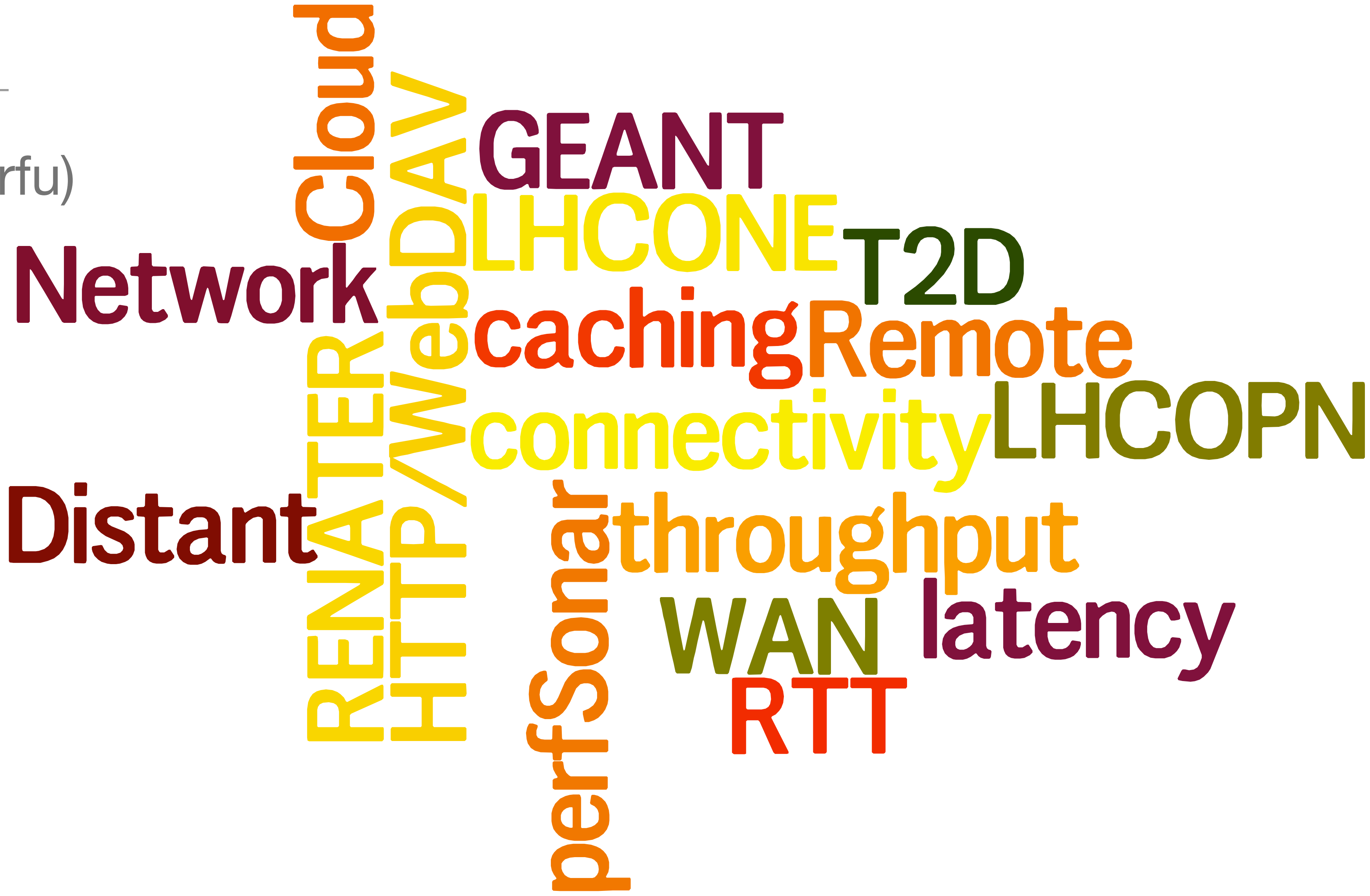
Network issues on FR cloud

Eric Lançon (CEA-Saclay/Irfu)



Data SIO, NOAA, U.S. Navy, NGA, GEBCO
© 2012 Cnes/Spot Image
Image © 2012 TerraMetrics
©2010 Google

Network Usage

- Data distribution
- MC production
- Analysis
- Distributed storage

Network used for

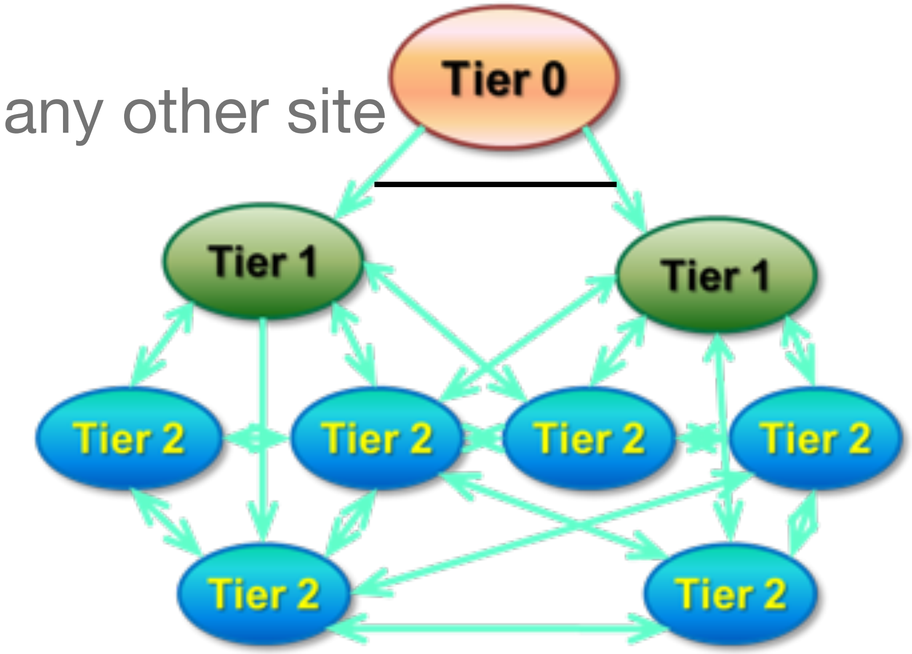
- **Data distribution**, 2 components :
  - Pre-placed data (a la MONARC)
  - Dynamic data distribution (popular data to available sites)
- **MC production**
  - Within a given cloud
  - Across clouds
- **Analysis**
  - Retrieving results
  - Small data sets
- **Distributed storage**
  - To optimize resources
  - Simplify data management

The 'old' computing model is dying

The ATLAS Data Model has changed

USE THE NETWORK

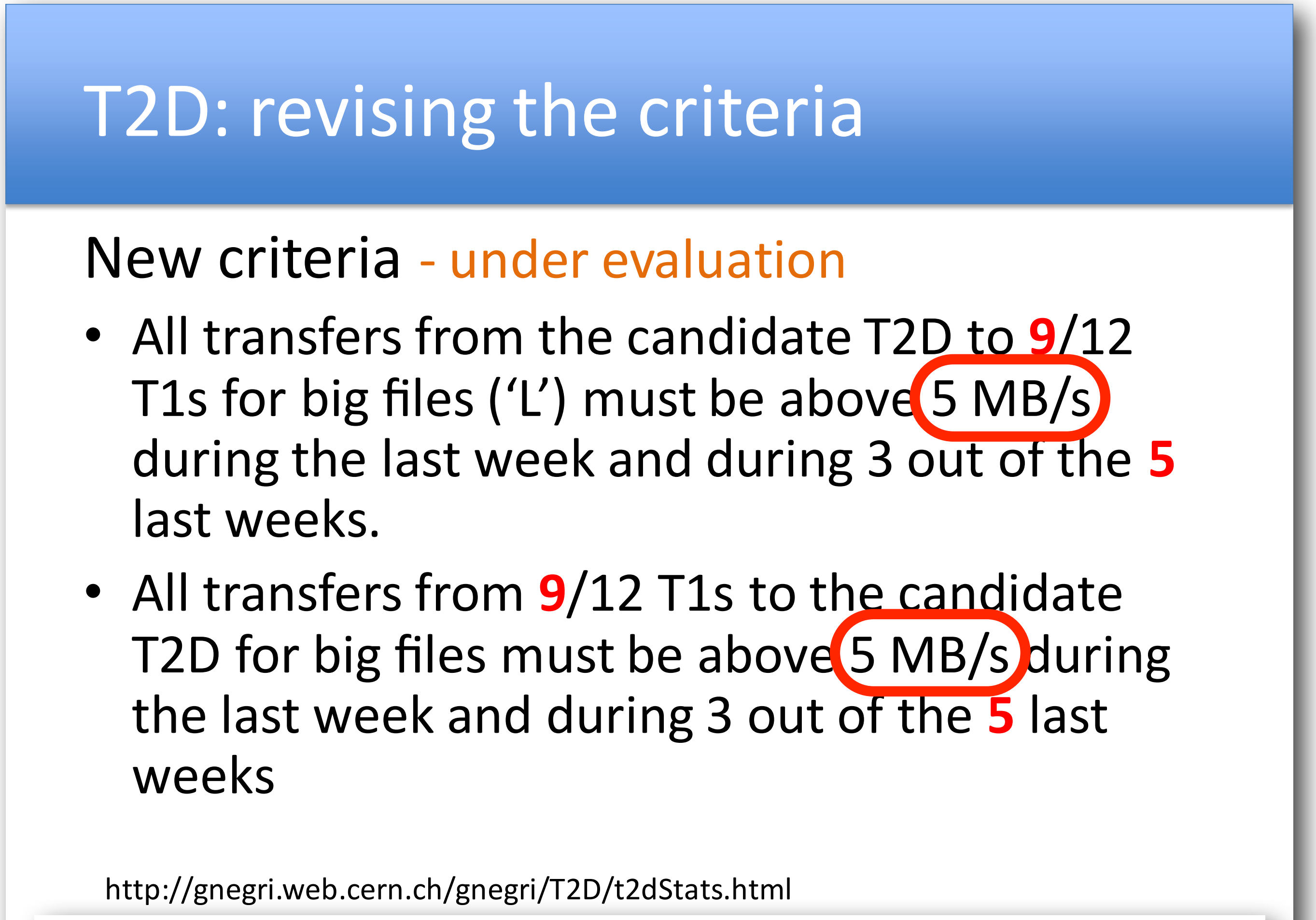
- Moved away from the historical model
- 4 recurring themes:
  - **Flat(ter) hierarchy**: Any site can replicate data from any other site
  - **Multi Cloud Production**
    - Need to replicate output files to remote Tier-1
  - **Dynamic data caching**: Analysis sites receive datasets from any other site "on demand" based on usage pattern
    - Possibly in combination with pre-placement of data sets by centrally managed replication of datasets
  - **Remote data access**: local jobs accessing data stored at remote sites



- **ATLAS is now heavily relying on multi-domain networks and needs decent e2e network monitoring**

ATLAS sites and connectivity

- ATLAS computing model has (will continue to) changed
  - More experience
  - More tools and monitoring
- New category of sites : Direct T2s (**T2Ds**)
  - Primary hosts for datasets (**analysis**) and for group analysis
  - Get and send data from different clouds
  - Participate in cross cloud production



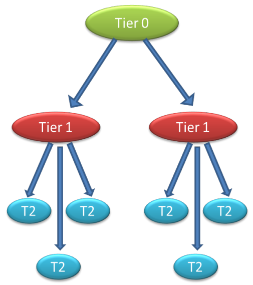
T2D: revising the criteria

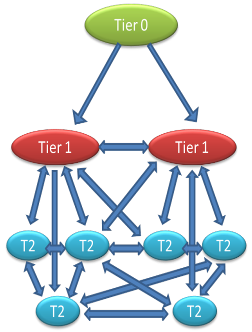
New criteria - under evaluation

- All transfers from the candidate T2D to 9/12 T1s for big files ('L') must be above 5 MB/s during the last week and during 3 out of the 5 last weeks.
- All transfers from 9/12 T1s to the candidate T2D for big files must be above 5 MB/s during the last week and during 3 out of the 5 last weeks

http://gnegri.web.cern.ch/gnegri/T2D/t2dStats.html

FR-cloud T2Ds : BEIJING, GRIF-LAL, GRIF-LPNHE, IN2P3-CPPM, IN2P3-LAPP, IN2P3-LPC, IN2P3-LPSC





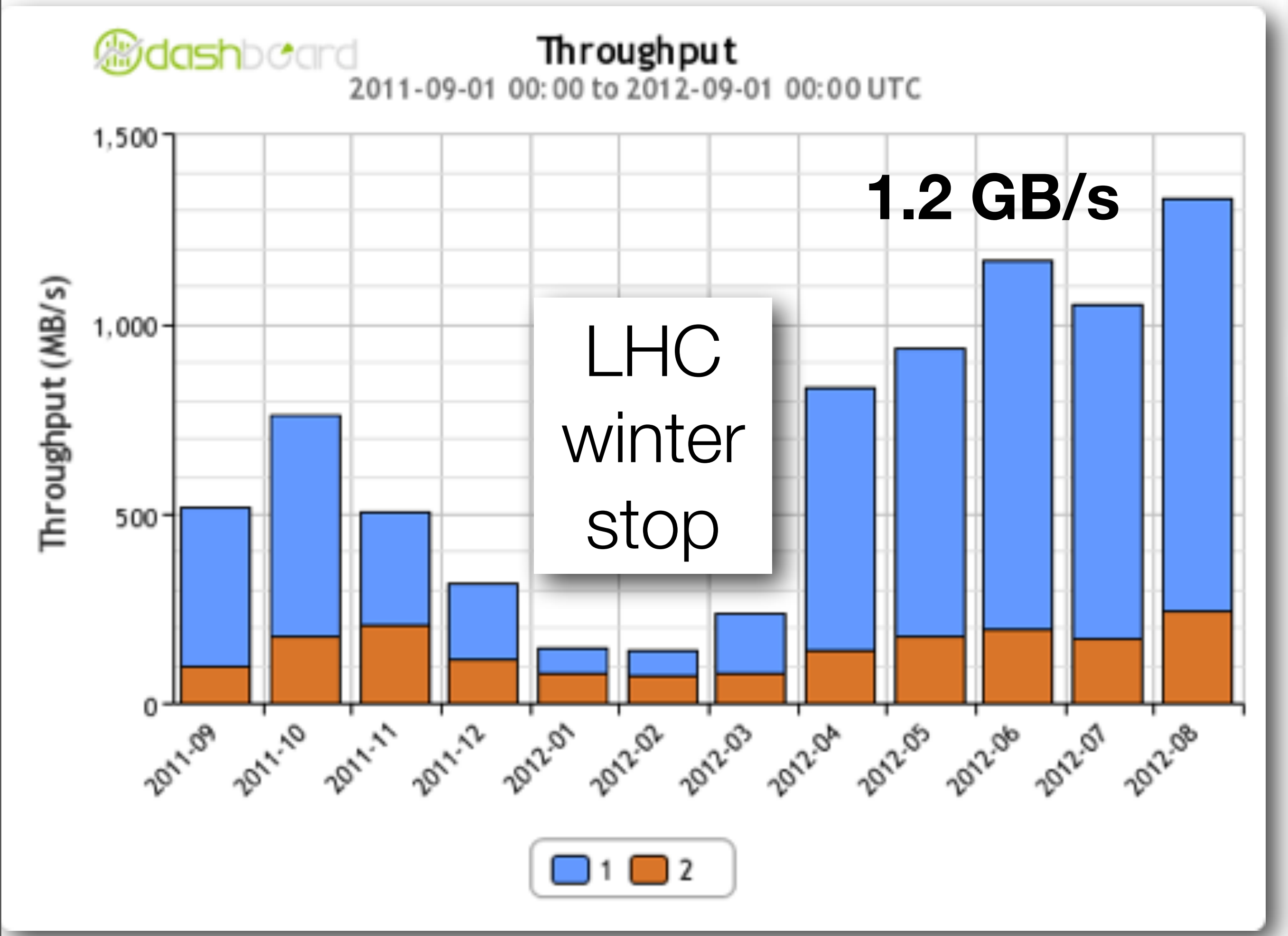
Network performance monitoring

- **Networking accounting :**
  - **Organized** (FTS) file transfers : http://dashb-atlas-data.cern.ch/ddm2/, not for direct transfers by users (dq2-get)
- **ATLAS ‘sonar’ :**
  - Calibrated file transfers by ATLAS Data Distribution system, from **storage to storage** : http://bourricot.cern.ch/dq2/ftsmon/
  - > 1 GB file transfers used to monitor and validate T2Ds
- **perfSONAR (PS) :**
  - **Network performance** (throughput, latency) : http://perfsonar.racf.bnl.gov:8080/exda/
  - Located as close as possible to storage at site and with similar connection hardware

T0 exports over a year

Over 1GB/s

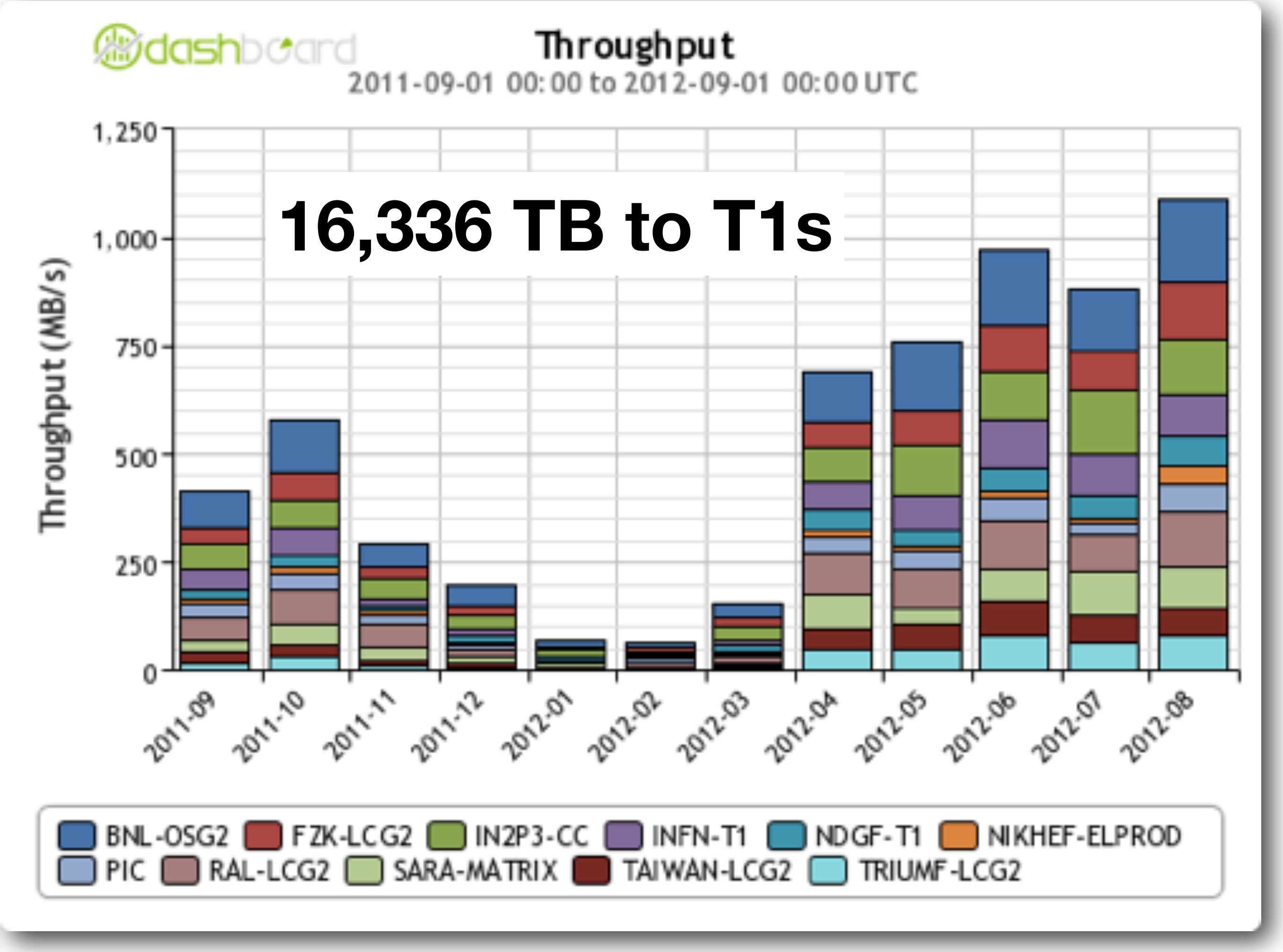
Better LHC efficiency and higher trigger rate



