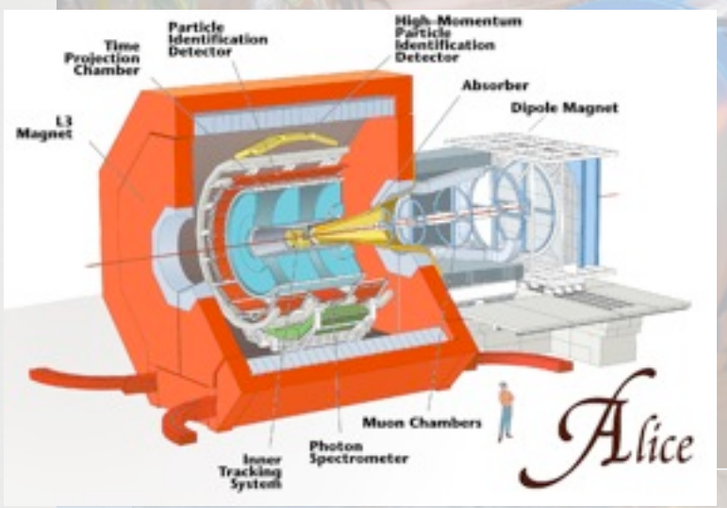
- LHC is a proton-proton collider
- center-of-mass energy = 14 TeV
- 2009

ATLAS

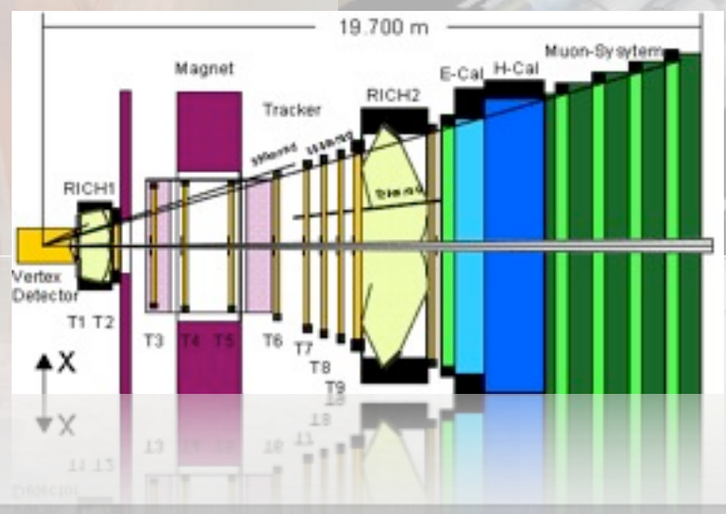


CMS

Alice



LHCb

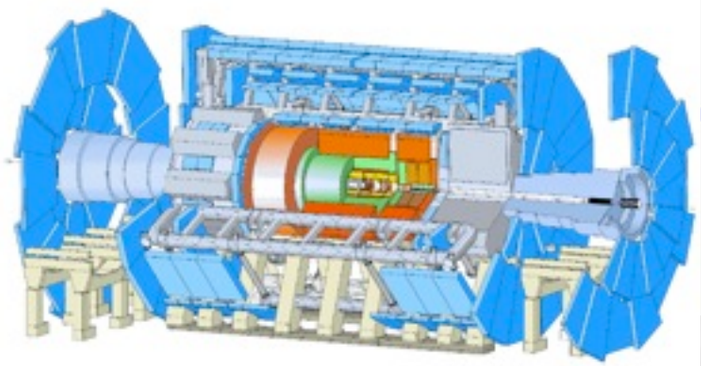


~8-10 PetaBytes/ year
~10^{10} events/year
~10^{3} batch and interactive users

THE CERN LARGE HADRON COLLIDER – LHC

- LHC is a proton-proton collider
- center-of-mass energy = 14 TeV
- 09

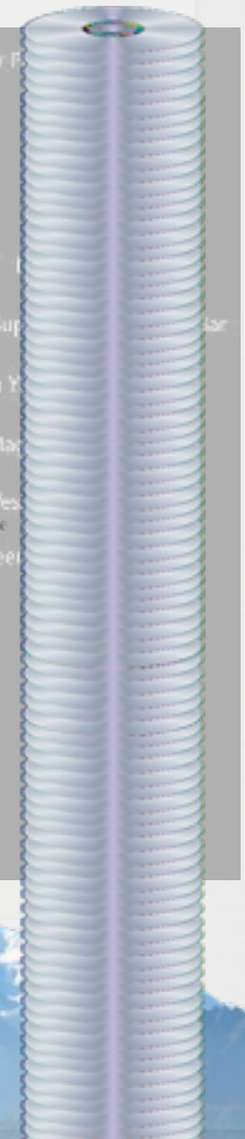
ATLAS



CMS

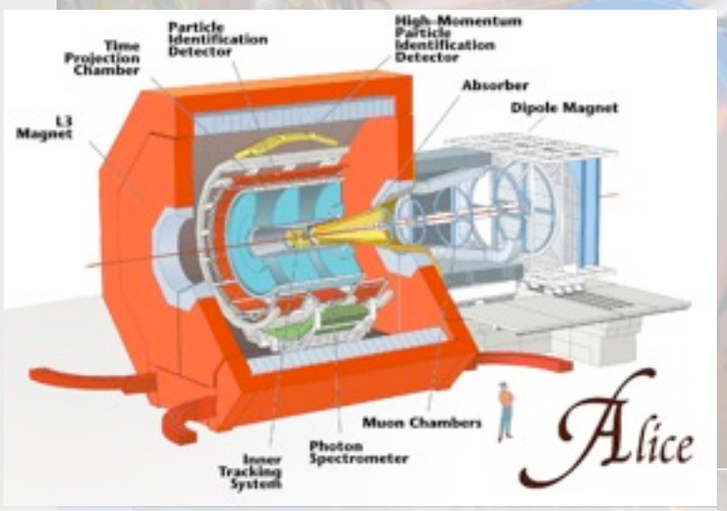
Balloon →
(30 Km)

CD stack with one year of LHC data!
(~ 20 Km)

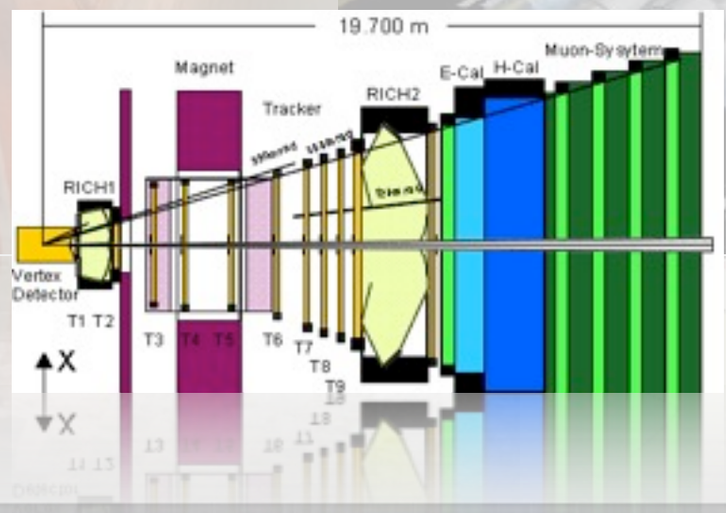


~8-10 PetaBytes/ year
~10^{10} events/year
~10^{3} batch and interactive users

Alice



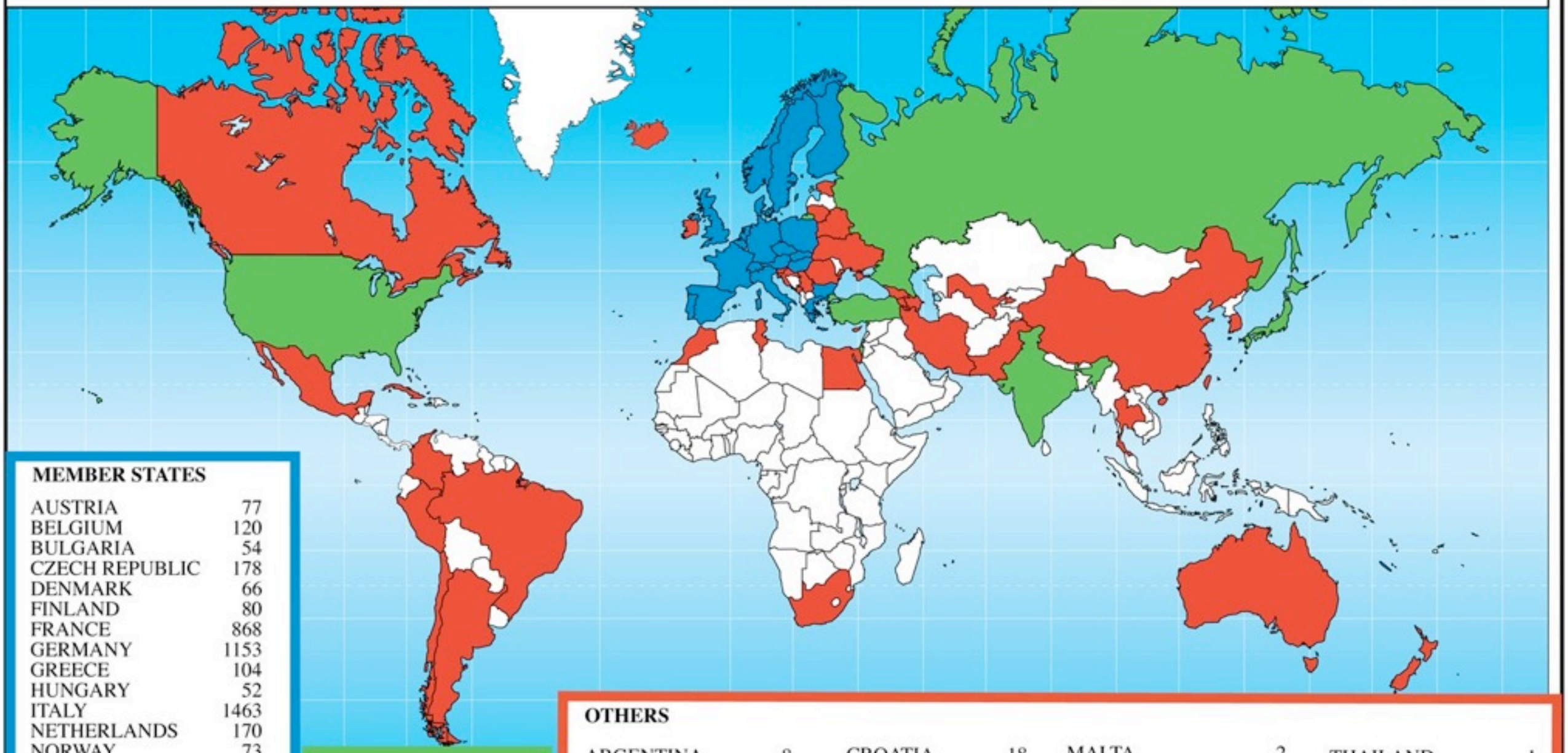
LHCb



Airplane →
(10 Km)

