CMS Experiment and GLORIAD

(CMS Data Transfer using Global Network)

Suh, Jun-Suhk
CHEP, KNU

Contents

- Large Hadron Collider Experiment
  - CMS Experiment
- LHC Computing Grid
  - LCG and Global Ring Network
- CMS Data Challenges
- Summary

Particle Physics

- Aim to answer the two following questions
  - What are the elementary constituents of matter?
  - What are the fundamental forces that control their behavior at the most basic level?

Particles & Forces

Leptons

	Electric Charge		Electric Charge
Tau	-1	Tau Neutrino	0
Muon	-1	Muon Neutrino	0
Electron	-1	Electron Neutrino	0

Quarks

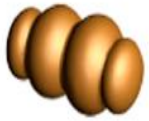
	Electric Charge		Electric Charge
Bottom	-1/3	Top	2/3
Strange	-1/3	Charm	2/3
Down	-1/3	Up	2/3

each quark: R, B, G 3 colors

The particle drawings are simple artistic representations

Strong

Gluons (8)



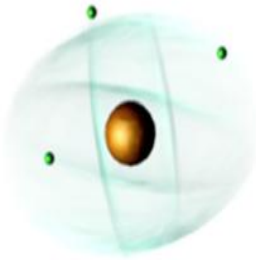
Quarks

Mesons
Baryons

Nuclei

Electromagnetic

Photon



Atoms
Light
Chemistry
Electronics

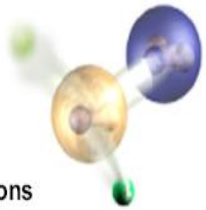
Gravitational

Graviton ?

Solar system
Galaxies
Black holes

Weak

Bosons (W,Z)

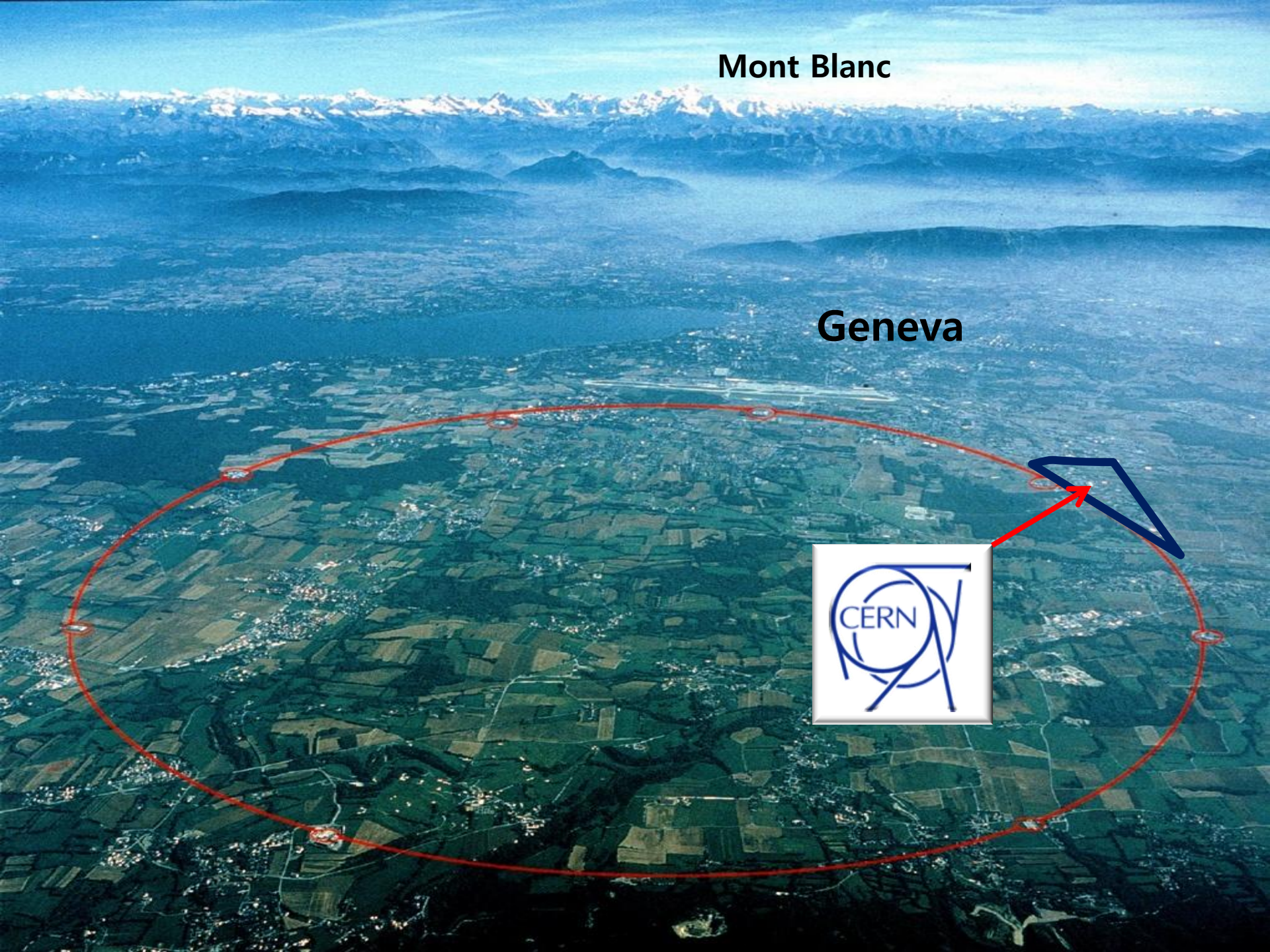


Neutron decay
Beta radioactivity
Neutrino interactions
Burning of the sun

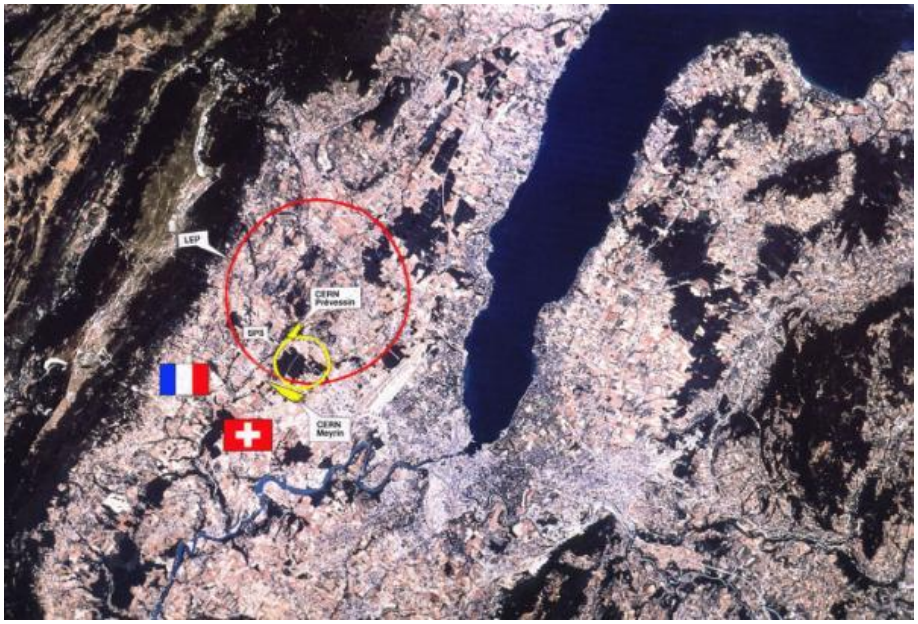
The particle drawings are simple artistic representations

LHC

- The Large Hadron Collider (LHC), currently being built at CERN near Geneva, is the largest scientific instrument on the planet.
- LHC will produce roughly 15 Petabytes of data annually.
- Thousands of scientists around the world will access and analyse.


Mont Blanc
Geneva
CERN

The LHC Machine



- **27 km ring of superconducting magnets operating at 1.9° Kelvin – colder than outer space**
- **Colliding proton beams travel at 99.999999991% the speed of light**
- **Collisions at 7 TeV + 7 TeV generate temperatures a billion times hotter than the heart of the sun**
- **The beam has the energy of a loaded jumbo jet at full speed**
- **Cost of the accelerator - ~3,000M €**

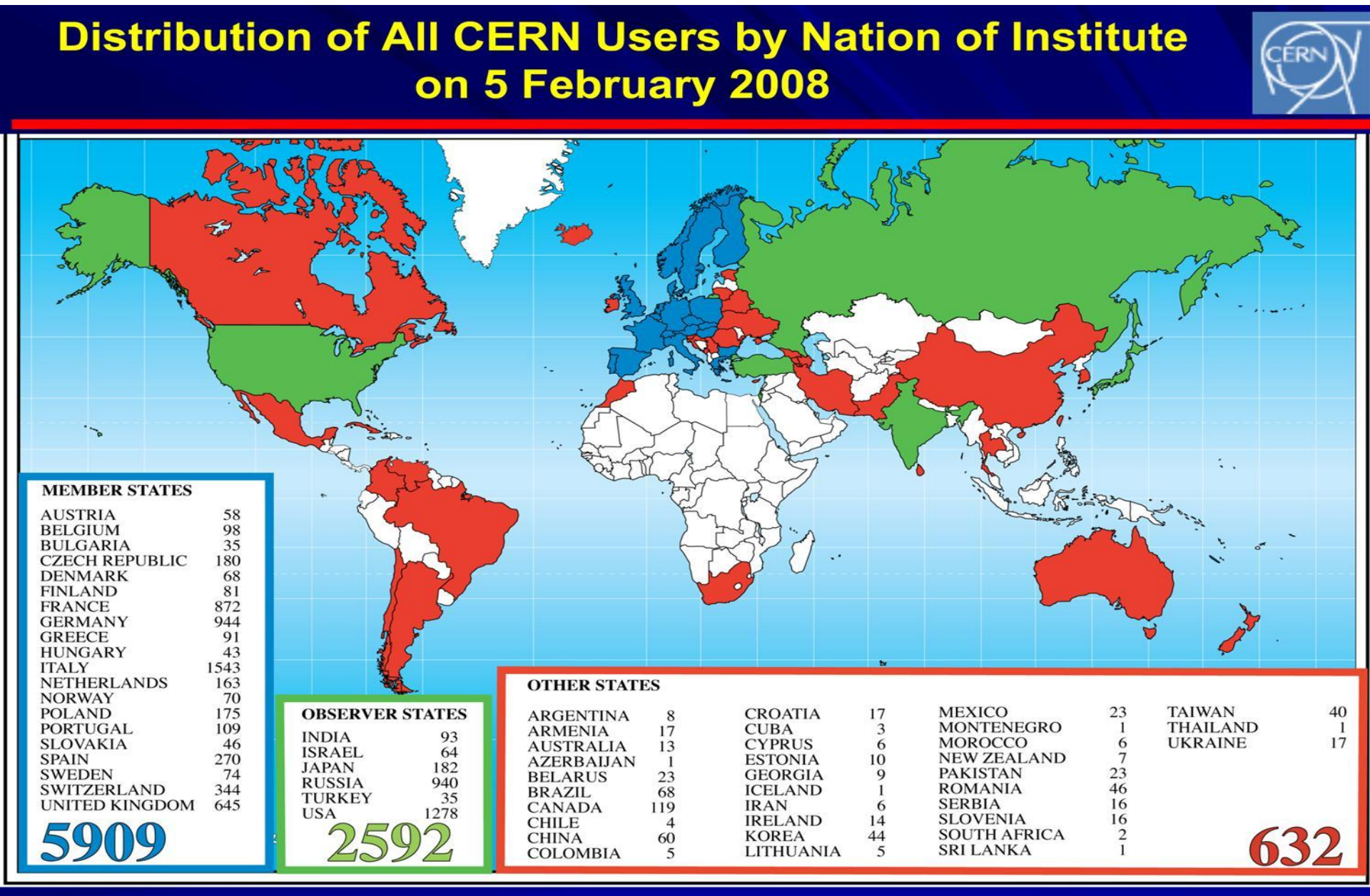
CERN - The European Organization for Nuclear Research

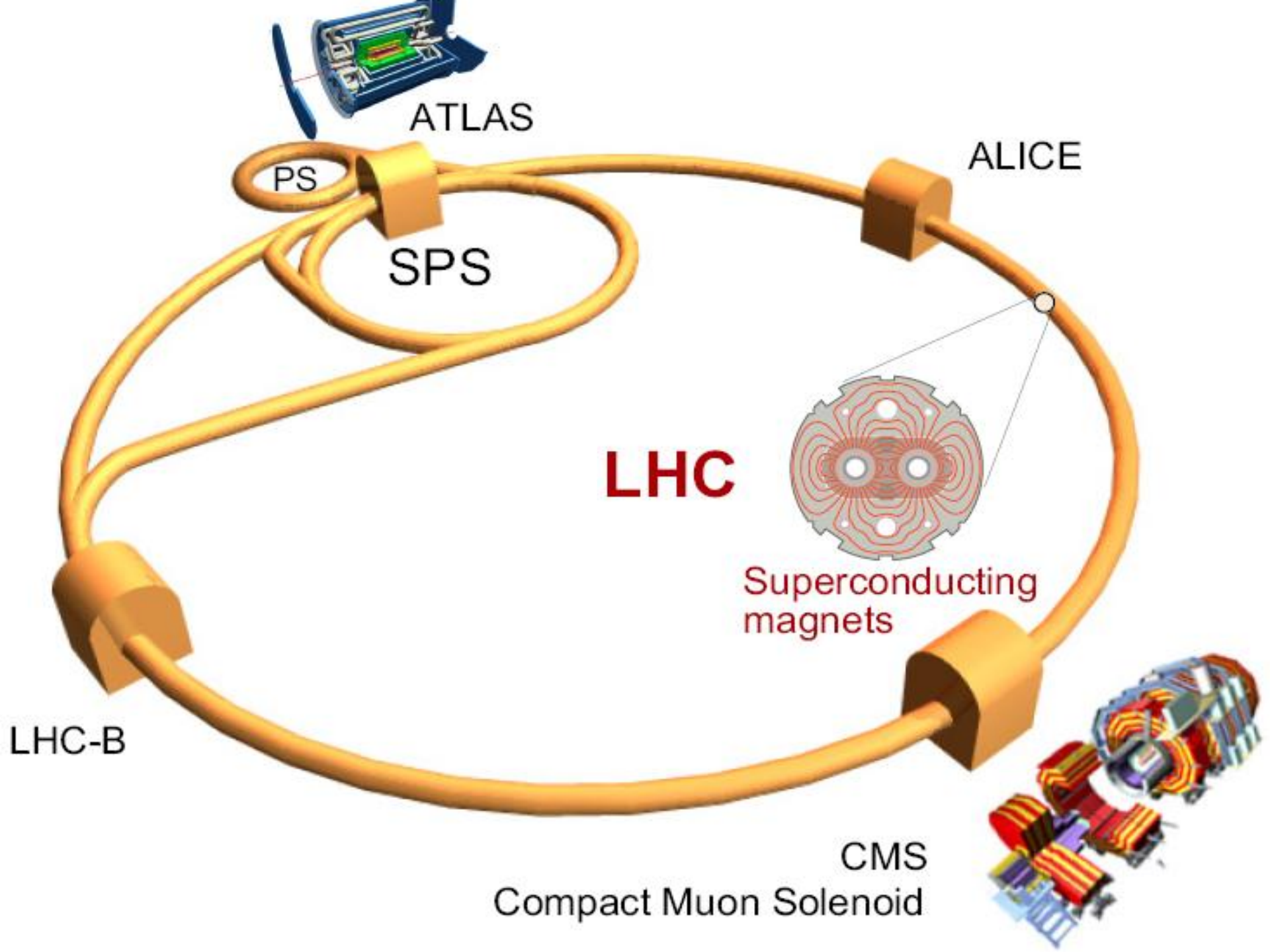
- Twenty Member States:

Austria	Belgium	Bulgaria	Czech Republic
Denmark	Finland	France	Germany
Greece	Hungary	Italy	Netherlands
Norway	Poland	Portugal	Slovak Republic
Spain	Sweden	Switzerland	United Kingdom

- Plus eight Observer States: European Commission, India, Israel, Japan, Russian Federation, Turkey, UNESCO and USA
- Budget: 1,200 MCHF (~800M €)
- Personnel: 2,600 Staff, 700 Fellows and Associates, 8,000 Users