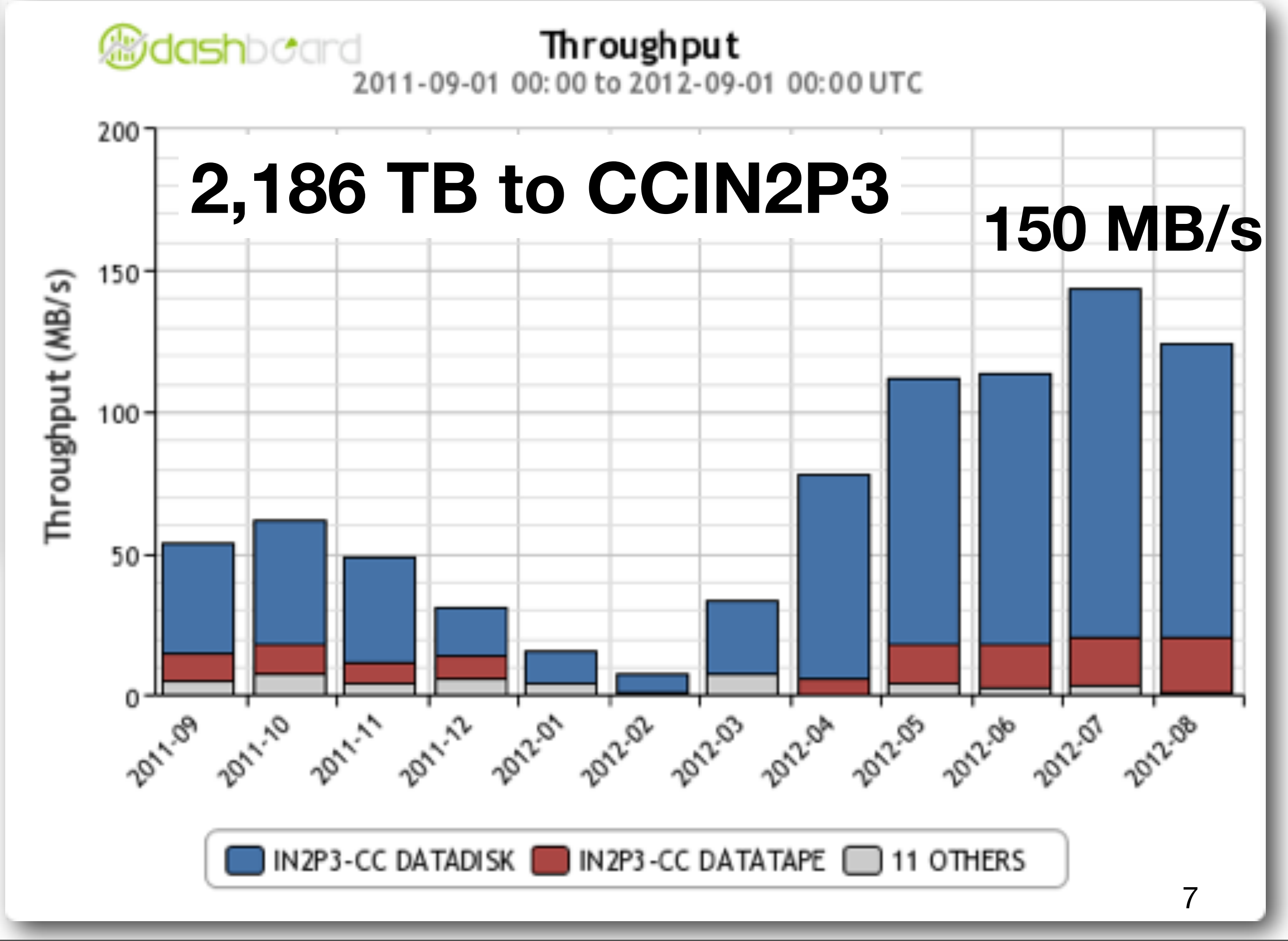
1.2 GB/s

LHC winter stop

78% to T1s



16,336 TB to T1s

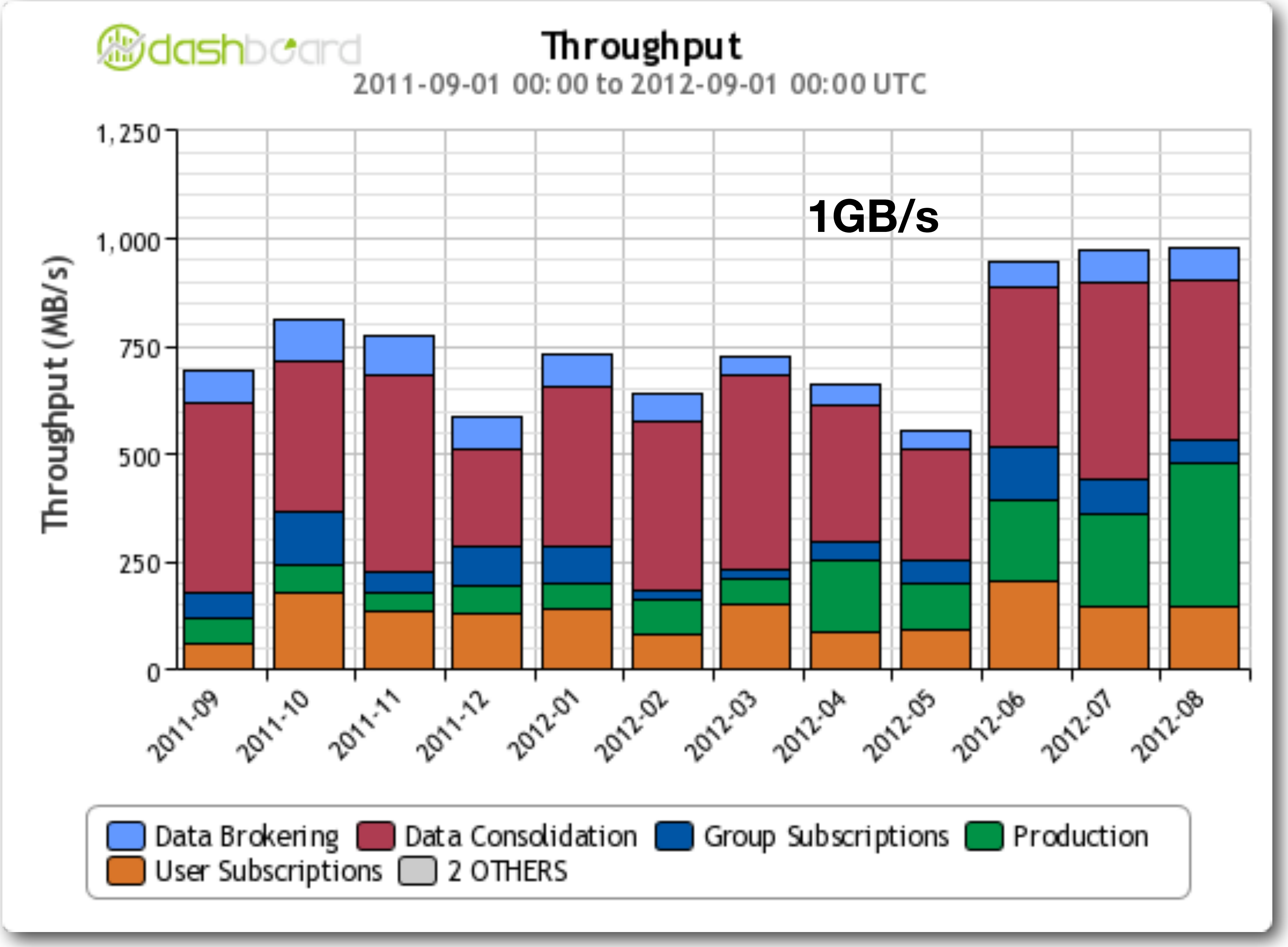


2,186 TB to CCIN2P3

150 MB/s

T1→T1

23,966 TB



1GB/s

Popularity based

Pre-placement

Group data

Cross-cloud MC production

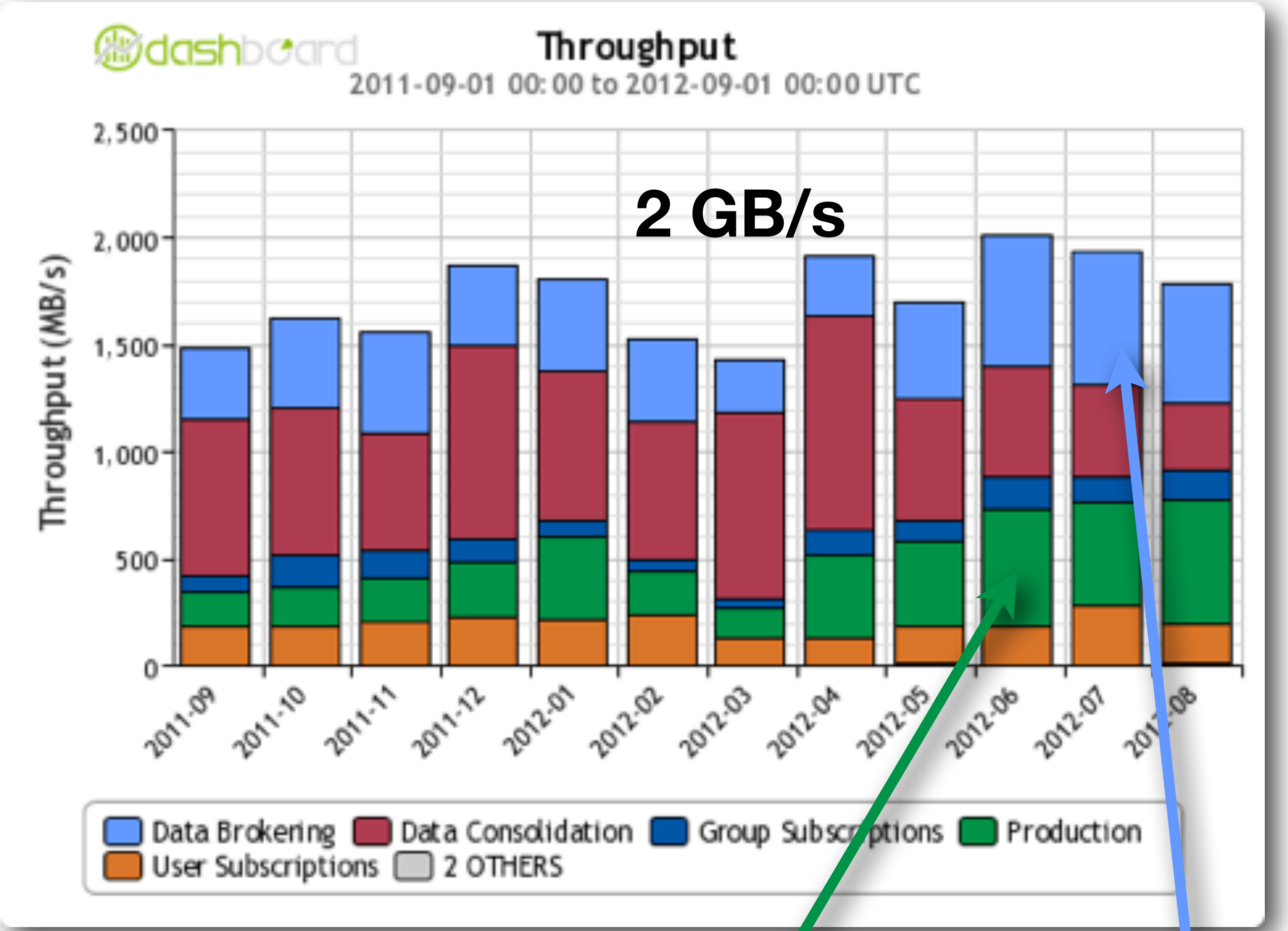
User requests

T1→T2

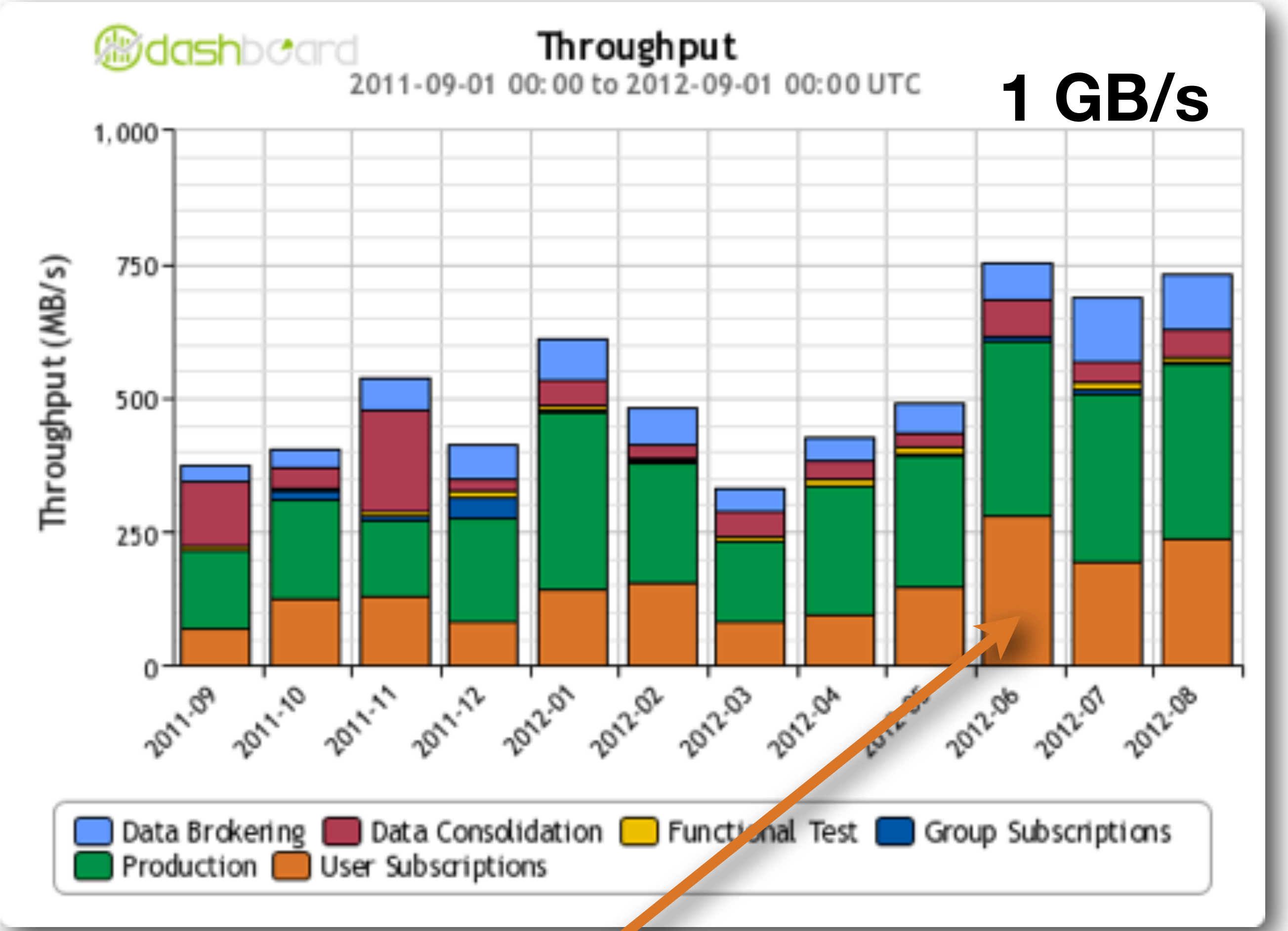
T2→T1

54,491 TB

16,487 TB



2 GB/s



1 GB/s

Reconstruction in T2s

User subscriptions!

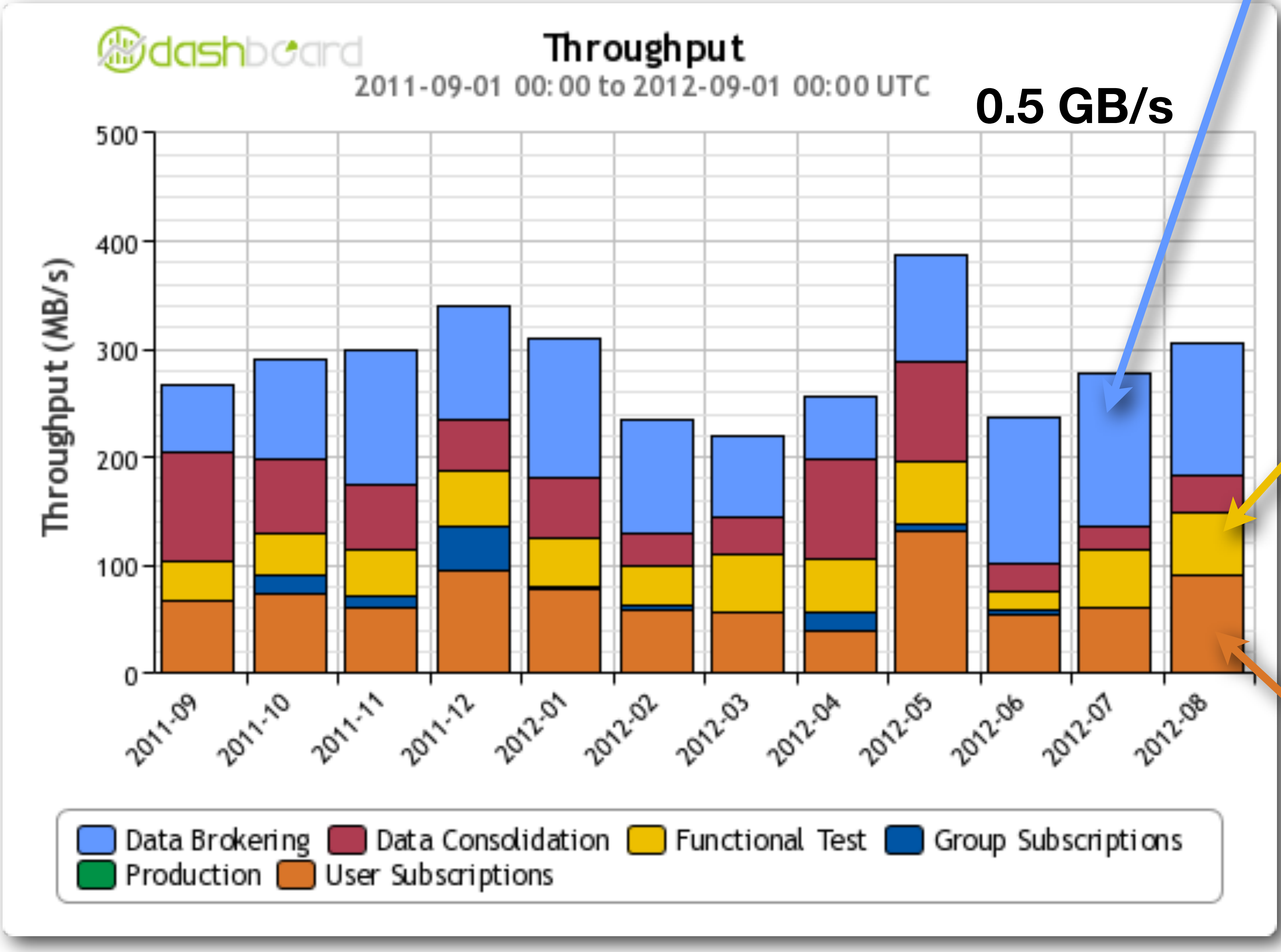
Dynamic data placement > pre-defined

Not in original computing model

T2→T2

Dynamic data placement > pre-defined

9,050 TB



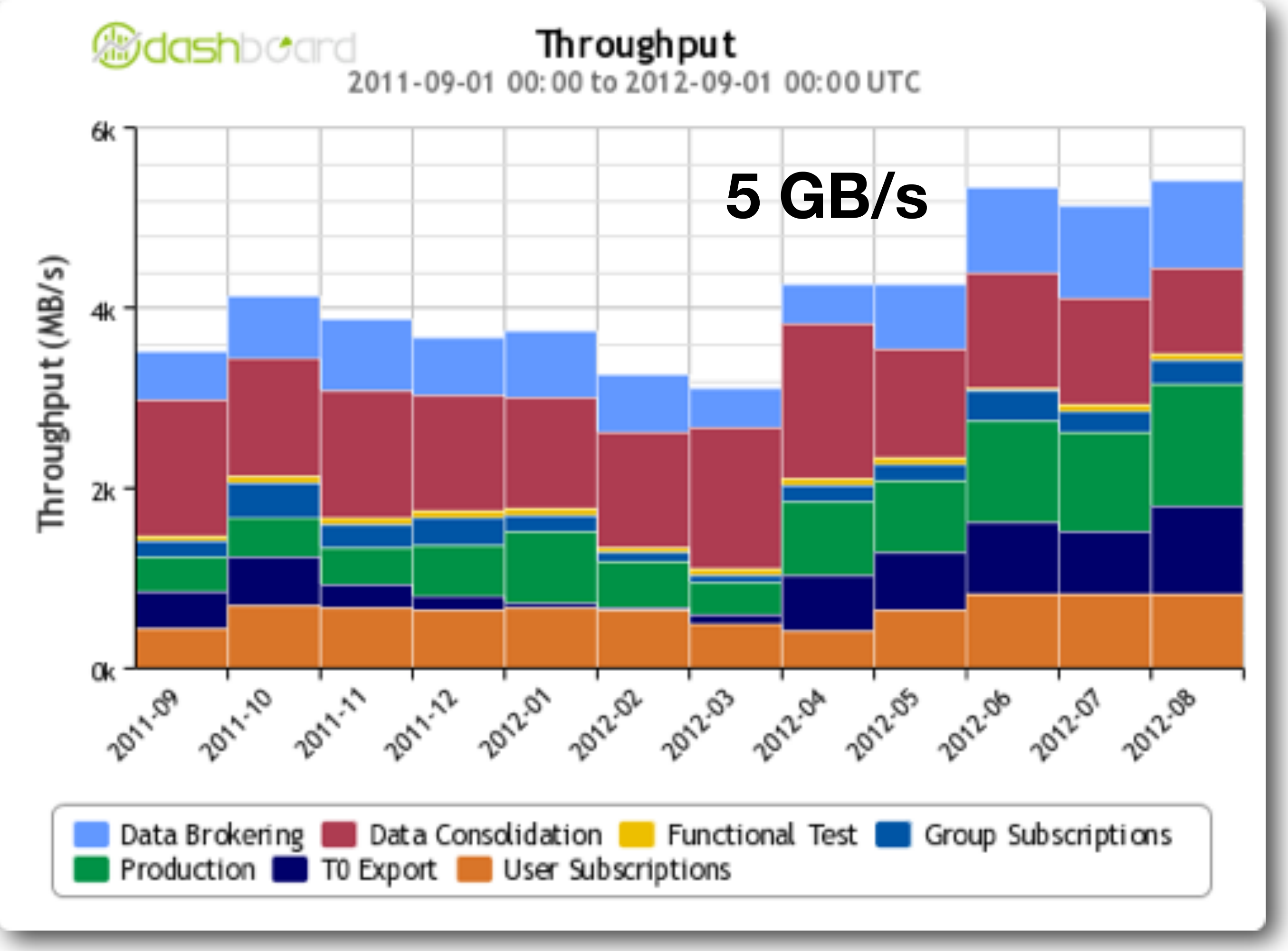
0.5 GB/s

Network mesh tests

User subscriptions
Group data at some T2s
+ Outputs of analysis

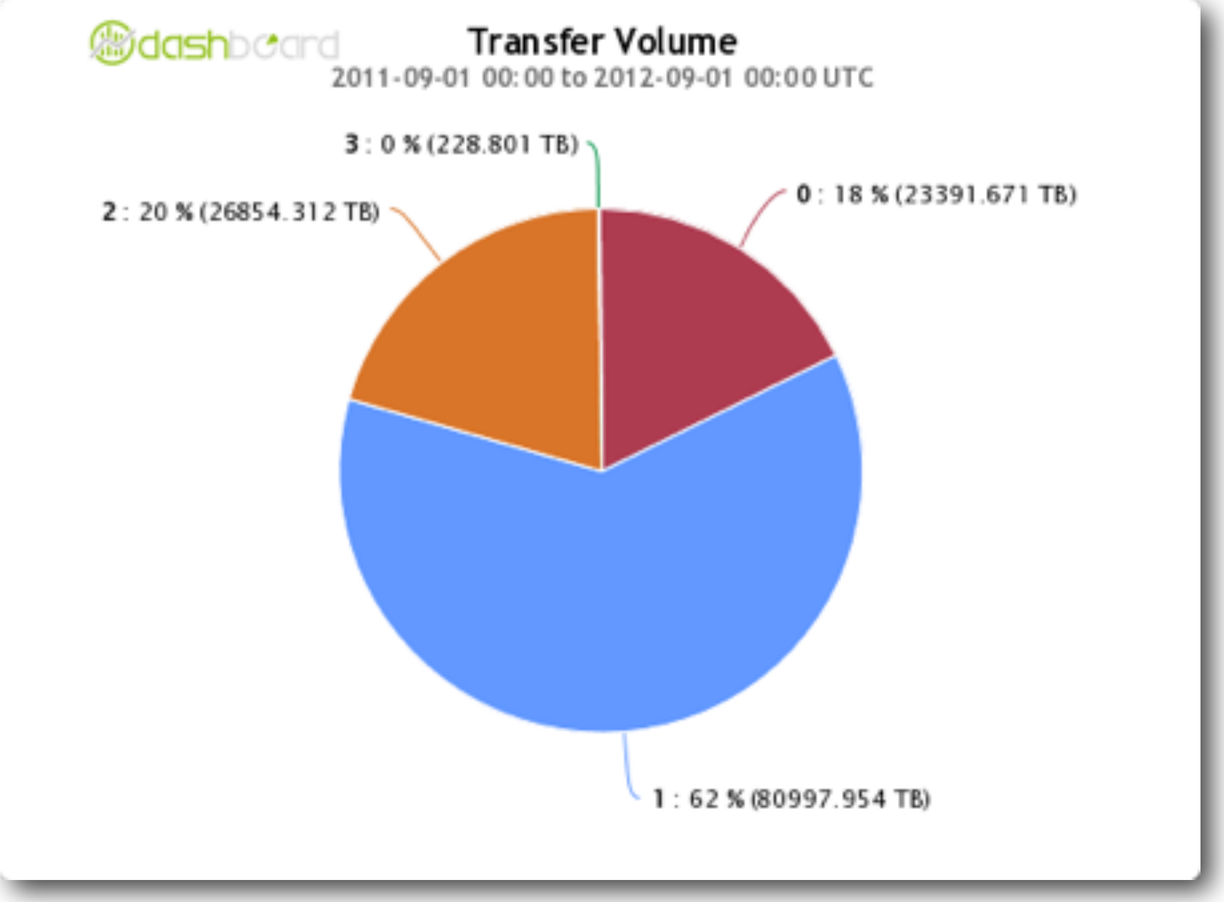
ALL together

131,473 TB

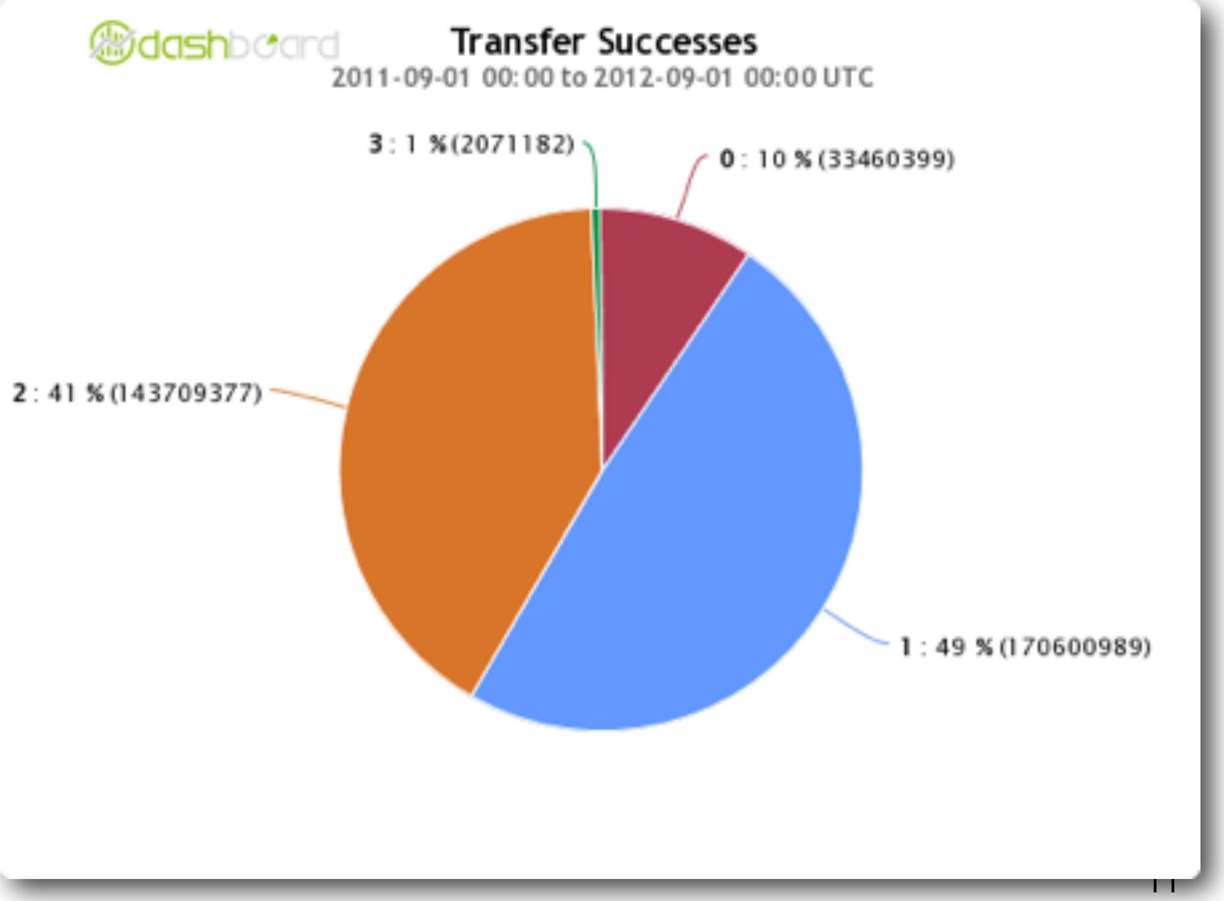


5 GB/s

Data volume: T1s 60% of sources

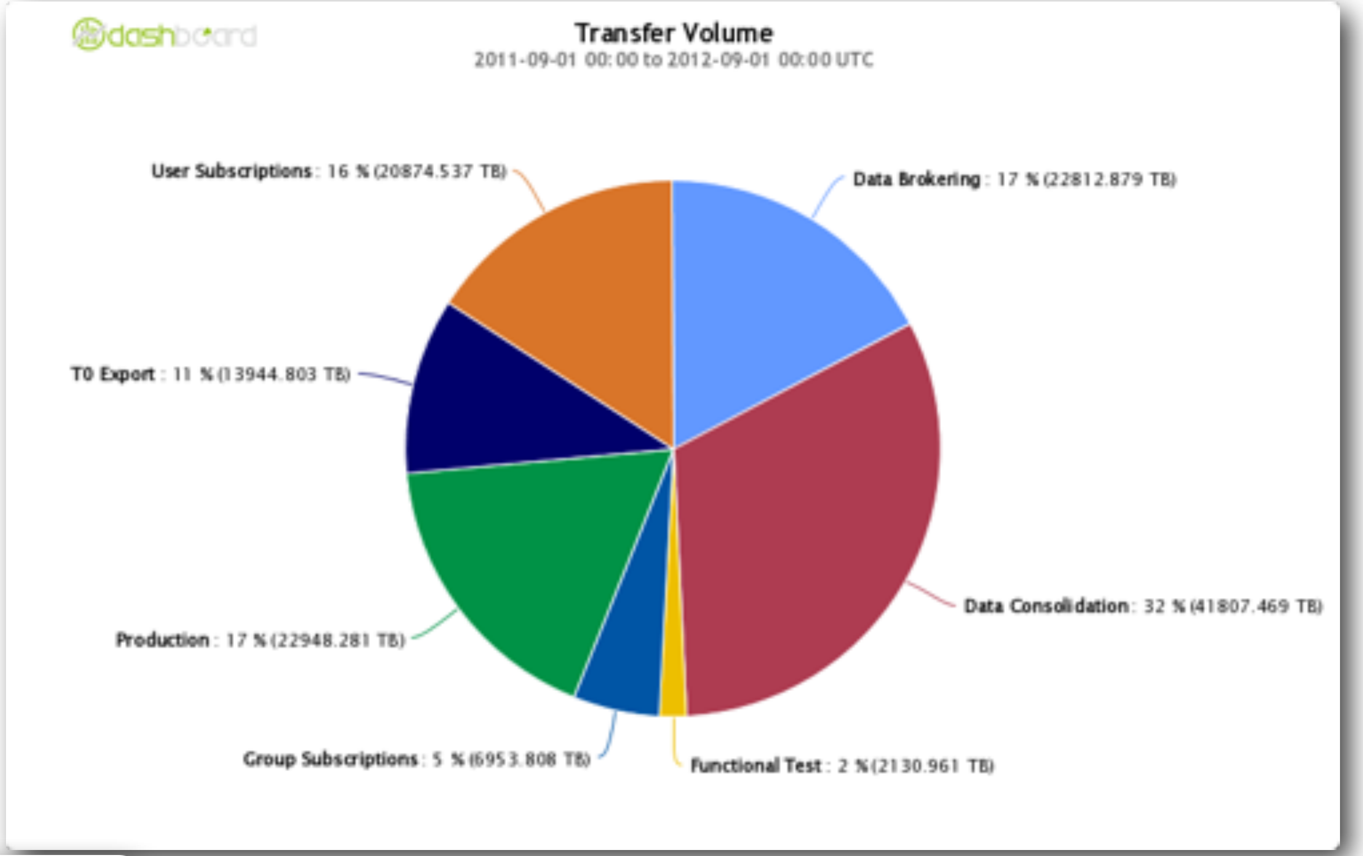


Activity: T2s 40% of sources

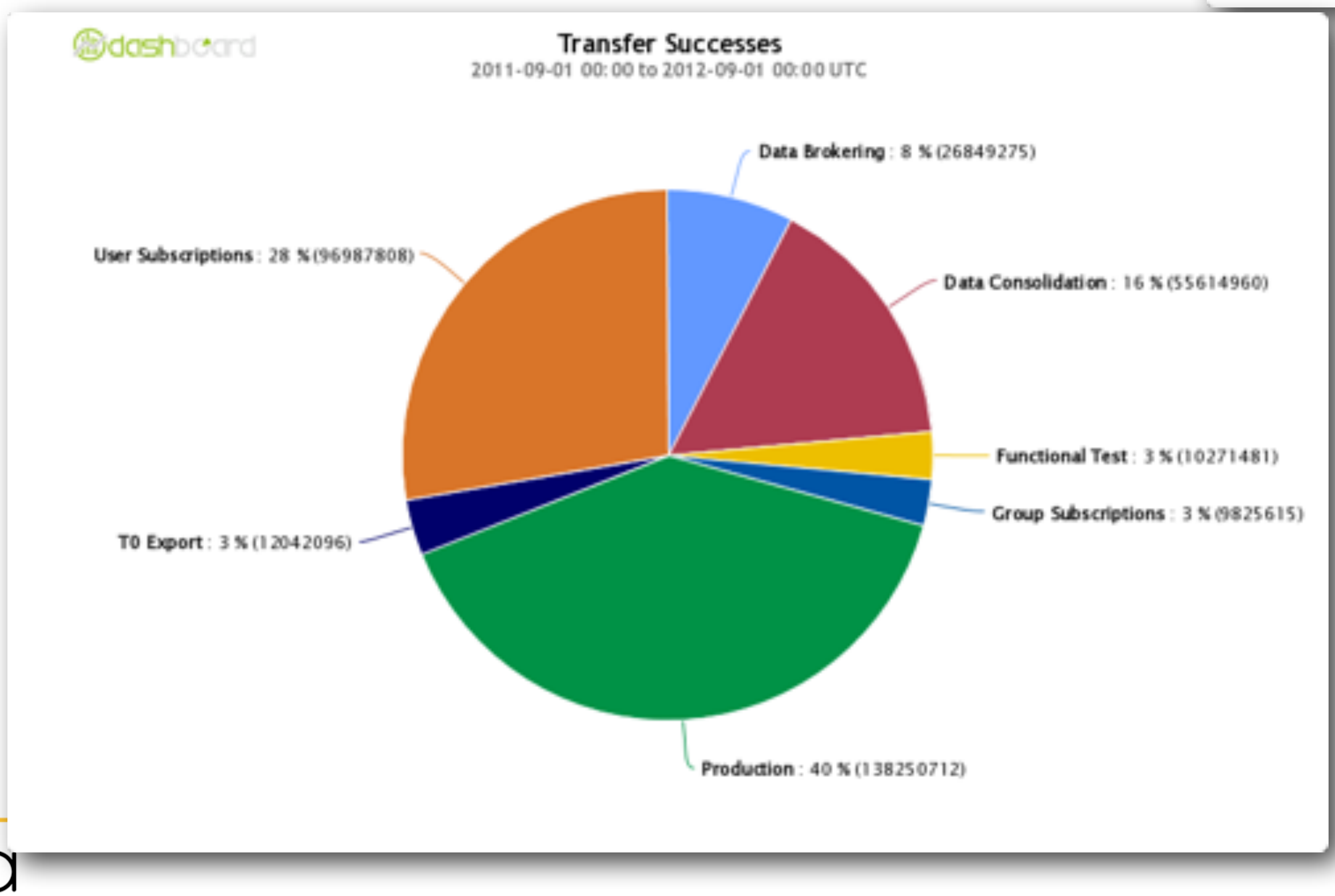


ALL together

Data volume: pre-placement ~2 times
dynamic placement room for improvements

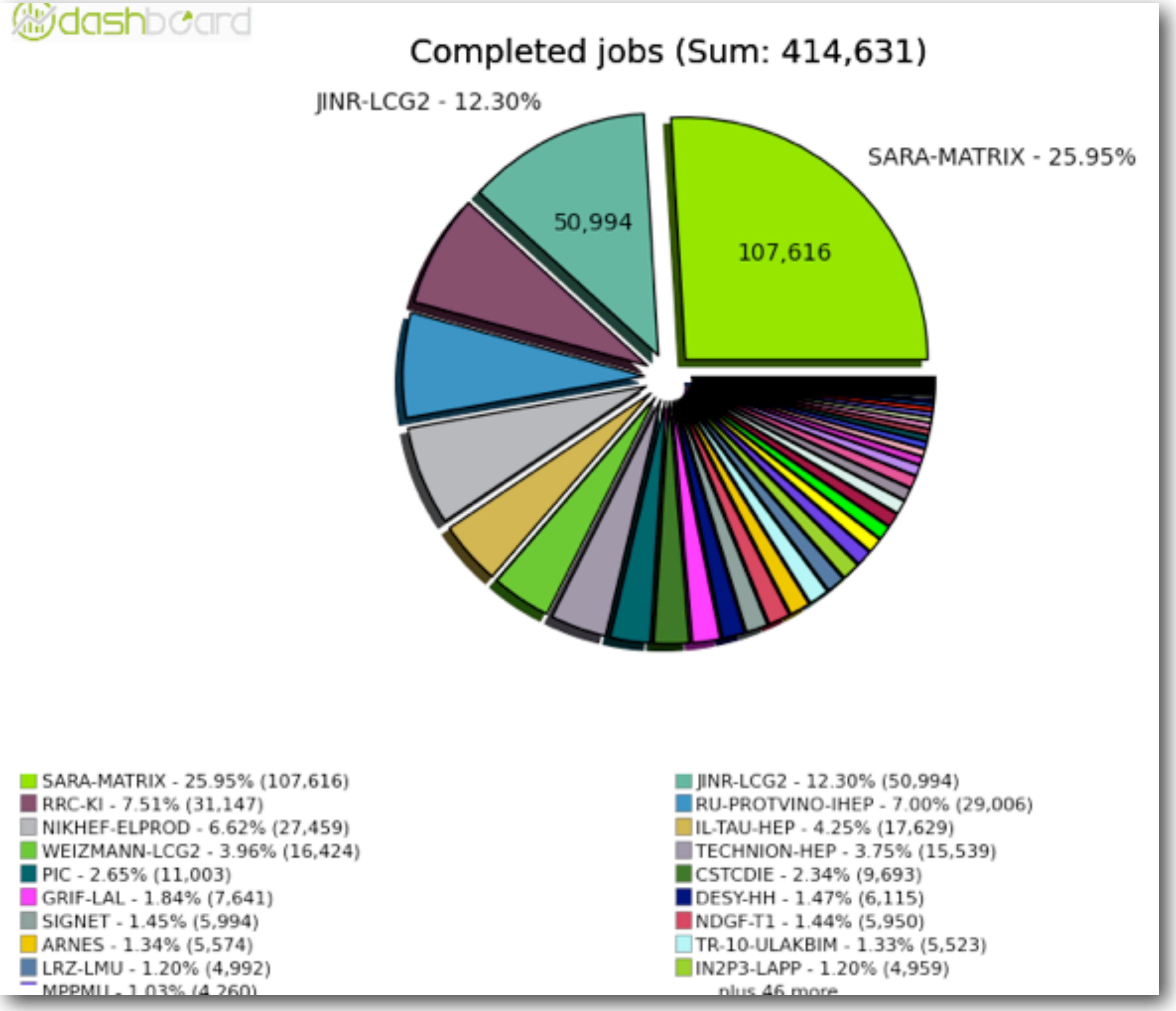


Activity: users ~30% of transfers



Cross-cloud MC production

- The ‘easy’ part
- No need to be a T2D
- Only connection to remote T1 needed
- Example : NL cloud
  - 65 ! sites contributing