Mt. Blanc
(4.8 Km)

Distribution of All CERN Users by Nation of Institute on 20 January 2010



MEMBER STATES

AUSTRIA	77
BELGIUM	120
BULGARIA	54
CZECH REPUBLIC	178
DENMARK	66
FINLAND	80
FRANCE	868
GERMANY	1153
GREECE	104
HUNGARY	52
ITALY	1463
NETHERLANDS	170
NORWAY	73
POLAND	191
PORTUGAL	122
SLOVAKIA	55
SPAIN	311
SWEDEN	71
SWITZERLAND	362
UNITED KINGDOM	732

6302

OBSERVER STATES

INDIA	91
ISRAEL	49
JAPAN	204
RUSSIA	901
TURKEY	60
USA	1618

2923

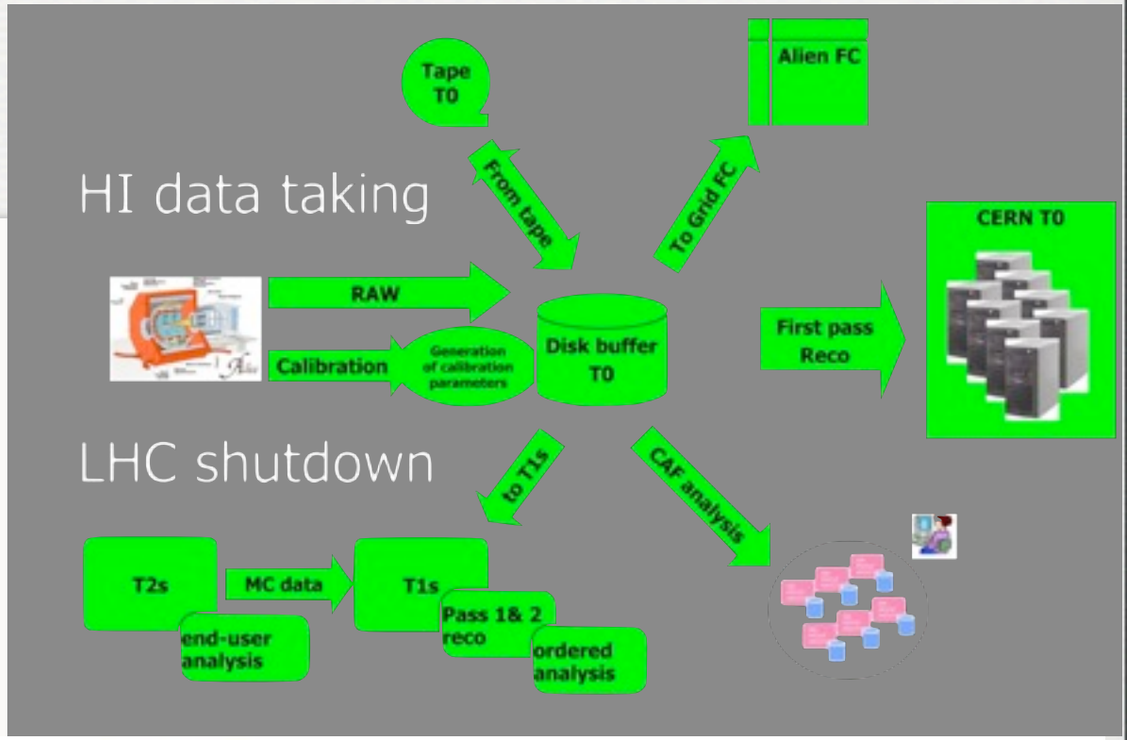
OTHERS

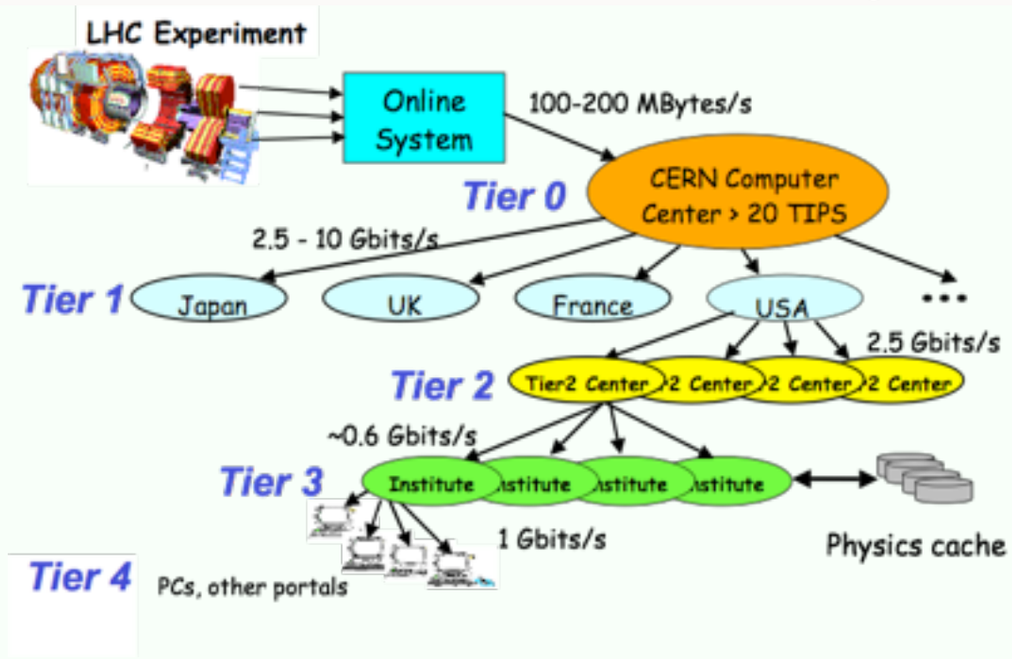
ARGENTINA	8	CROATIA	18	MALTA	2	THAILAND	1
ARMENIA	16	CUBA	4	MEXICO	33	TUNISIA	1
AUSTRALIA	17	CYPRUS	8	MONTENEGRO	1	UKRAINE	17
AZERBAIJAN	1	EGYPT	3	MOROCCO	6	UZBEKISTAN	1
BELARUS	19	ESTONIA	9	NEW ZEALAND	8		
BRAZIL	77	GEORGIA	10	PAKISTAN	15		
CANADA	141	ICELAND	1	PERU	1		
CHILE	2	IRAN	15	ROMANIA	59		
CHINA	78	IRELAND	14	SERBIA	20		
CHINA (TAIPEI)	53	KOREA	64	SLOVENIA	17		
COLOMBIA	9	LITHUANIA	5	SOUTH AFRICA	8		