High Energy Physics: a Global Community




ATLAS
PS
SPS
ALICE
LHC
Superconducting
magnets
LHC-B
CMS
Compact Muon Solenoid

LHC experiments

ATLAS A Toroidal LHC Apparatus

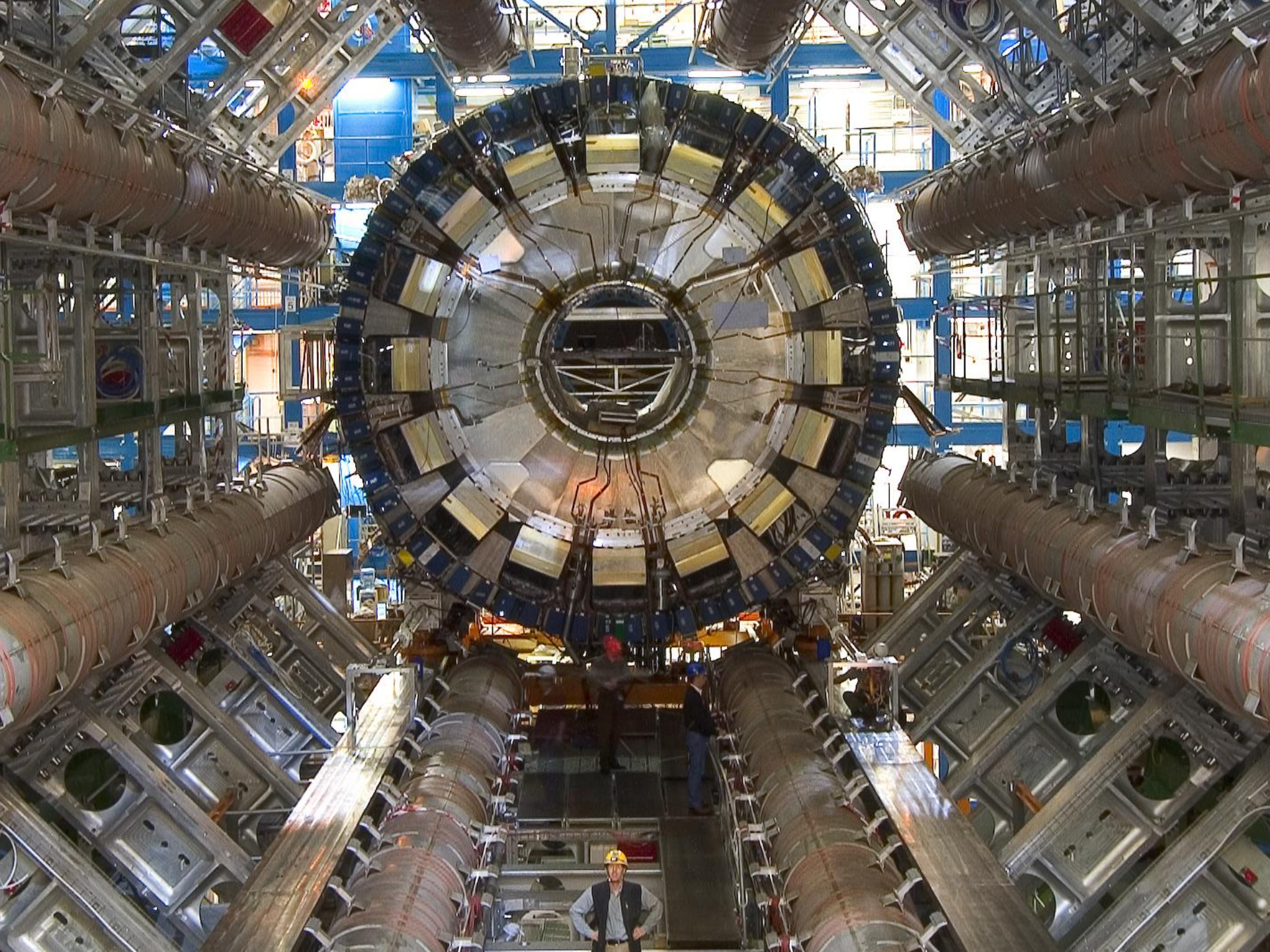
CMS Compact Muon Solenoid

LHCb LHC-beauty

ALICE A Large Ion Collider Experiment

TOTEM Total Cross Section, Elastic Scattering and Diffraction Dissociation

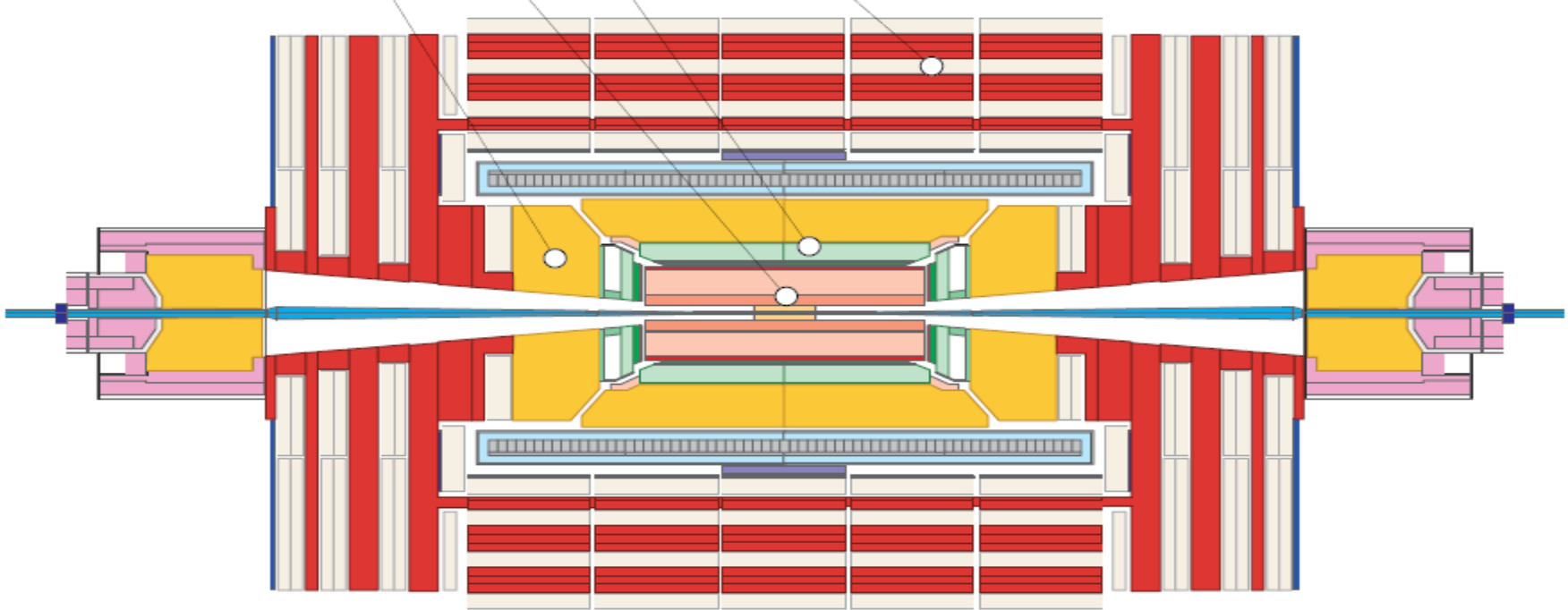
LHCf LHC-forward



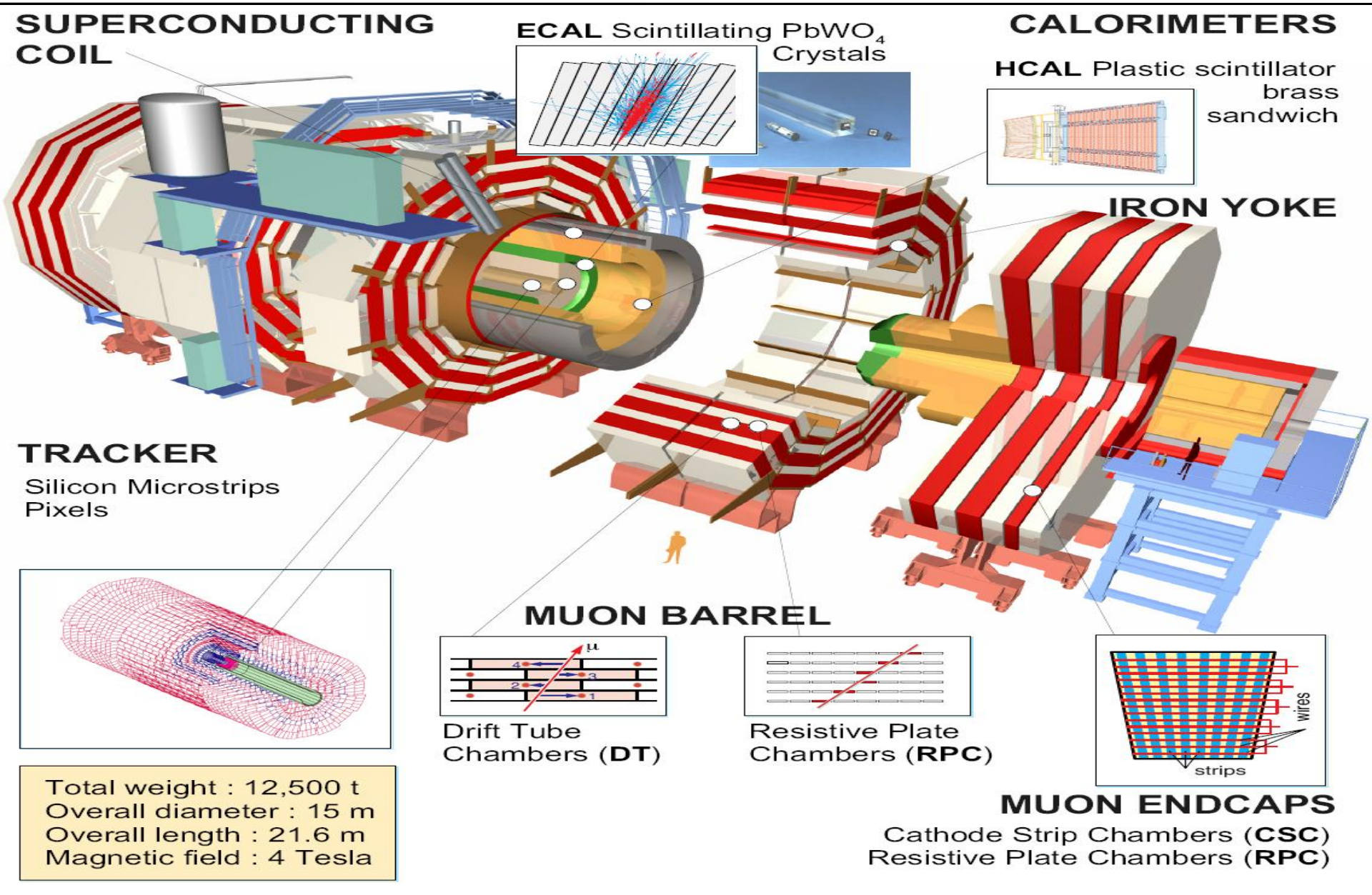
CMS Detectors

CMS is a general purpose proton-proton detector designed to run at the highest luminosity at the LHC. It is also well adapted for studies at the initially lower luminosities. The main design goals of CMS are:

1) a highly performant **muon system**
2) the best possible **electromagnetic calorimeter**
3) a high quality **central tracking**
4) a hermetic **hadron calorimeter**



CMS Detectors



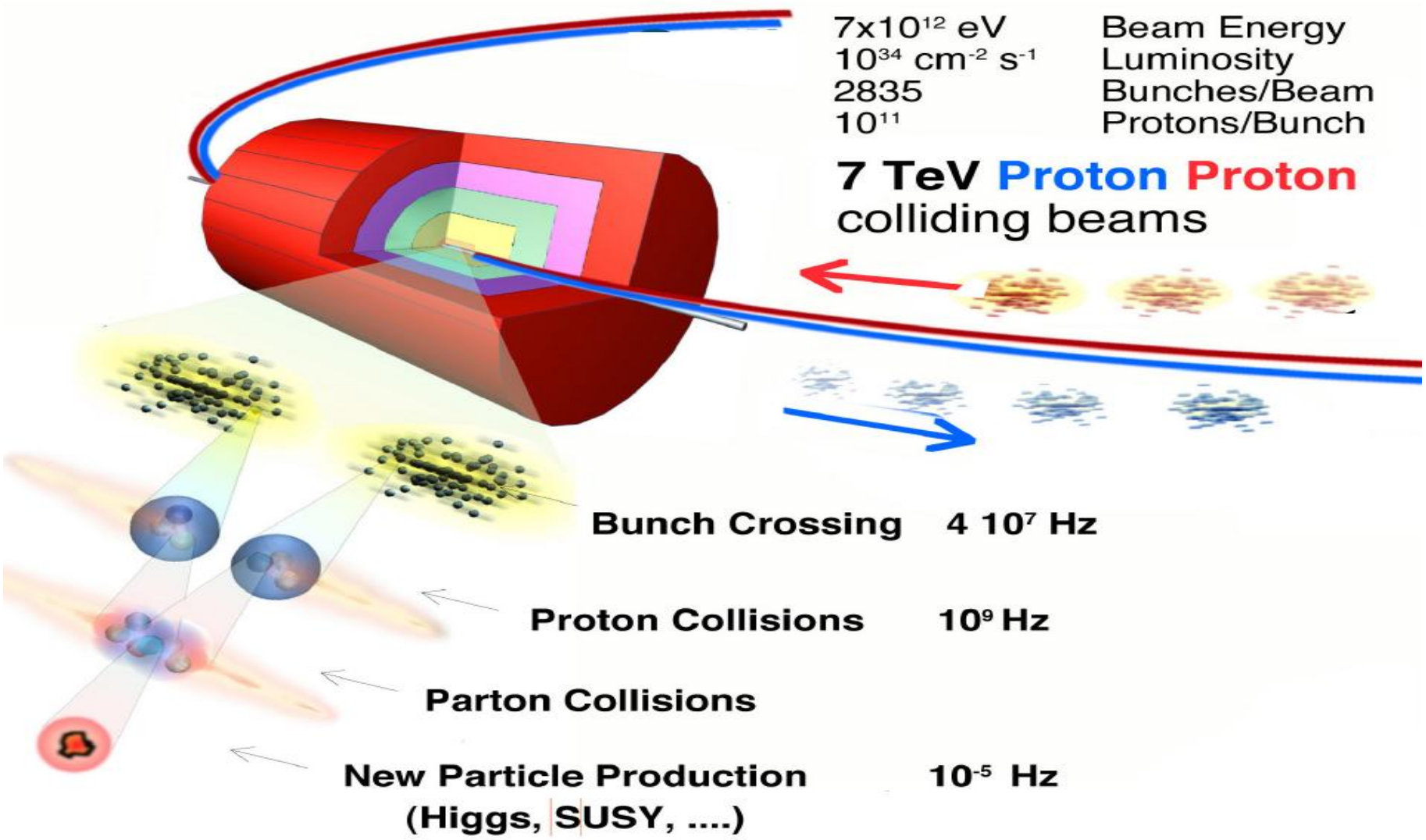
CMS Experiment

- The main goals of the CMS experiment are
  - to explore physics at TeV scale
  - to discover the Higgs boson
  - to look for evidence of physics beyond the standard model, such as supersymmetry, or extra dimensions
  - to study aspects of heavy ion collisions

CMS Physics Analysis Groups

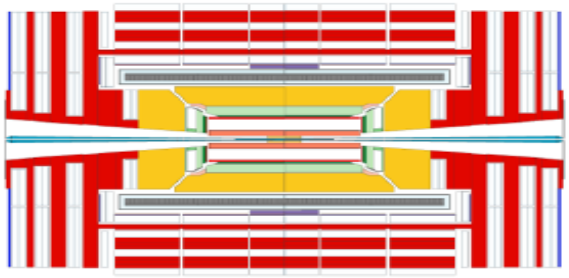
- Forward physics
- QCD
- Electroweak
- Top
- Higgs
- SUSY
- EXOTICA
- B-Physics
- Heavy Ions
- Generator Tools

Collisions at LHC



Selections of 1 in 10,000,000,000,000

CMS Trigger and Data Acquisition



COMMUNICATION

40 MHz
COLLISION RATE

Energy Tracks

100 kHz
LEVEL-1 TRIGGER

1 Terabit/s
(50000 DATA CHANNELS)

500 Gigabit/s

100 Hz
FILTERED EVENT

Gigabit/s
SERVICE LAN

PROCESSING

16 Million channels
3 Gigacell buffers

1 Megabyte EVENT DATA

200 Gigabyte BUFFERS
500 Readout memories

EVENT BUILDER. A large switching network (512+512 ports) with a total throughput of approximately 500 Gbit/s forms the interconnection between the sources (Readout Dual Port Memory) and the destinations (switch to Farm Interface). The Event Manager collects the status and request of event filters and distributes event building commands (read/clear) to RDPMs