The French cloud

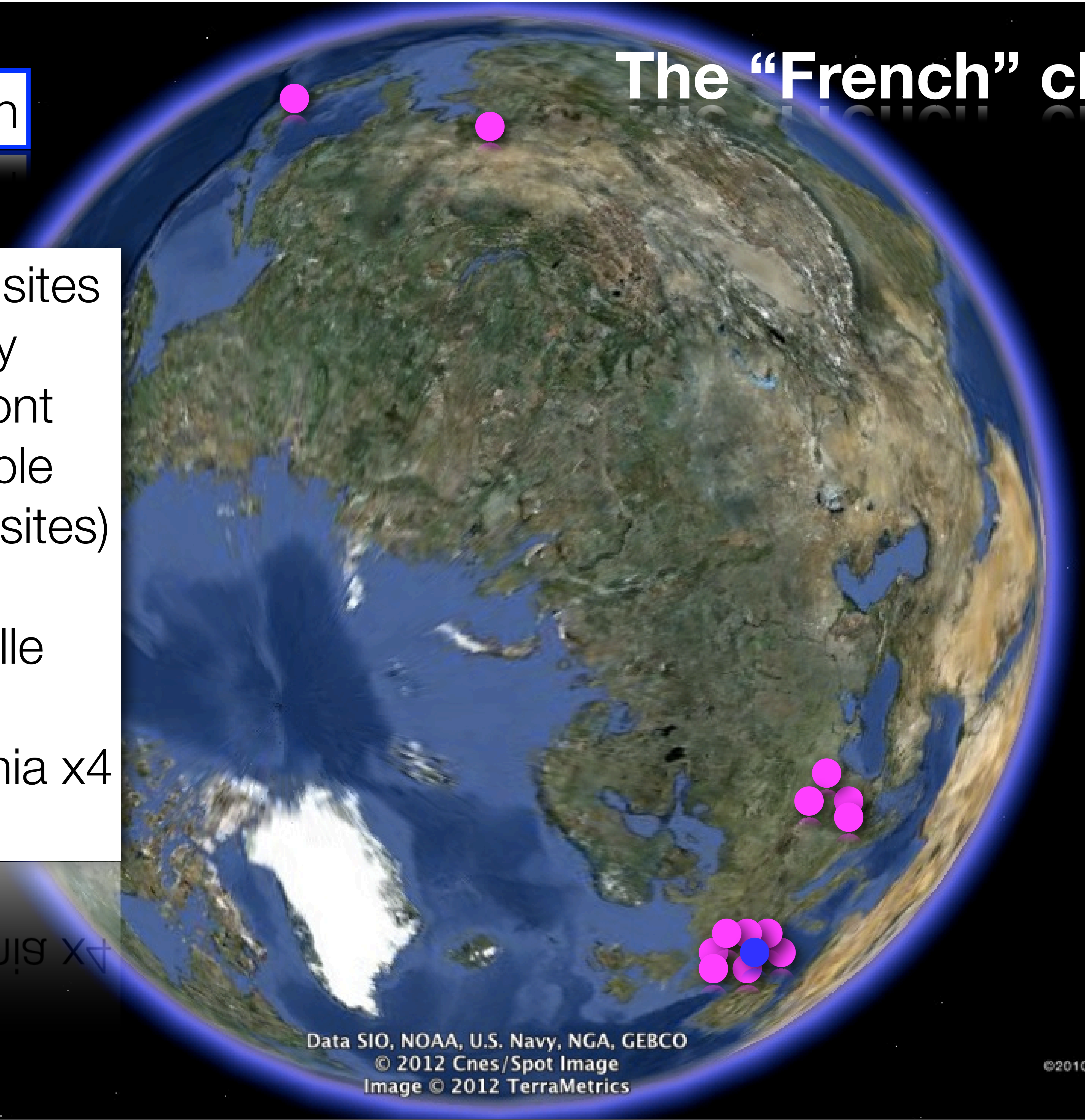
- The most ‘exploded’ cloud of ATLAS
- 4 Romanians sites at the far end of GEANT
- 2 sites in far east Beijing & Tokyo connected to CCIN2P3 via different paths

The “French” cloud

T1 : Lyon

T2s : 14 sites

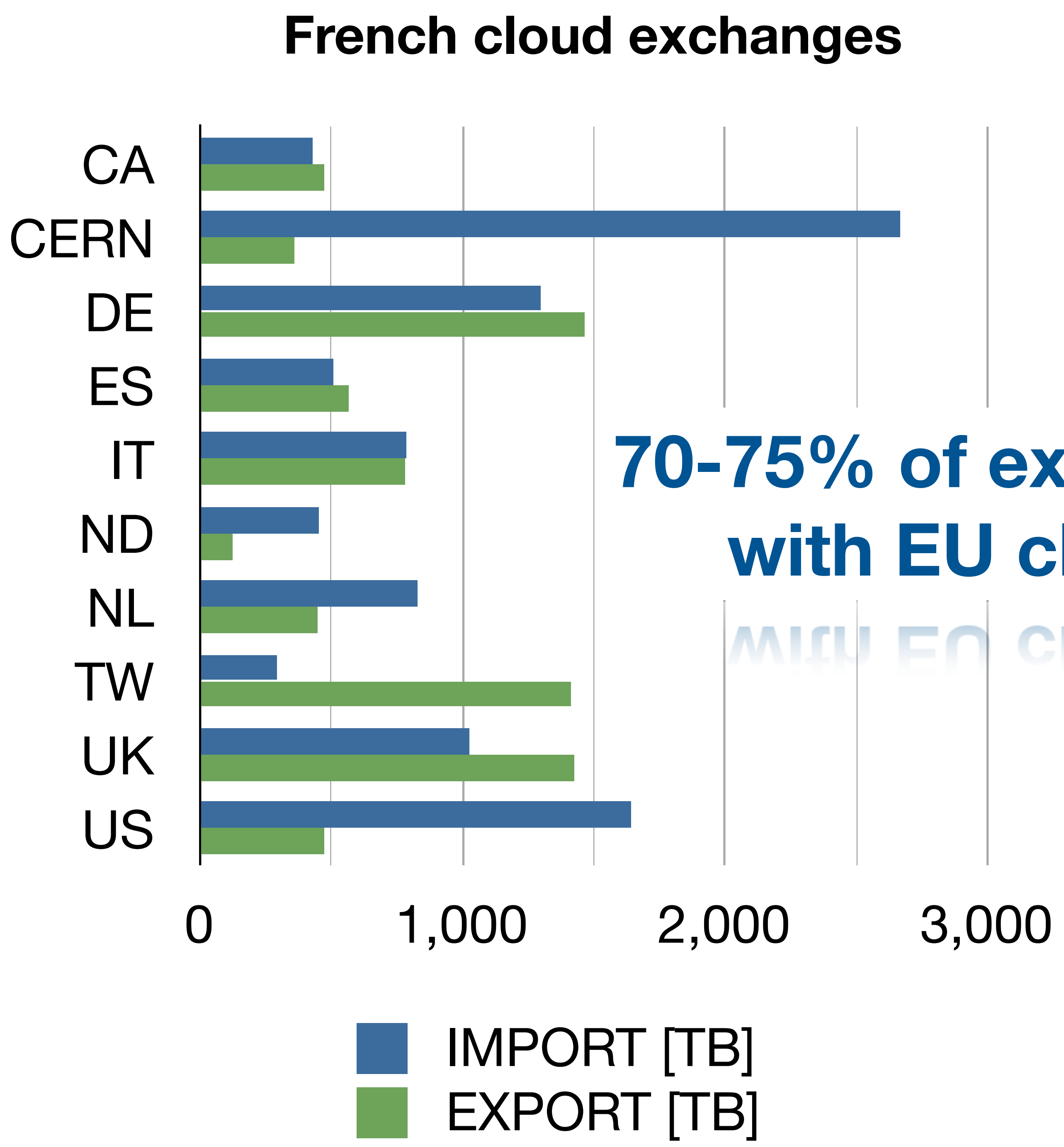
- Annecy
- Clermont
- Grenoble
- Grif (3 sites)
- Lyon
- Marseille
- Beijing
- Romania x4
- Tokyo



Data SIO, NOAA, U.S. Navy, NGA, GEBCO
© 2012 Cnes/Spot Image
Image © 2012 TerraMetrics

©2010 Google

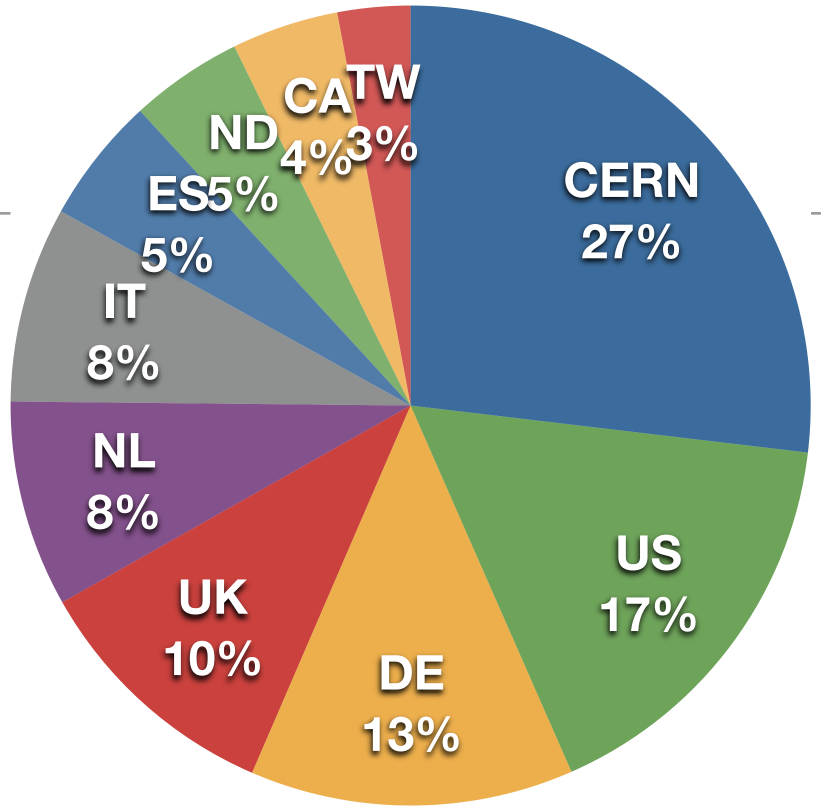
Data exchanges for FR-cloud



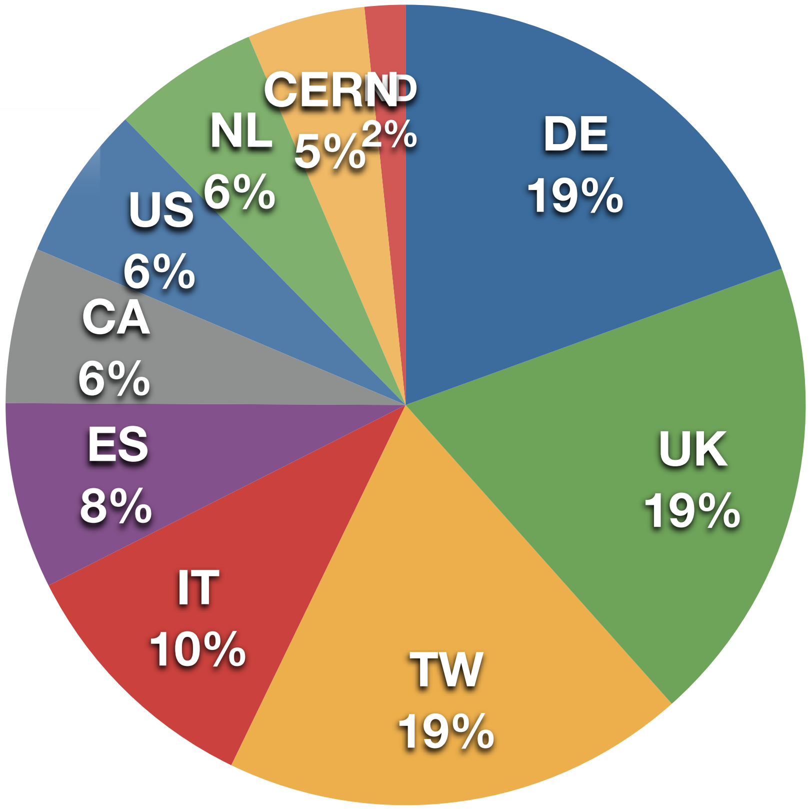
70-75% of exchanges with EU clouds

[Sep. 2011 - Sep. - 2012]

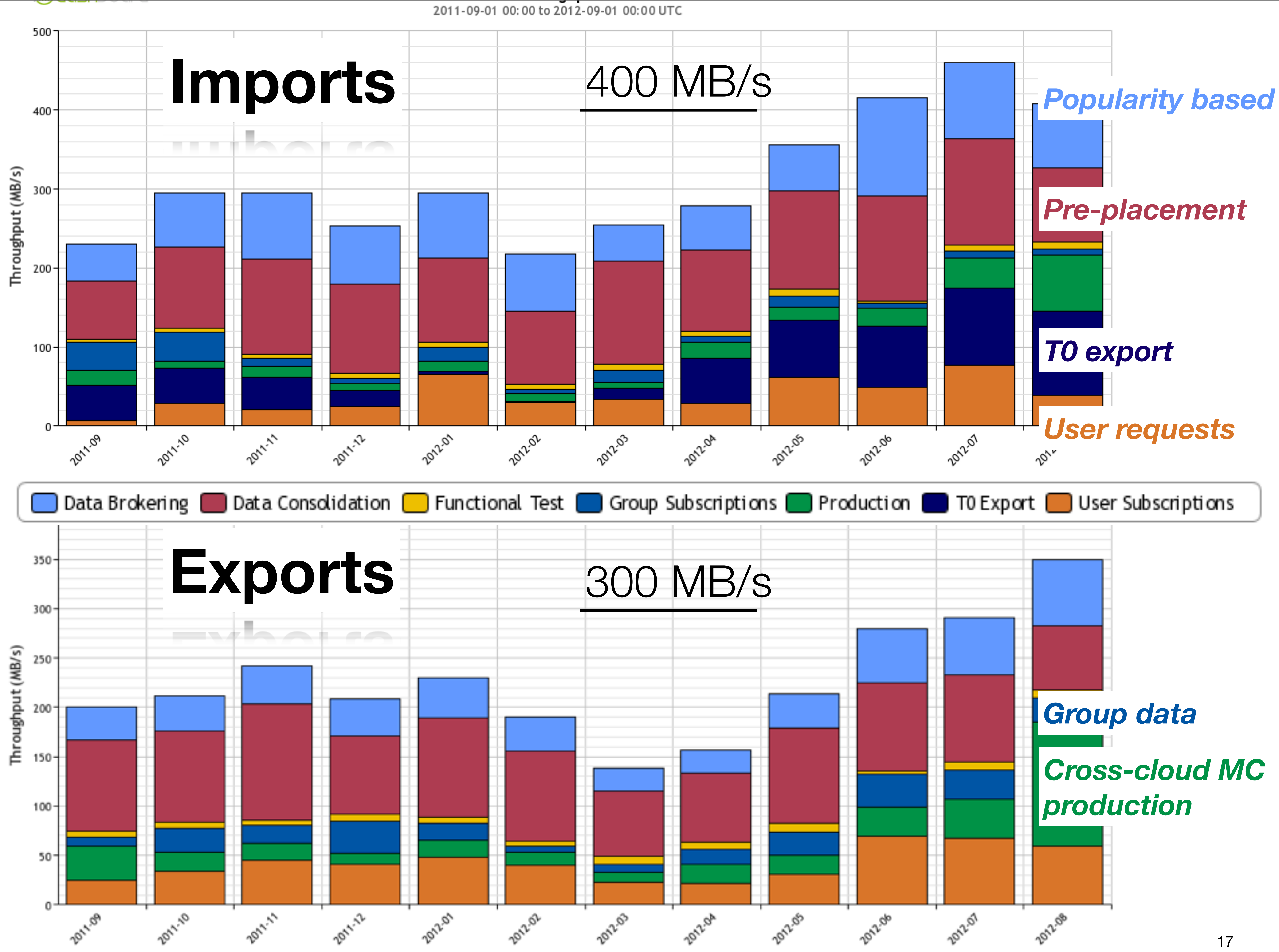
Imports : 9,9 PB



Exports : 7,5 PB



cross-cloud production



Imports

400 MB/s

Popularity based

Pre-placement

T0 export

User requests

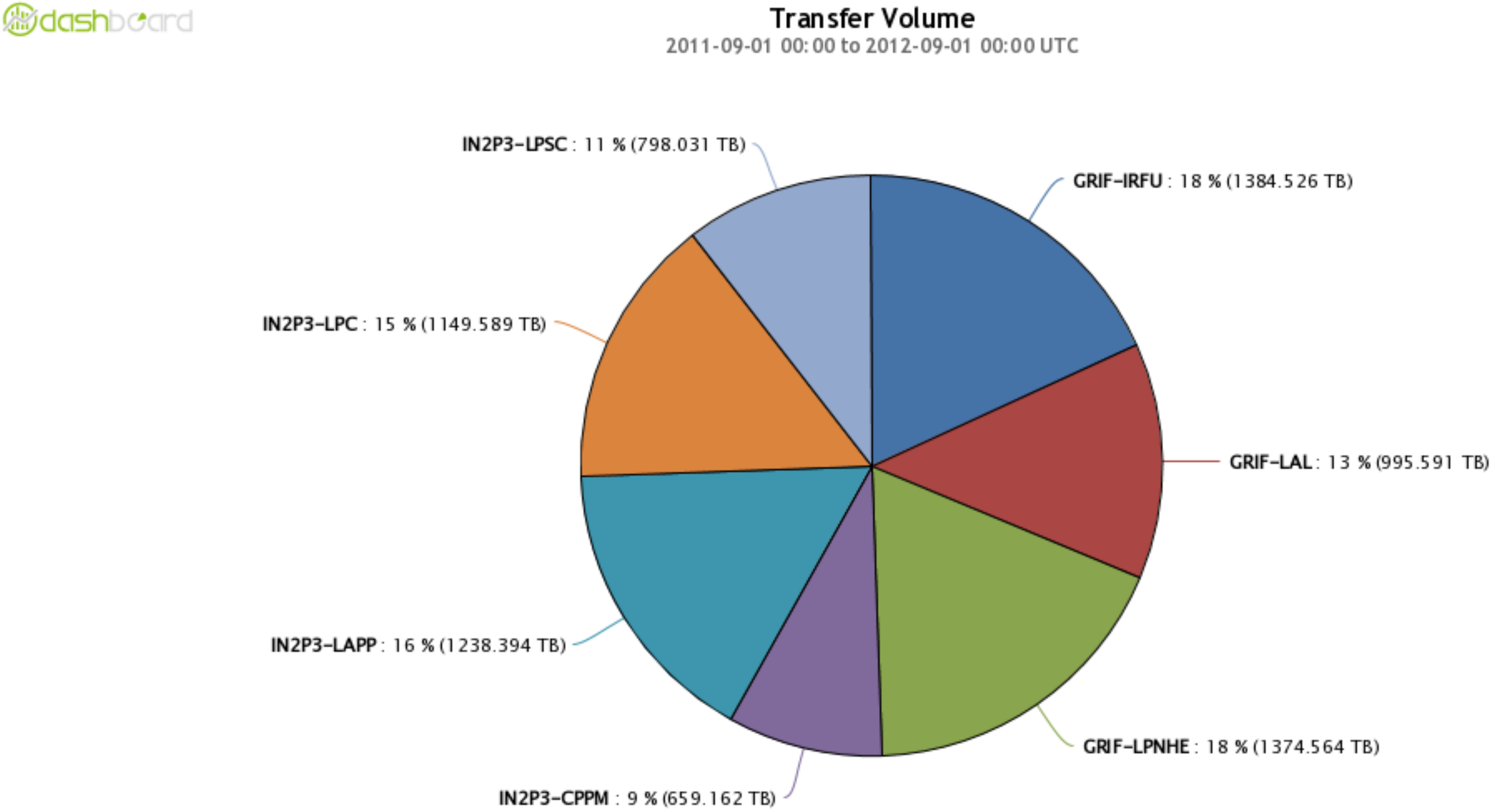
Exports

300 MB/s

Group data

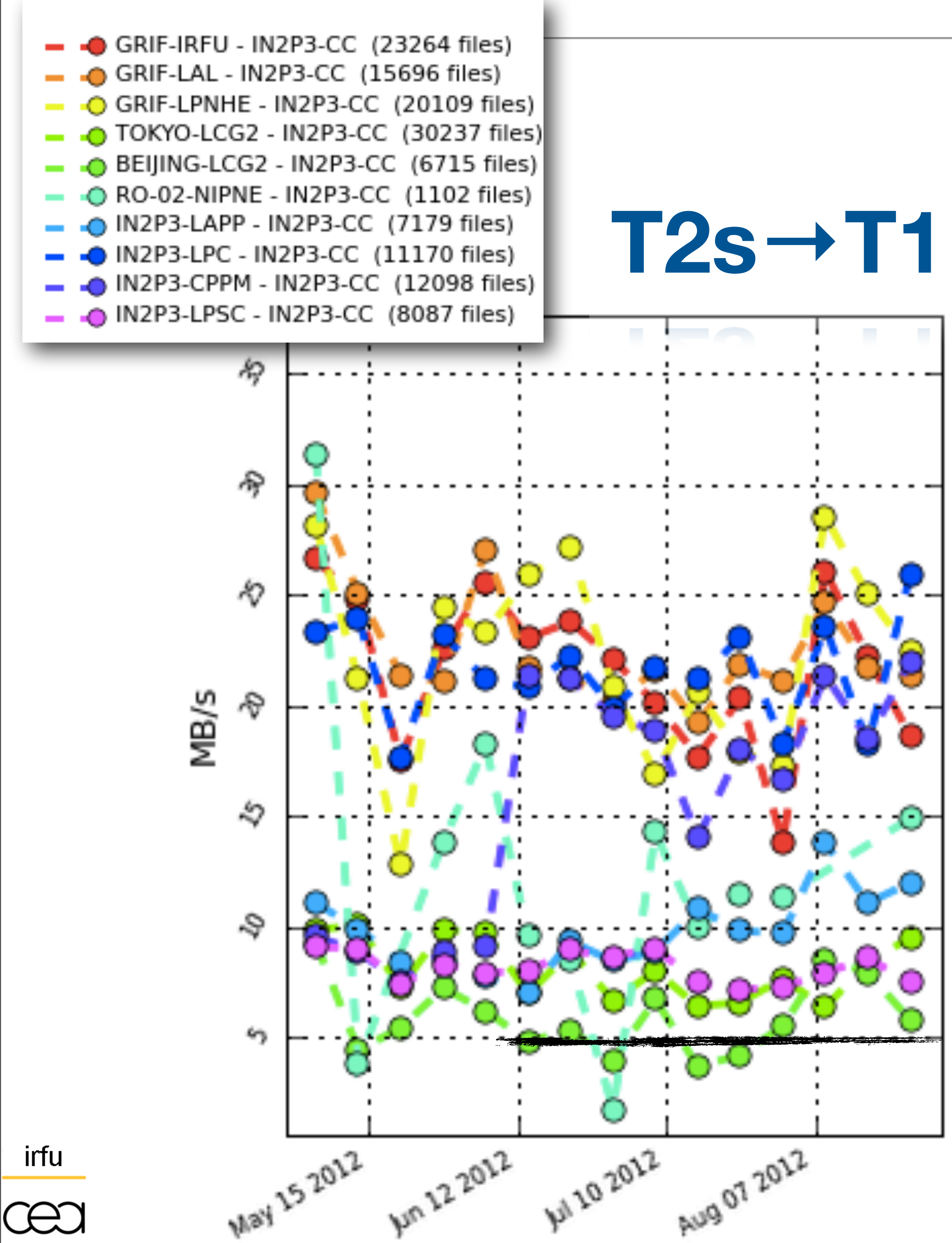
Cross-cloud MC production

Data volume transferred to French T2s

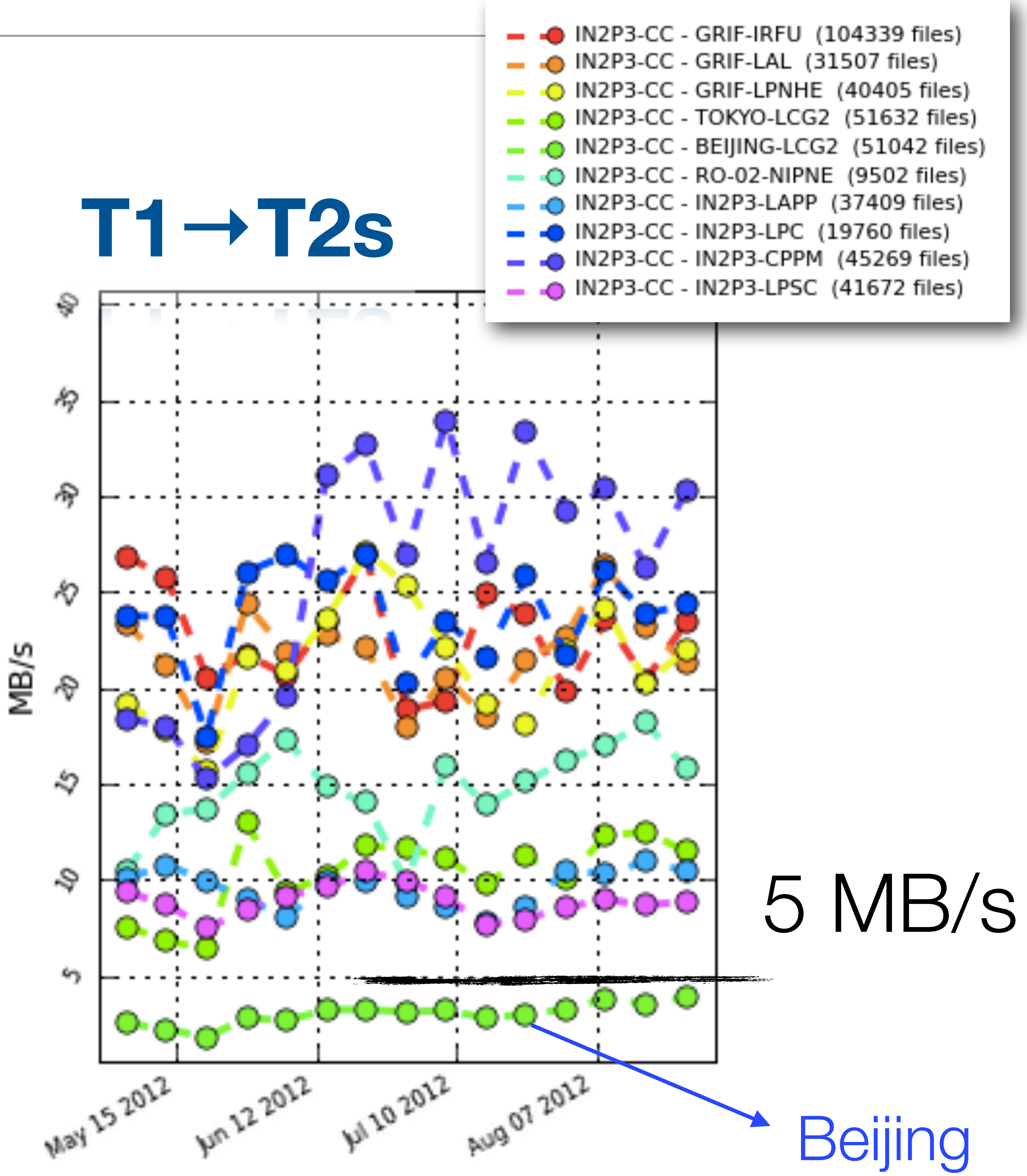


not proportional to number physicist nor CPUs

Connectivity within French cloud (ATLAS sonar)