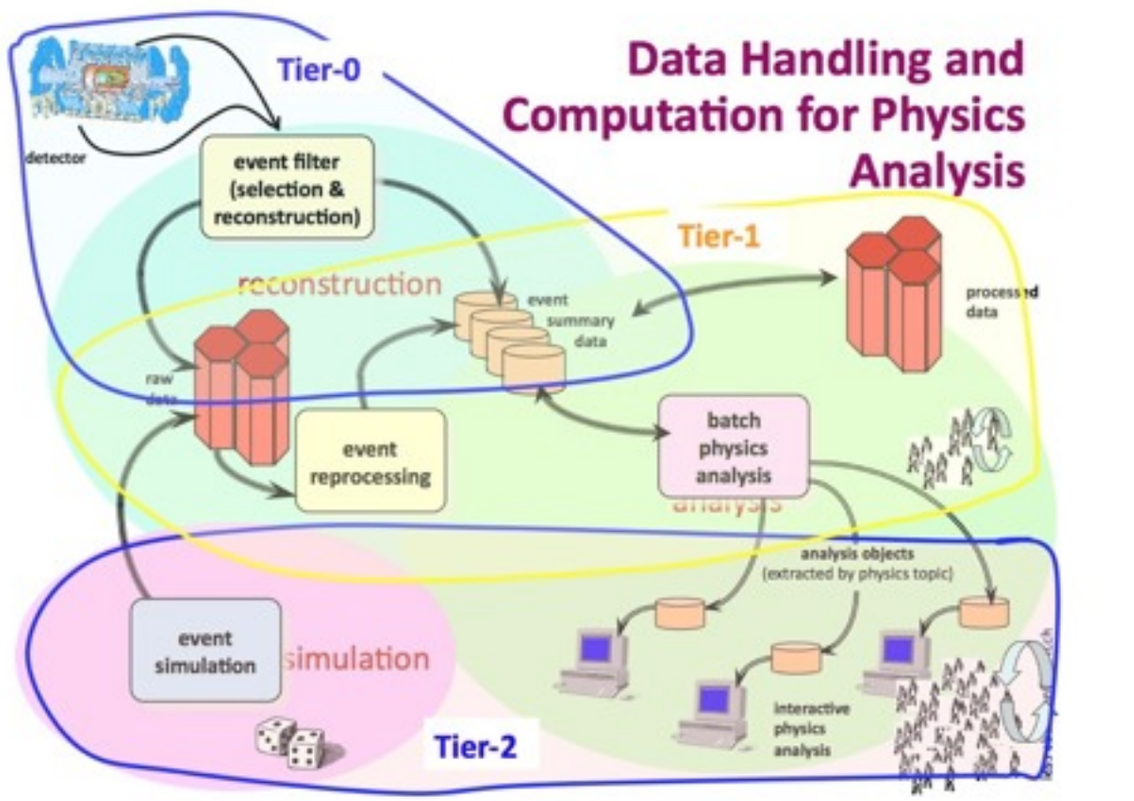
762

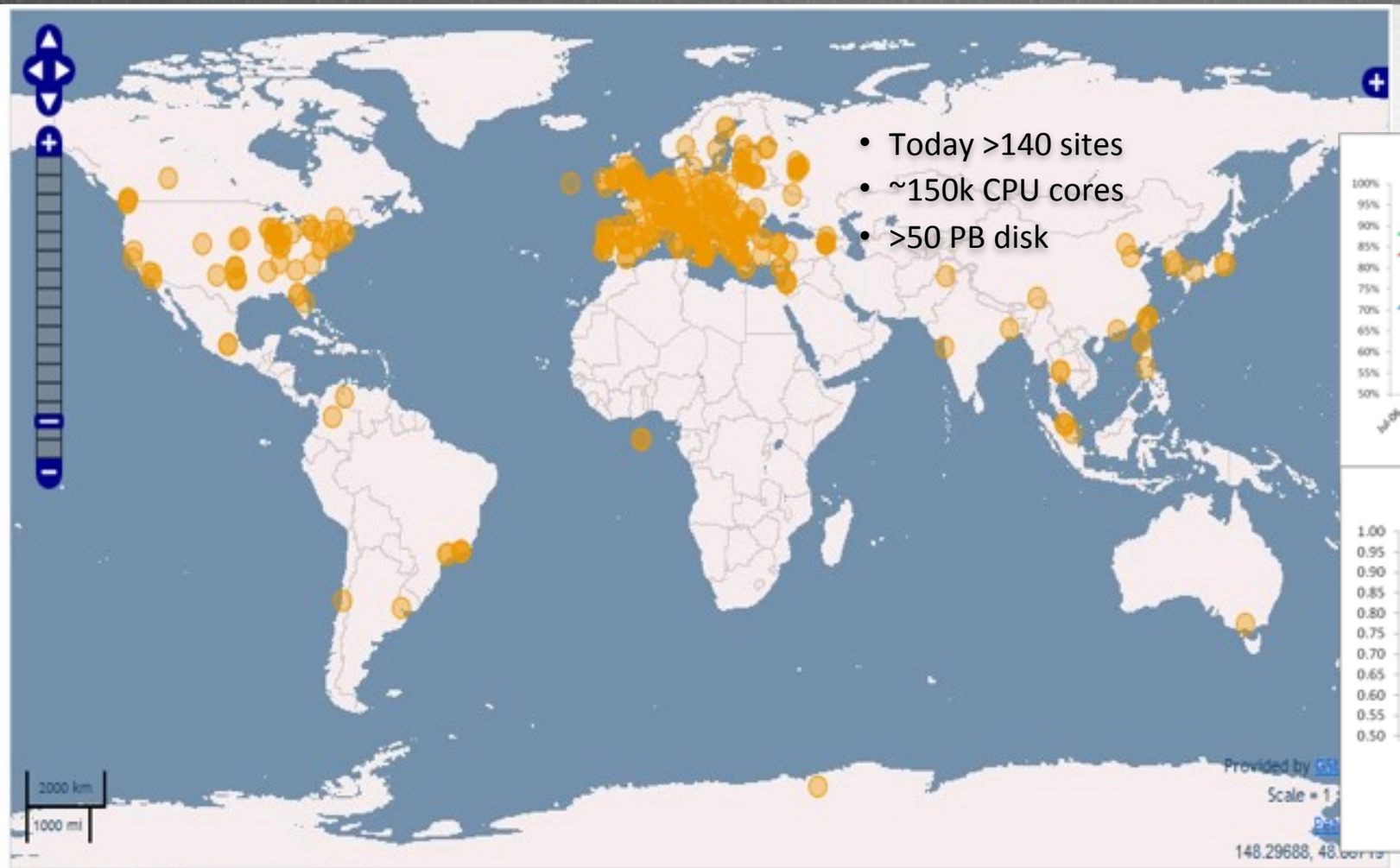
THE GRID – IT ALL STARTED WITH MONARC

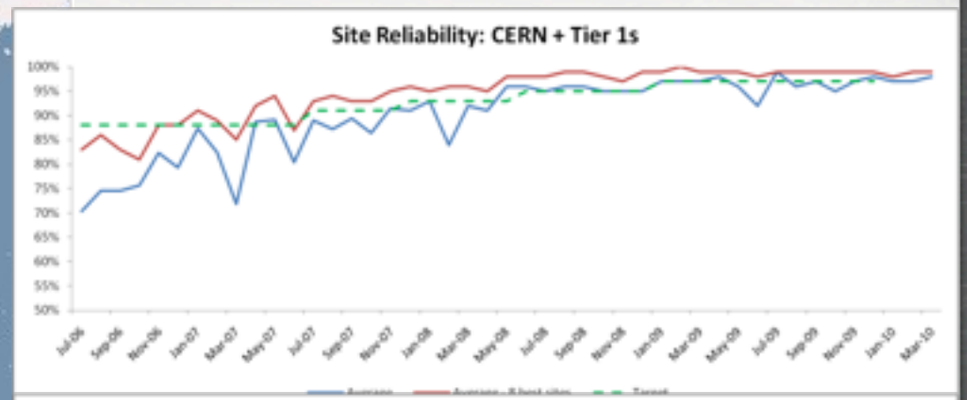
- The computing models are remarkably similar
  - T0: First reco, data storage
  - T1: Subsequent recos, MC reco, ordered analysis
  - T2: MC, Chaotic analysis
  - T3: End user analysis
- ALICE does not have (or admit having…) T3s