5 TeraFLOP
EVENT FILTER. It consists of a set of high performance commercial processors organized into many farms convenient for on-line and off-line applications. The farm architecture is such that a single CPU processes one event

Petabyte ARCHIVE

LHC computing

- Signal/Background = 10^{-9}

- Data volume:

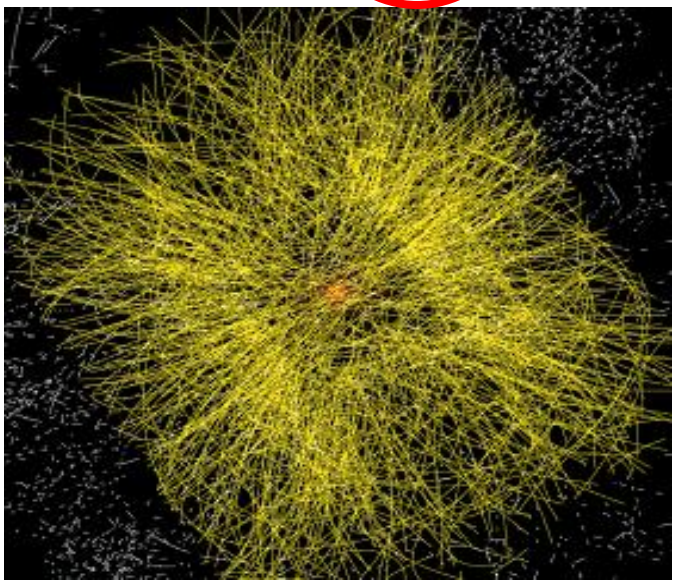
  High rate * large number of channels * 4 experiments

  ➔ **15 PetaBytes of new data each year**

- Compute power

  Event complexity * Number of events * thousands of users

  ➔ **100 k processor cores**



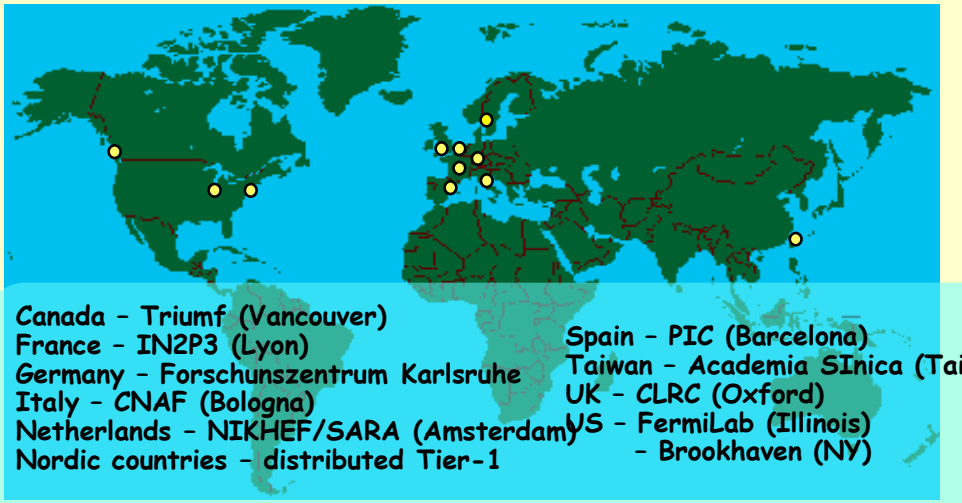
Worldwide LHC Computing Grid

- Discovering new fundamental particles and fields and analysing their properties with the LHC accelerator is possible only through statistical analysis of the massive amounts of data gathered by the LHC detectors and detail ed comparison with compute-intensive theoretical simulations.
- The mission of the Worldwide LHC Computing Grid (LCG) project is to build and maintain a data storage and analysis infrastructure for the entire high energy physics community that will use the LHC.
- The data from the LHC experiments will be distributed around the globe, according to a four-tiered model.

LCG: 4 Experiments + ~140 Computer Centres

Tier-0 – the accelerator centre

- Data acquisition & initial processing
- Long-term data curation
- Distribution of data → Tier-1 centres



11 Tier-1 Centres – "online" to the data acquisition process → high availability

- Managed Mass Storage – → grid-enabled data service
- Data-heavy analysis
- National, regional support

Tier-2 – 120 Centres in 60 Federations in 35 countries

- **End-user (physicist, research group) analysis**
  – where the discoveries are made
- Produce MC data ~ same amount data from detector.

Motivations for using a Grid

- CERN's budget for physics computing was insufficient
- Easy parallelism, use of simple PCs, availability of high bandwidth international networking .... make it ***possible*** to extend the distributed architecture to the wide area ....

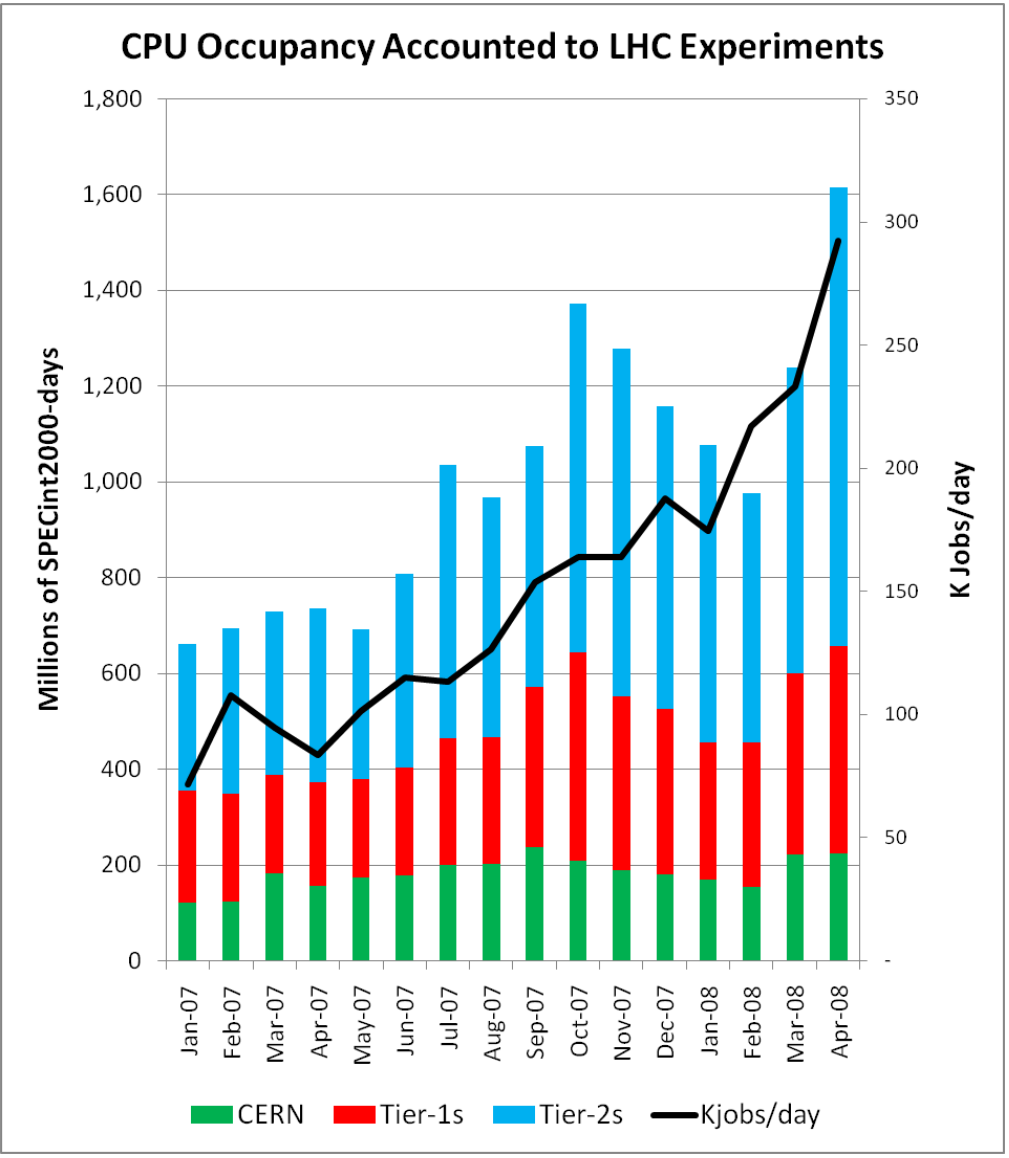
AND

- The ~5,000 LHC collaborators are distributed across institutes all around the world with access to local computing facilities, ...
  ... and funding agencies prefer to spend at home if they can
- Mitigates the risks inherent in the computing being controlled at CERN, subject to the lab's funding priorities and with access and usage policies set by central groups within the experiments

ALSO

- Active participation in the LHC computing service gives the institute (not just the physicist) a continuing and key role in the data analysis
  -- which is where the physics discovery happens
- Encourages novel approaches to analysis ....
  ... and to the provision of computing resources

Growth of LHC Grid CPU Usage



150% growth since Jan 07

~300K jobs/day

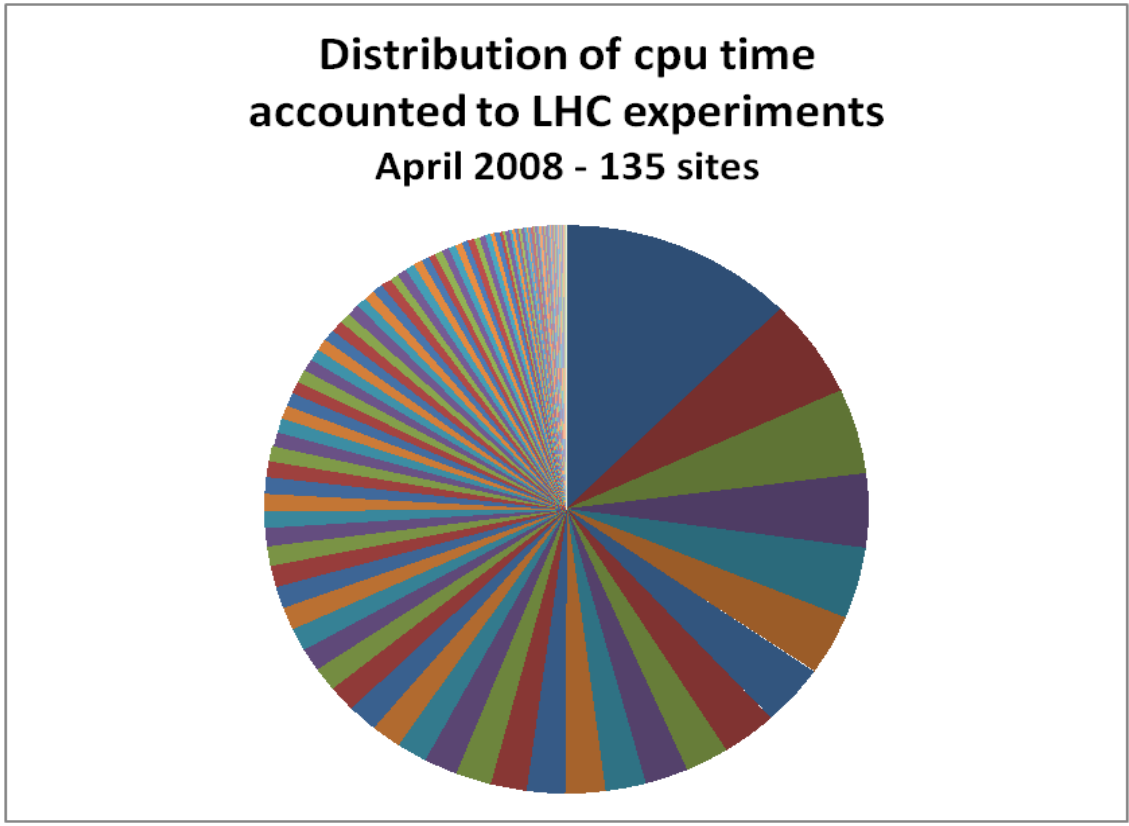
Equivalent of ~35K permanently occupied Intel cores

~50% of the full capacity committed for the 2008 run

Experiment services are still in test mode – awaiting the real data

LCG Accounting April 2008

CPU April 2008	
CERN	13%
Tier-1s	31%
Tier-2s	56%

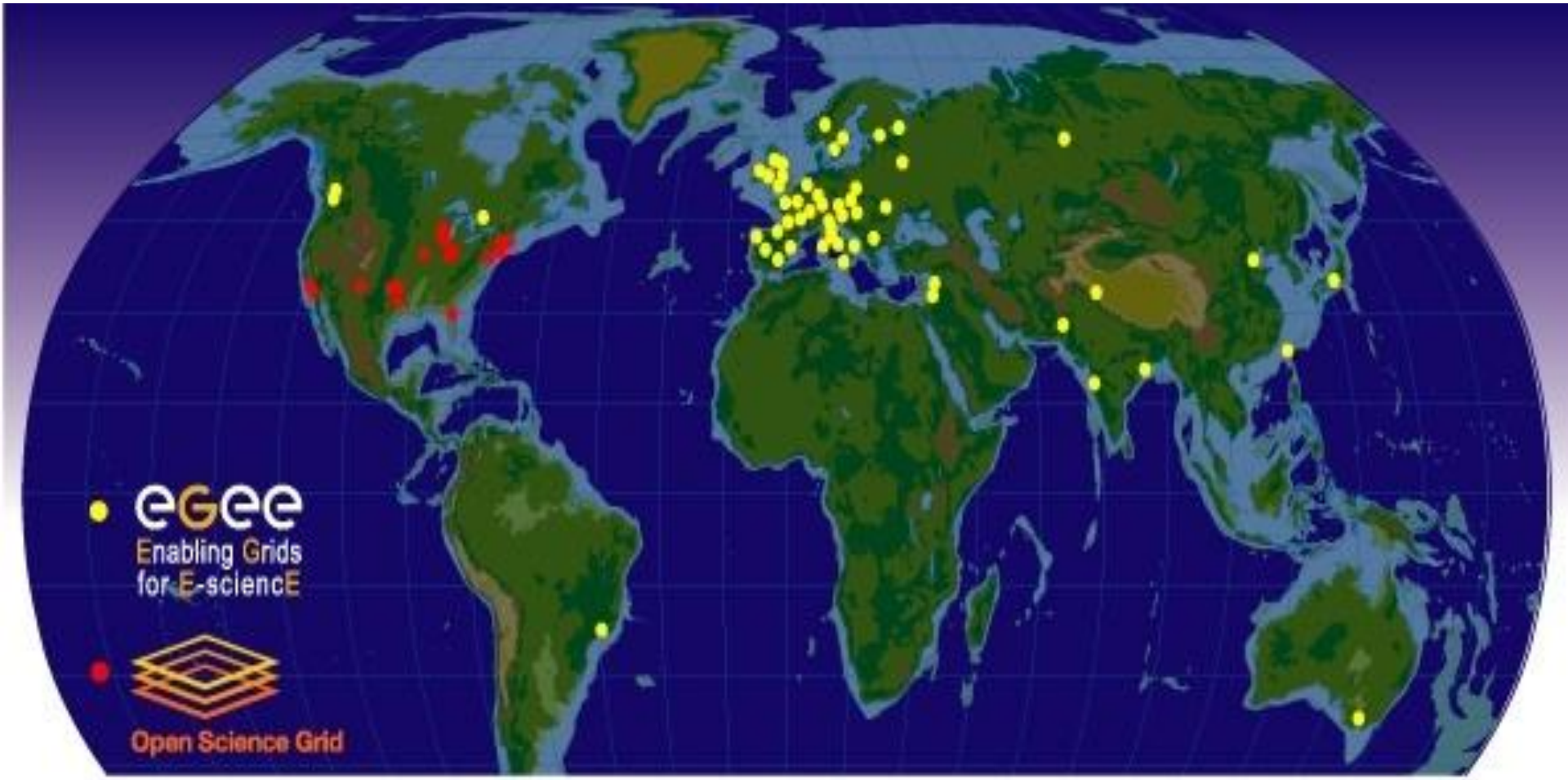


- Small number of very large contributors ..
  .. and a large number of very small contributors
- This was the aim
  - access capacity wherever it is available
- Also suggests that joining the Grid is not too difficult