T2s→T1



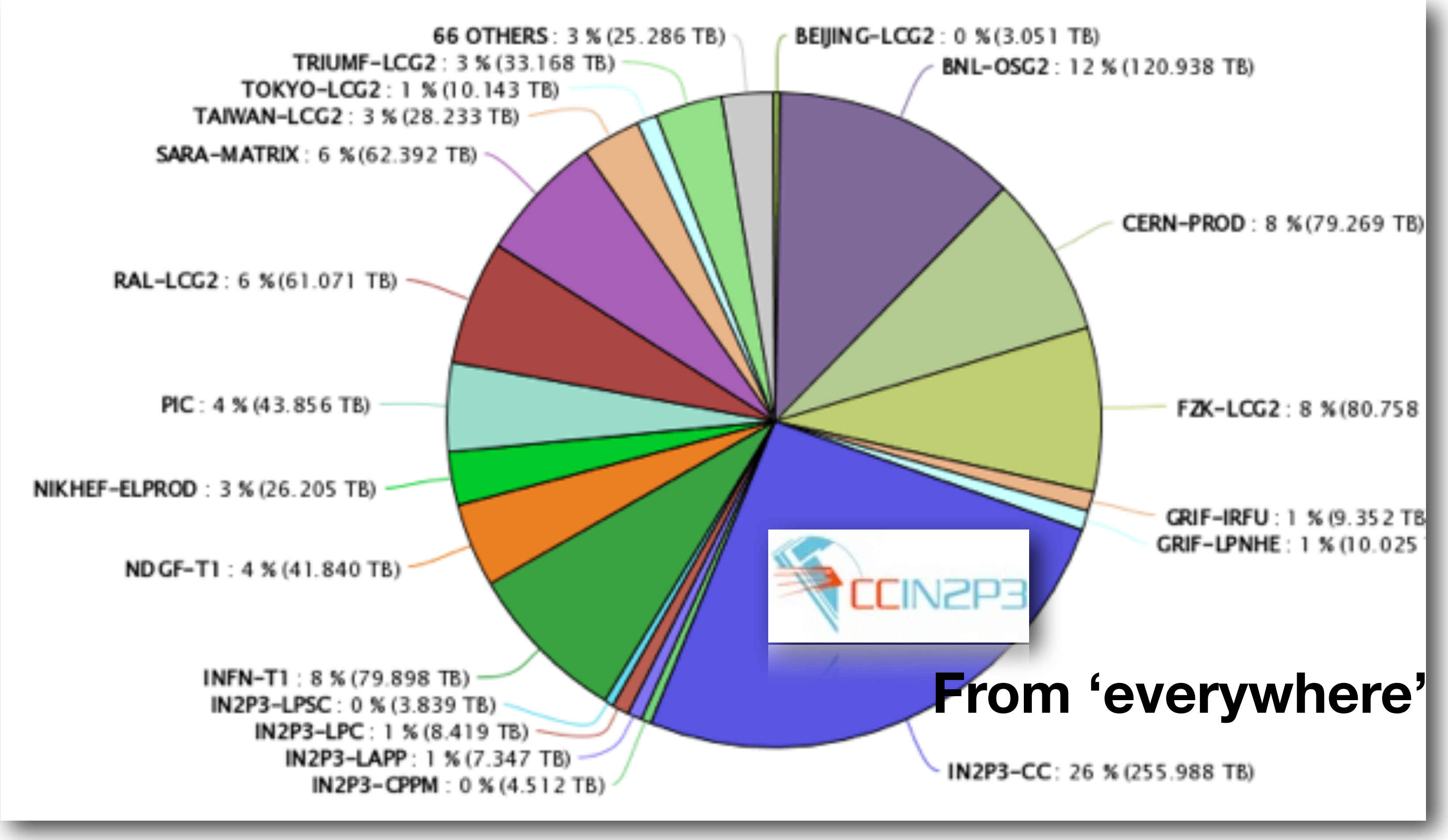
T1→T2s

5 MB/s

Beijing

Origin of data transferred to 2 GRIF sites

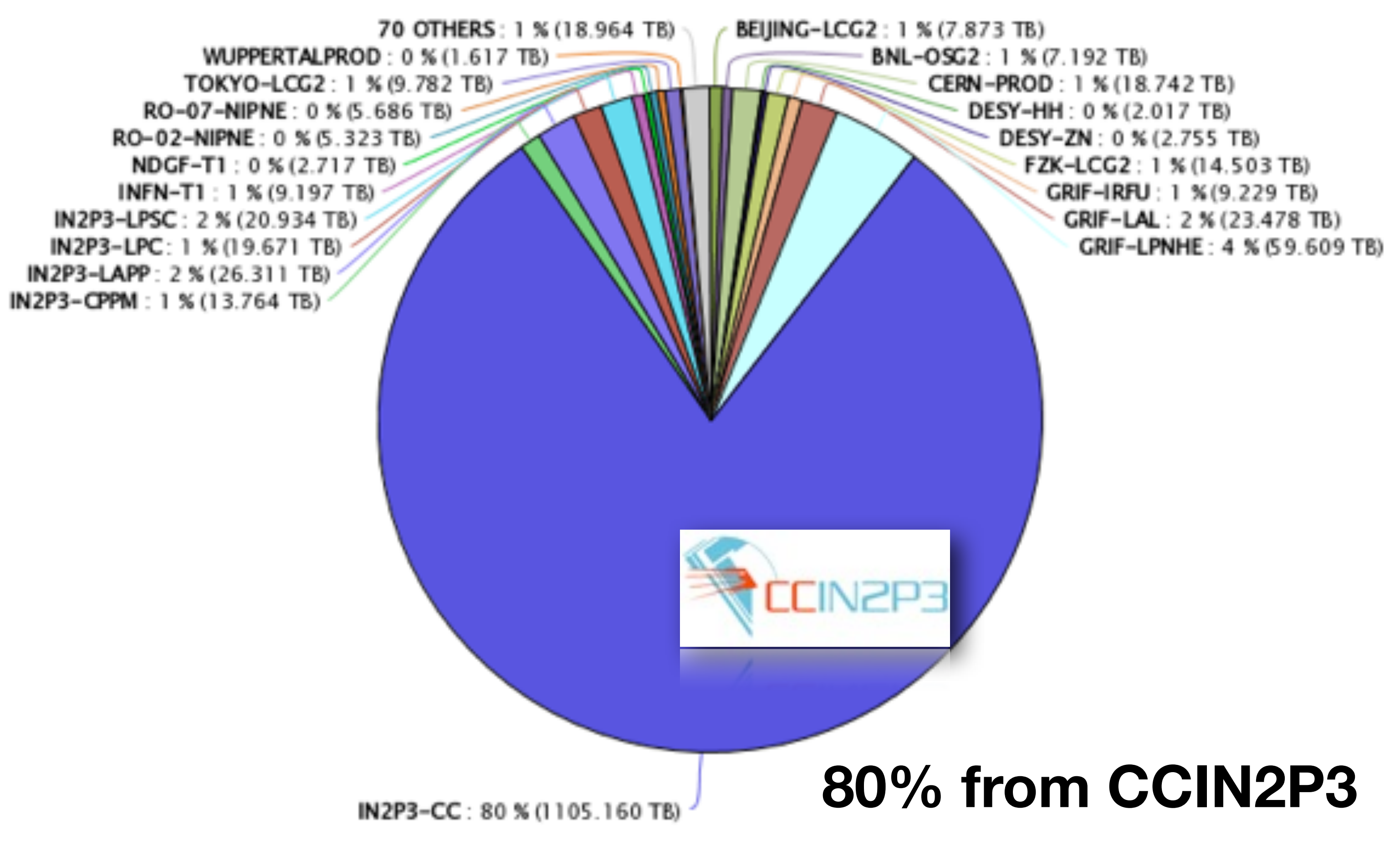
LAL : T2D connected to LHCONE



From ‘everywhere’

IRFU : ~~T2D~~

~~connected to LHCONE~~



80% from CCIN2P3

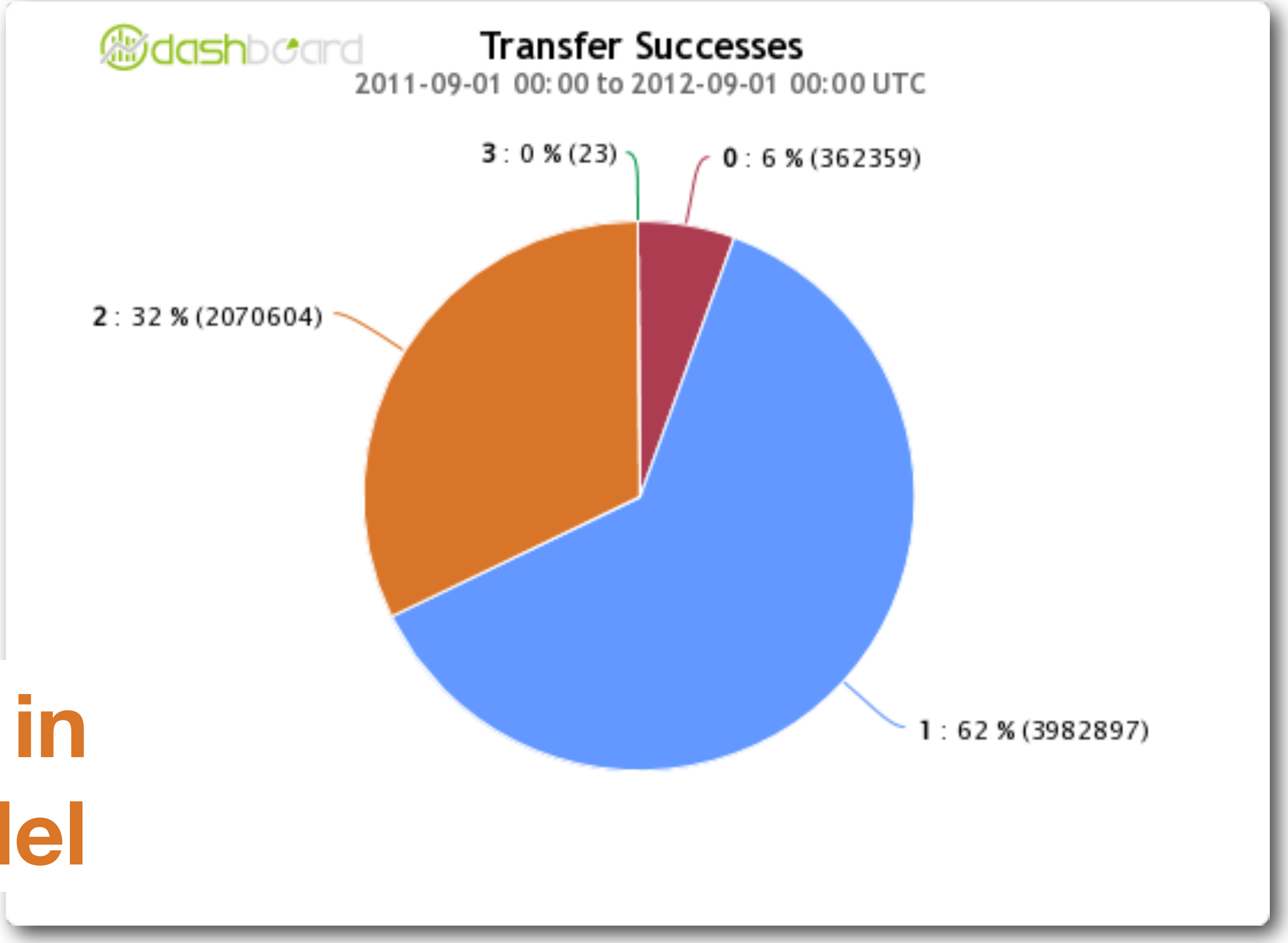
User & group analysis

- Most of users run final analysis at their local site : delivery of data or analysis job outputs to users **very sensitive** (the last transferred file determines the efficiency)
- 2 ways to get (reduced format) data
  - The majority : Let PanDA decide where to run the jobs ; where data are (in most of the cases). Outputs stored on SCRATCHDISK (buffer area) at remote sites have to be shipped back to user local site
  - Get limited volume of data at local site to run locally analysis
- Both imply data transfers from remote site to local site
  - Direct transfers for T2Ds
  - Through T1 for other T2s

User + Group transfers to FR-cloud sites

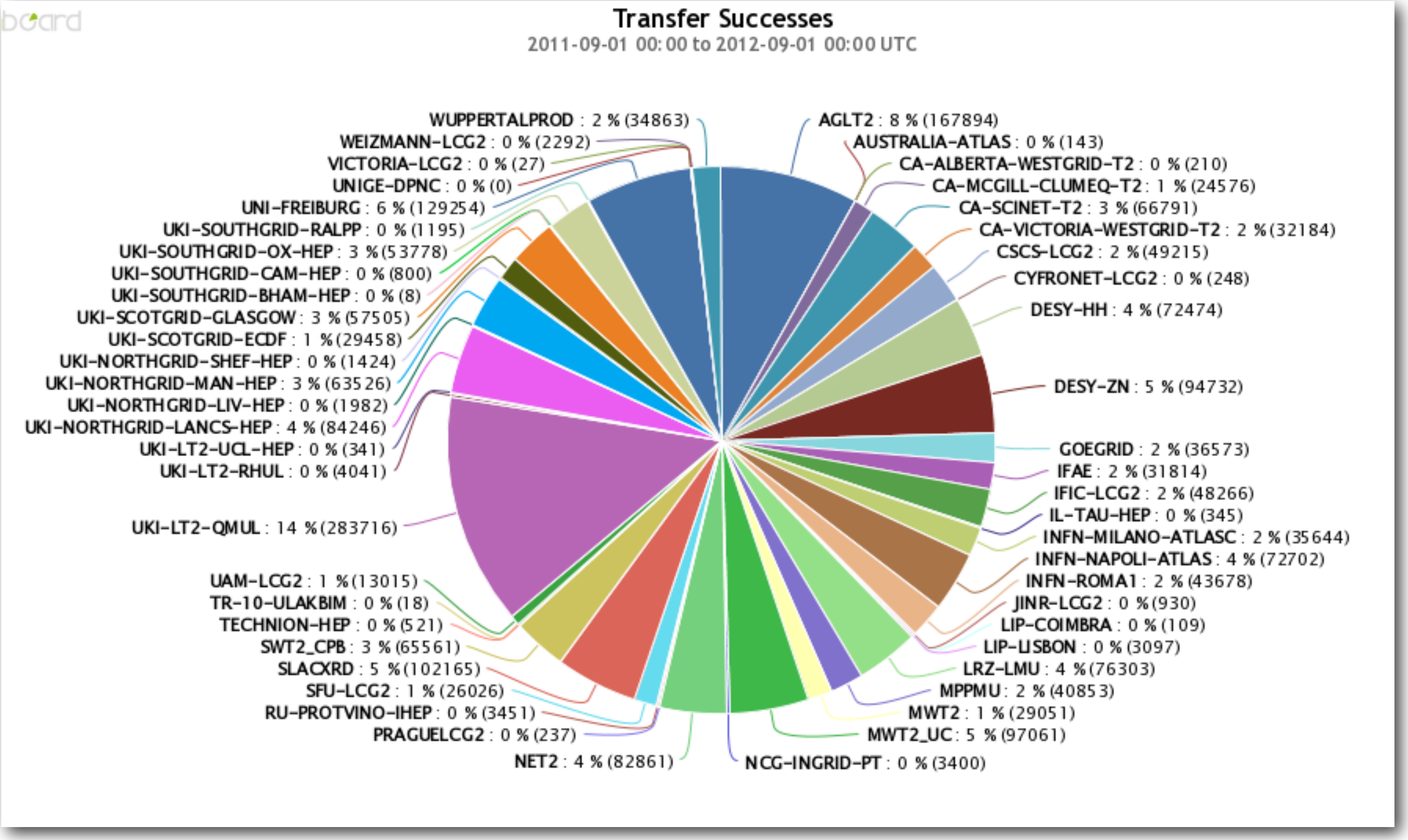
⅓ of data transfers (not data volume) used for final analysis from **T2 sites outside** FR-cloud

1,828 TB

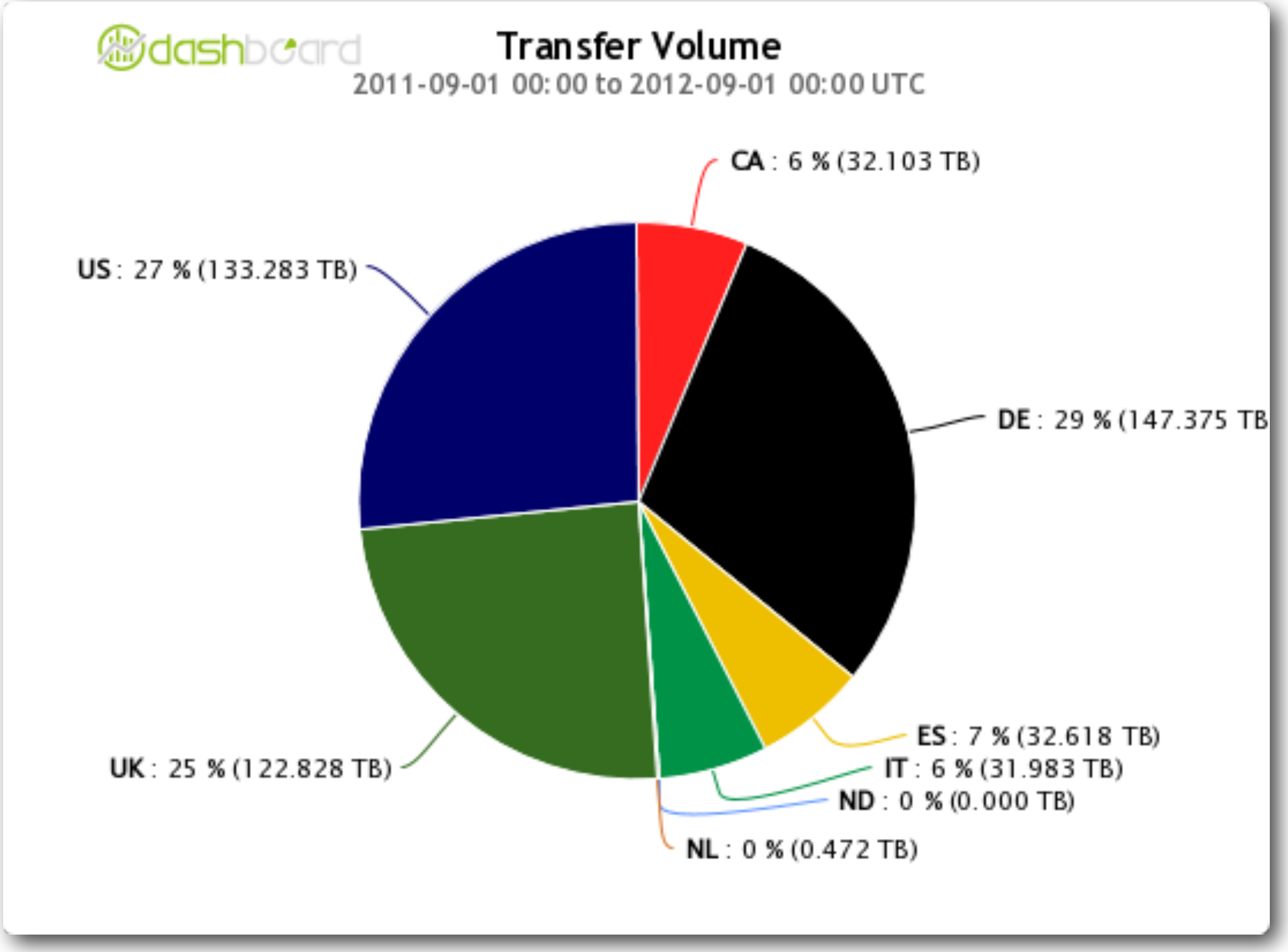


Not allowed in original model

Data come from 52 T2 sites



⅓ from non EU T2s



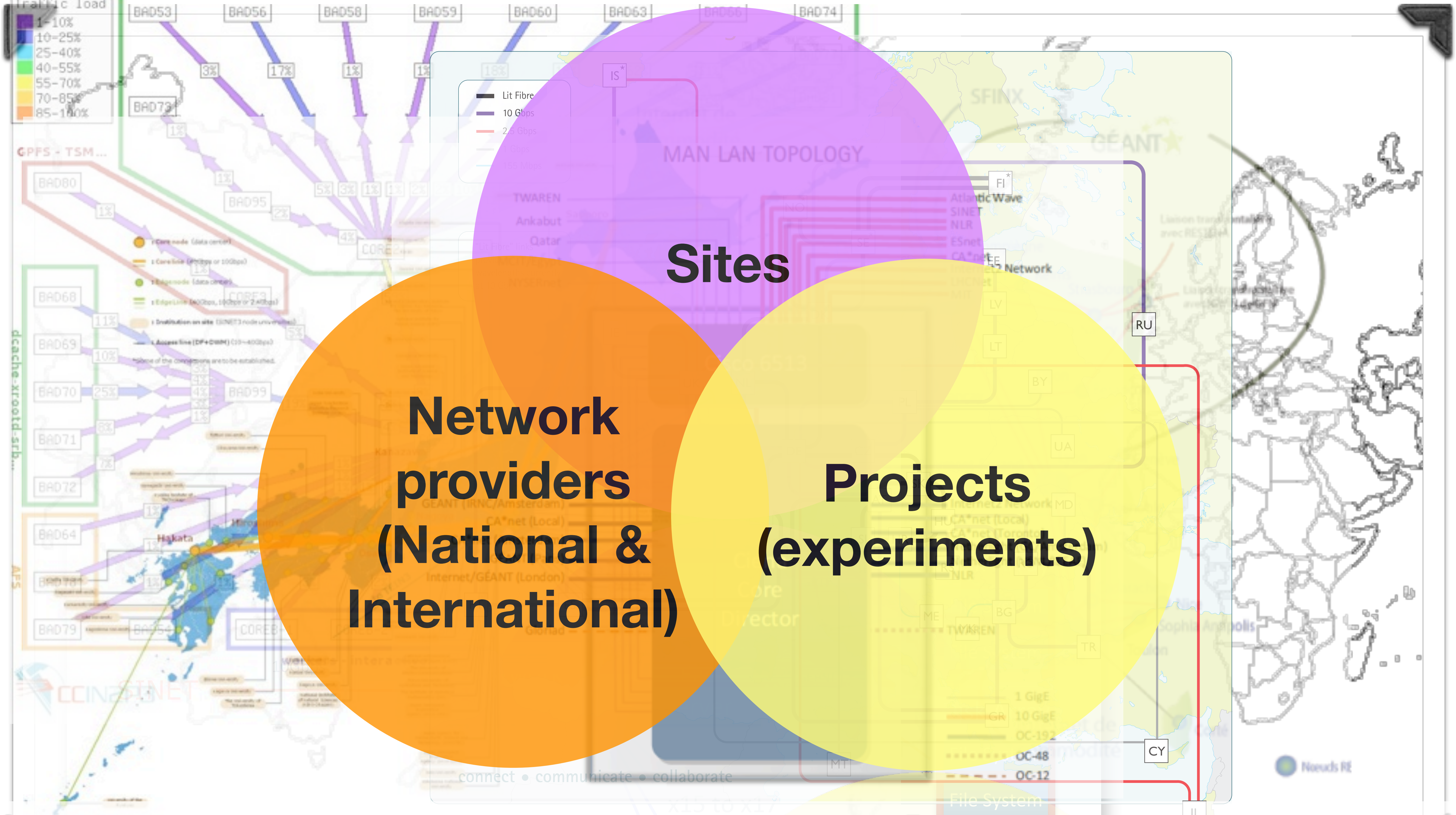
1/4 from UK (not on LHCONE)

Issues with distant T2s

Beijing

- **Beijing** :
  - Connected to Europe via GEANT/TEIN3 (except CERN : GLORIAD/KREONET)
  - RTT ~190 ms
- **Tokyo** :
  - Connected to Europe via GEANT/MANLAN/SINET4
  - RTT ~ 300 ms
- Several network operators on the path (Nationals, GEANT, …)

The difficulty to solve network issues



Sites

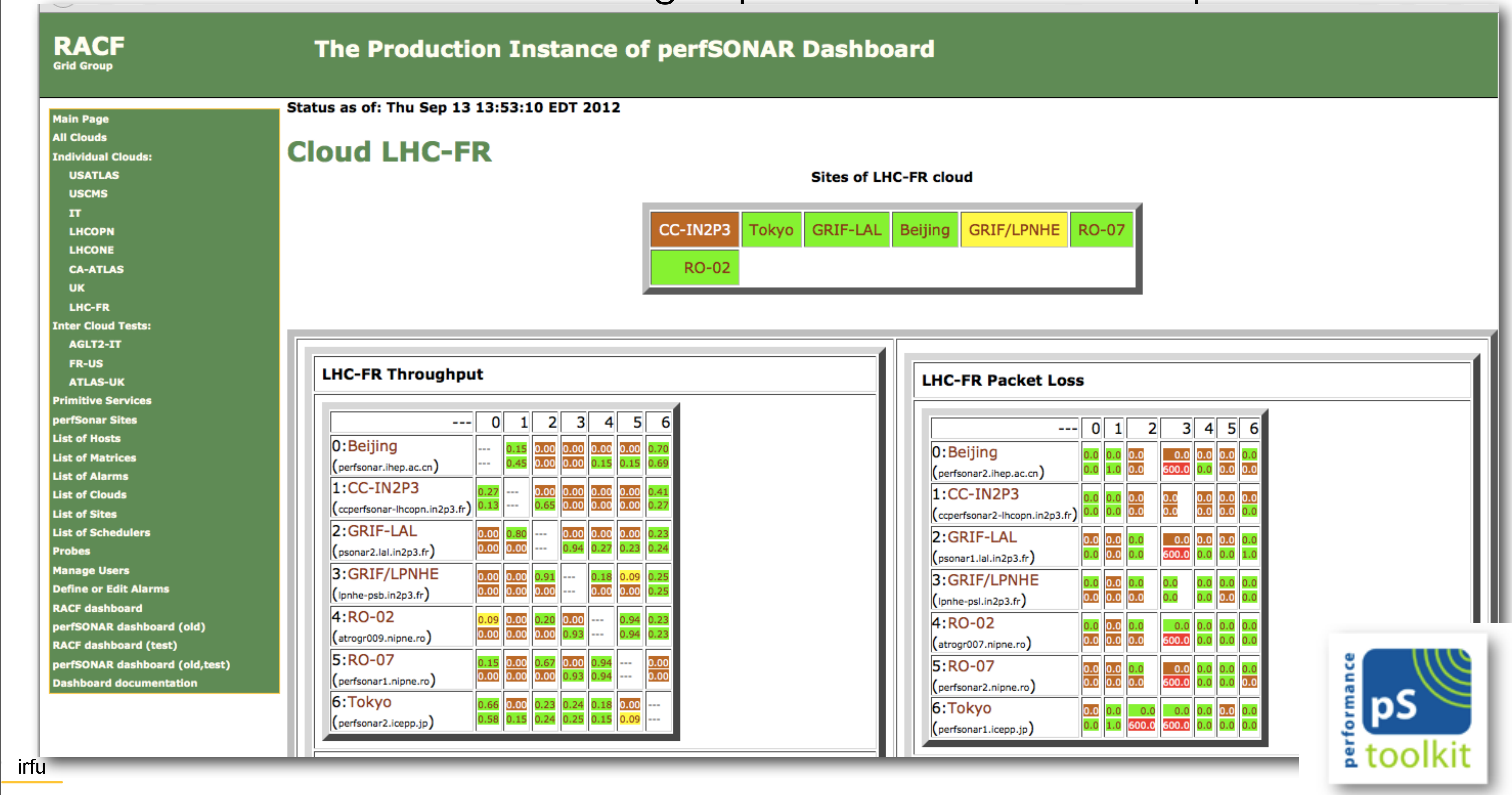
Network providers (National & International)

Projects (experiments)

Various approaches and complementary tools needed

perfSonar dashboard of FR-cloud

Being expanded as sites install perfSonar



RACF
Grid Group

The Production Instance of perfSONAR Dashboard

- Main Page
- All Clouds
- Individual Clouds:
  - USATLAS
  - USCMS
  - IT
  - LHCOPN
  - LHCONE
  - CA-ATLAS
  - UK
  - LHC-FR
- Inter Cloud Tests:
  - AGLT2-IT
  - FR-US
  - ATLAS-UK
- Primitive Services
- perfSonar Sites
- List of Hosts
- List of Matrices
- List of Alarms
- List of Clouds
- List of Sites
- List of Schedulers
- Probes
- Manage Users
- Define or Edit Alarms
- RACF dashboard
- perfSONAR dashboard (old)
- RACF dashboard (test)
- perfSONAR dashboard (old,test)
- Dashboard documentation