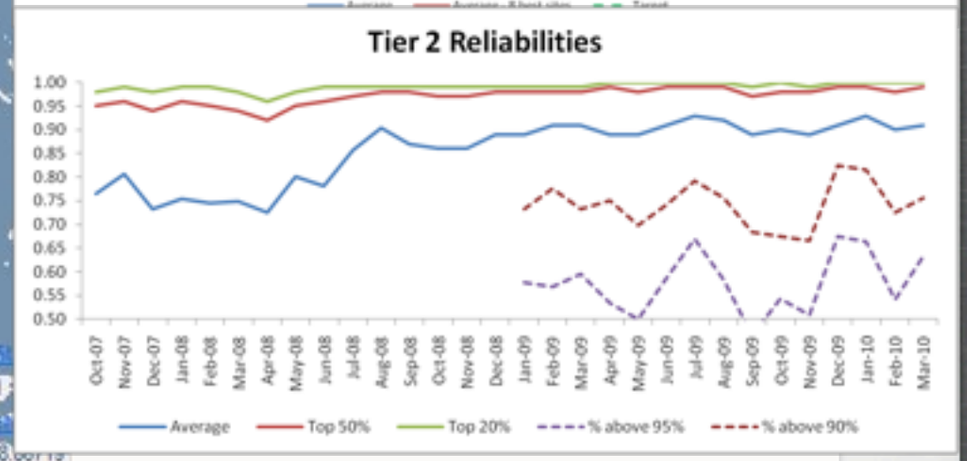




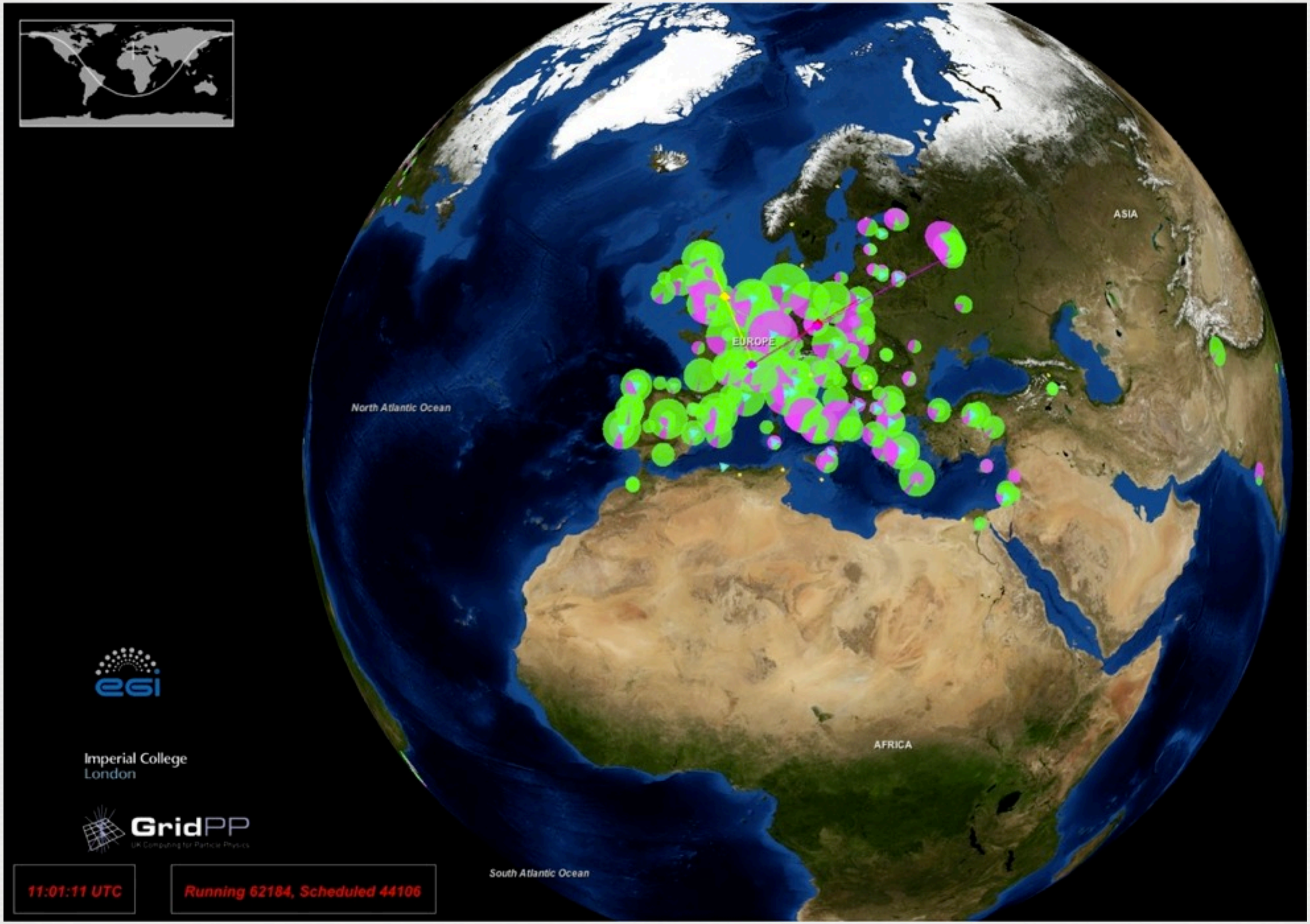


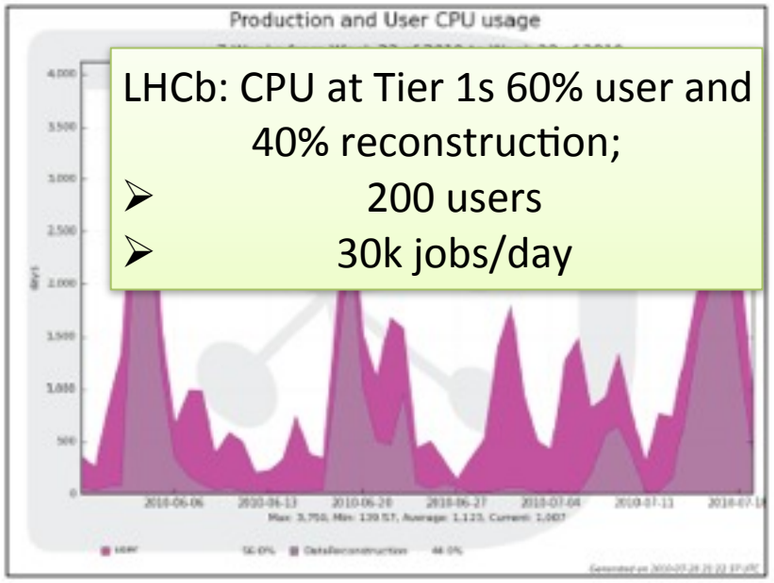






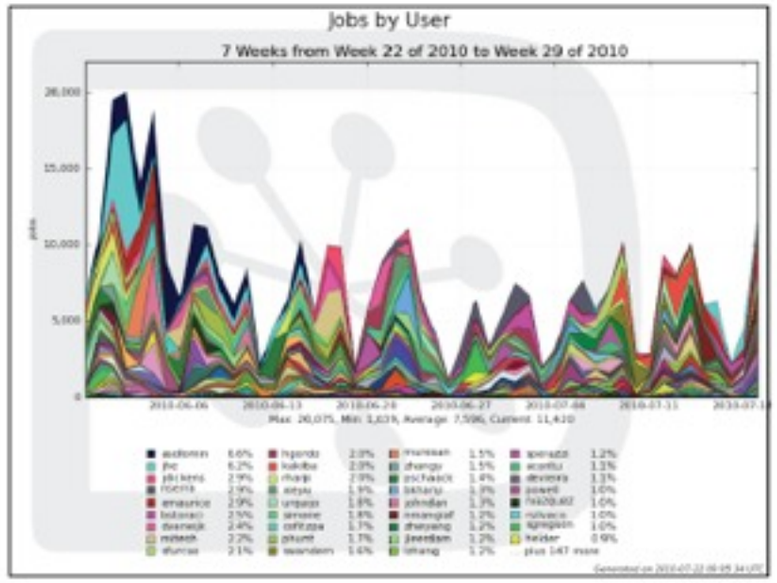
- The idea of using the Grid for HEP was launched at CHEP2001 at Padova
- The Grid has been one of the greatest successes of LHC
- At the first MONARCH meetings we had no idea of how we could bring different sites to work together
- Now it is an everyday reality, and it is quickly becoming “simply” a physicist’s working tool

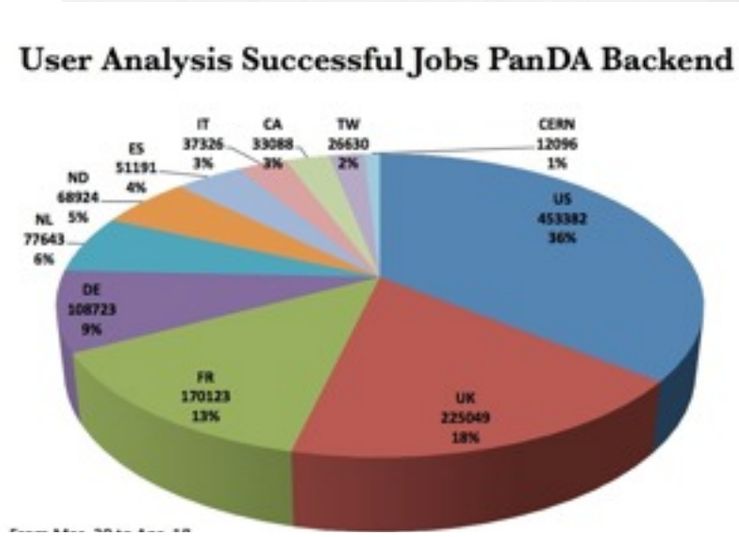

EUROPE
ASIA
North Atlantic Ocean
AFRICA
South Atlantic Ocean
Imperial College
London
GridPP
UK Computing for Particle Physics
11:01:11 UTC
Running 62184, Scheduled 44106

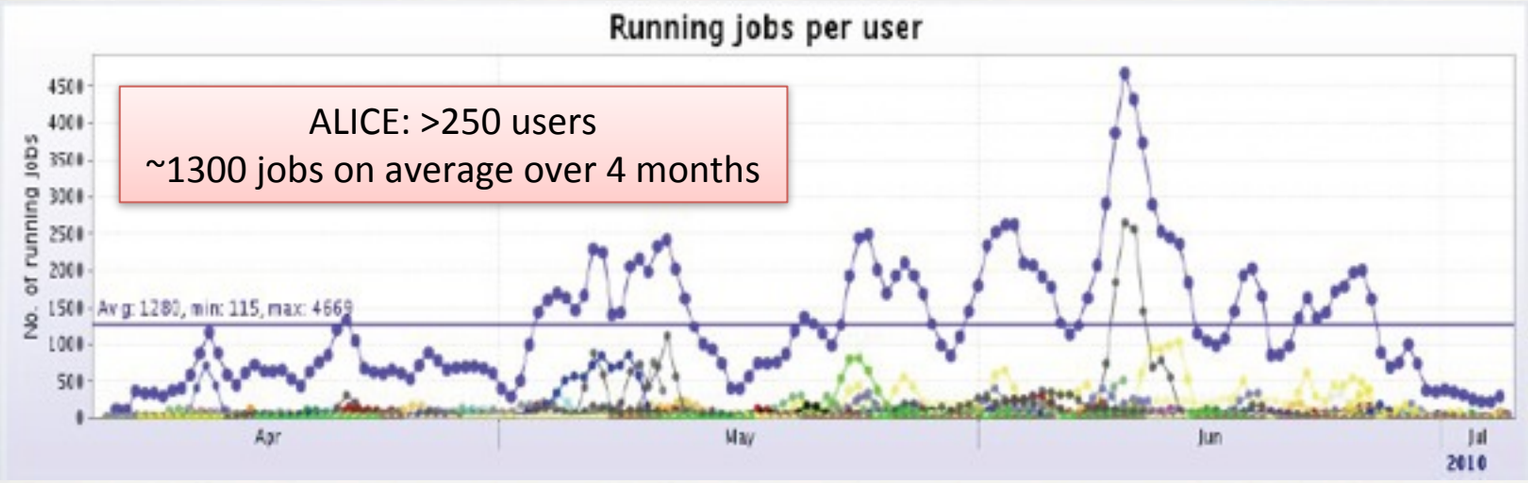


LHCb: CPU at Tier 1s 60% user and 40% reconstruction;