% total	# sites
25%	4
50%	12
75%	29
90%	50
100%	135

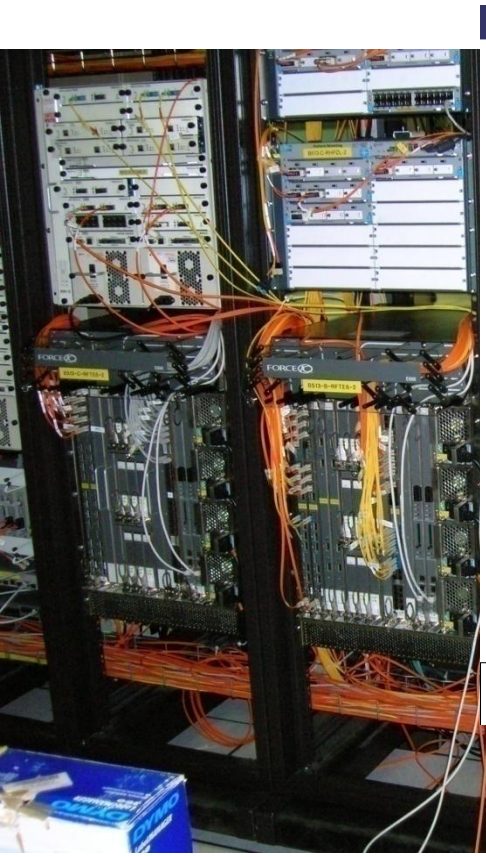
LCG is built on two infrastructure grids:

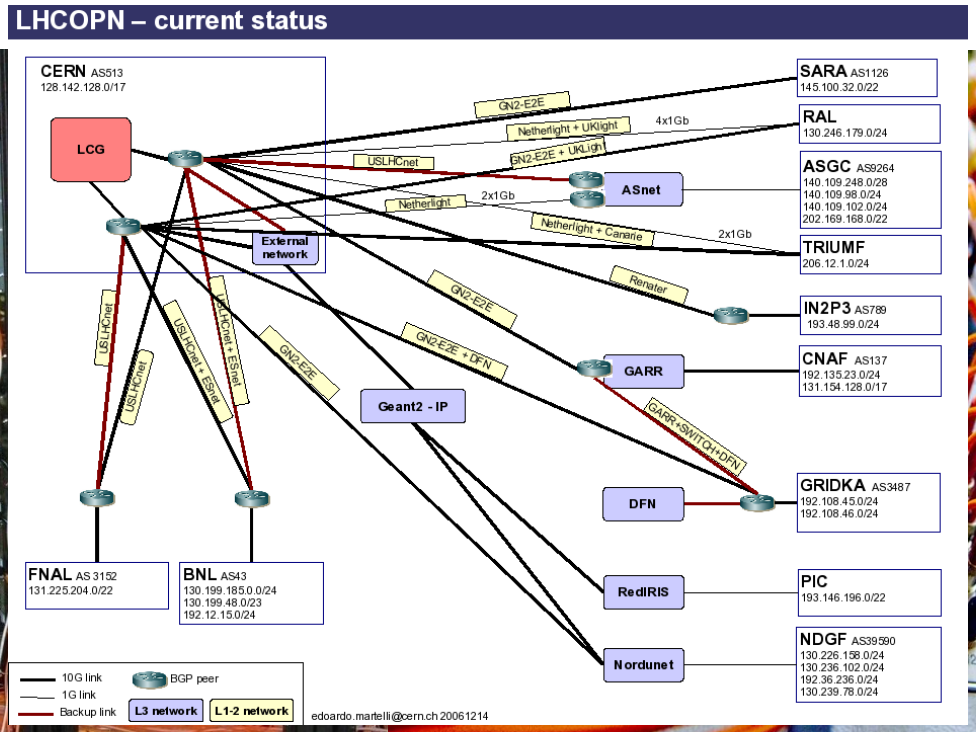
EGEE - Enabling Grids for E-Science
OSG - US Open Science Grid



LCG and Global Ring Network

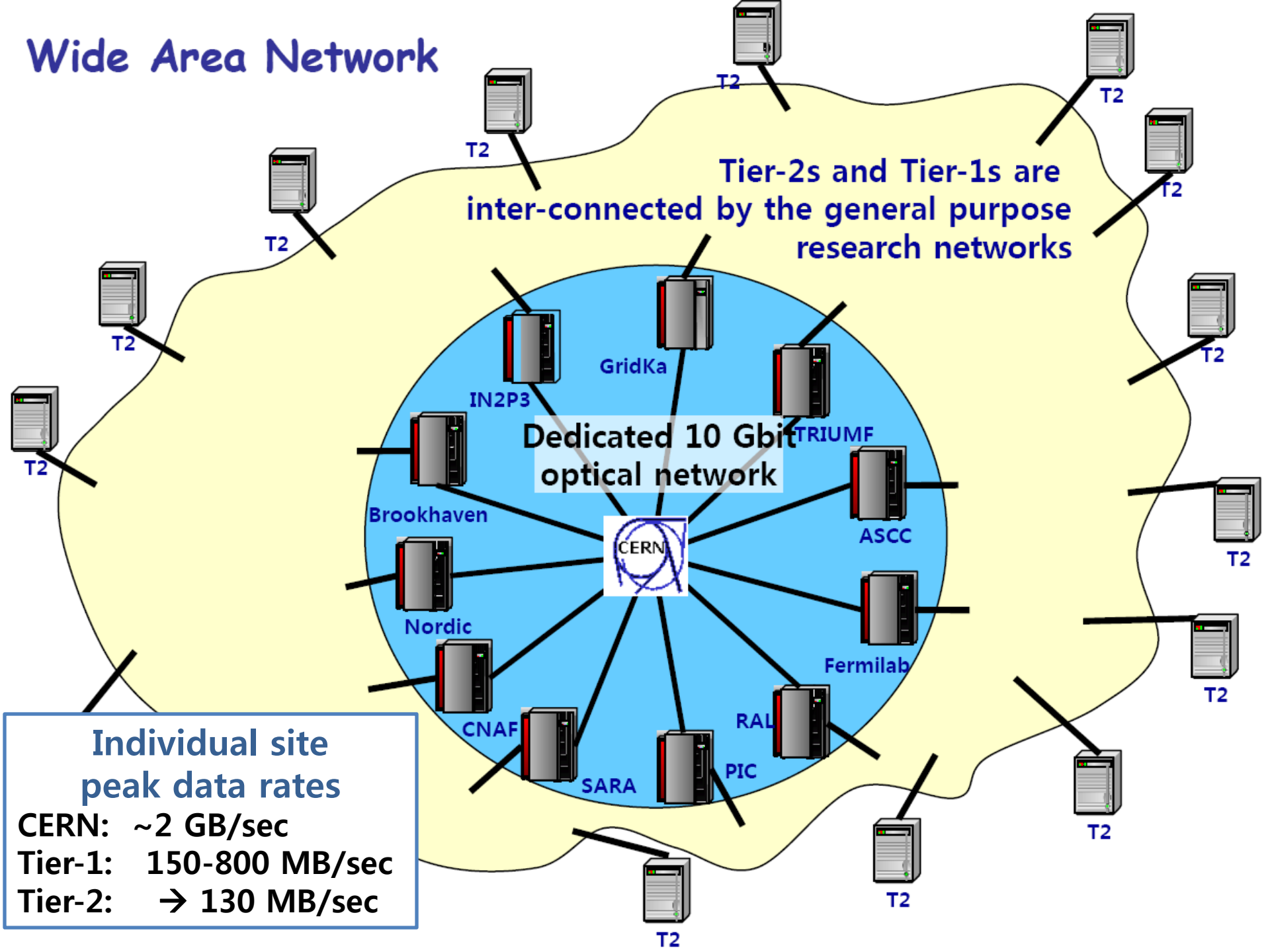
LHC → Networks → Physics





The critical role of networks for LCG

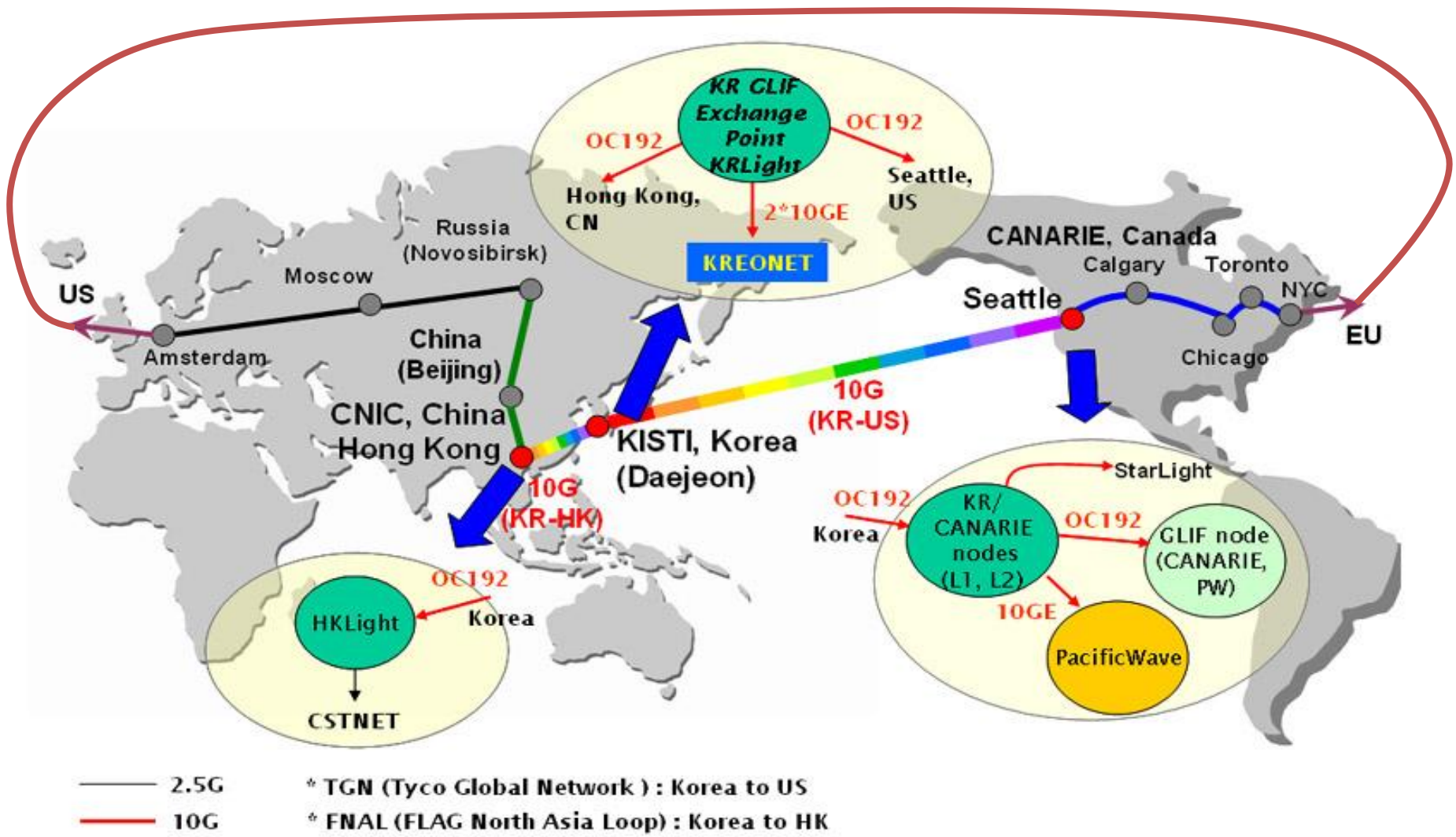
Wide Area Network



Individual site peak data rates

CERN: ~2 GB/sec
Tier-1: 150-800 MB/sec
Tier-2: → 130 MB/sec

Global Ring Network for Advanced Application Development



Global ring network plays an important role for the Worldwide LHC Computing Grid.

CMS Data Challenges

- **SC4**(2006. Jun.-Sep.): Service Challenge 4
- **CSA06**(Oct.-Nov.): Computing, Software, and Analysis challenge
- **Load Test 2007**(Feb.-May)
- **CSA07**(Sep.-Oct.)
- **CSA08**(**CCRC08** or **iCSA08**: Feb. & May) The Common Computing Readiness Chall enge of 2008

Service Challenge 4 Goals

- **Data Transferred from CERN to Tier-1 centers with PhEDEx through FTS**
  - ASGC: 10MB/s to tape
  - CNAF: 25MB/s to tape
  - FNAL: 50MB/s to tape
  - GridKa: 20MB/s to tape
  - IN2P3: 25MB/s to tape
  - PIC: 20MB/s to tape
  - RAL: 10MB/s to tape

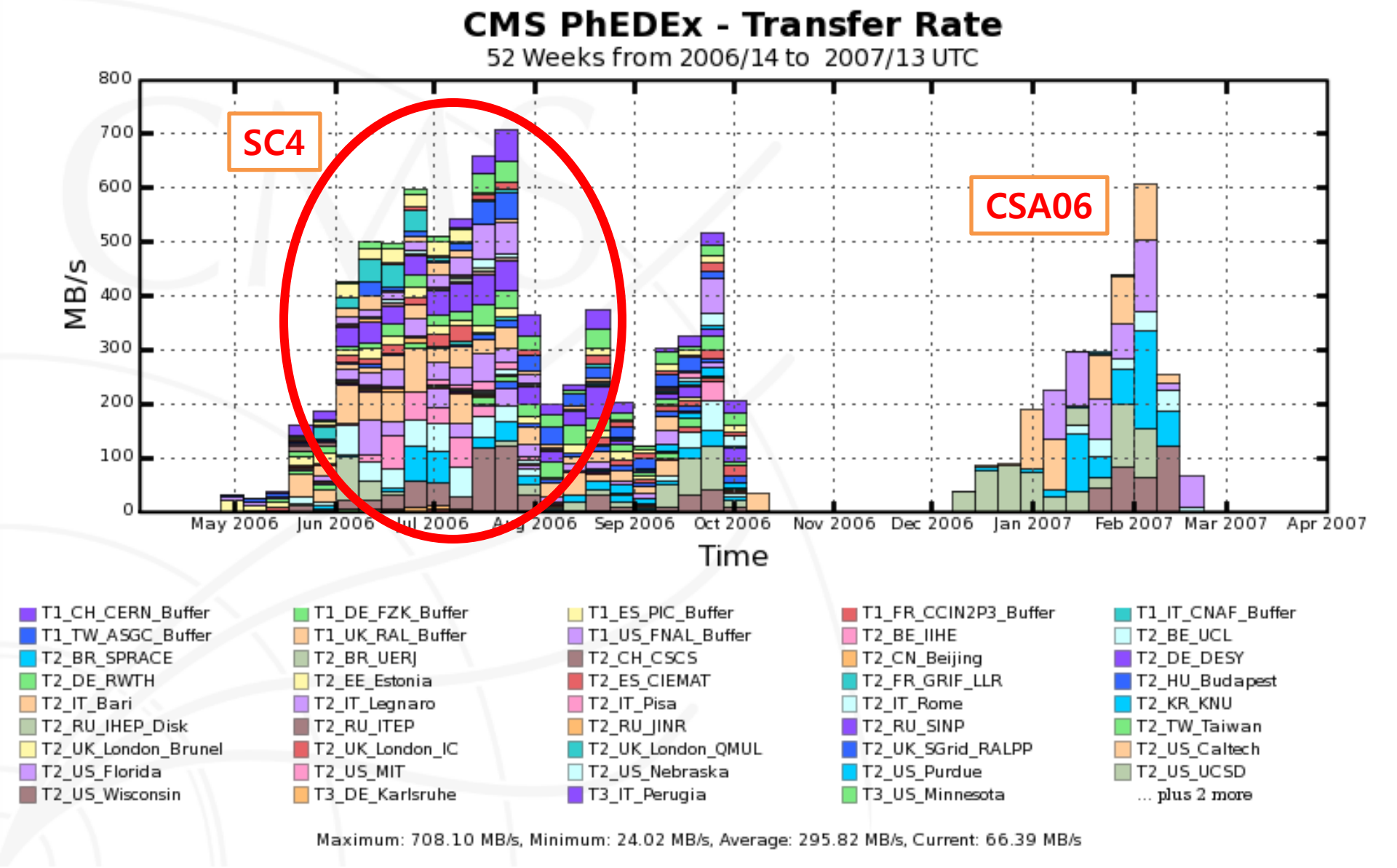
Service Challenge 4 Goals

- Data Transferred **Tier-1** ⇒ **Tier-2** with PhEDEx through FTS or srmcp
  - **10 MB/s for worst** connected **Tier-2** sites
  - **100 MB/s for best** connected **Tier-2** sites