Status as of: Thu Sep 13 13:53:10 EDT 2012

Cloud LHC-FR

Sites of LHC-FR cloud

CC-IN2P3	Tokyo	GRIF-LAL	Beijing	GRIF/LPNHE	RO-07
RO-02					

LHC-FR Throughput

---	0	1	2	3	4	5	6
0:Beijing (perfsonar.ihep.ac.cn)	--- ---	0.15 0.45	0.00 0.00	0.00 0.00	0.00 0.15	0.00 0.15	0.70 0.69
1:CC-IN2P3 (ccperfsonar-lhcopn.in2p3.fr)	0.27 0.13	--- ---	0.00 0.65	0.00 0.00	0.00 0.00	0.00 0.00	0.41 0.27
2:GRIF-LAL (psonar2.lal.in2p3.fr)	0.00 0.00	0.80 0.00	--- ---	0.00 0.94	0.00 0.27	0.00 0.23	0.23 0.24
3:GRIF/LPNHE (lpnhe-psb.in2p3.fr)	0.00 0.00	0.00 0.00	0.91 0.00	--- ---	0.18 0.00	0.09 0.00	0.25 0.25
4:RO-02 (atrogr009.nipne.ro)	0.09 0.00	0.00 0.00	0.20 0.00	0.00 0.93	--- ---	0.94 0.94	0.23 0.23
5:RO-07 (perfsonar1.nipne.ro)	0.15 0.00	0.00 0.00	0.67 0.00	0.00 0.93	0.94 0.94	--- ---	0.00 0.00
6:Tokyo (perfsonar2.icepp.jp)	0.66 0.58	0.00 0.15	0.23 0.24	0.24 0.25	0.18 0.15	0.00 0.09	--- ---

LHC-FR Packet Loss

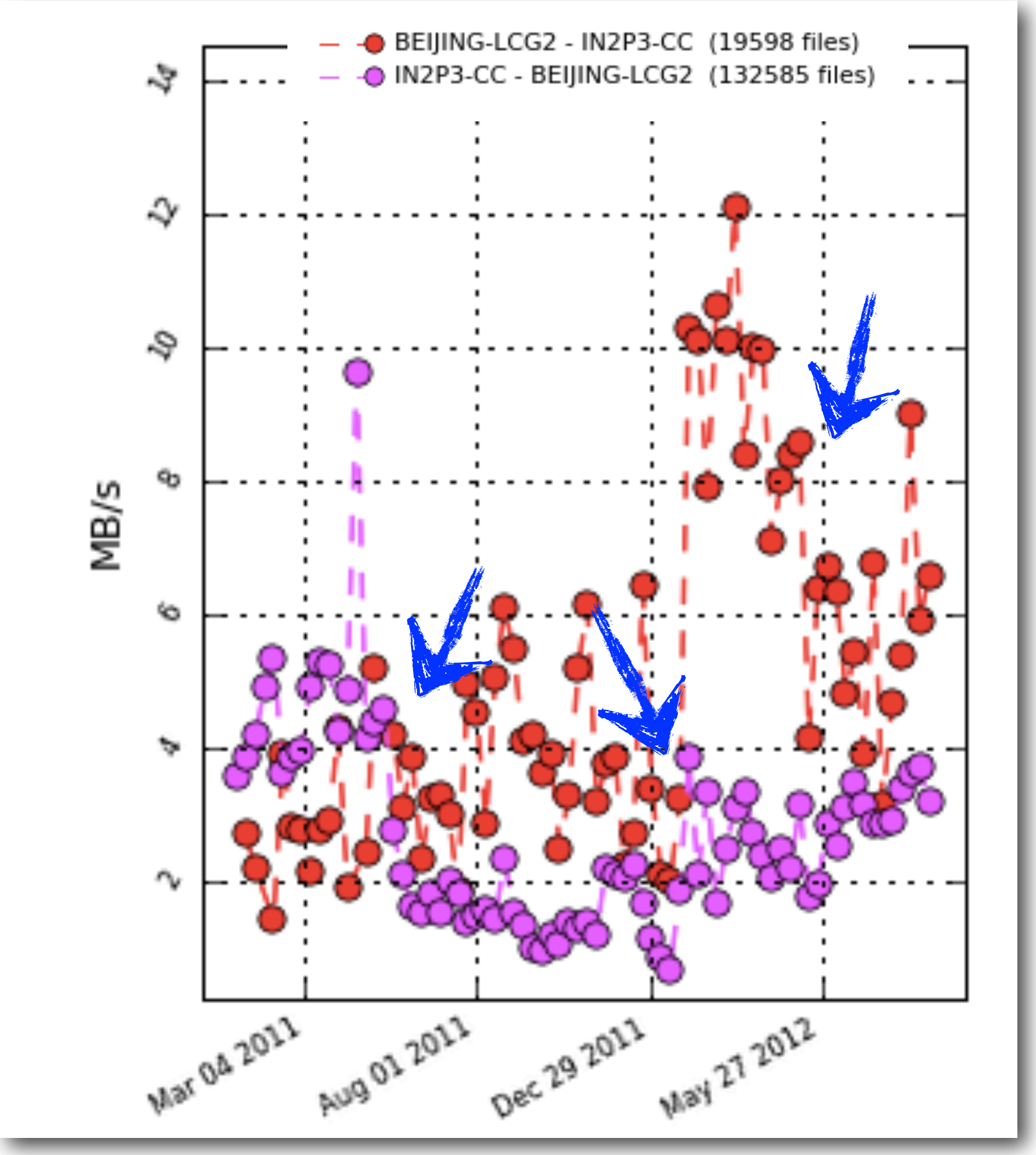
---	0	1	2	3	4	5	6
0:Beijing (perfsonar2.ihep.ac.cn)	0.0 0.0	0.0 1.0	0.0 0.0	0.0 600.0	0.0 0.0	0.0 0.0	0.0 0.0
1:CC-IN2P3 (ccperfsonar2-lhcopn.in2p3.fr)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
2:GRIF-LAL (psonar1.lal.in2p3.fr)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 600.0	0.0 0.0	0.0 0.0	0.0 1.0
3:GRIF/LPNHE (lpnhe-psl.in2p3.fr)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0
4:RO-02 (atrogr007.nipne.ro)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 600.0	0.0 0.0	0.0 0.0	0.0 0.0
5:RO-07 (perfsonar2.nipne.ro)	0.0 0.0	0.0 0.0	0.0 0.0	0.0 600.0	0.0 0.0	0.0 0.0	0.0 0.0
6:Tokyo (perfsonar1.icepp.jp)	0.0 0.0	0.0 1.0	0.0 600.0	0.0 600.0	0.0 0.0	0.0 0.0	0.0 0.0

http://psps.perfsonar.net/

ATLAS transfers to Beijing since beg. 2011

Beijing→ CCIN2P3

CCIN2P3 → Beijing



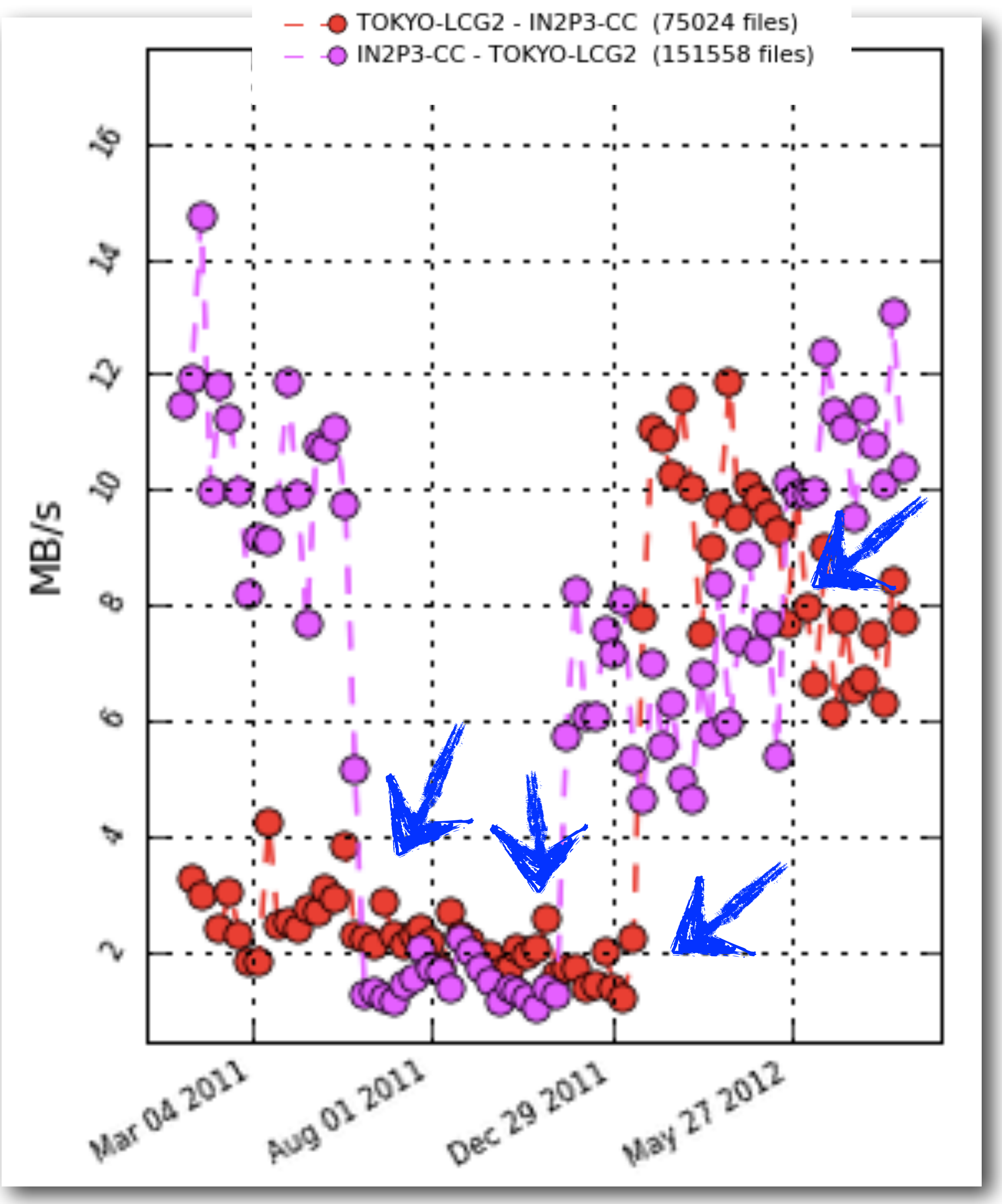
Performances changed over last year

- *Asymmetry in transfer rate : why?*
- *Asymmetry reversed*

Each 'event' explained sometime after some delay...

ATLAS transfers to Tokyo since beg. 2011

Tokyo→ CCIN2P3
CCIN2P3 → Tokyo

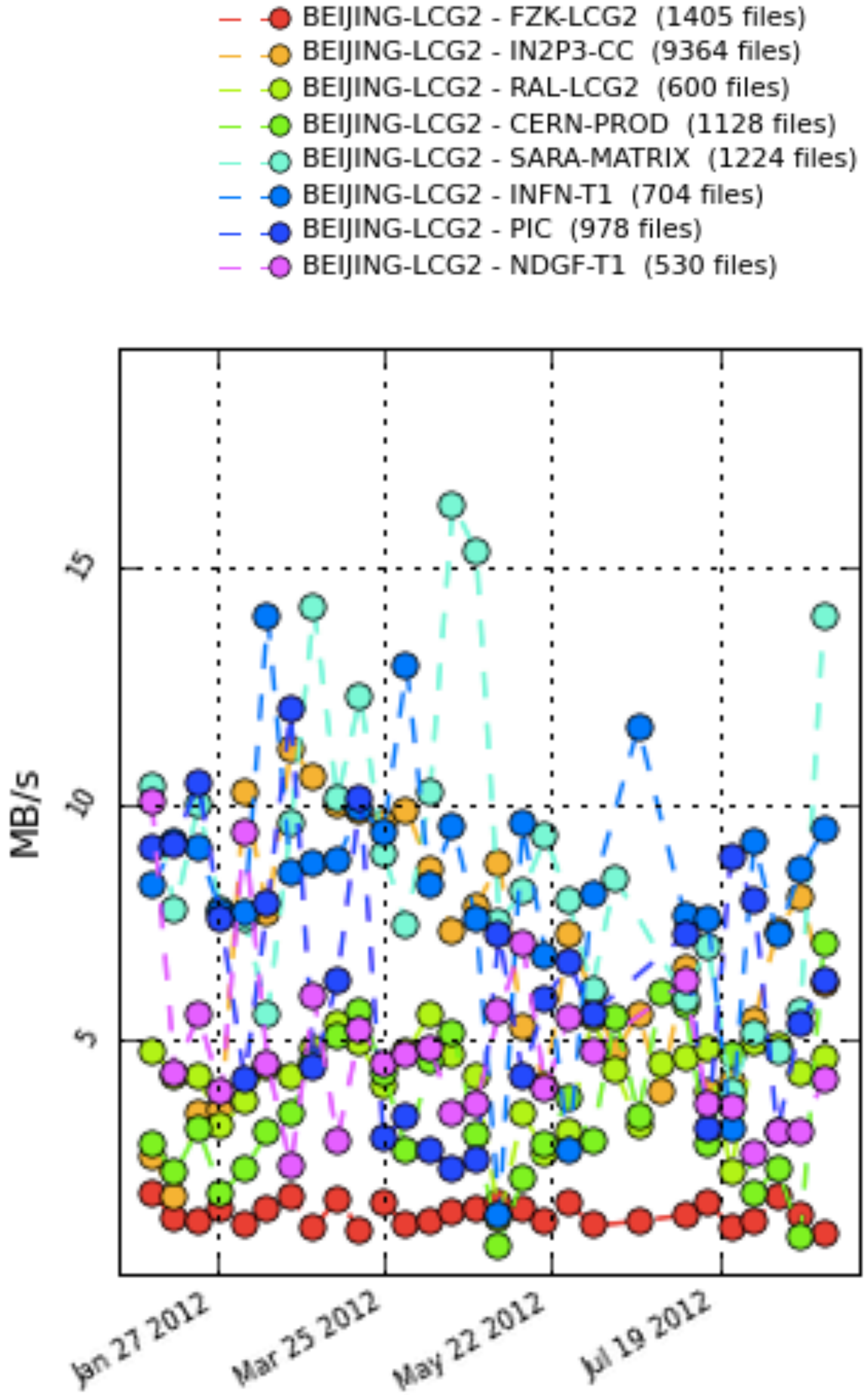


Performances changed over last year

- *Asymmetry in transfer rate : why?*
- *Asymmetry reversed*

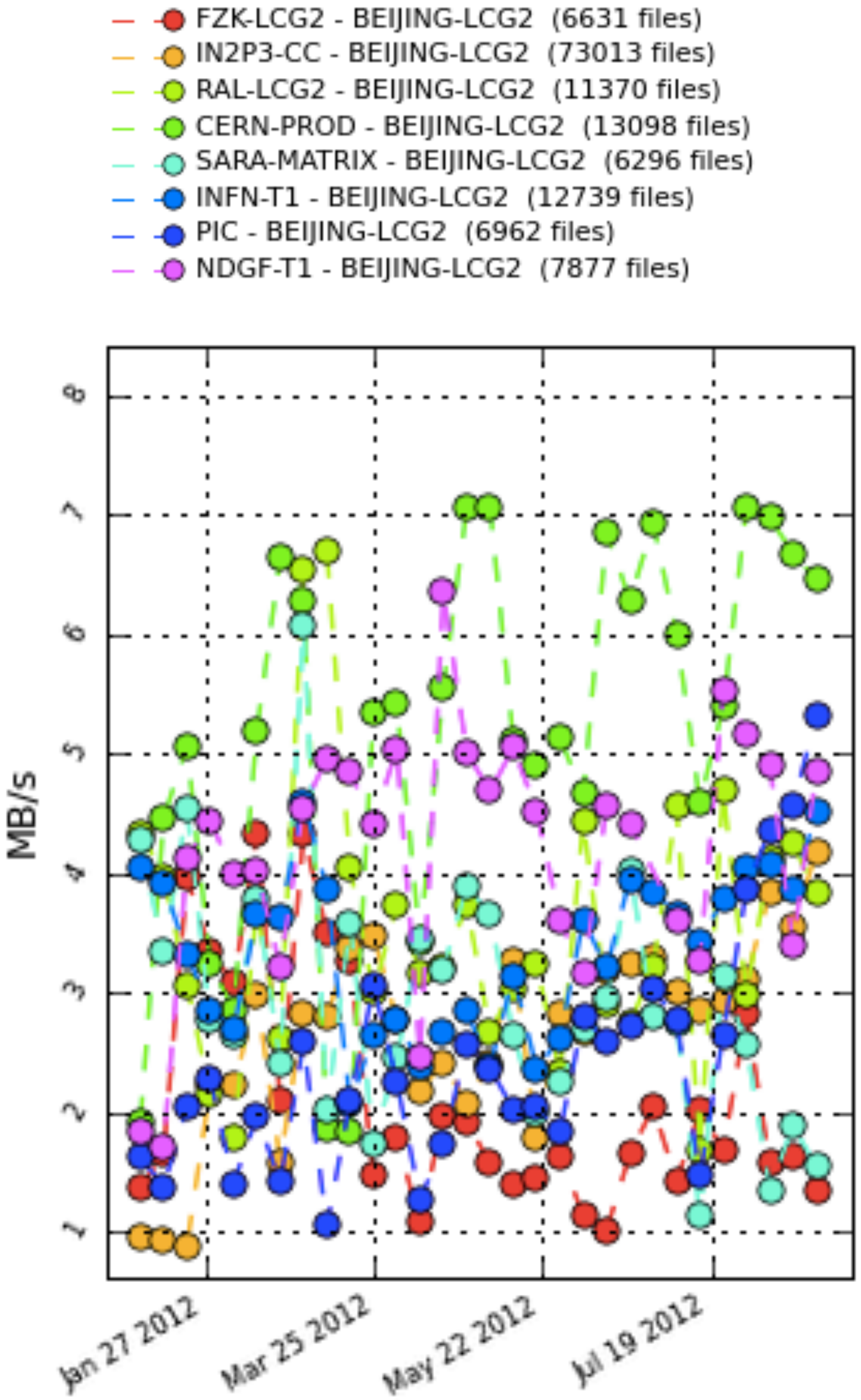
Each 'event' explained sometime after some delay...