- 200 users
- 30k jobs/day

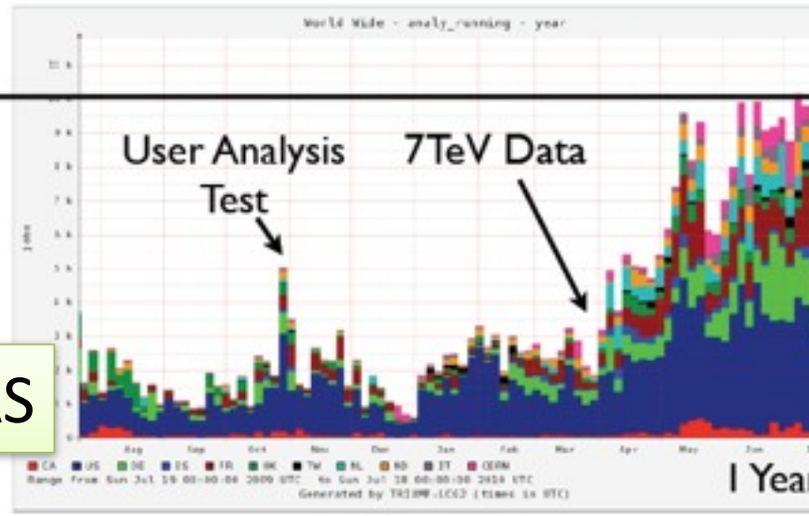




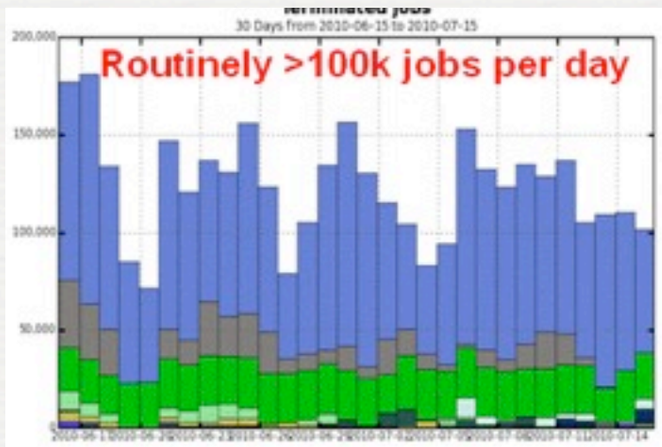


ALICE: >250 users
~1300 jobs on average over 4 months

10k running jobs



ATLAS

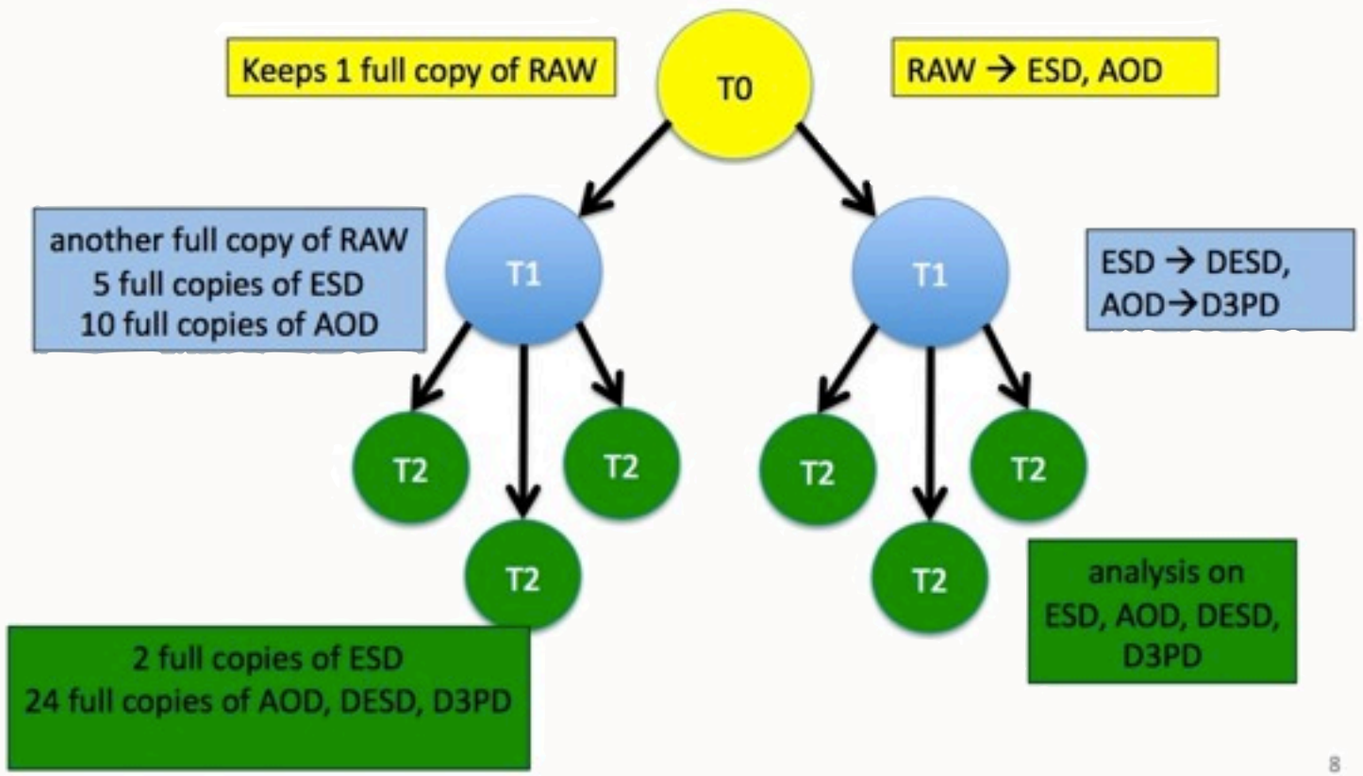


CMS

GRID-based analysis in June-July 2010:
>1000 different users, ~ 11 million analysis jobs processed

THE MONARC MODEL

- The Monarc model was designed at the end of the last century based on a "rigid" distribution of tasks between centres of different size and role



THE ROLE OF THE T1 IN MONARC

- The T1 has an important role as
  - Data custody
  - Data serving to a number of “dependent” T2-3 centres
  - Support and consultancy for “dependent” T2-3 centres
- Most of the LHC Grid has been build this way