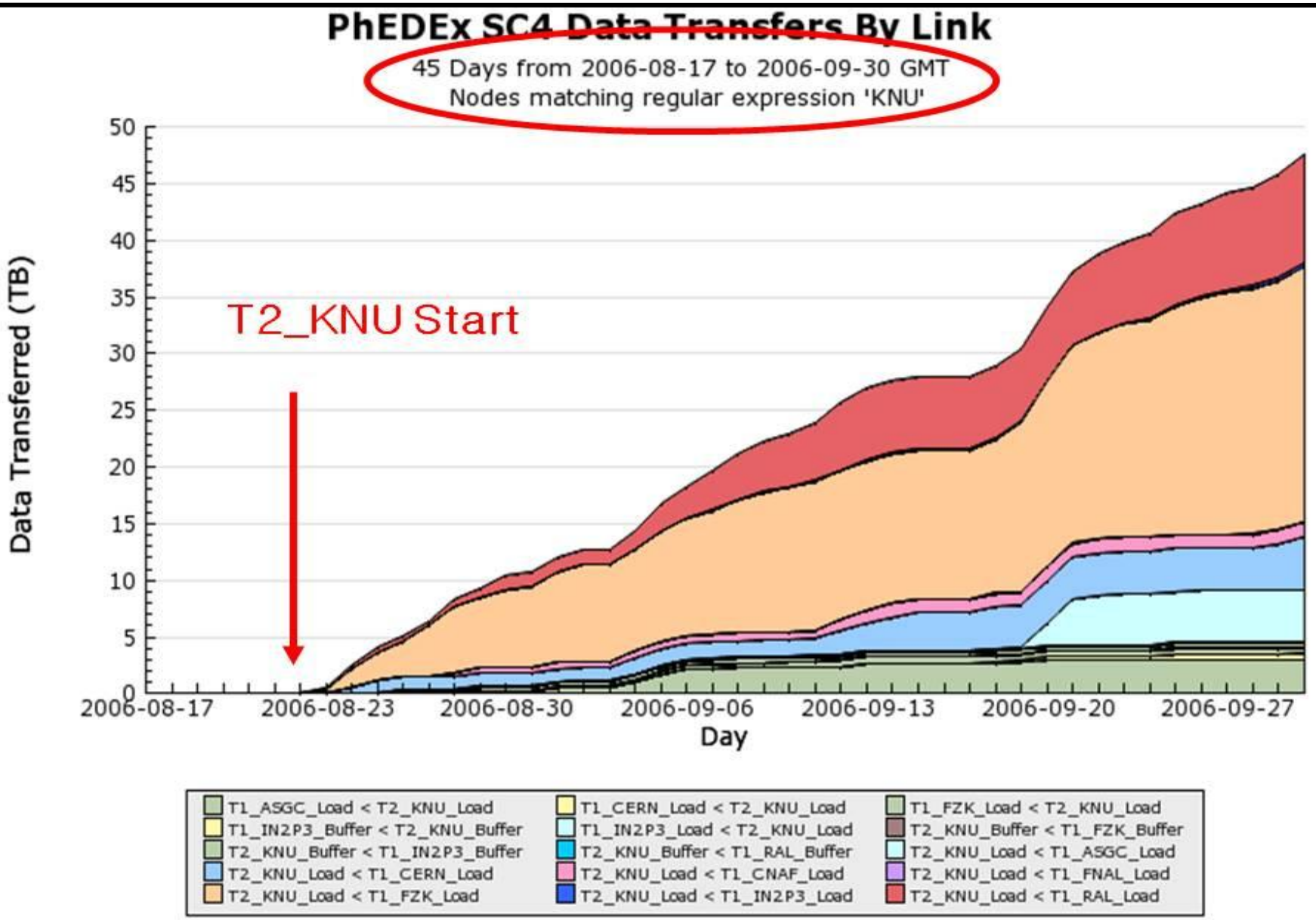
- Data Transferred **Tier-2** ⇒ **Tier-1** with PhEDEx through FTS or srmcp
  - **10 MB/s** for all sites to archive data

8 Tier-1 centers &
24 Tier-2 centers

Data Transfer SC4



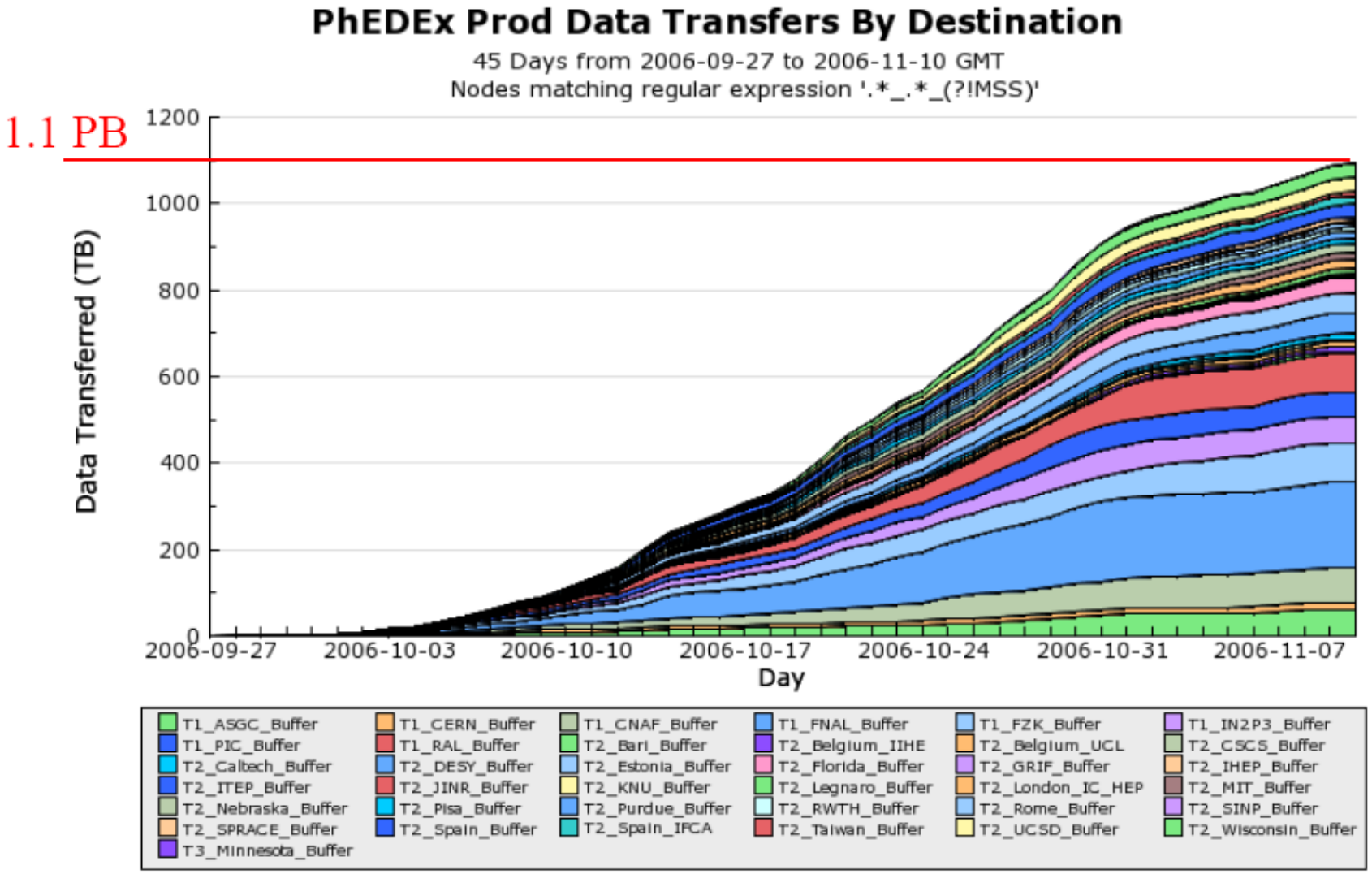
SC4


PhEDEx SC4 Data Transfers By Link
45 Days from 2006-08-17 to 2006-09-30 GMT
Nodes matching regular expression 'KNU'
T2_KNU Start
Data Transferred (TB)
Day
0
5
10
15
20
25
30
35
40
45
50
2006-08-17
2006-08-23
2006-08-30
2006-09-06
2006-09-13
2006-09-20
2006-09-27
T1_ASGC_Load < T2_KNU_Load
T1_CERN_Load < T2_KNU_Load
T1_FZK_Load < T2_KNU_Load
T1_IN2P3_Buffer < T2_KNU_Buffer
T1_IN2P3_Load < T2_KNU_Load
T2_KNU_Buffer < T1_FZK_Buffer
T2_KNU_Buffer < T1_IN2P3_Buffer
T2_KNU_Buffer < T1_RAL_Buffer
T2_KNU_Load < T1_ASGC_Load
T2_KNU_Load < T1_CERN_Load
T2_KNU_Load < T1_CNAF_Load
T2_KNU_Load < T1_FNAL_Load
T2_KNU_Load < T1_FZK_Load
T2_KNU_Load < T1_IN2P3_Load
T2_KNU_Load < T1_RAL_Load

CSA06 Data Transfer

- Tier-0 → Tier-1 (to Tape)
  - Individual goals for each Tier-1, sum to 150 MB/s
    - 25% of 2008 goal
    - 2X the actual sustained rate needed for 40 Hz
      - Exercise clearance of backlog after failure
  - Successful Transfers for 90% of CSA06 interval
    - According to availability metric as defined by WLCG for 2006
  - Not all Tier-1 Centers were prepared to write to Tape
- Tier-1 → Tier-1
  - No such dataflow in CSA06
    - Though exercised to some extent
- Tier-1 →Tier-2
  - Goal: 20MB/s into each Tier-2
  - Threshold: 5MB/s
  - Overall "success" is to have 50% of participants at or above goal and 90% above threshold
  - Successful Transfers for 80% of CSA06 interval

CSA06 Data Transfer



37 Sites, 1.1 PB transferred via the Wide Area Network!

Load Test 2007

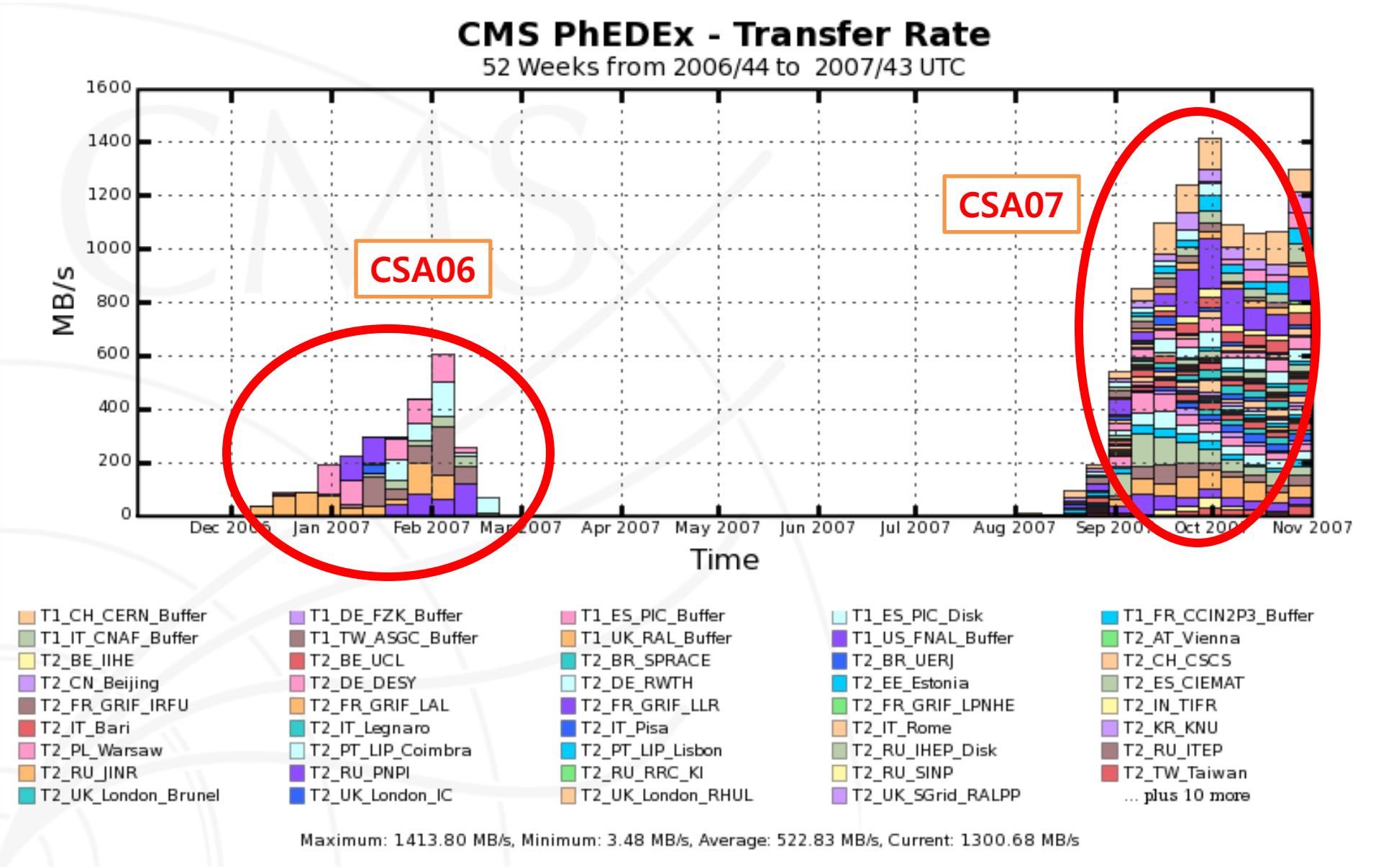
Load Test 2007 targets are to demonstrate the following:

- demontrate 65% T0 -> T1 **peak** rate, for 1 week;
- demontrate *simultaneously*
  - a) T0->T1 50% average rate
  - b) T1->all T2 50% sum of average rate
  - c) all T2->T1 50% sum of average rate, for 12 hrs
- demonstrate T1 -> each T2 sustainable at 10 MB/s , for 12 hrs;
- demonstrate each T2 -> T1 sustainable at 5 MB/s, for 12 hrs;

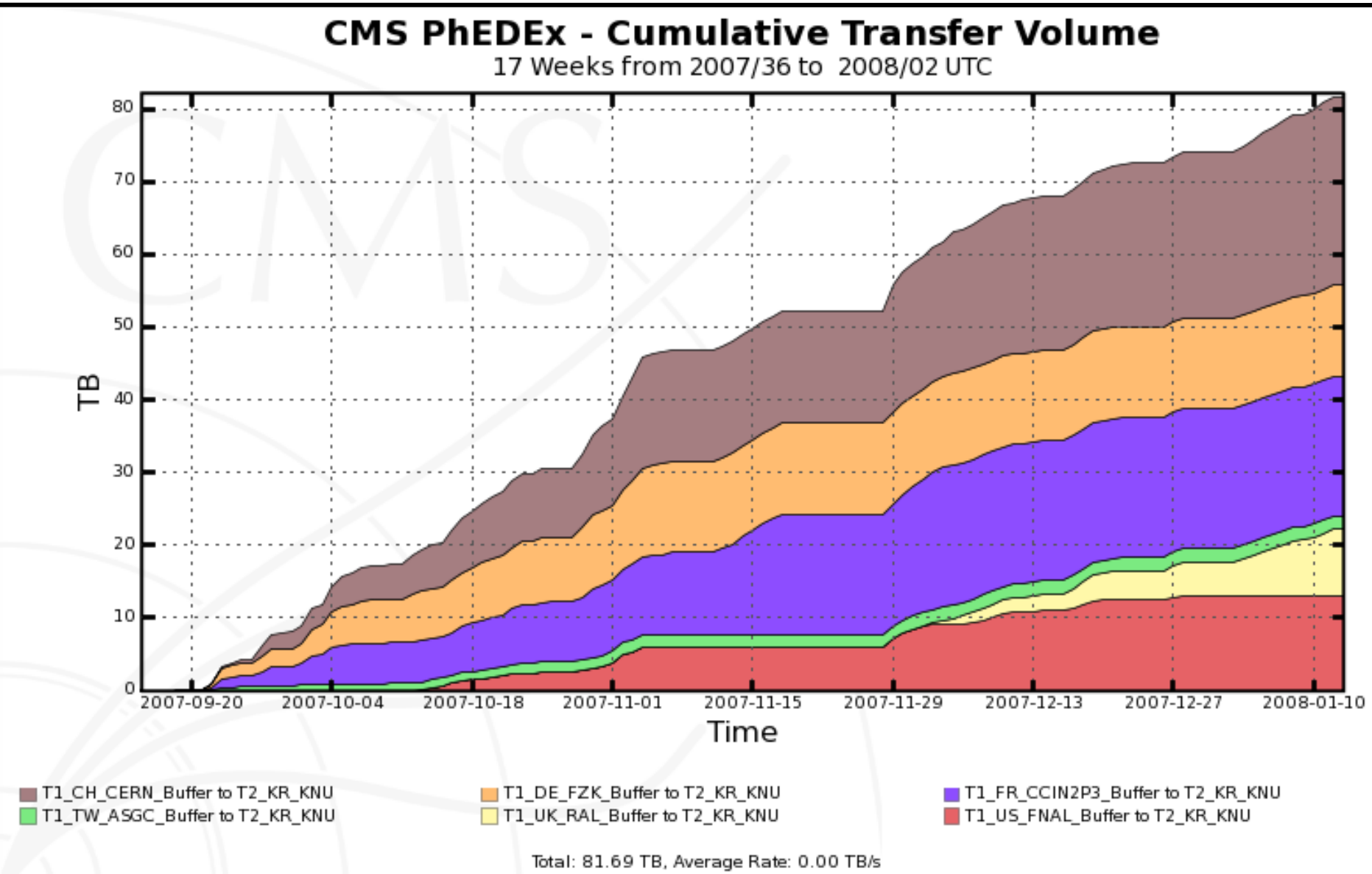
CSA07

- CSA07 Expected Average Transfer Rates CERN to Tier-1
-
  - ASGC 26 MB/s
  - CNAF 37 MB/s
  - FNAL 105 MB/s
  - FZK 26 MB/s
  - IN2P3 32 MB/s
  - PIC 13 MB/s
  - RAL 26 MB/s