Beijing from/to EU T1s

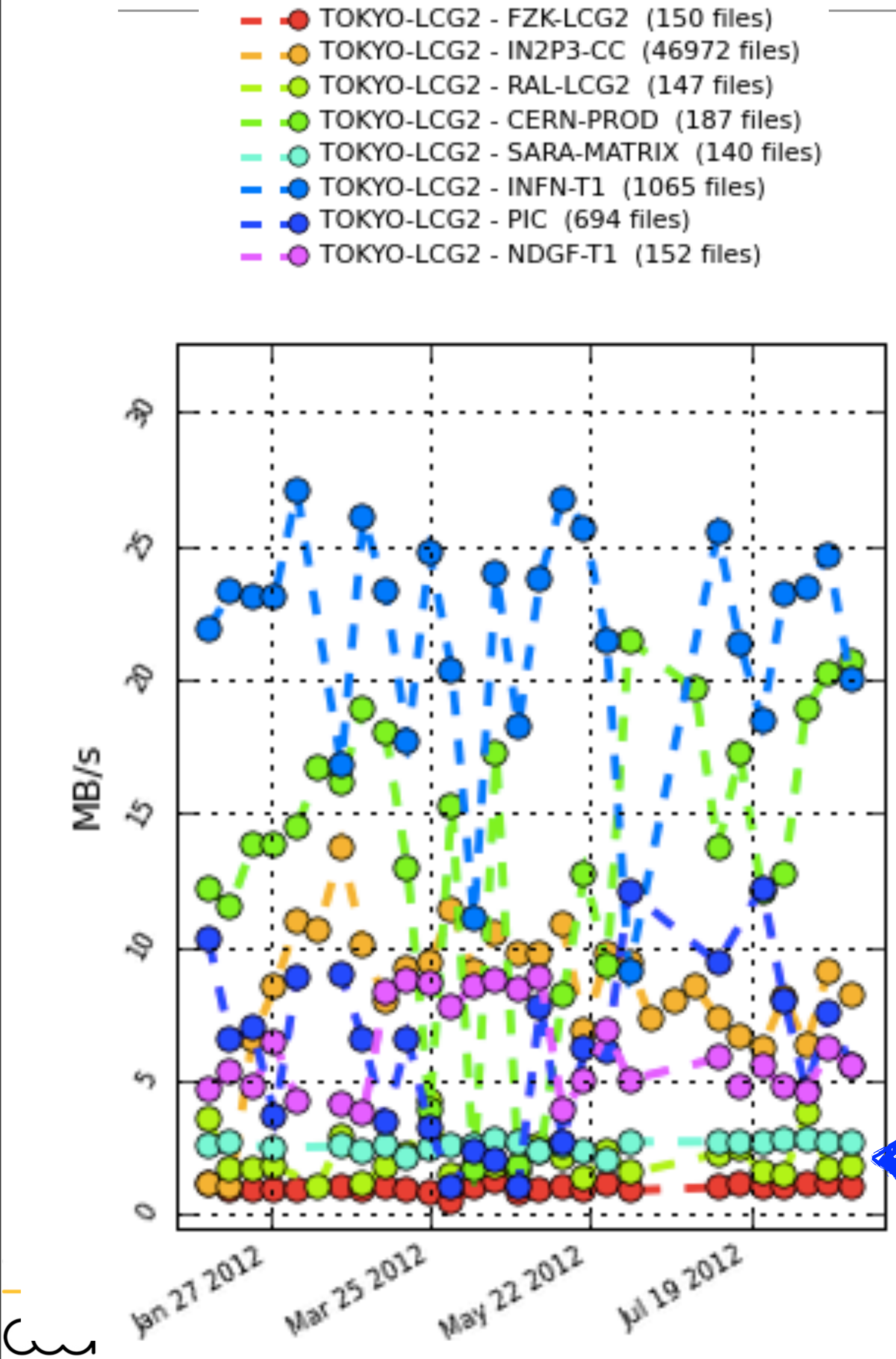


Each T1 is different

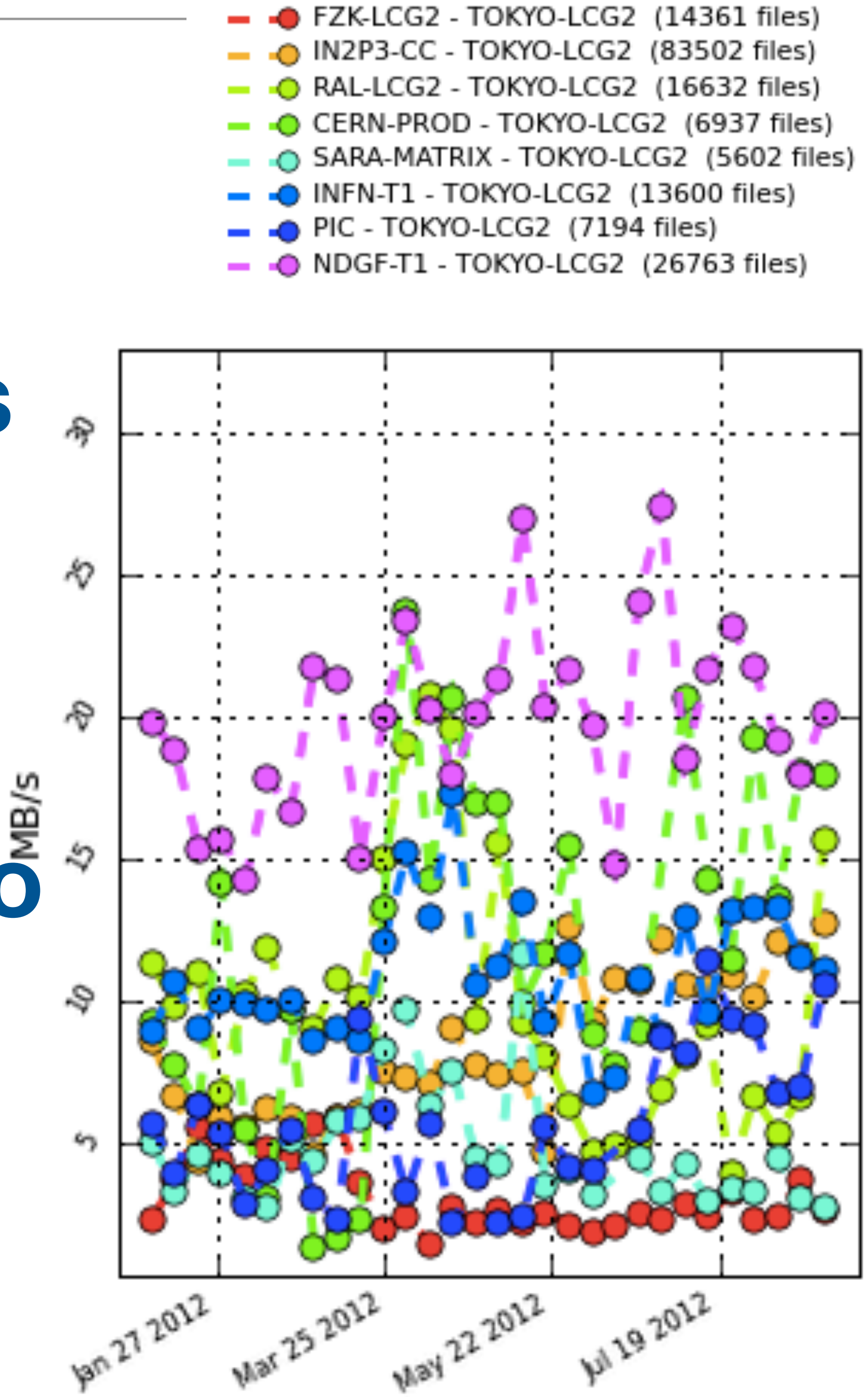
Beijing →EU better for most T1s



Tokyo from/to EU T1s

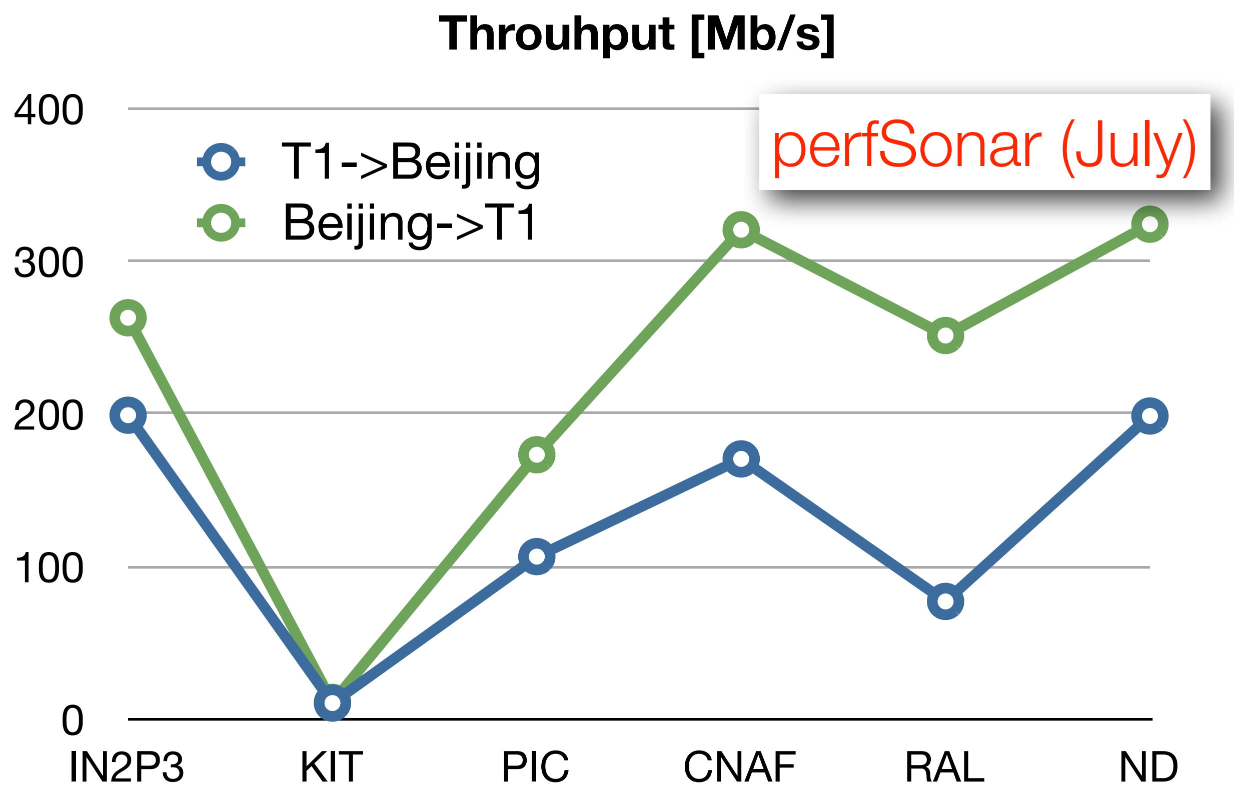


Each T1 is different



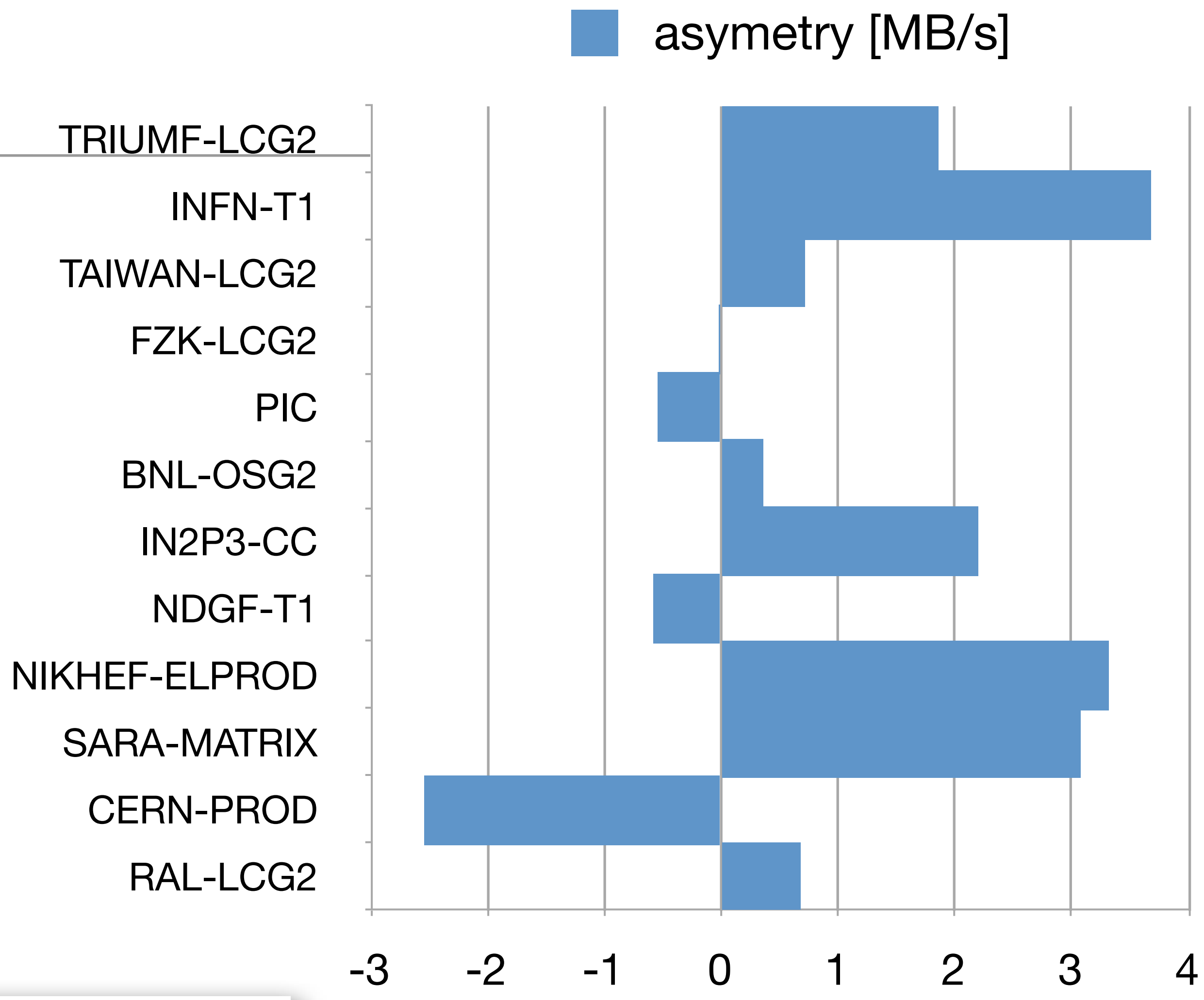
EU → Tokyo better for most T1s

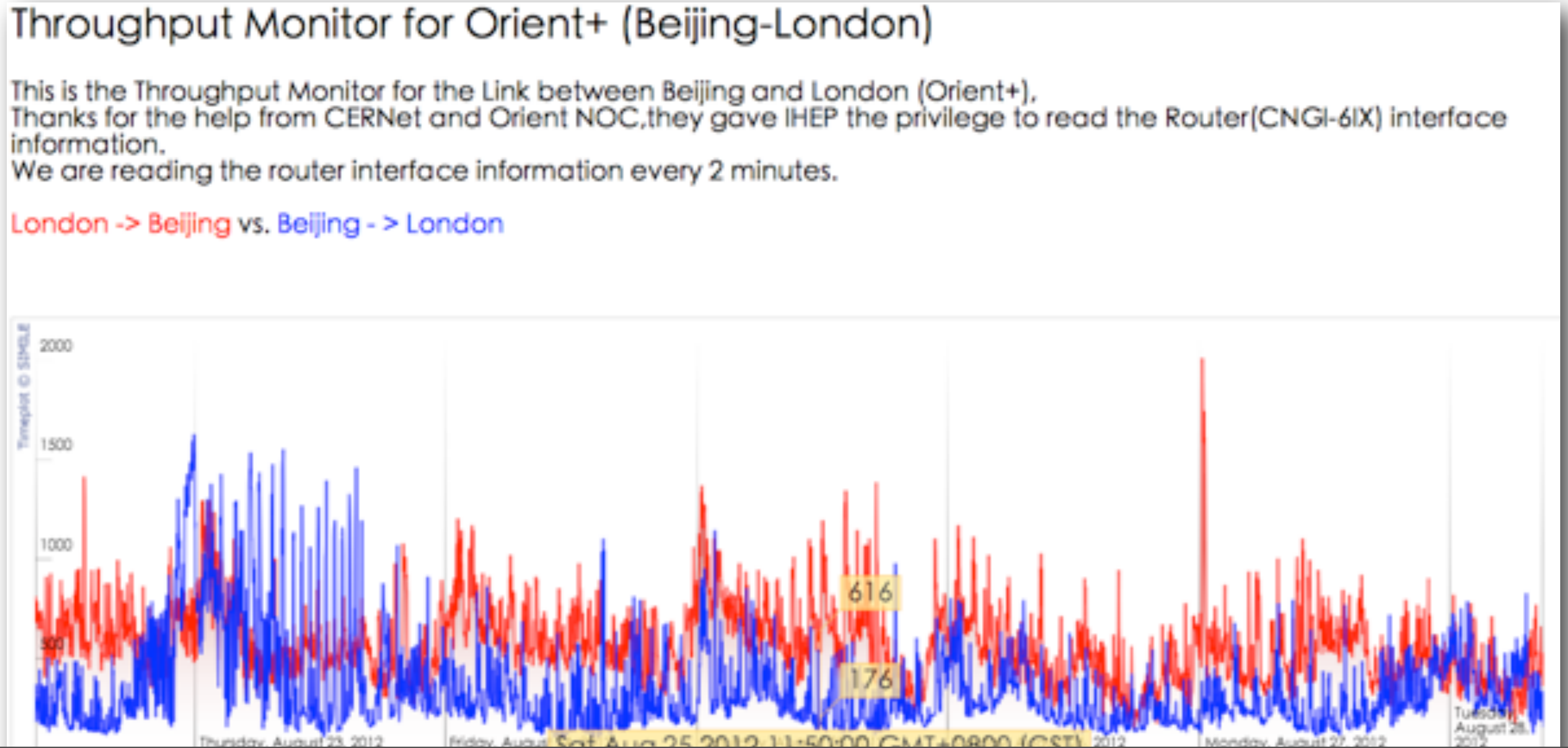
Beijing - T1s asymmetry



perfSonar (July)

FTSmon
(last 5 weeks)





Throughput Monitor for Orient+ (Beijing-London)

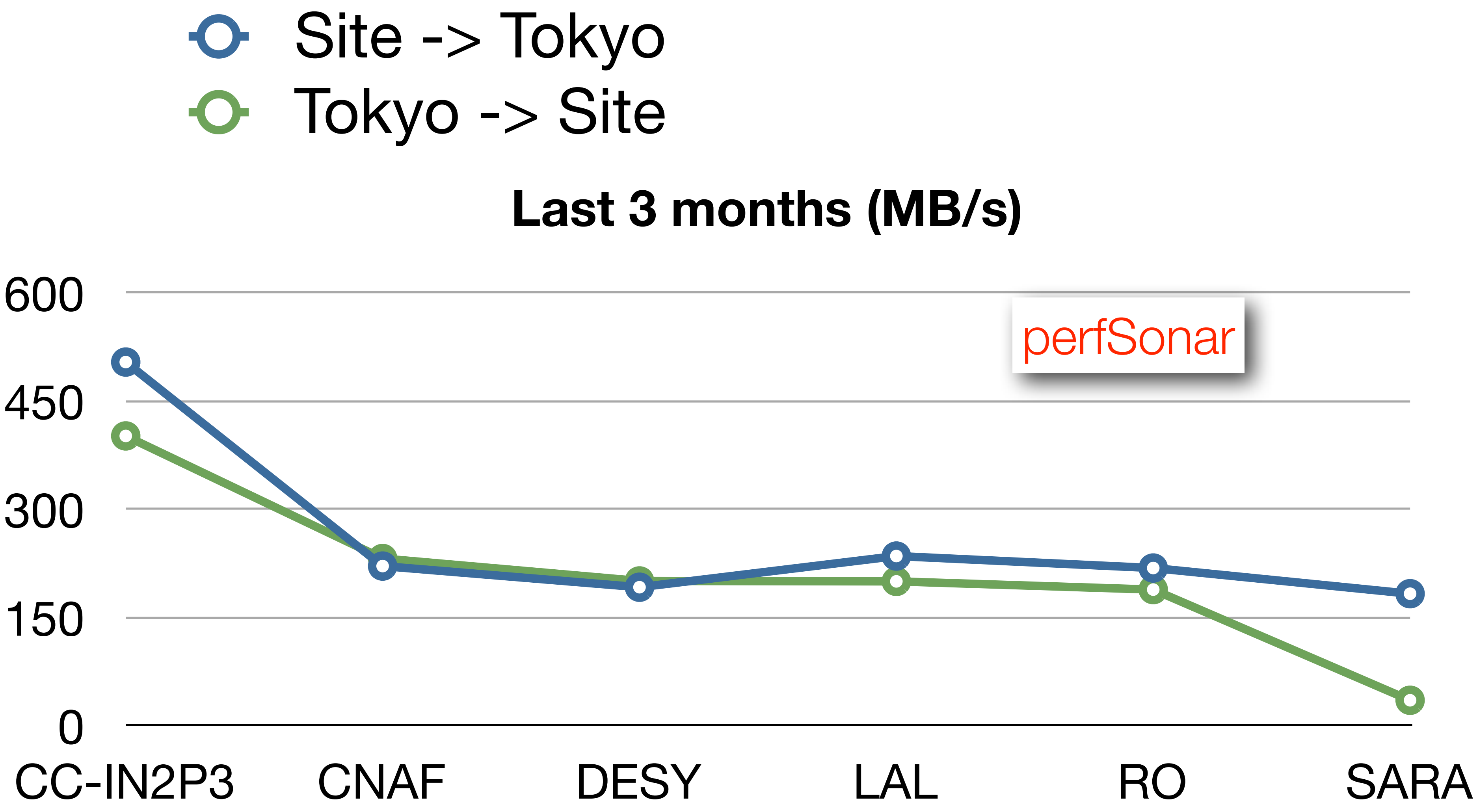
This is the Throughput Monitor for the Link between Beijing and London (Orient+).
Thanks for the help from CERNet and Orient NOC,they gave IHEP the privilege to read the Router(CNGI-6IX) interface information.
We are reading the router interface information every 2 minutes.

London -> Beijing vs. Beijing - > London

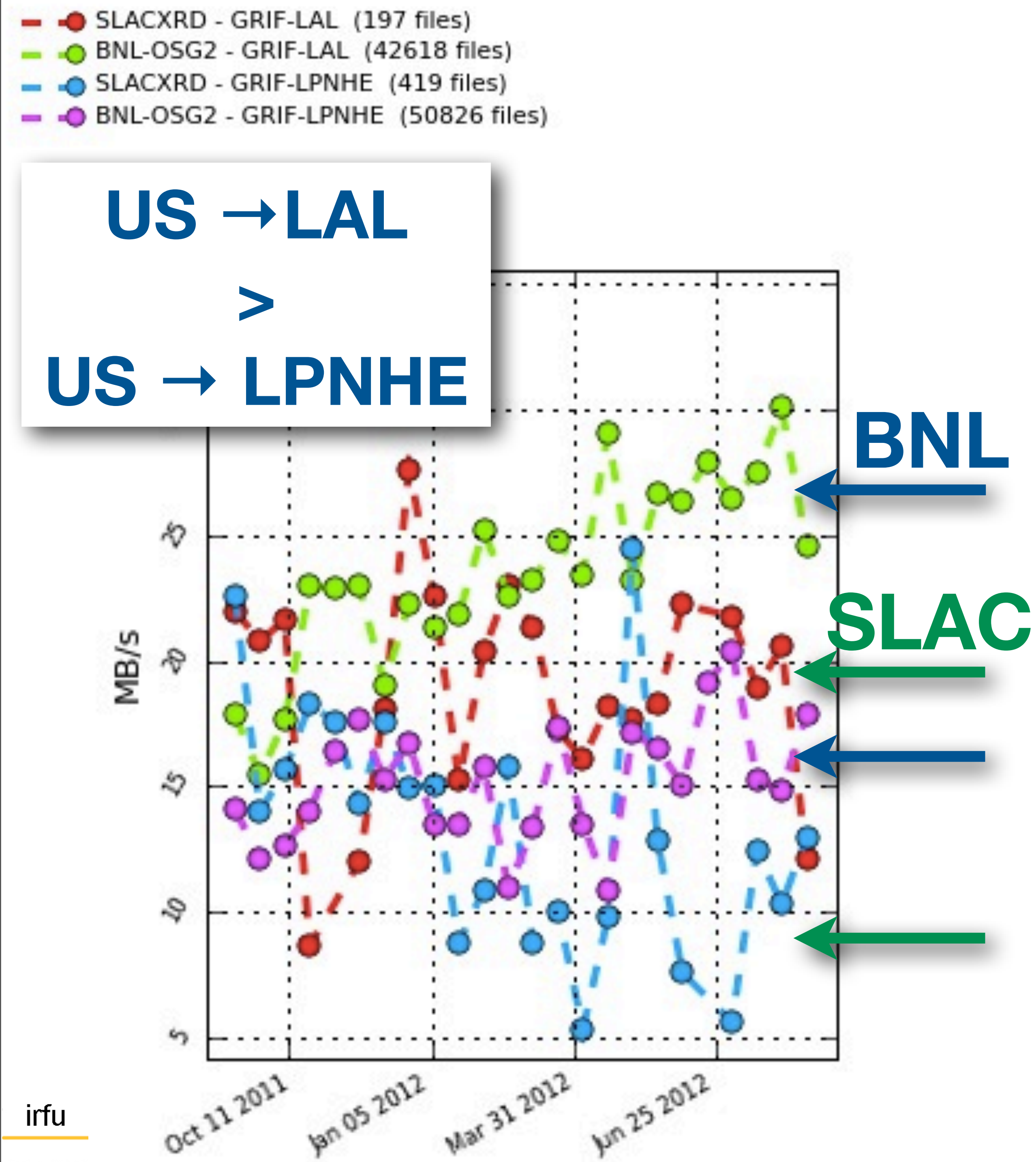
2000
1500
1000
616
176
Thursday, August 23, 2012
Sat Aug 25 2012 11:50:00 GMT+0800 (CST)
Monday, August 27, 2012

Beijing -> T1s
>
T1s -> Beijing

Tokyo ↔ EU as seen by perfSonar



US (LHCONE) ↔ GRIF (LHCONE)



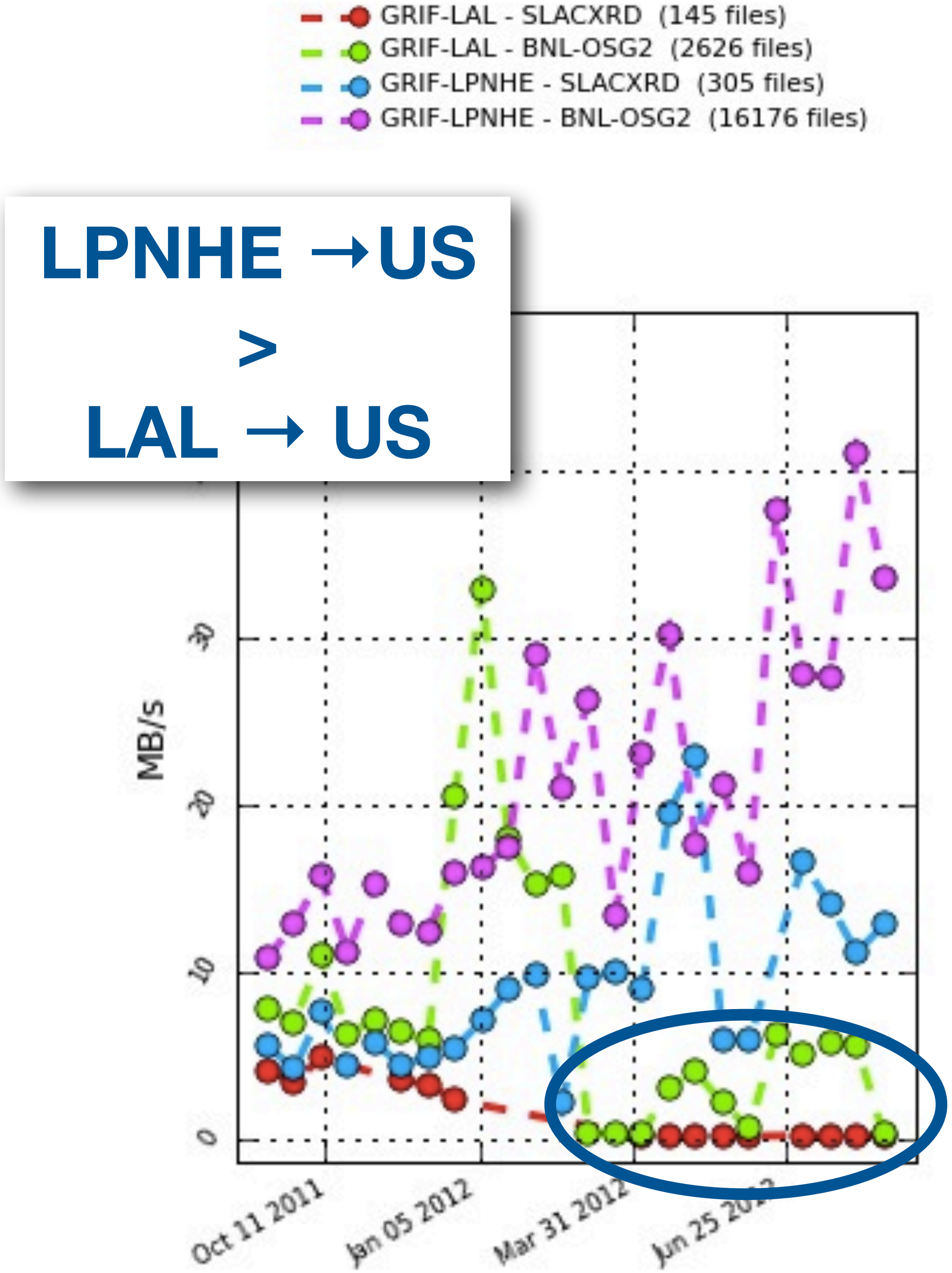
US →LAL

>

US → LPNHE

BNL

SLAC



LPNHE →US

>

LAL → US

DISTRIBUTED STORAGE / REMOTE ACCESS

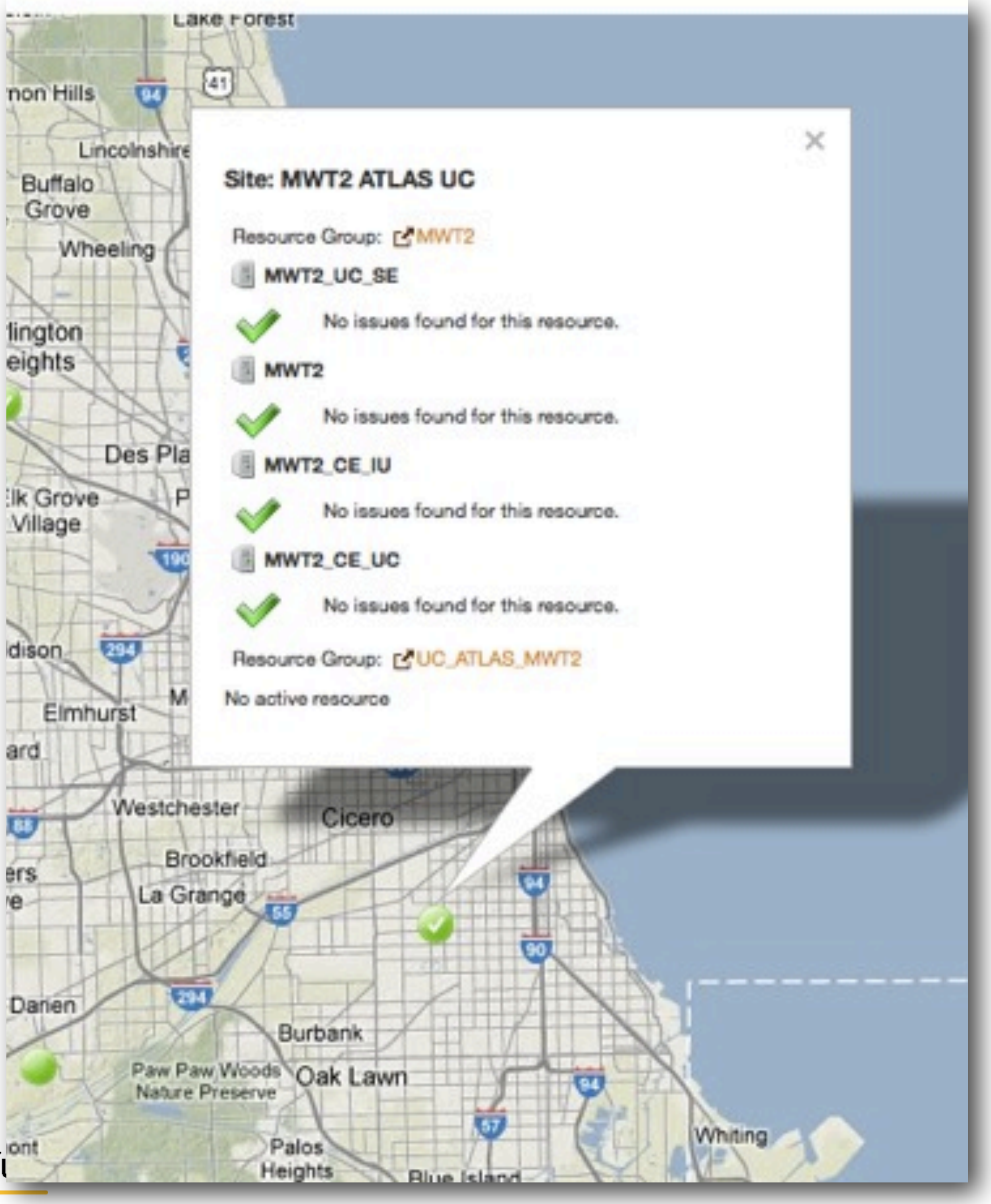
- Better used of storage resources (disk prices!)
- Simplification of data management
- Eventually remote access (with caching at both ends); direct reading or file copy
- Bandwidth and stability needed

On going projects

- Storage : dCache, dpm,...
- Protocol : Xrootd, HTTP/WebDAV

Existing distributed dCache systems : 2 examples

MWT2 (Chicago)



Site: MWT2 ATLAS UC

Resource Group: MWT2

MWT2_UC_SE

✓ No issues found for this resource.

MWT2

✓ No issues found for this resource.

MWT2_CE_IU

✓ No issues found for this resource.

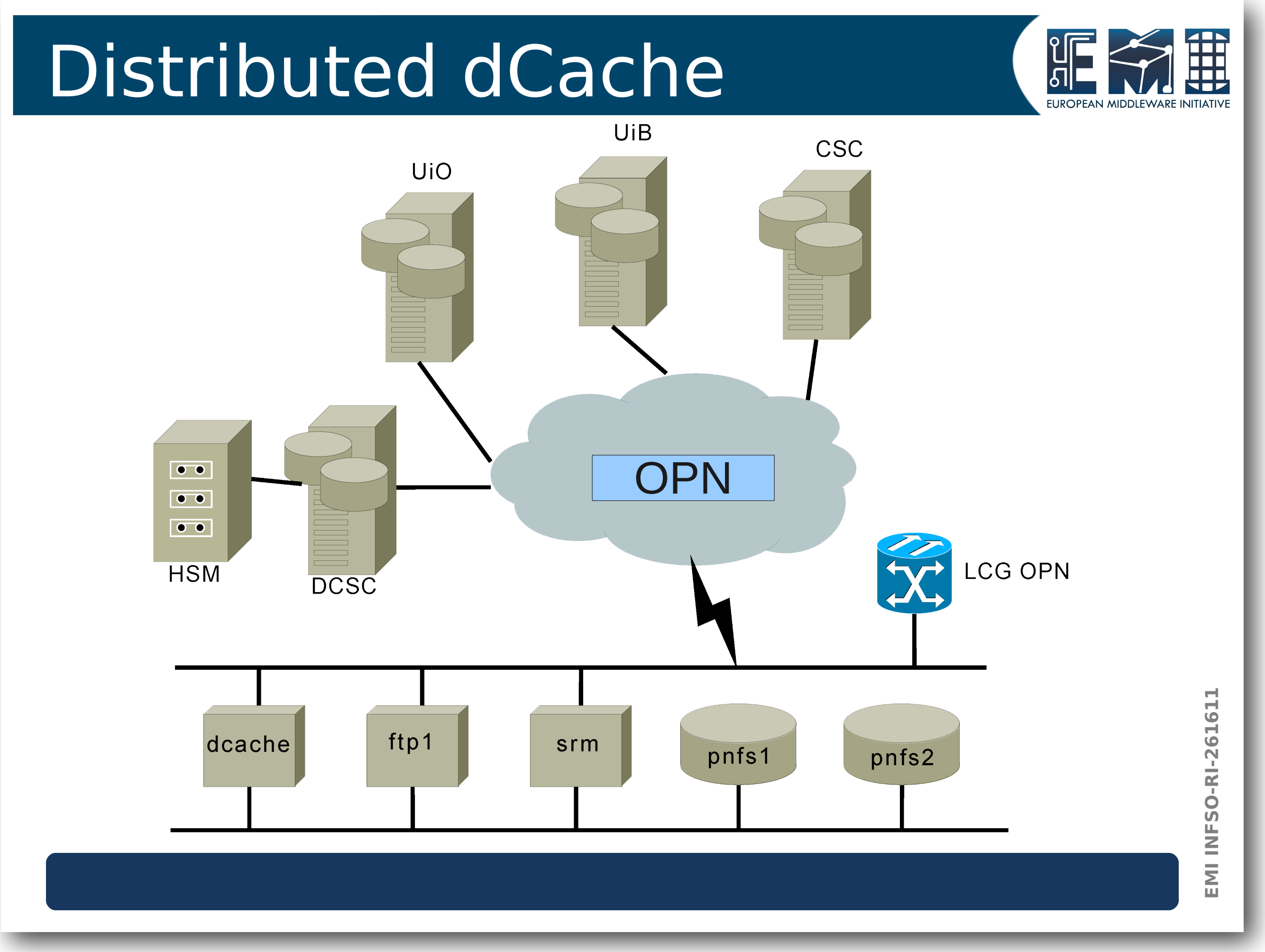
MWT2_CE_UC

✓ No issues found for this resource.