The ALICE Grid

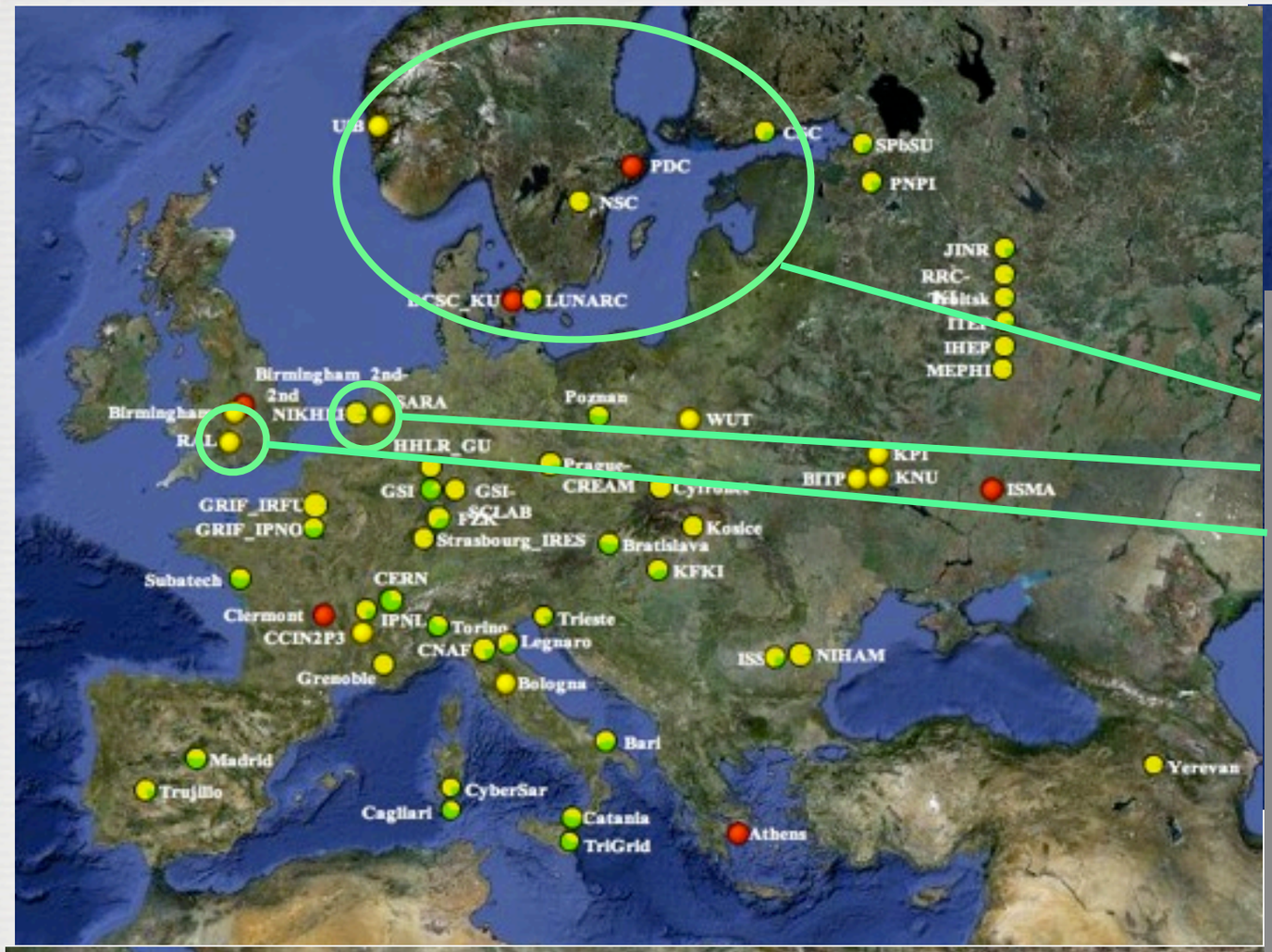
The ALICE Grid

The ALICE Grid

T1
NorduGrid
NIKHEF/SARA

The ALICE Grid

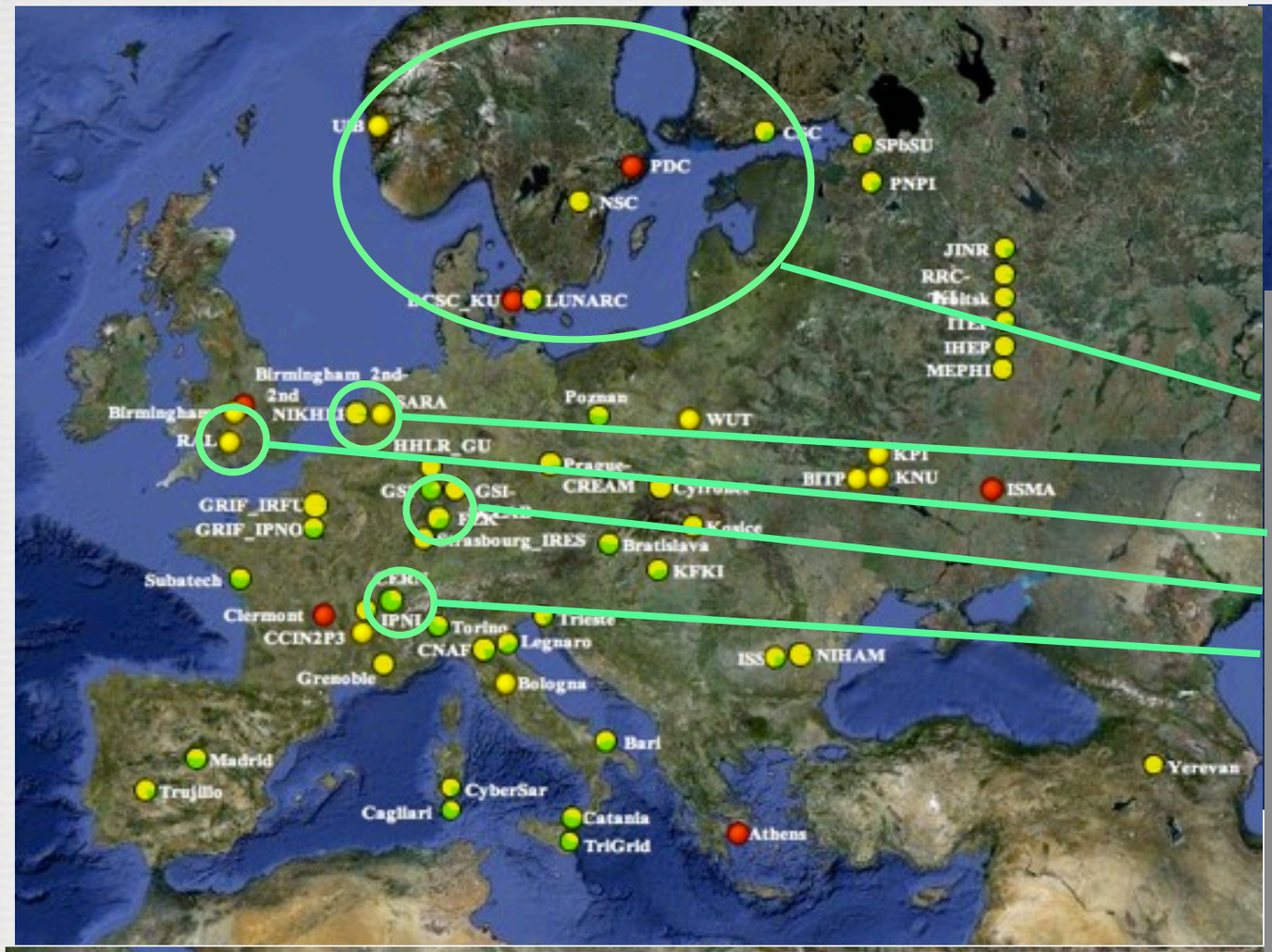
T1
NorduGrid
NIKHEF/SARA
RAL



The ALICE Grid

T1
NorduGrid
NIKHEF/SARA
RAL
FZK

The ALICE Grid



T1
NorduGrid
NIKHEF/SARA
RAL
FZK
CERN

The ALICE Grid

T1
NorduGrid
NIKHEF/SARA
RAL
FZK
CERN
CCIN2P3

The ALICE Grid

T1
NorduGrid
NIKHEF/SARA
RAL
FZK
CERN
CCIN2P3
CNAF

The ALICE Grid

T1
NorduGrid
NIKHEF/SARA
RAL
FZK
CERN
CCIN2P3
CNAF
UNAM

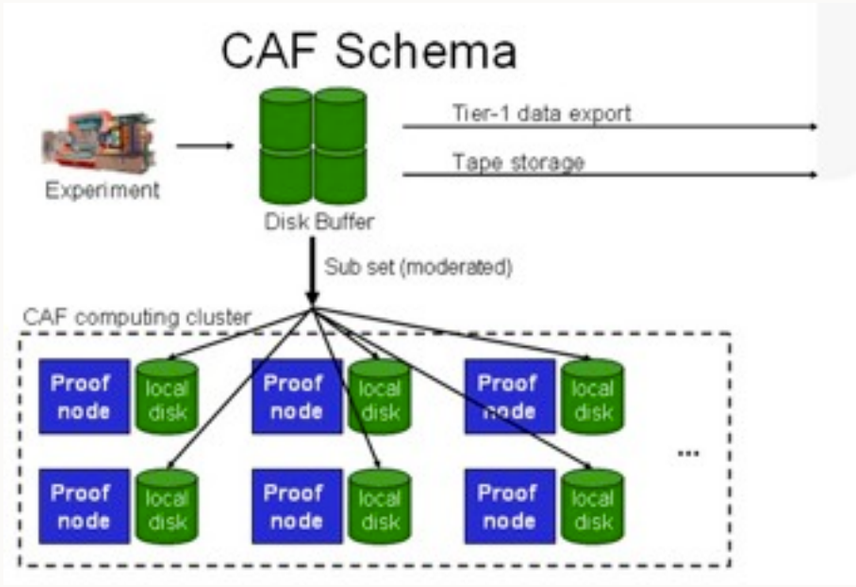
The ALICE Grid

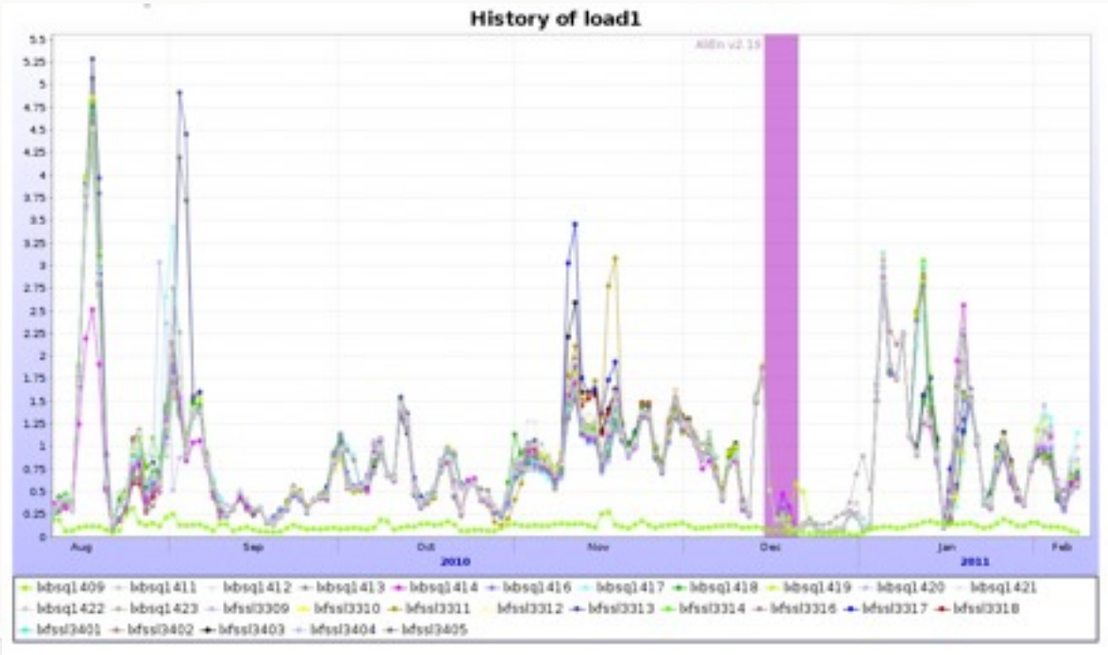
T1
NorduGrid
NIKHEF/SARA
RAL
FZK
CERN
CCIN2P3
CNAF
UNAM
KISTI

The ALICE Grid

THE ALICE ANALYSIS FACILITIES

- Proof-enabled, Grid-aware parallel computing platform
- Used for early discovery physics, calibration
- “Victim of its own success” has doubled twice in the last year at CERN, 480 cores in few days





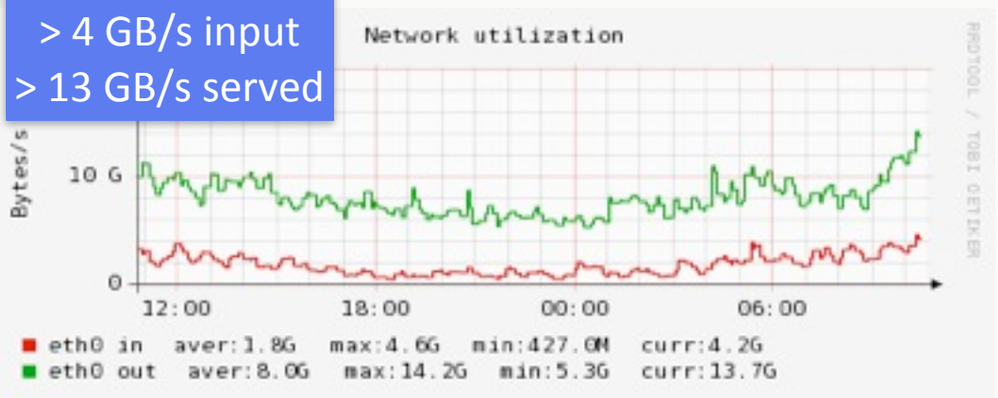
T1 OBLIGATIONS (WLCG MOU)