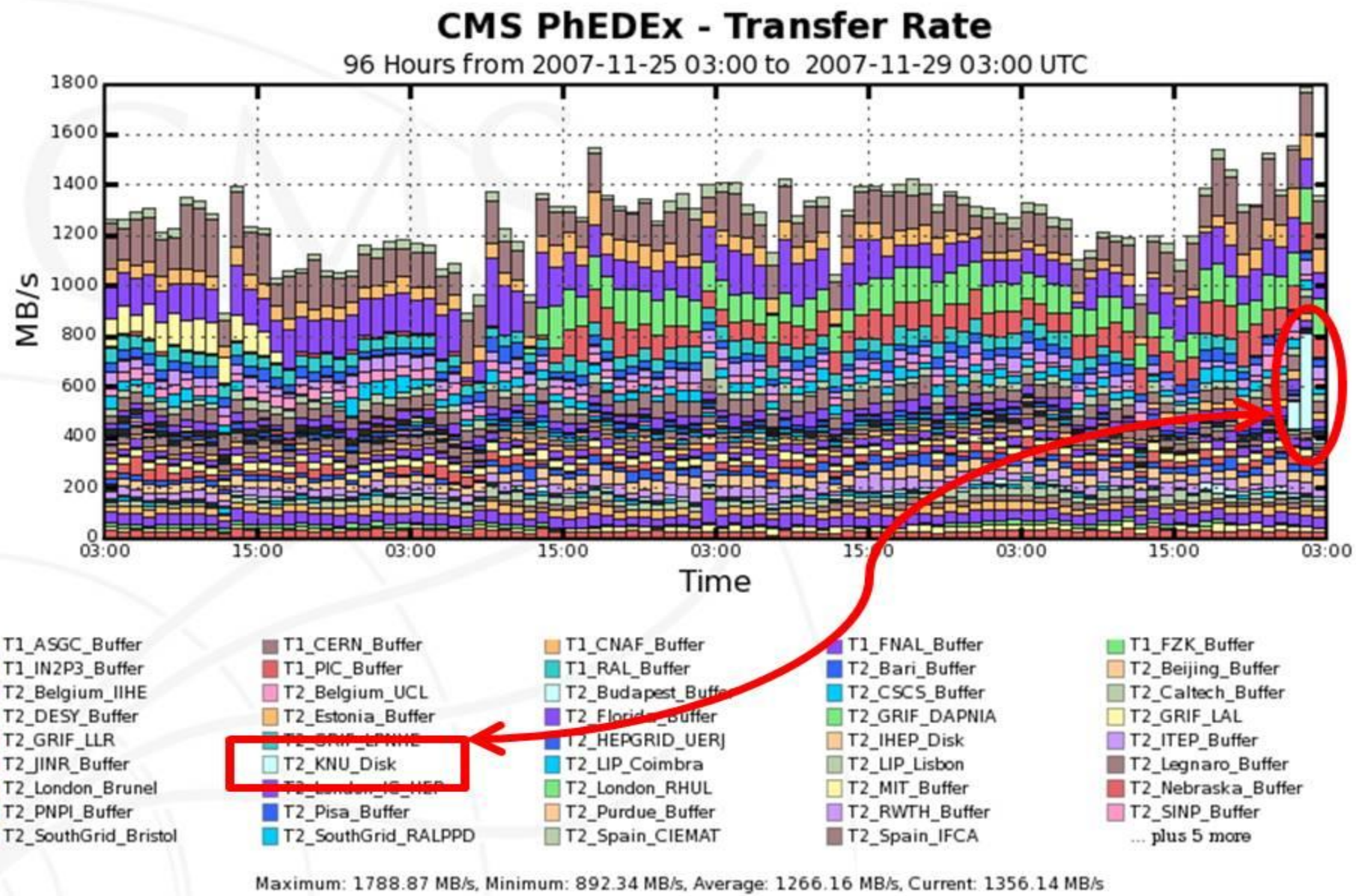
CSA06 & CSA07


CMS PhEDEx - Transfer Rate
52 Weeks from 2006/44 to 2007/43 UTC
MB/s
1600
1400
1200
1000
800
600
400
200
0
CSA06
CSA07
Dec 2006
Jan 2007
Feb 2007
Mar 2007
Apr 2007
May 2007
Jun 2007
Jul 2007
Aug 2007
Sep 2007
Oct 2007
Nov 2007
Time
T1_CH_CERN_Buffer
T1_DE_FZK_Buffer
T1_ES_PIC_Buffer
T1_ES_PIC_Disk
T1_FR_CCIN2P3_Buffer
T1_IT_CNAF_Buffer
T1_TW_ASGC_Buffer
T1_UK_RAL_Buffer
T1_US_FNAL_Buffer
T2_AT_Vienna
T2_BE_IIHE
T2_BE_UCL
T2_BR_SPRACE
T2_BR_UERJ
T2_CH_CSCS
T2_CN_Beijing
T2_DE_DESY
T2_DE_RWTH
T2_EE_Estonia
T2_ES_CIEMAT
T2_FR_GRIF_IRFU
T2_FR_GRIF_LAL
T2_FR_GRIF_LLR
T2_FR_GRIF_LPNHE
T2_IN_TIFR
T2_IT_Bari
T2_IT_Legnaro
T2_IT_Pisa
T2_IT_Rome
T2_KR_KNU
T2_PL_Warsaw
T2_PT_LIP_Coimbra
T2_PT_LIP_Lisbon
T2_RU_IHEP_Disk
T2_RU_ITEP
T2_RU_JINR
T2_RU_PNPI
T2_RU_RRC_KI
T2_RU_SINP
T2_TW_Taiwan
T2_UK_London_Brunel
T2_UK_London_IC
T2_UK_London_RHUL
T2_UK_SGrid_RALPP
... plus 10 more
Maximum: 1413.80 MB/s, Minimum: 3.48 MB/s, Average: 522.83 MB/s, Current: 1300.68 MB/s

CSA07 Data Transfer - KNU



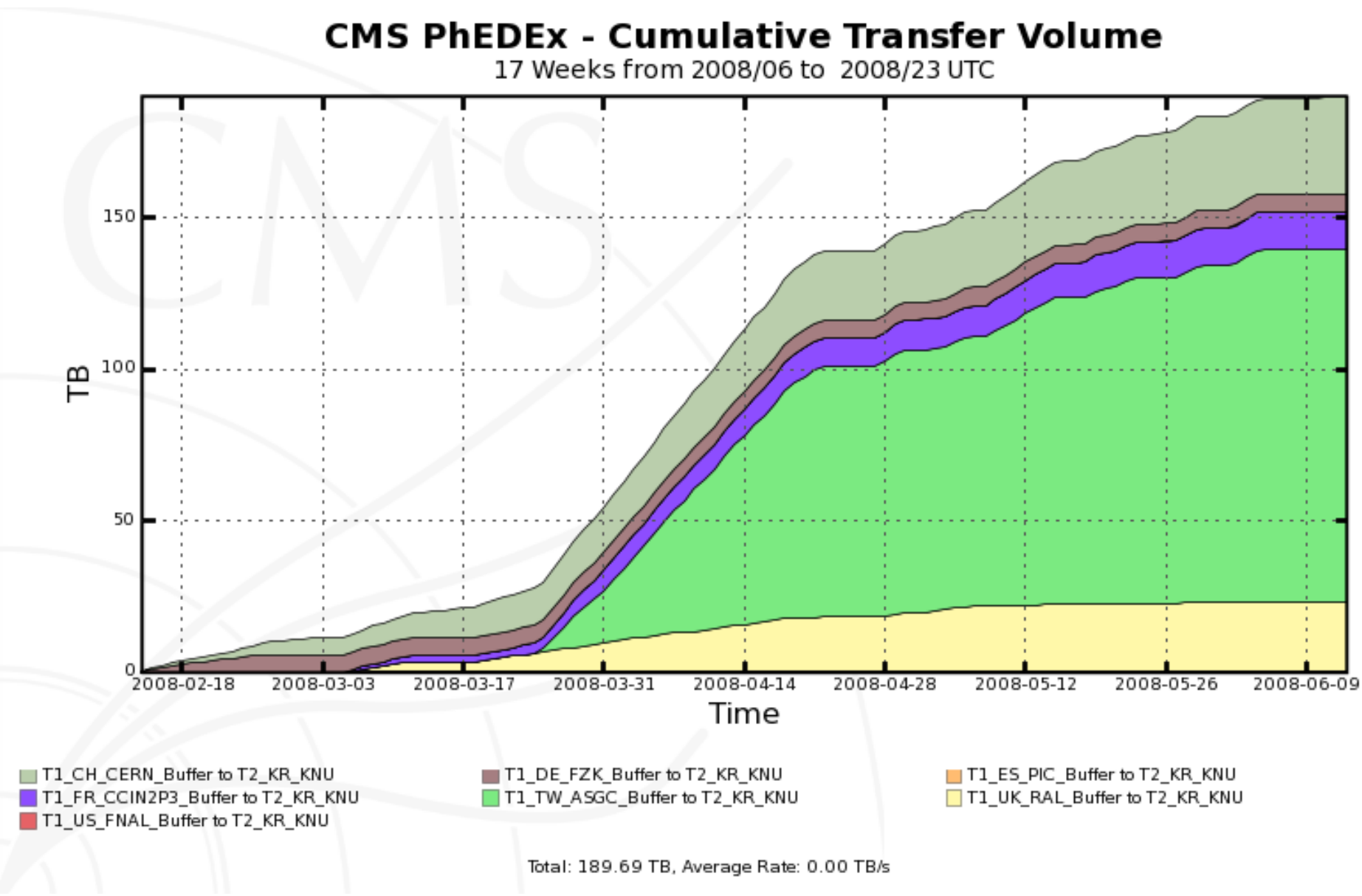
Something happened after CSA07



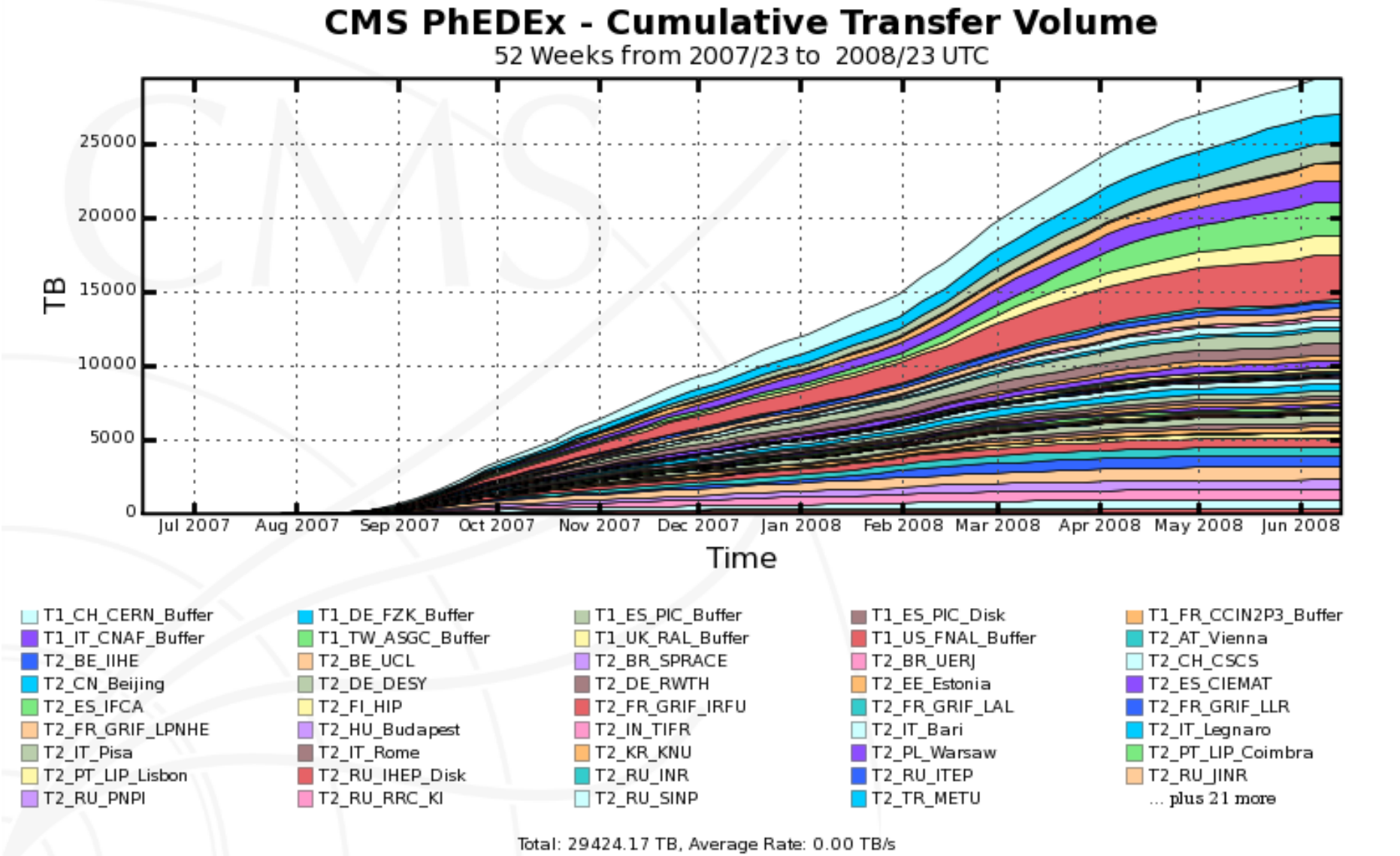
CCRC08

- The Common Computing Readiness Challenge of 2008 refers to an LHC-wide computing challenge.
- CCRC08 has a preparatory Phase 1 series of tests in February, and a main Phase 2 in May 2008.
- The combined Computing Software and Analysis challenge of 2008 (CSA08) refers to the collection of CMS tests aimed to test the full scope of the data handling and analysis activities needed for LHC data-taking operations in 2008.
- It has components in overlap with CCRC Phase 2 (the so- called “initial CSA08”, or “iCSA08”).