Resource Group: UC_ATLAS_MWT2

No active resource

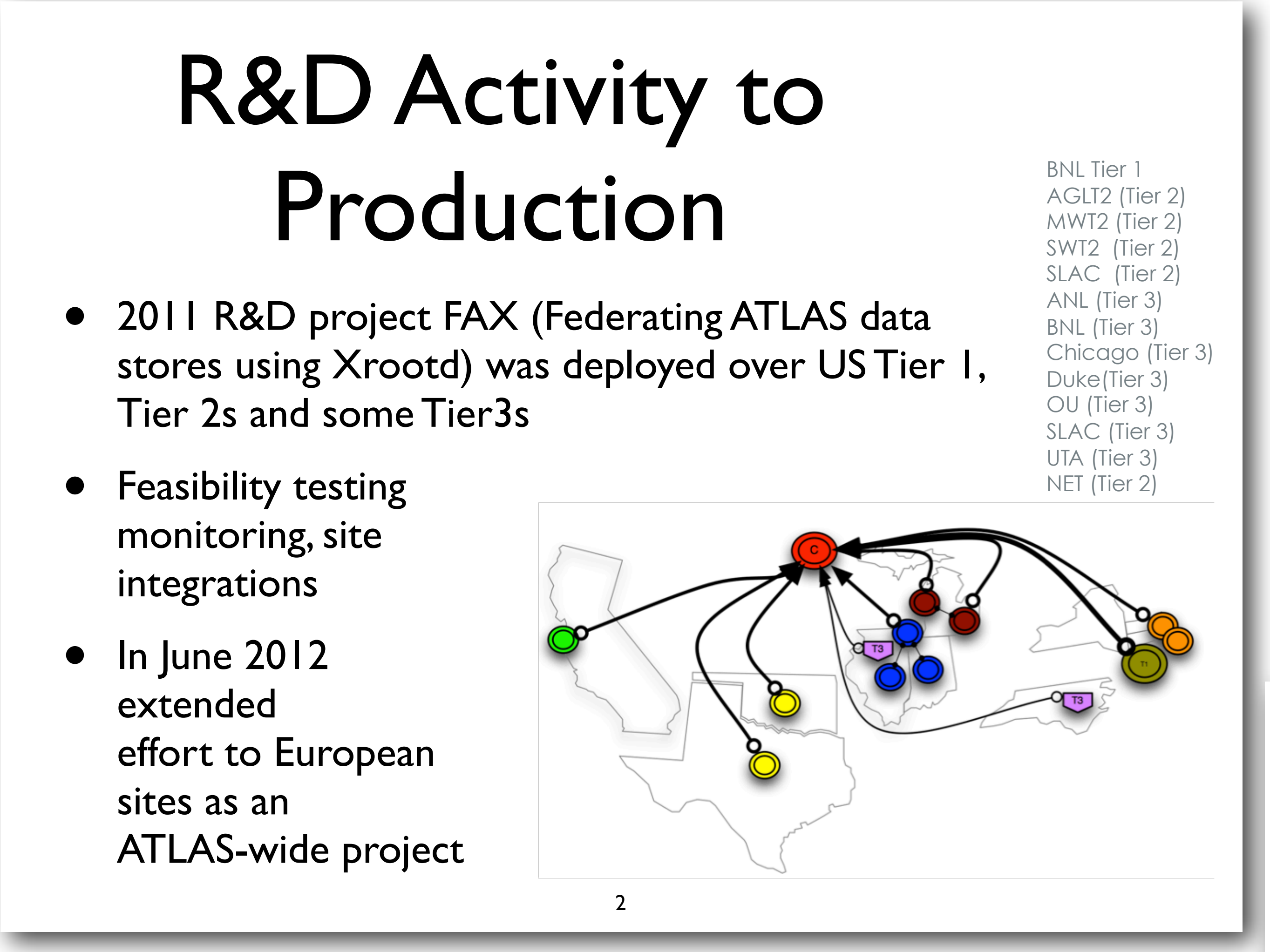
NORDUGRID



Distributed dCache

Xrootd federation project in US

REMOTE ACCESS



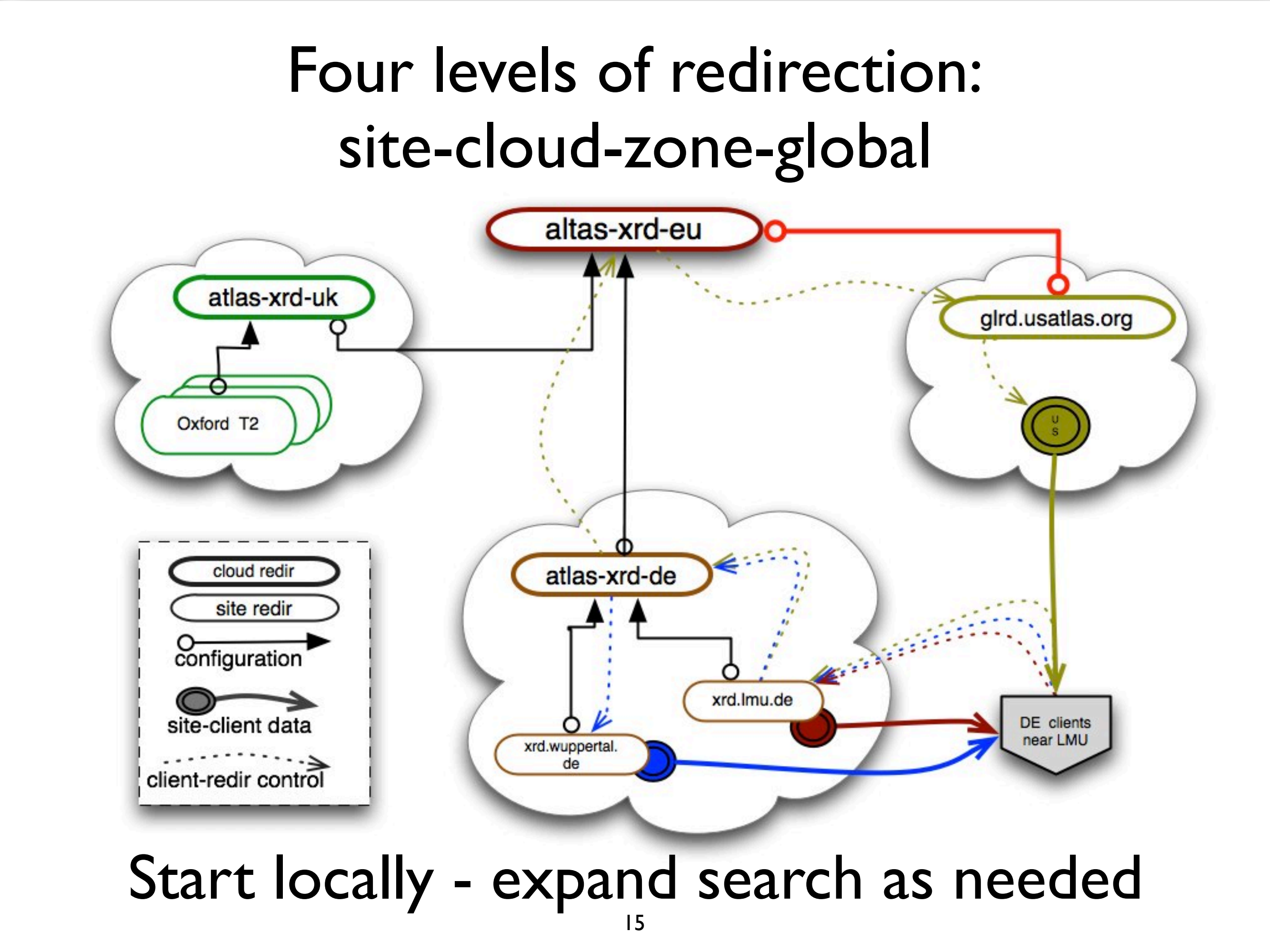
R&D Activity to Production

- 2011 R&D project FAX (Federating ATLAS data stores using Xrootd) was deployed over US Tier 1, Tier 2s and some Tier3s
- Feasibility testing monitoring, site integrations
- In June 2012 extended effort to European sites as an ATLAS-wide project

BNL Tier 1
AGLT2 (Tier 2)
MWT2 (Tier 2)
SWT2 (Tier 2)
SLAC (Tier 2)
ANL (Tier 3)
BNL (Tier 3)
Chicago (Tier 3)
Duke(Tier 3)
OU (Tier 3)
SLAC (Tier 3)
UTA (Tier 3)
NET (Tier 2)

2

EU federation tests

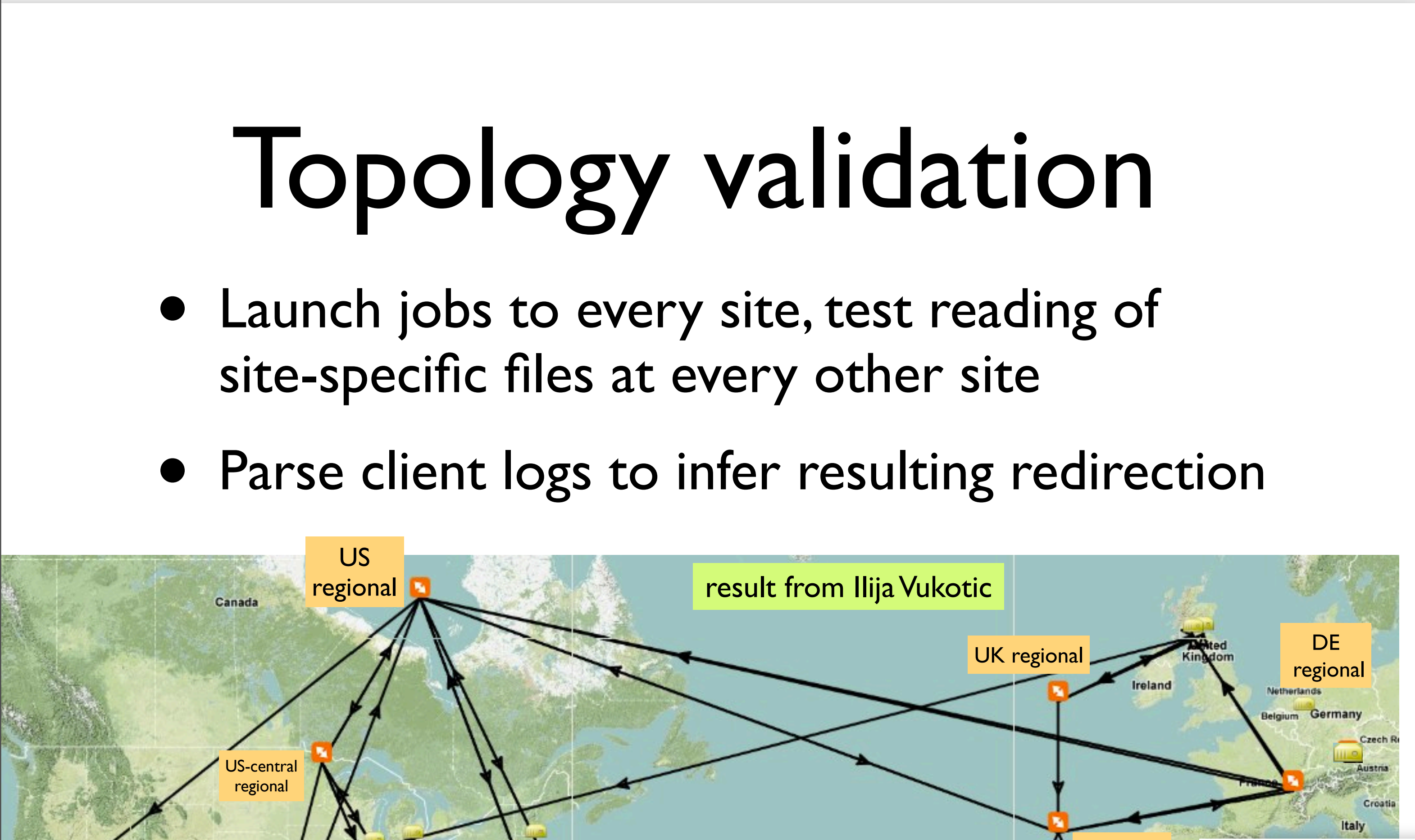


Four levels of redirection:
site-cloud-zone-global

Start locally - expand search as needed

15

REMOTE ACCESS



Topology validation

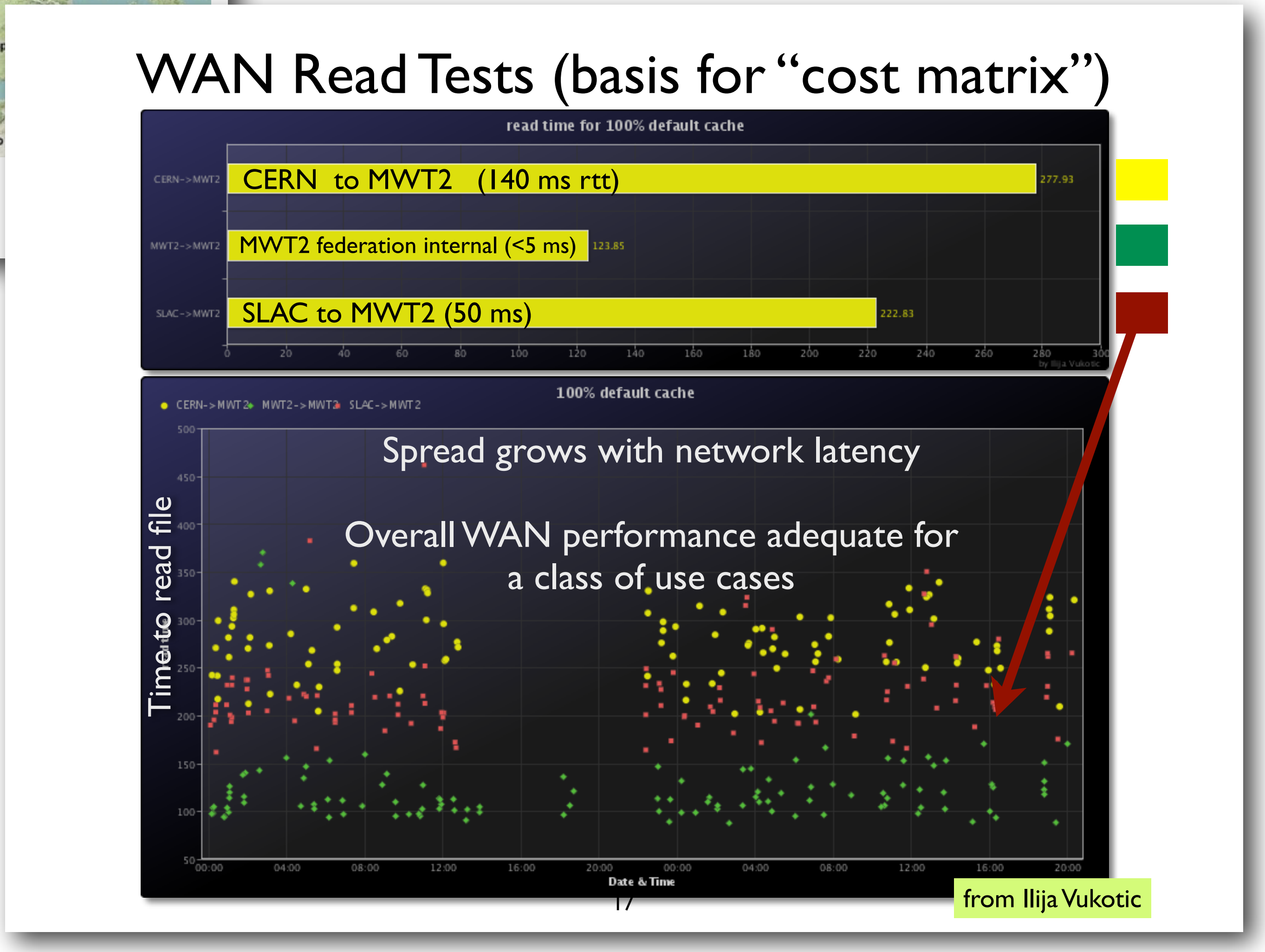
- Launch jobs to every site, test reading of site-specific files at every other site
- Parse client logs to infer resulting redirection

http://ivukotic.web.cern.ch/ivukotic/FAX/index.asp

XRD redirector

federating site

16



WAN Read Tests (basis for "cost matrix")

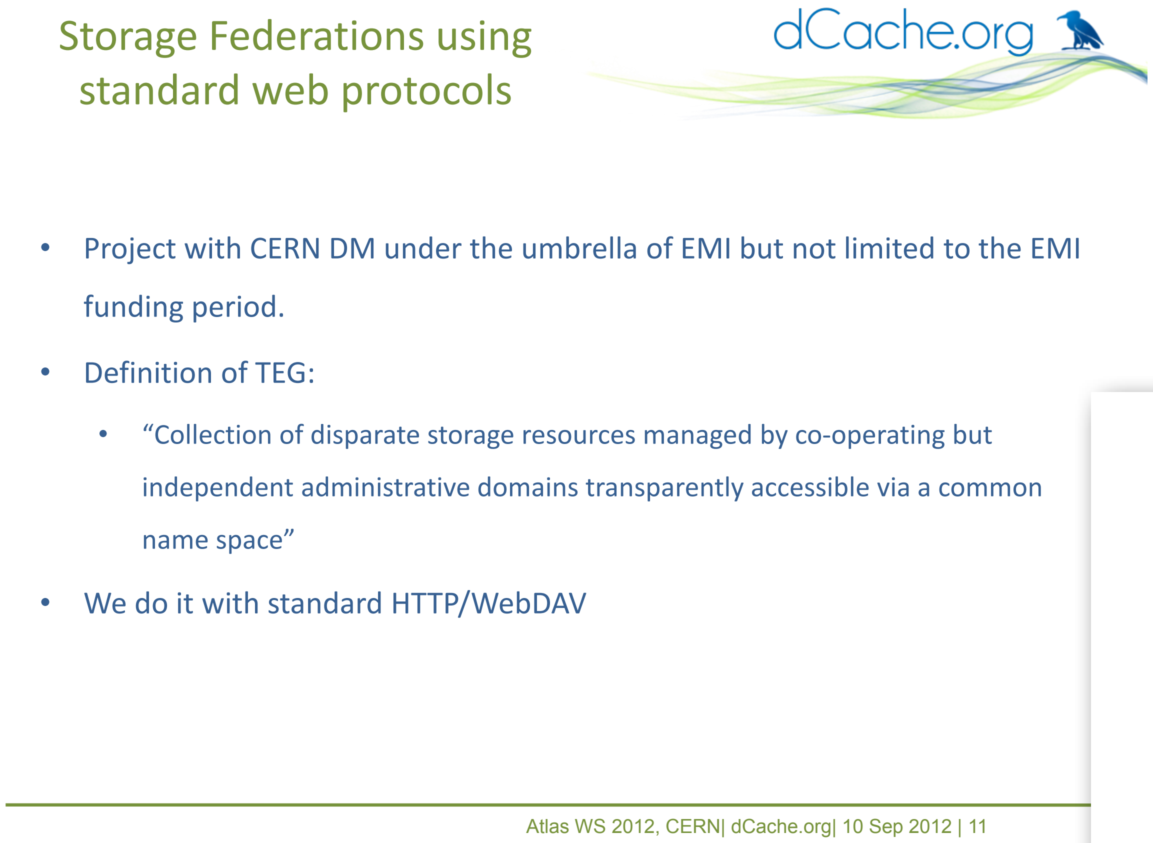
Spread grows with network latency

Overall WAN performance adequate for a class of use cases

from Ilija Vukotic

17

HTTP/WebDAV

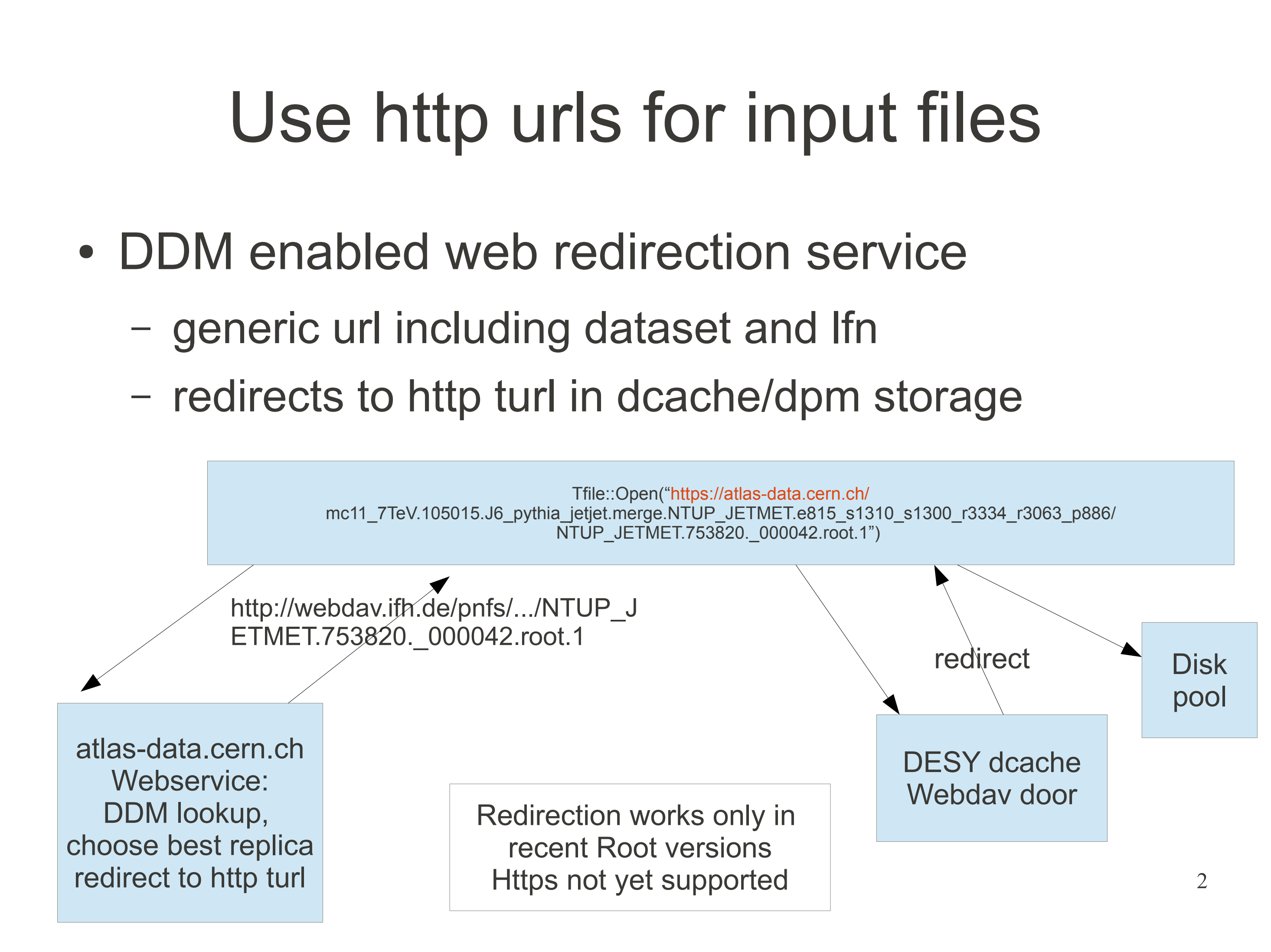


Storage Federations using standard web protocols

dCache.org

- Project with CERN DM under the umbrella of EMI but not limited to the EMI funding period.
- Definition of TEG:
  - "Collection of disparate storage resources managed by co-operating but independent administrative domains transparently accessible via a common name space"
- We do it with standard HTTP/WebDAV

Atlas WS 2012, CERN| dCache.org| 10 Sep 2012 | 11



Use http urls for input files

- DDM enabled web redirection service
  - generic url including dataset and lfn
  - redirects to http turl in dcache/dpm storage

Tfile::Open("https://atlas-data.cern.ch/mc11_7TeV.105015.J6_pythia_jetjet.merge.NTUP_JETMET.e815_s1310_s1300_r3334_r3063_p886/NTUP_JETMET.753820._000042.root.1")

http://webdav.ifh.de/pnfs/.../NTUP_JETMET.753820._000042.root.1

atlas-data.cern.ch Webservice: DDM lookup, choose best replica redirect to http turl

redirect

DESY dcache Webdav door

Disk pool

Redirection works only in recent Root versions Https not yet supported

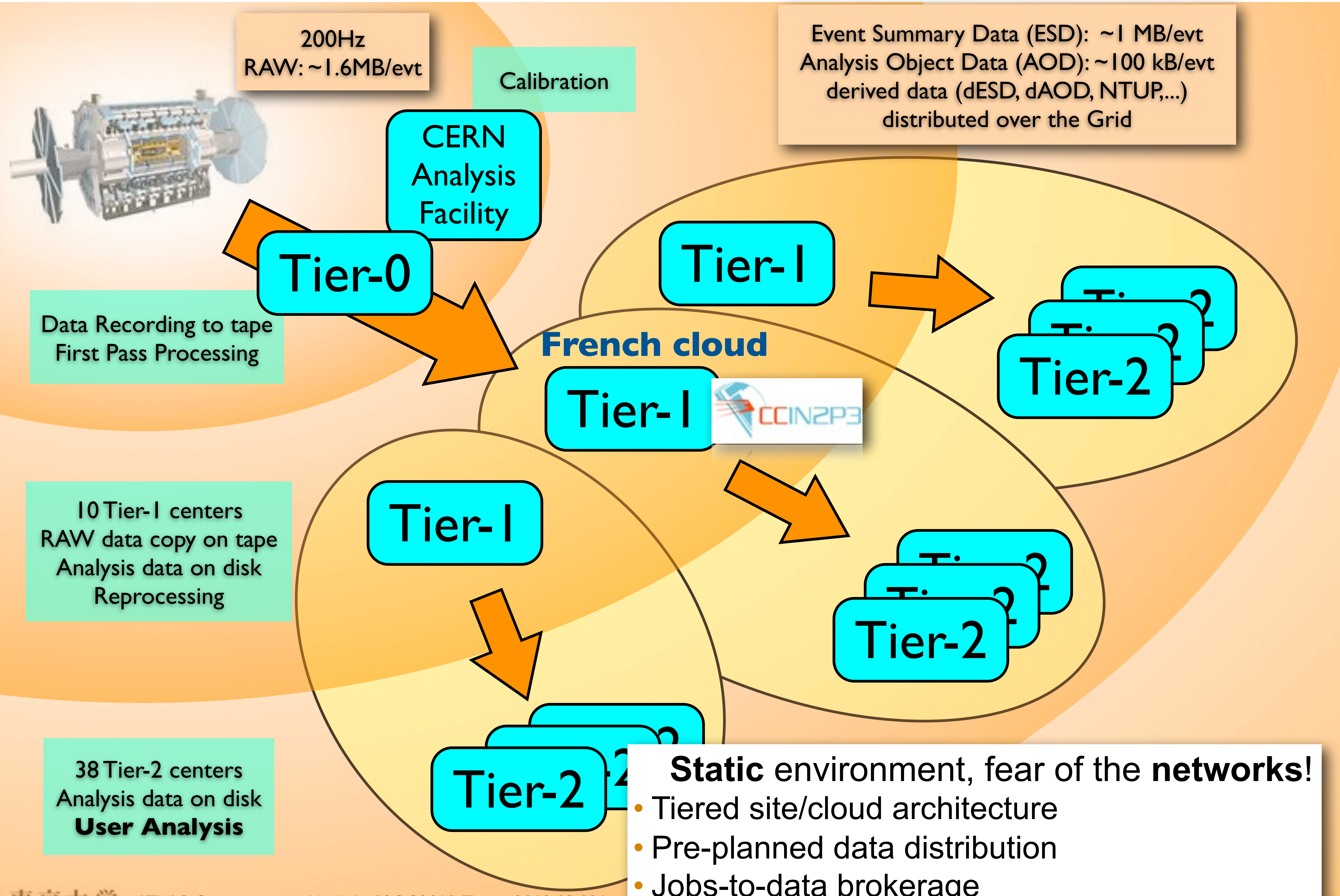
2

Summary

- Thanks to the high performances of networks
- ATLAS computing model has changed significantly: simplification of data and workflow management
- Would have been impossible to handle current data volume (LHC performing beyond expectations) and LHC running extension up to spring 2013 with initial model
- More efficient use of storage resources (reduce replica counts; direct sharing of replicas across sites)
- Ongoing projects (distributed storage, remote access) will further change the landscape

BACKUP

ATLAS Computing Model: T0 Data Flow



200Hz
RAW: ~1.6MB/evt

Calibration

CERN Analysis Facility

Event Summary Data (ESD): ~1 MB/evt
Analysis Object Data (AOD): ~100 kB/evt
derived data (dESD, dAOD, NTUP,...)
distributed over the Grid

Tier-0

Data Recording to tape
First Pass Processing

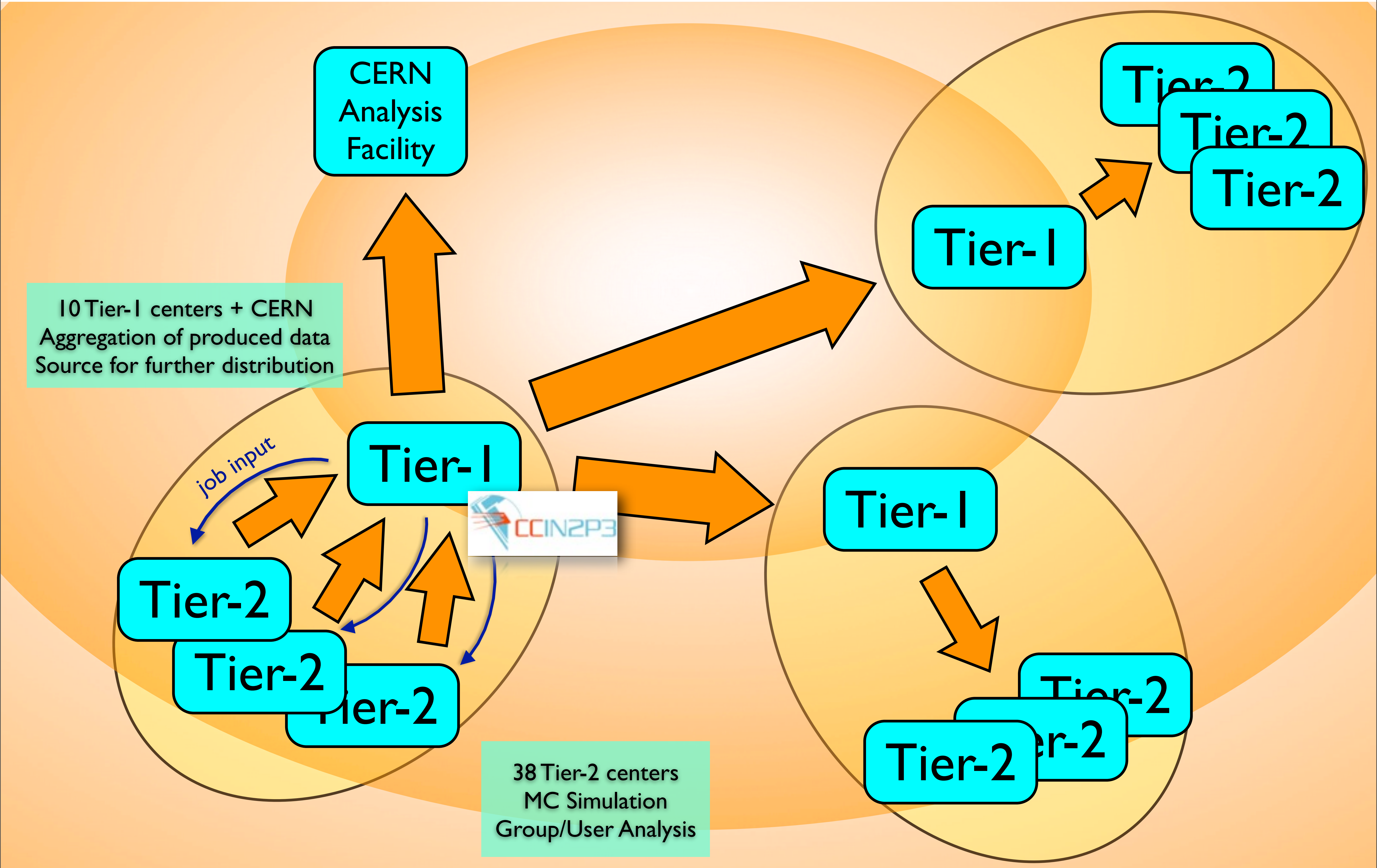
10 Tier-1 centers
RAW data copy on tape
Analysis data on disk
Reprocessing

38 Tier-2 centers
Analysis data on disk
User Analysis

Static environment, fear of the **networks**!