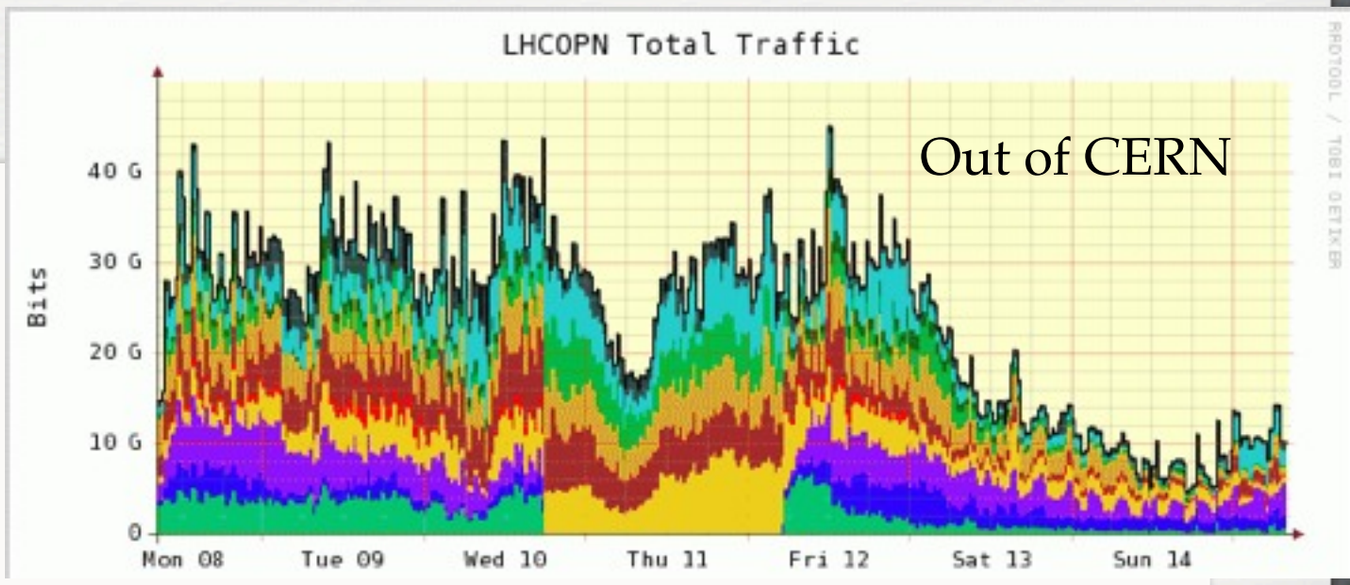
Service	*Maximum delay in responding to operational problems*			*Average availability measured on an annual basis*	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outwith accelerator operation	24 hours	48 hours	48 hours	n/a	98%
All other services – prime service hours[6]	2 hour	2 hour	4 hours	98%	98%
All other services – outwith prime service hours	24 hours	48 hours	48 hours	97%	97%

THE GRID – DATA TRANSFER

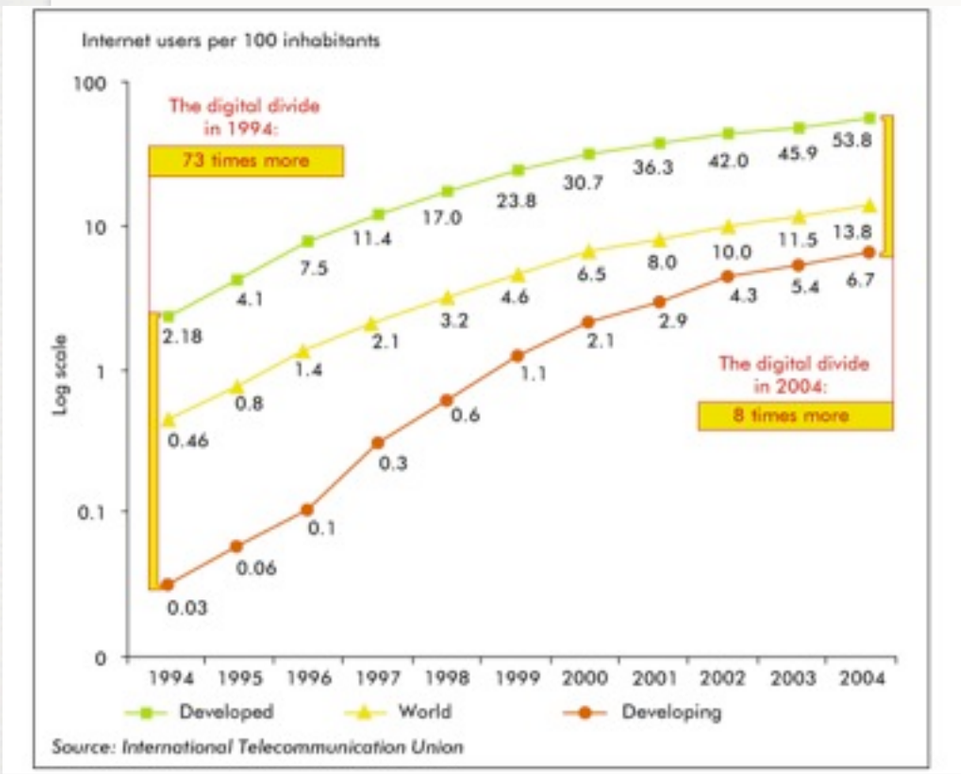
- Data transfer has been especially successful
  - Out of CERN has peaked above 1GB
- Transfer between centres also very good

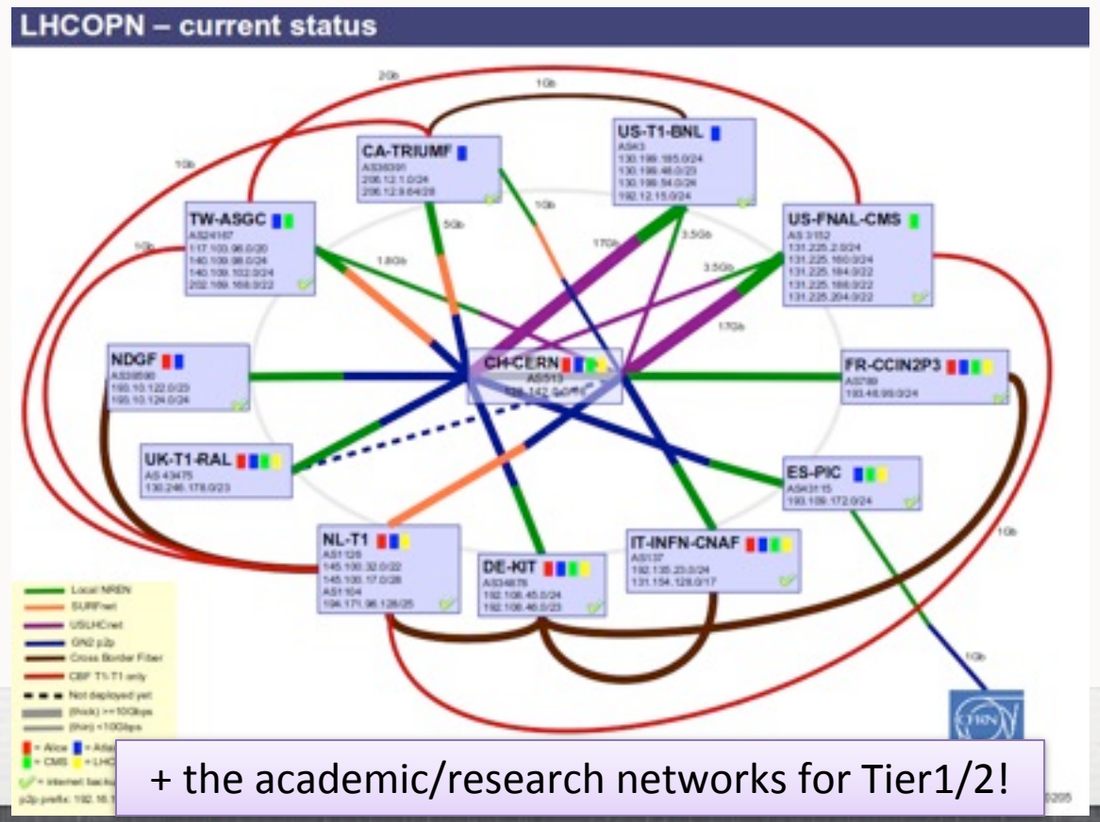
Tier 0 traffic:
> 4 GB/s input
> 13 GB/s served





- The network is probably the best surprise here
- Still the least oversubscribed resource we have





+ the academic/research networks for Tier1/2!

At present, institutes in Europe typically have a ~ 1 Mb/s access to CERN. In some places the available bandwidth is already as high as 622 Mb/s. We expect Gb/s networks to be available by the beginning of LHC operation. This assumes an increase of a factor of ~ 100, which is typical of the improvements in the technology over a ten-year period. However, current price trends would imply that achieving this performance would require an increase in network funding.