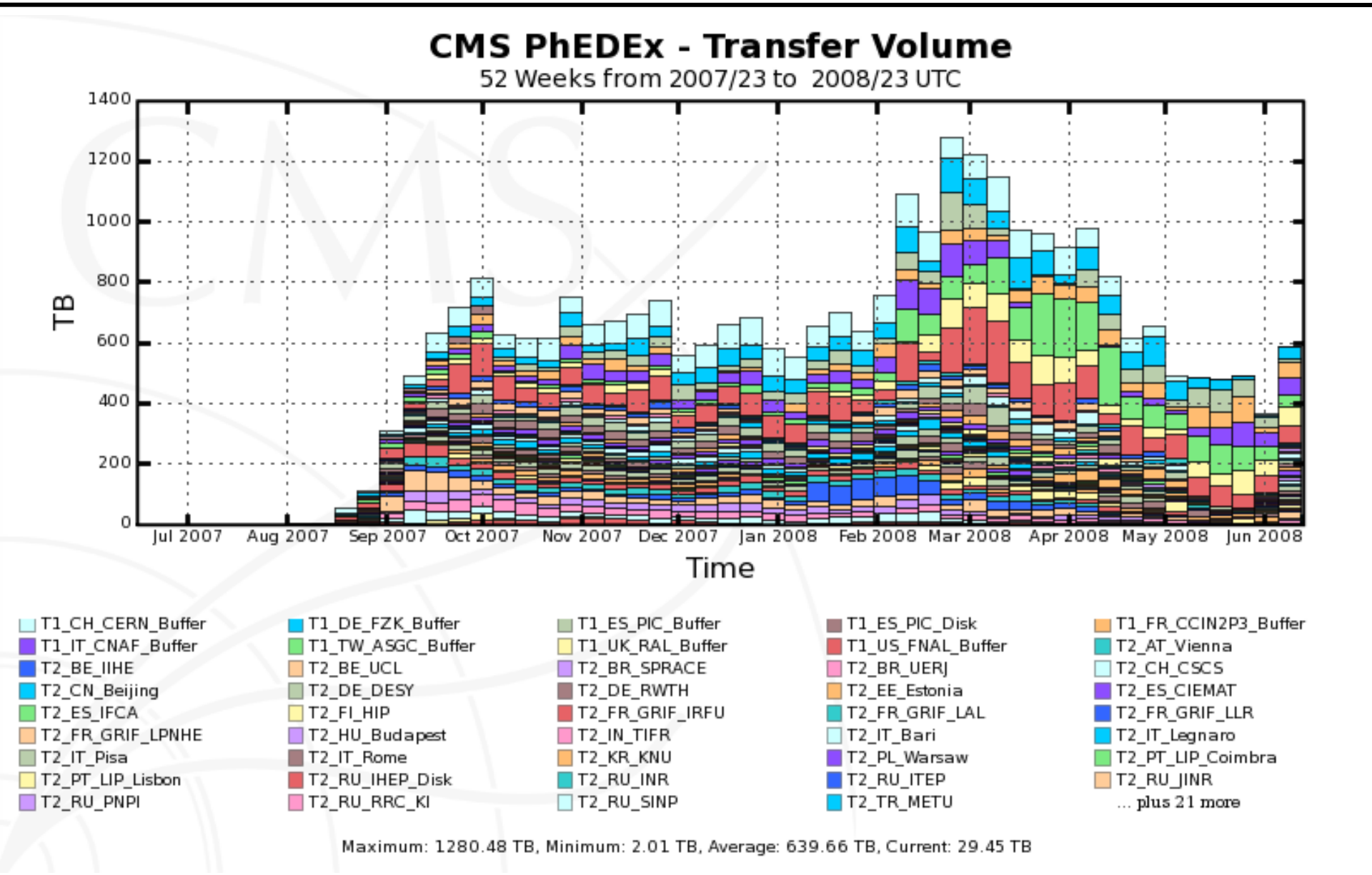
CCRC08 Data Transfer -KNU

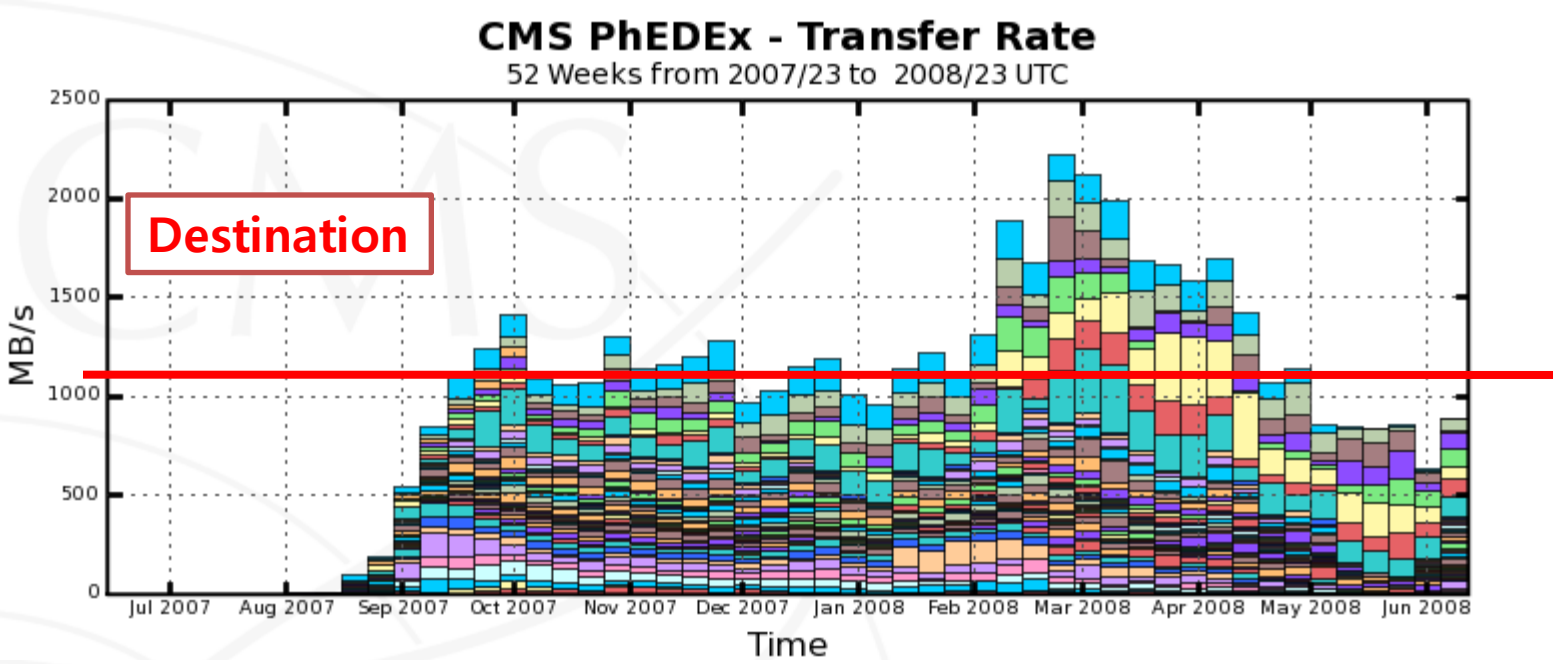


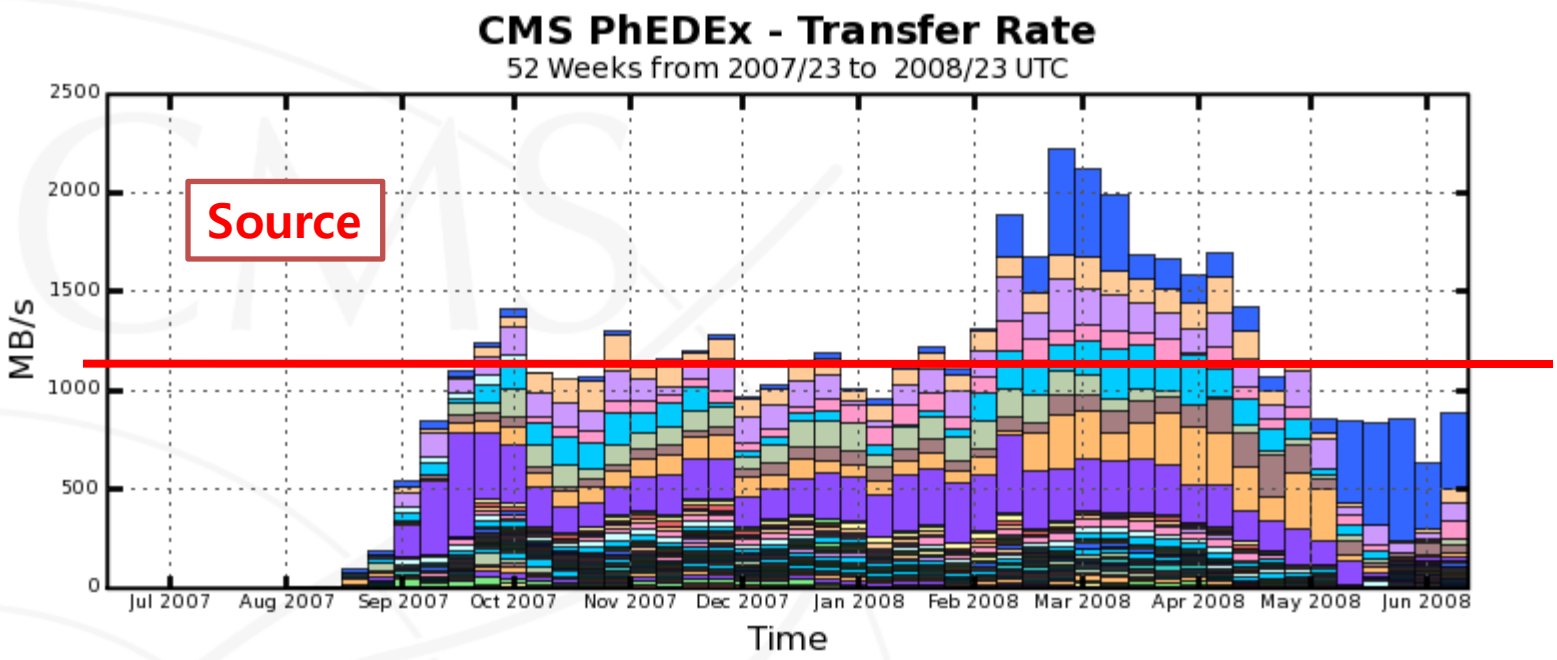
Cumulative Transferred Data Volume for all CMS Tiers (Last 52 Weeks)

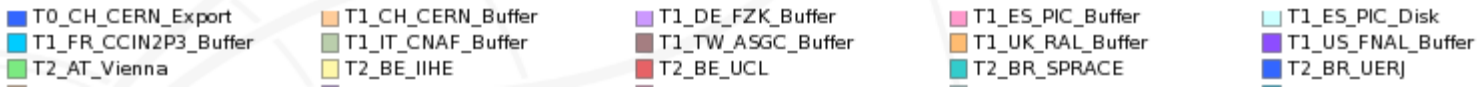


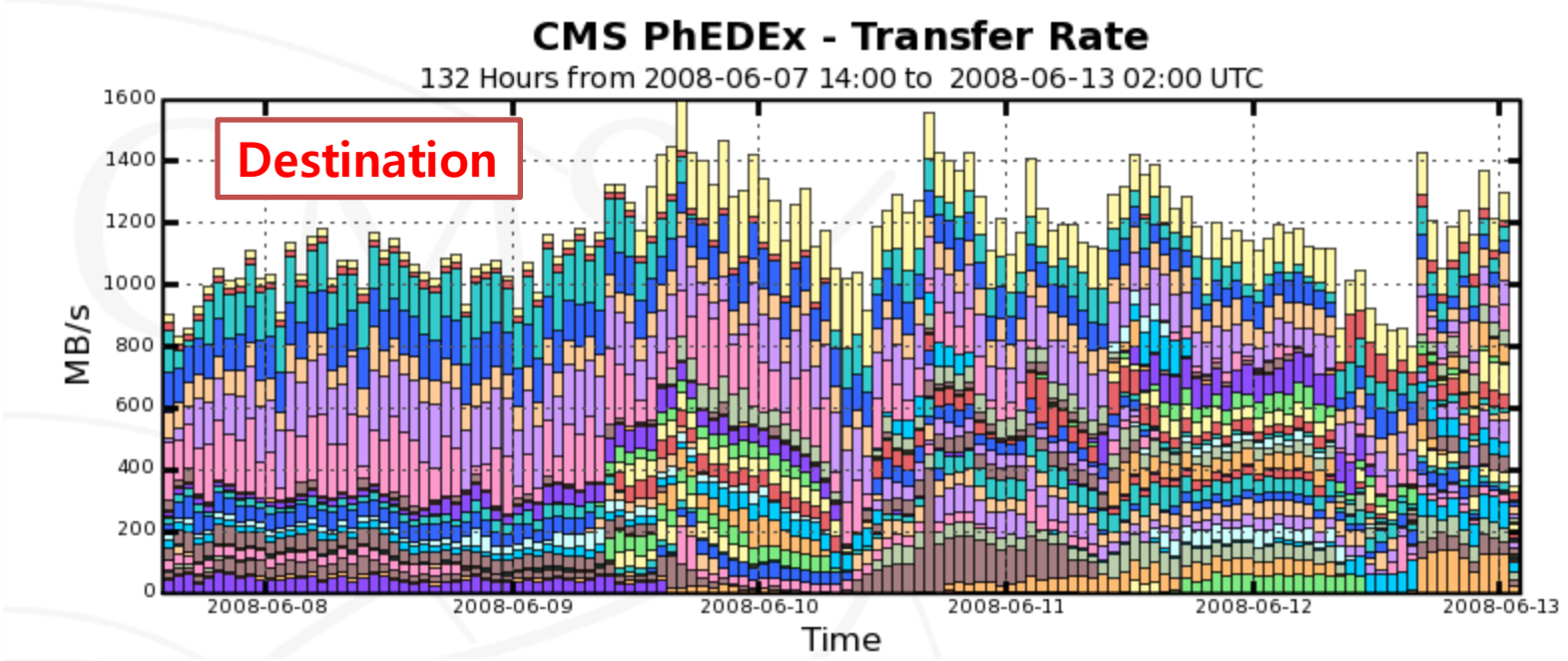
Transferred Data Volume (Last 52 Weeks)

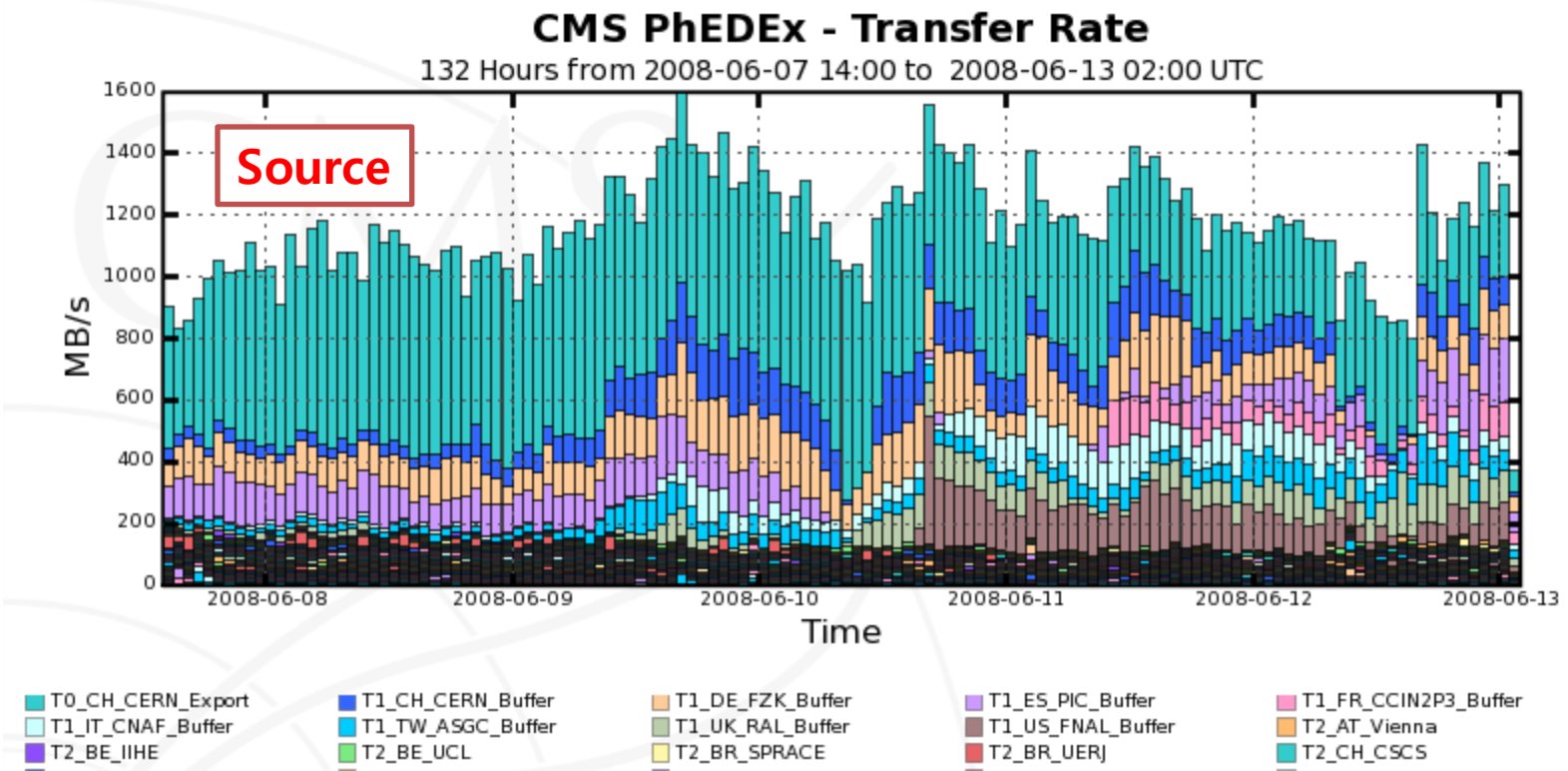



CMS PhEDEx - Transfer Rate
52 Weeks from 2007/23 to 2008/23 UTC
Destination
MB/s
2500
2000
1500
1000
500
0
Jul 2007
Aug 2007
Sep 2007
Oct 2007
Nov 2007
Dec 2007
Jan 2008
Feb 2008
Mar 2008
Apr 2008
May 2008
Jun 2008
Time