CERN/LHCC/94 43
LHCC/P2
15 December 1994

Technical Proposal

for a

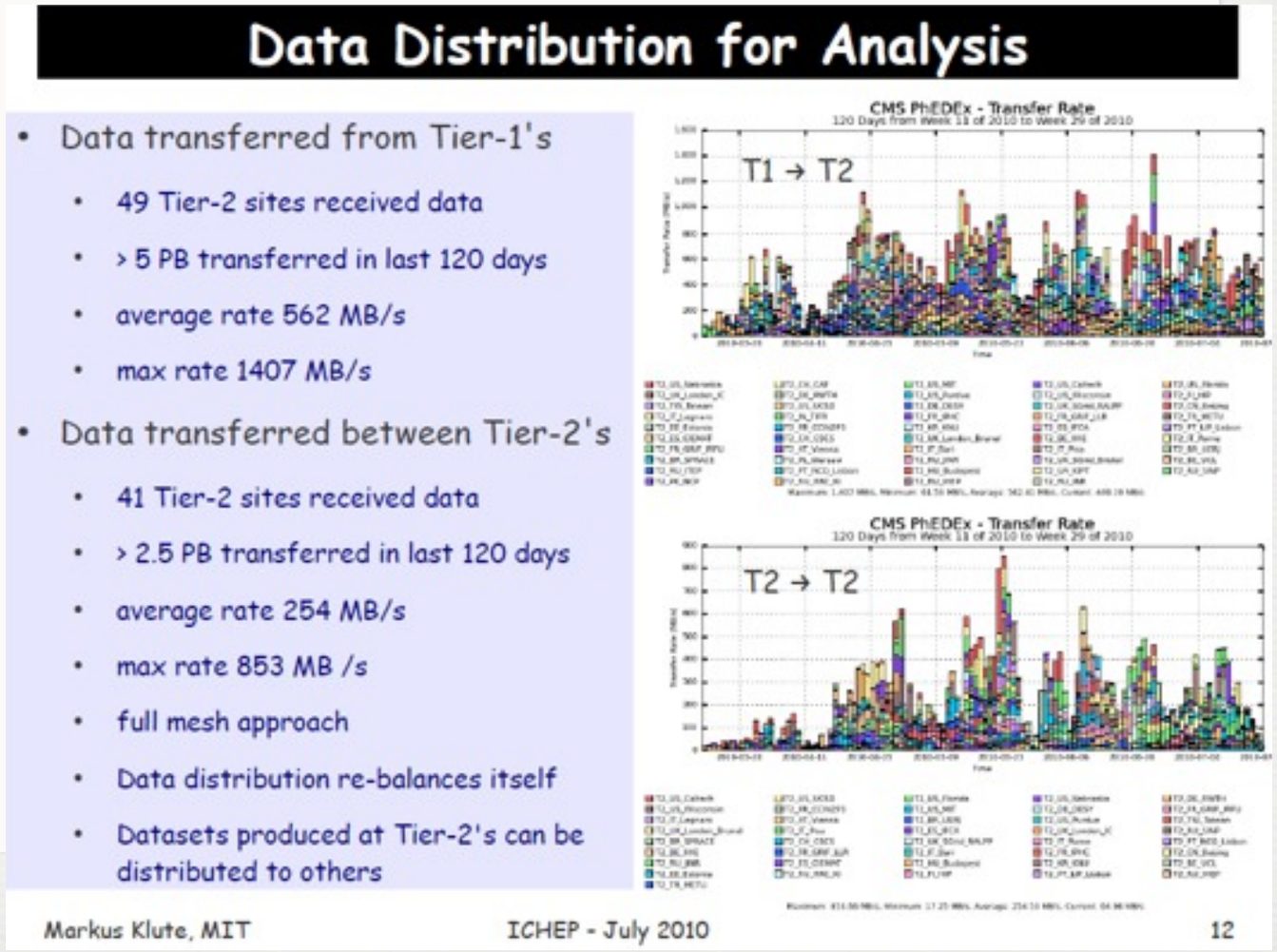
General-Purpose pp Experiment

at the

Large Hadron Collider at CERN

DATA IS STILL THE PROBLEM

- Data placement is the main problem, particularly for analysis
  - "predictive" data placement (ATLAS & CMS) or "opportunistic" (ALICE – need single catalogue)
  - Data distribution "per se" works very well
- With "infinite" disk space the two are equivalent
- If we distributed data "generously", deleting them is a real trouble
- Increasing the disk more difficult than increasing CPU
- Quotas & monitoring more difficult for data than CPU

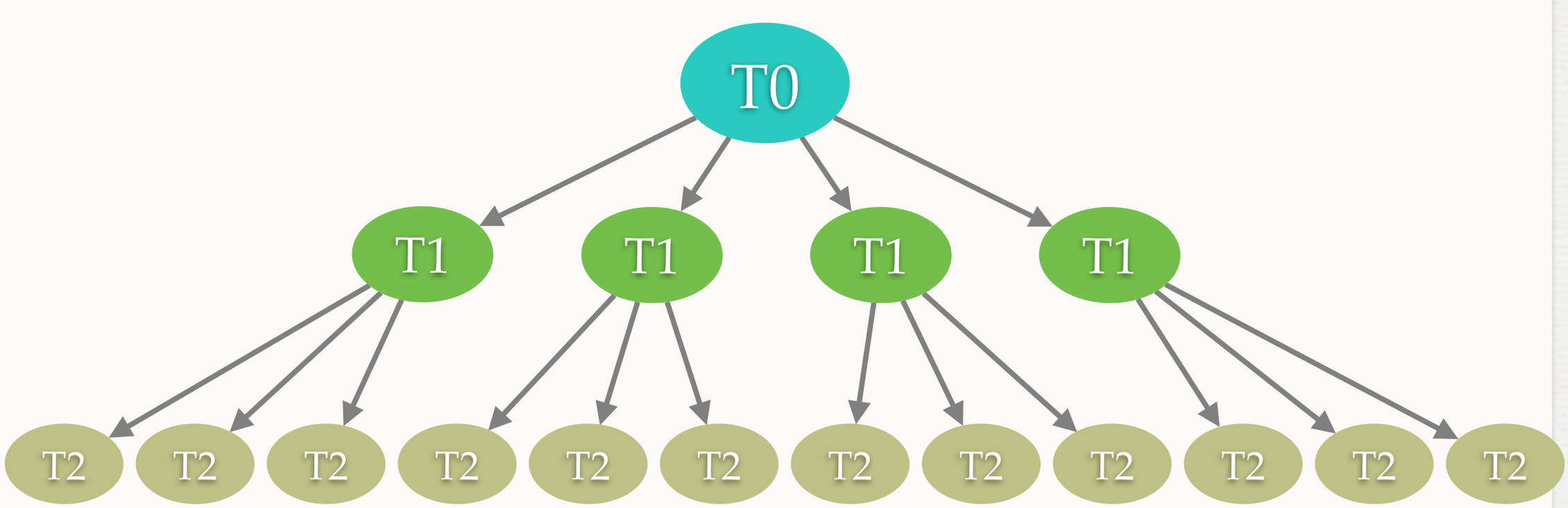


DESTITUTION OF THE MONARC

- Given the good performance of the network and the issues with data placement, the Monarc model is evolving from Grid to Cloud

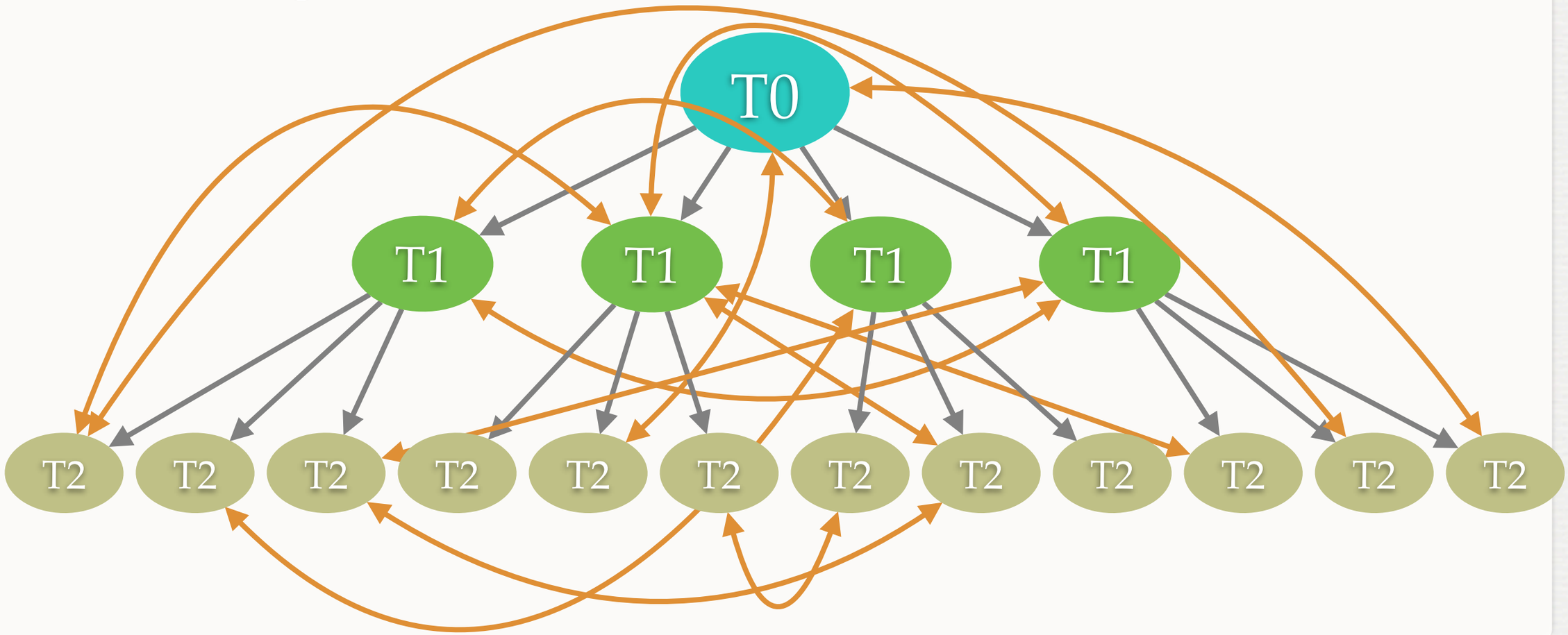
DESTITUTION OF THE MONARC

- Given the good performance of the network and the issues with data placement, the Monarc model is evolving from Grid to Cloud



DESTITUTION OF THE MONARC

- Given the good performance of the network and the issues with data placement, the Monarc model is evolving from Grid to Cloud



A CLOUD OVER THE GRID

- T0-1-2-3 hierarchy tends to be softened by user-driven data placement and transfer
  - T1 and T2 are becoming equivalent in the network (OPNng)
- No longer disk space but network bandwidth will scale with #users and #data

SO WHY A T1?

- Running a T1 means tackle important challenges of providing reliable and continual service to the worldwide community
- Being well maintained and highly available centres, the T1s are at the forefront of the current evolution of the Grid toward the Cloud and beyond
- Running at T1 is being at the edge of this evolution, in an ideal position to observe and influence it

WHAT ABOUT THE LOCAL COMMUNITY

- In the ALICE model resources are not dedicated
  - All jobs run in all the centre
- In this sense a T1 is a “contribution” to the whole community
  - However the proximity of highly experienced and skilled professionals is an important advantage for the local community
  - The establishment of an AAF system supported by a T1 infrastructure will be a big advantage for the local community

WHAT ABOUT RESOURCES

- A T1 is not about resources, it is about competence and service
  - It is more about people than machines
- However a T1 service makes sense only beyond a certain level
  - We expect a T1 to have few thousands cores
- And our experience is that 4 FTE/1000 cores is a reasonable order of magnitude
- Network provisioning will be of course of paramount importance

CONCLUSIONS

- We are at the beginning of a nice adventure, and I ALICE is looking forward to work with you
- Commitment and development of human potential are key to the success
- So good luck and let's go!


"That's
all
folks!"