CMS PhEDEx - Transfer Rate
52 Weeks from 2007/23 to 2008/23 UTC
Source
MB/s
2500
2000
1500
1000
500
0
Jul 2007
Aug 2007
Sep 2007
Oct 2007
Nov 2007
Dec 2007
Jan 2008
Feb 2008
Mar 2008
Apr 2008
May 2008
Jun 2008
Time

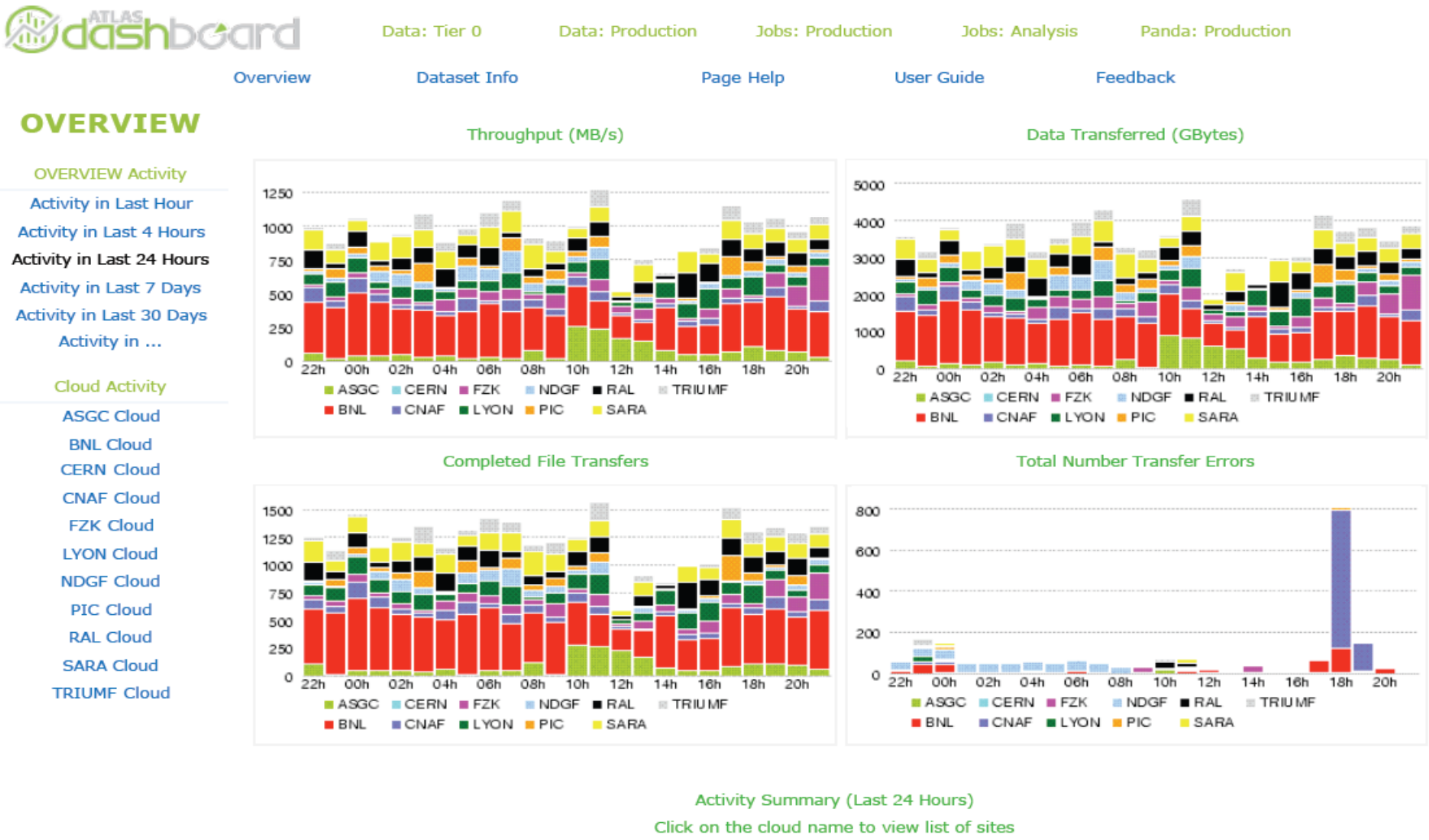

T0_CH_CERN_Export
T1_CH_CERN_Buffer
T1_DE_FZK_Buffer
T1_ES_PIC_Buffer
T1_ES_PIC_Disk
T1_FR_CCIN2P3_Buffer
T1_IT_CNAF_Buffer
T1_TW_ASGC_Buffer
T1_UK_RAL_Buffer
T1_US_FNAL_Buffer
T2_AT_Vienna
T2_BE_IIHE
T2_BE_UCL
T2_BR_SPRACE
T2_BR_UERJ


CMS PhEDEx - Transfer Rate
132 Hours from 2008-06-07 14:00 to 2008-06-13 02:00 UTC
Destination
MB/s
0
200
400
600
800
1000
1200
1400
1600
2008-06-08
2008-06-09
2008-06-10
2008-06-11
2008-06-12
2008-06-13
Time


CMS PhEDEx - Transfer Rate
132 Hours from 2008-06-07 14:00 to 2008-06-13 02:00 UTC
Source
MB/s
0
200
400
600
800
1000
1200
1400
1600
2008-06-08
2008-06-09
2008-06-10
2008-06-11
2008-06-12
2008-06-13
Time
T0_CH_CERN_Export
T1_CH_CERN_Buffer
T1_DE_FZK_Buffer
T1_ES_PIC_Buffer
T1_FR_CCIN2P3_Buffer
T1_IT_CNAF_Buffer
T1_TW_ASGC_Buffer
T1_UK_RAL_Buffer
T1_US_FNAL_Buffer
T2_AT_Vienna
T2_BE_IIHE
T2_BE_UCL
T2_BR_SPRACE
T2_BR_UERJ
T2_CH_CSCS

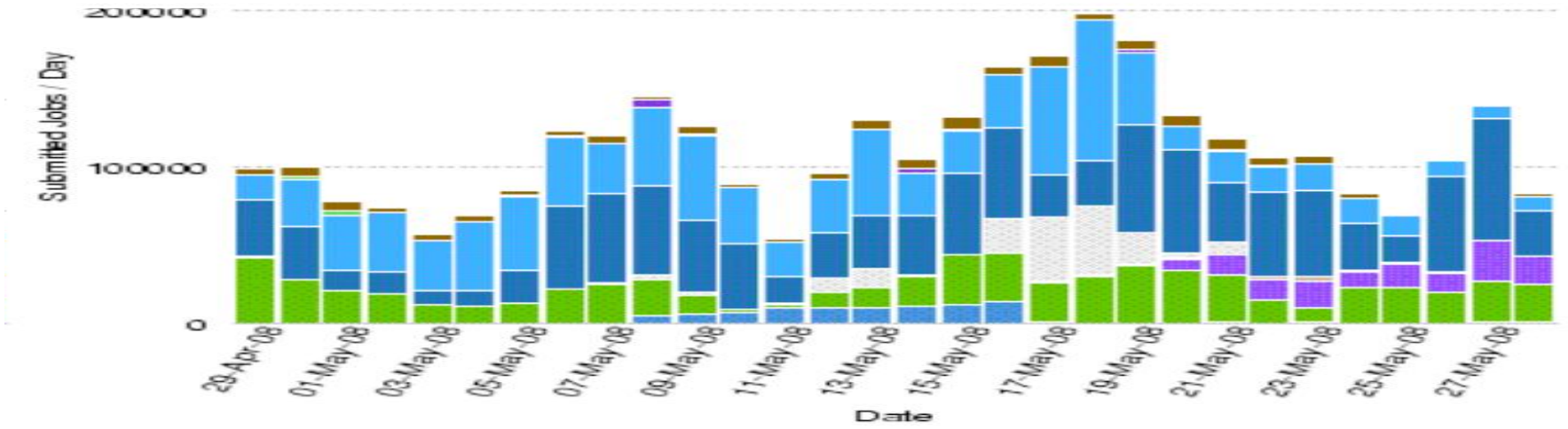
There are 3 more Experiments …

ATLAS data transfer status - 28 May 2008
Throughput ~1100 MB/s



Dashboards provide tools for monitoring and debugging

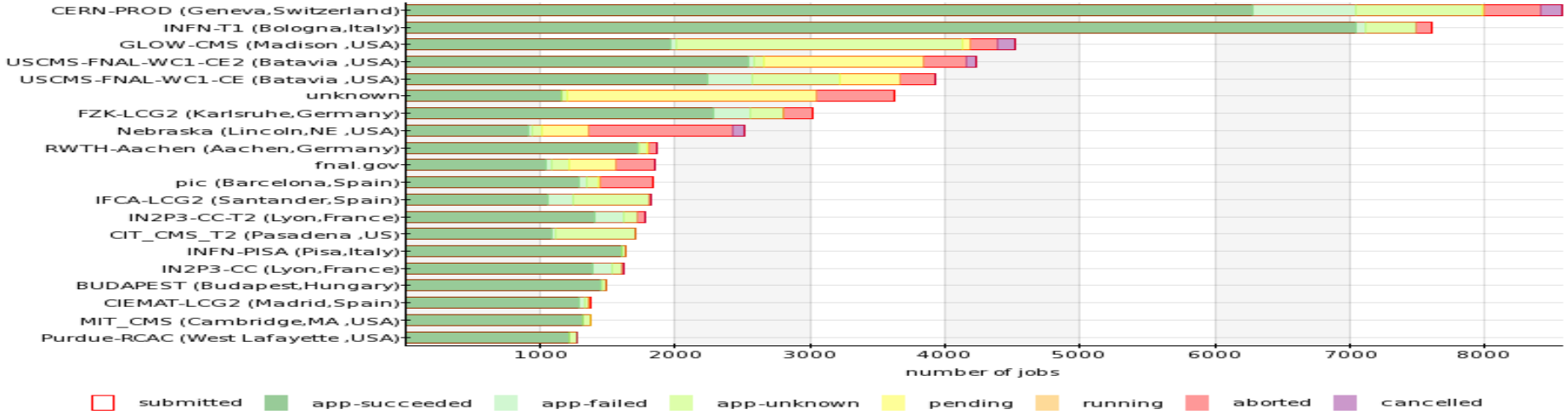
CMS jobs May 2008 sorted by activities
Up to 200K jobs per day - 35% end-user analysis



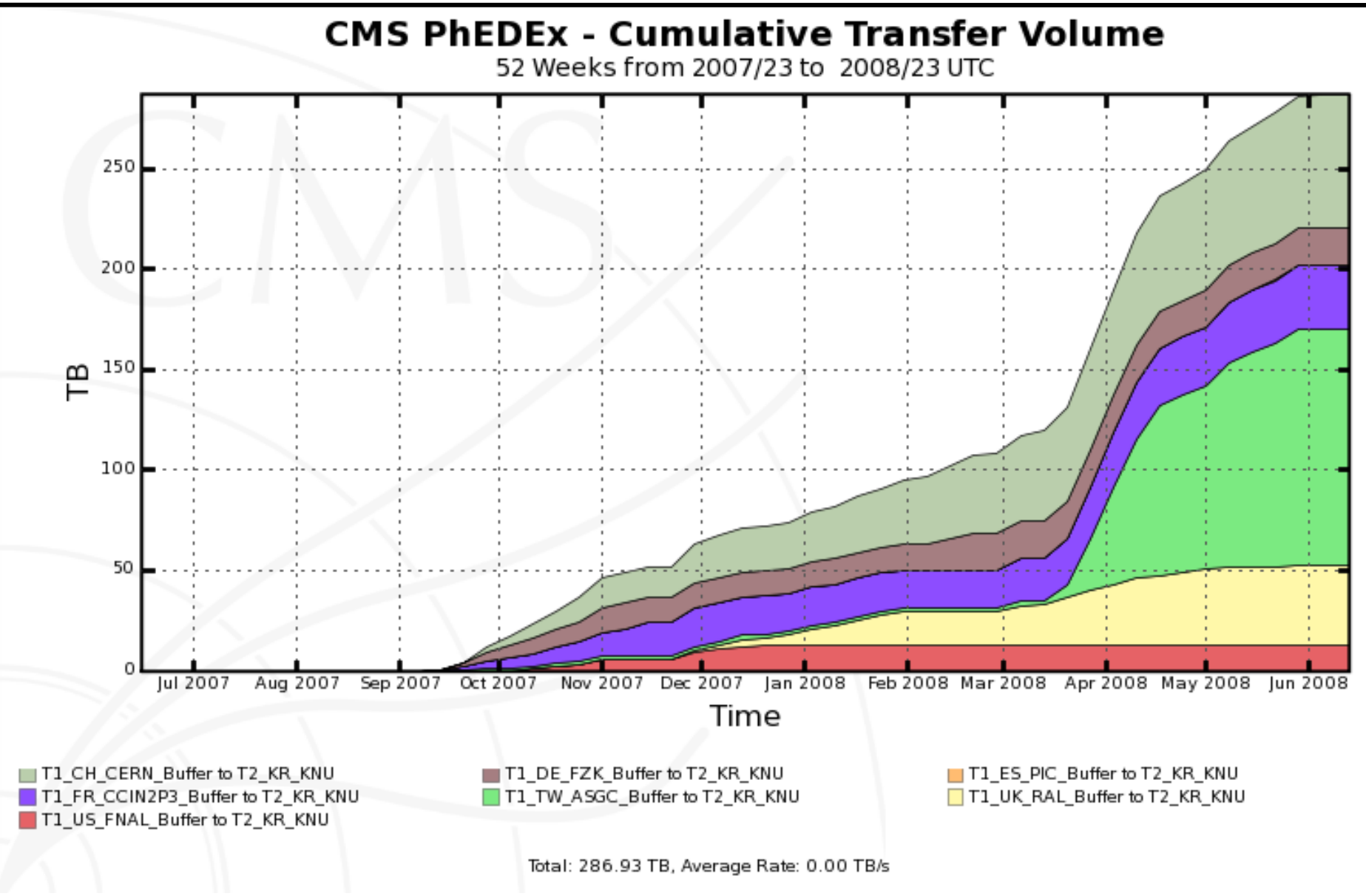
suggestion?

ningtime

job status by site – user can drill down to find details of the errors



Data Transfer for KNU (Last 52 Weeks)



Summary

- LHC will produce massive amounts of data (15 PB/year)
- The data from the LHC experiments will be distributed around the globe Tier centers.
- The same amount of MC will be produced by Tier-2 centers and transferred to Tier-1.
- CSA08 refers to the collection of CMS tests aimed to test the full scope of the data handling and analysis activities needed for LHC data-taking operations in 2008.
- Global ring network plays an important role for the Worldwide LHC Computing Grid.

References

- Les Robertson, “The LHC Computing Grid” 2008
- Worldwide LHC Computing GRID, http://lcg.web.cern.ch/LCG/
- Compact Muon Solenoid, http://cms.cern.ch/
- PhEDEx - CMS Data Transfers, http://cmsdoc.cern.ch/cms/aprom/phedex/
- KNU Data Center for CMS, http://lcg.knu.ac.kr/