- Tiered site/cloud architecture
- Pre-planned data distribution
- Jobs-to-data brokerage

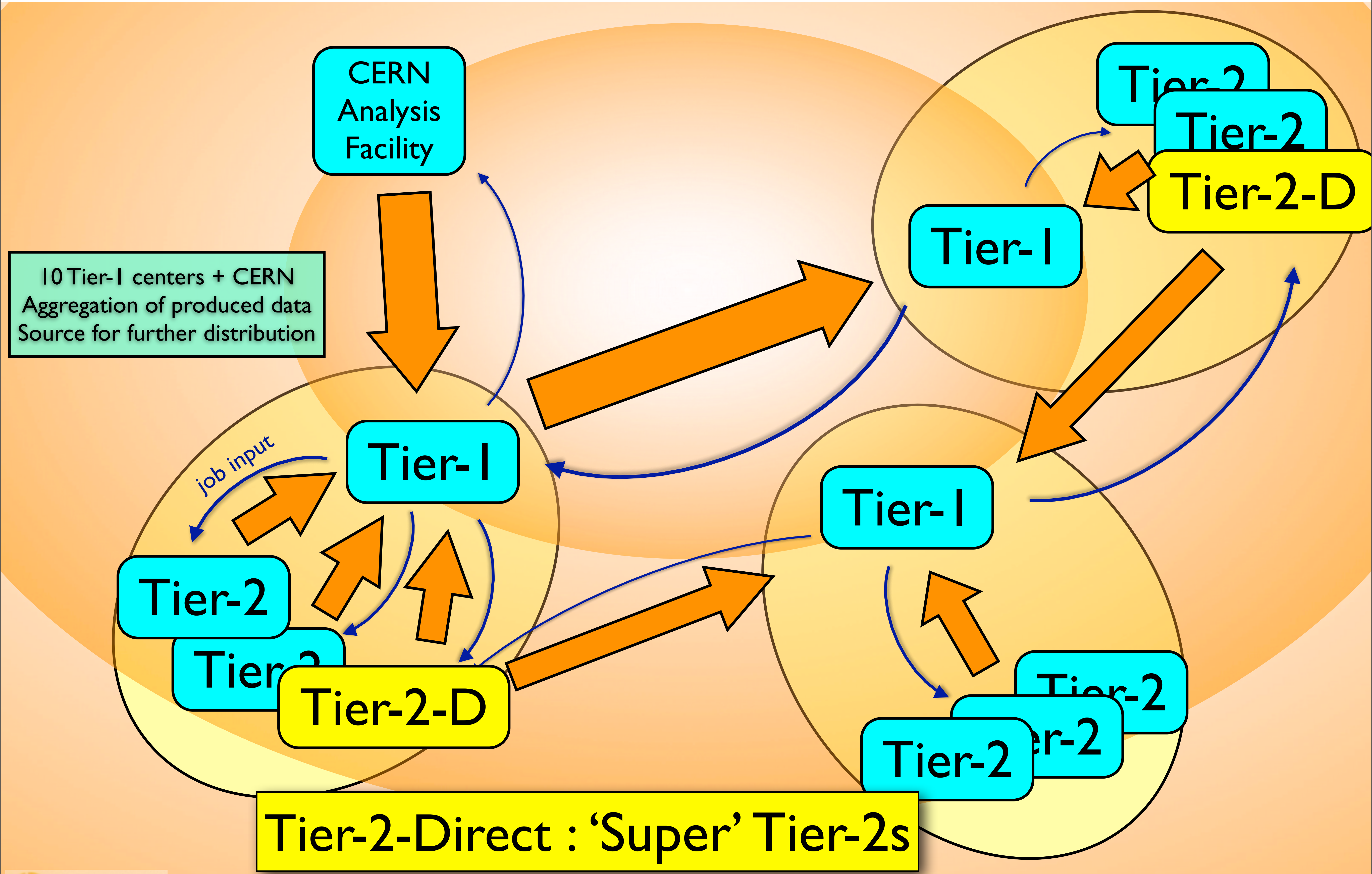
ATLAS Computing Model: MC Data Flow



10 Tier-1 centers + CERN
Aggregation of produced data
Source for further distribution

38 Tier-2 centers
MC Simulation
Group/User Analysis

Data Processing Model Revised

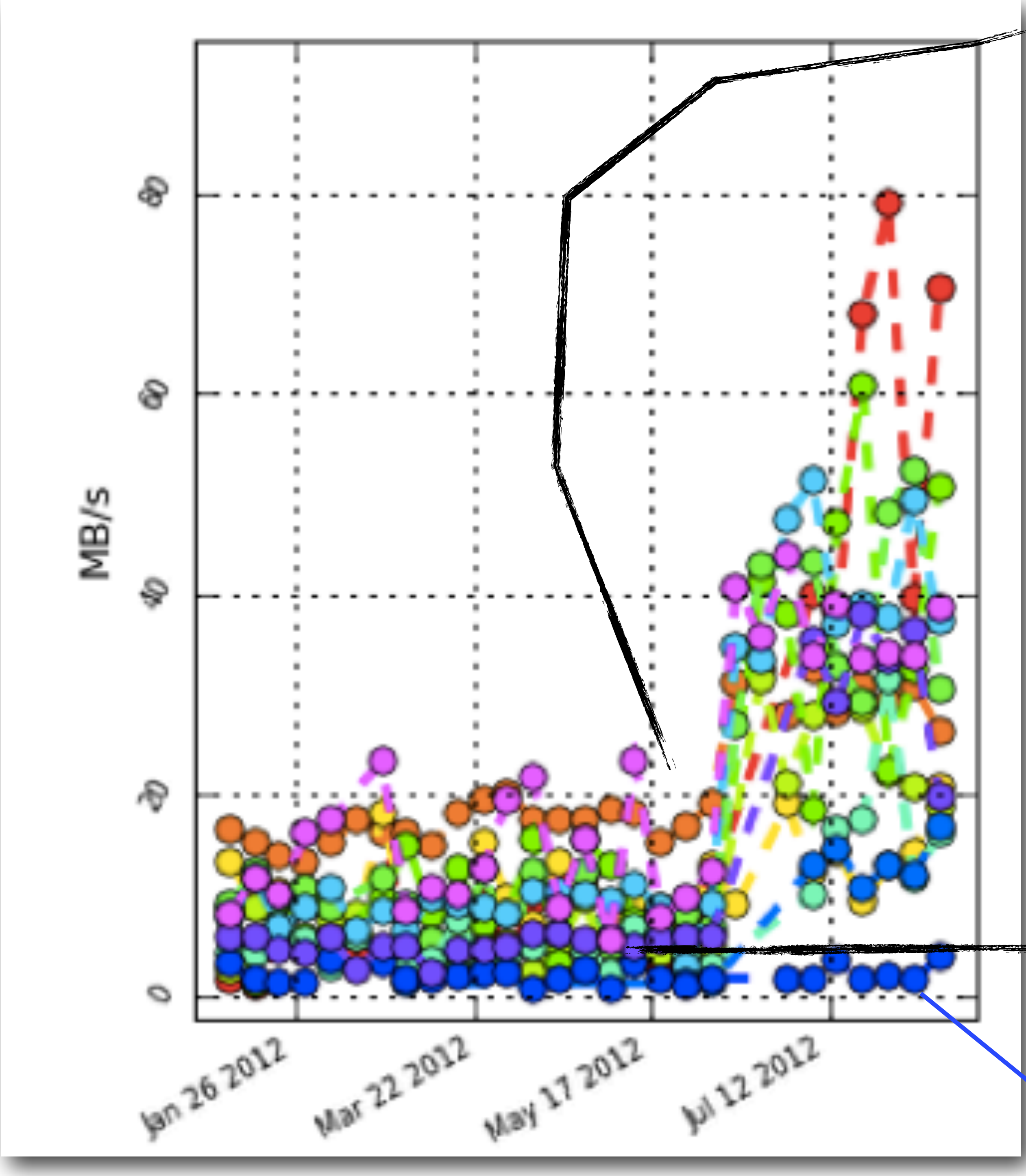


10 Tier-1 centers + CERN
Aggregation of produced data
Source for further distribution

Tier-2-Direct : 'Super' Tier-2s

T1s -> IN2P3-CPPM

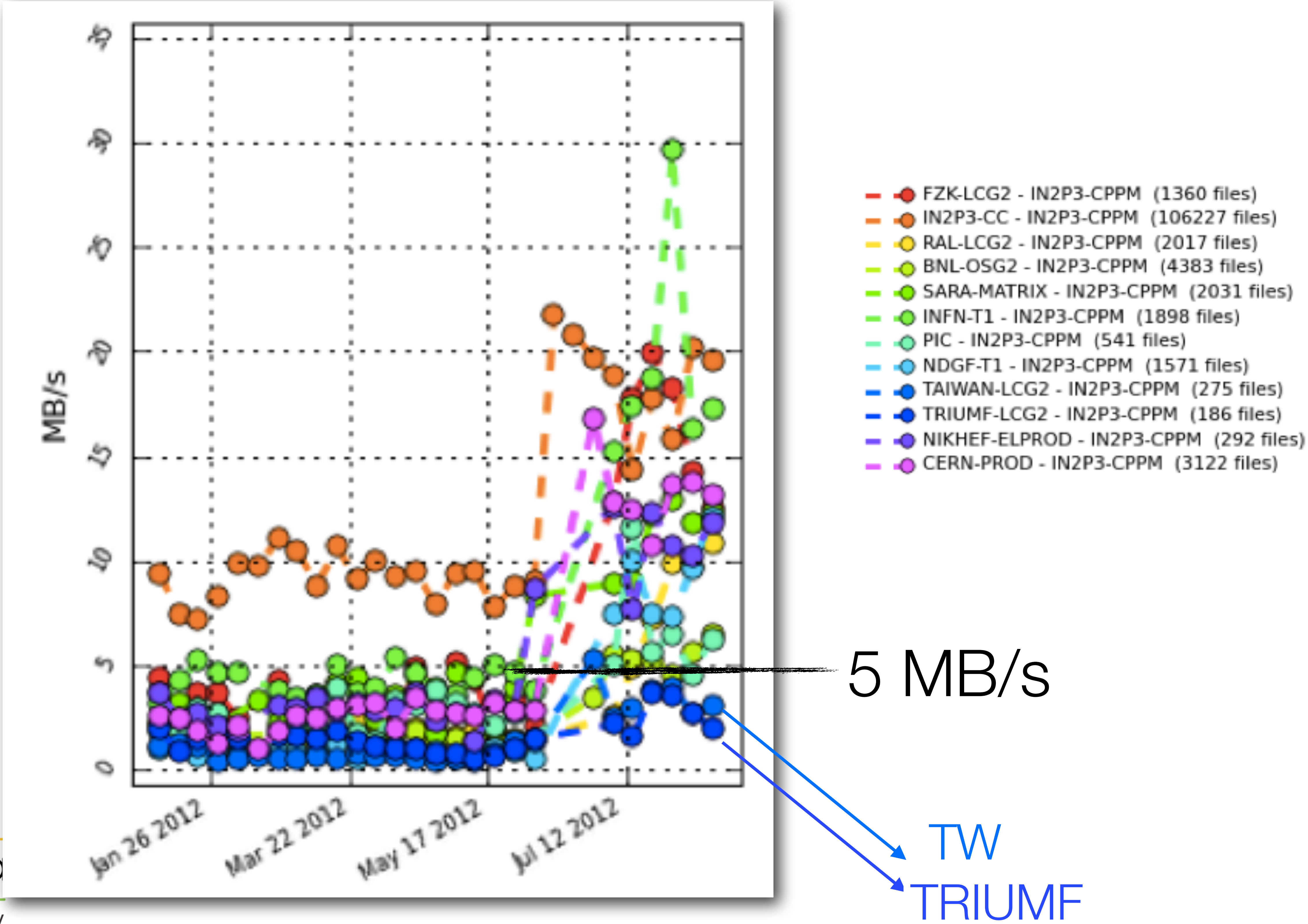
June 25th connected to LHCONE @ 10 Gb/s



5 MB/s

TRIUMF

IN2P3-CPPM -> T1s



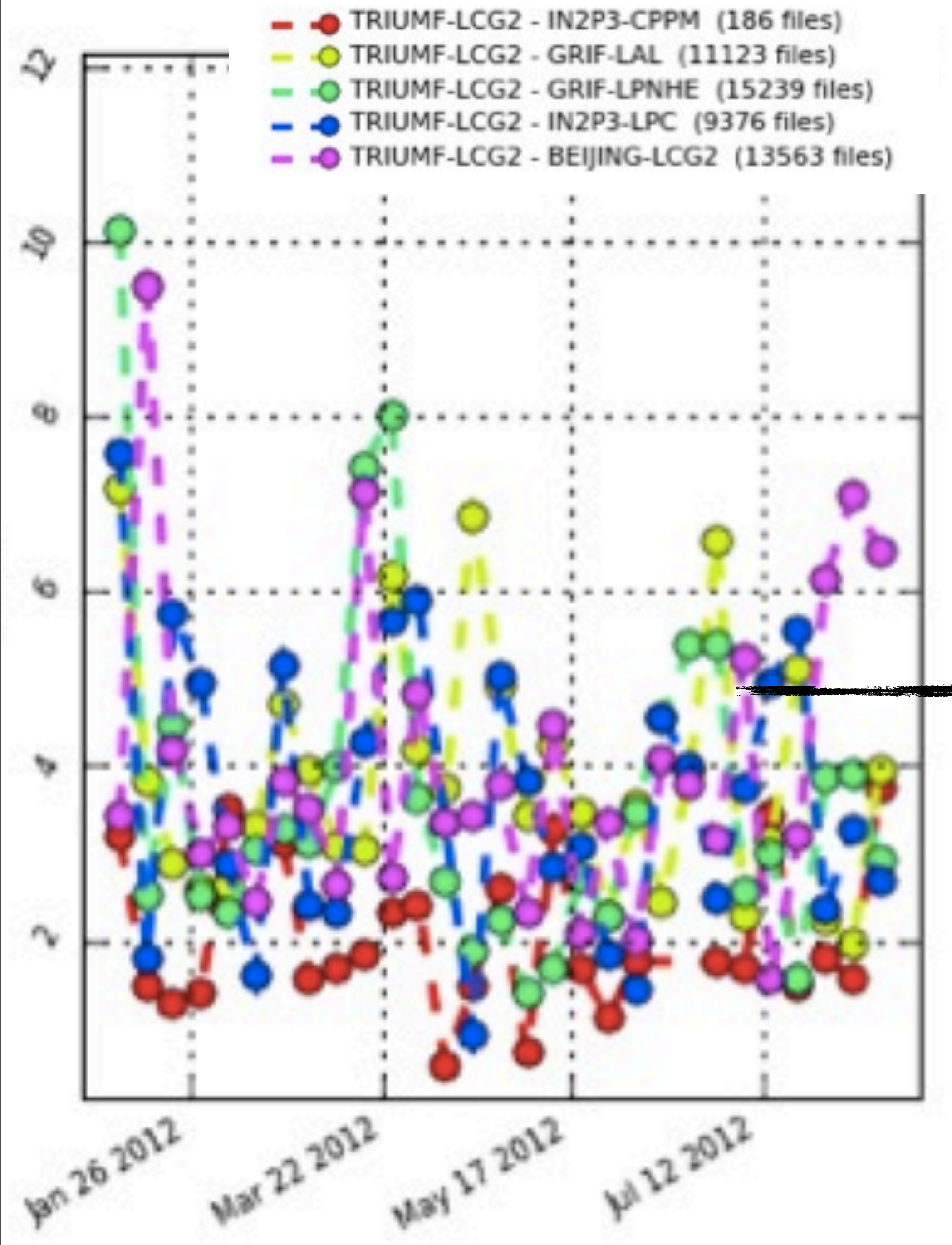
5 MB/s

TW

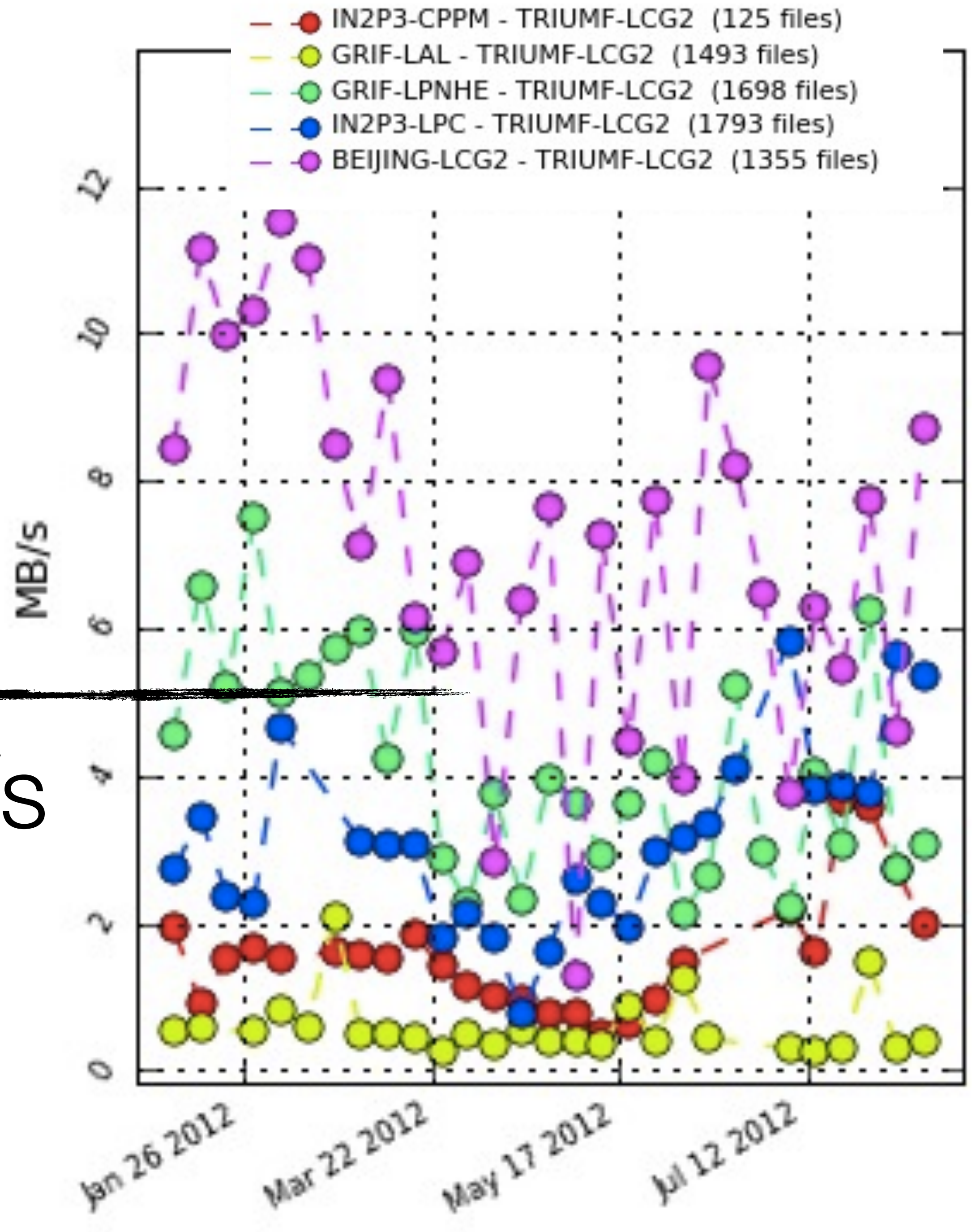
TRIUMF

TRIUMF <-> FR T2Ds

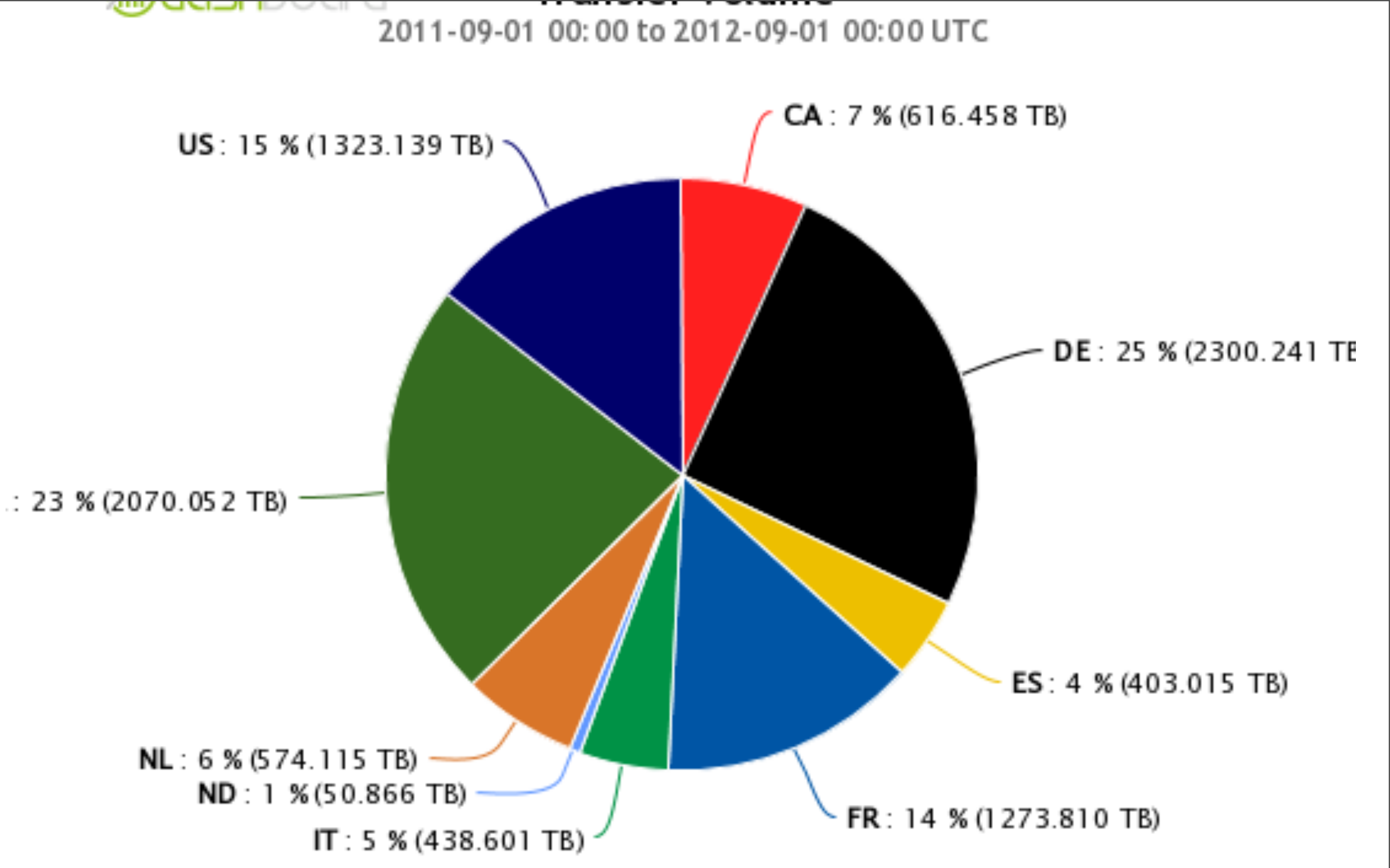
~None of T2Ds ever reaches the 5 MB/s canonical parameter value



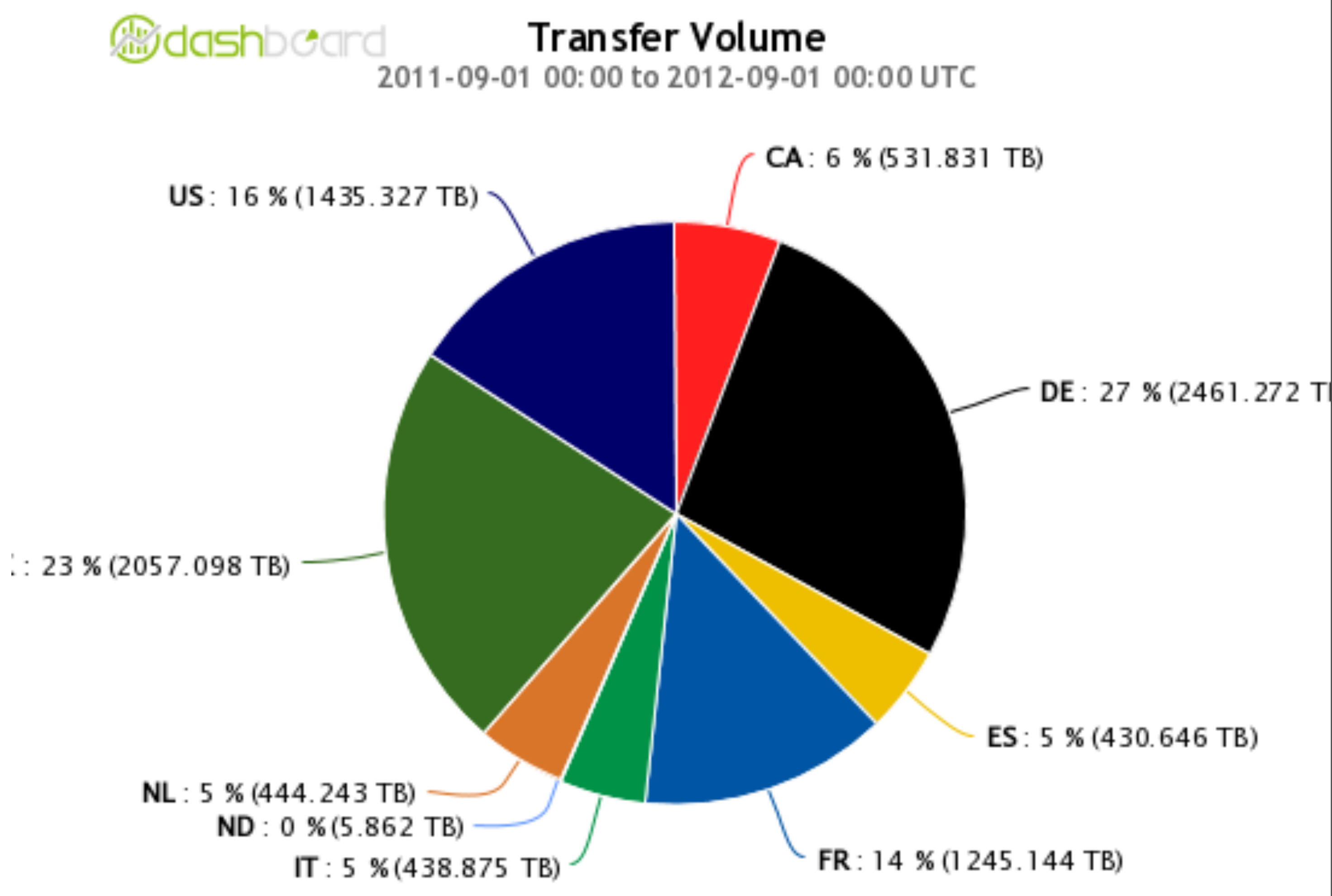
5 MB/s



T2-T2 destination

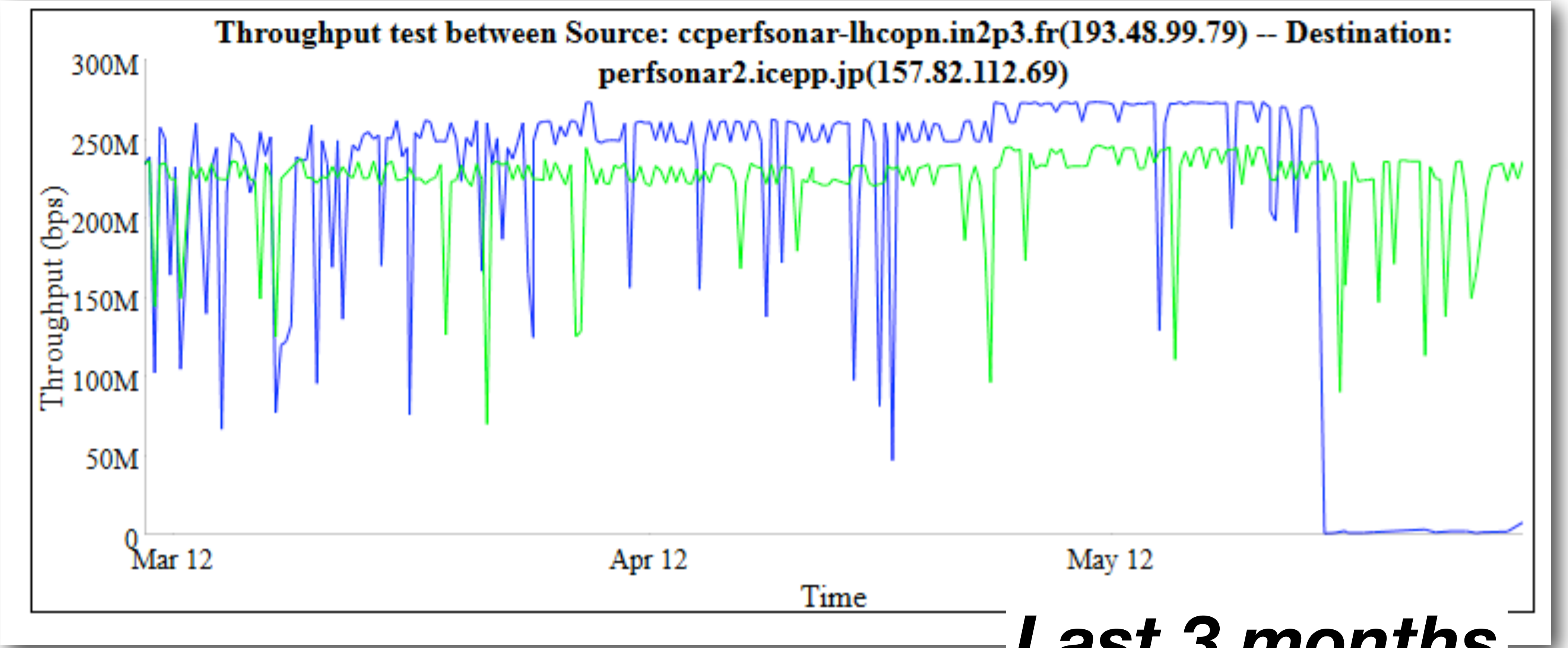


T2-T2 source



Network throughput measured with perfSONAR

CCIN2P3 → Tokyo **Tokyo → CCIN2P3**



5%

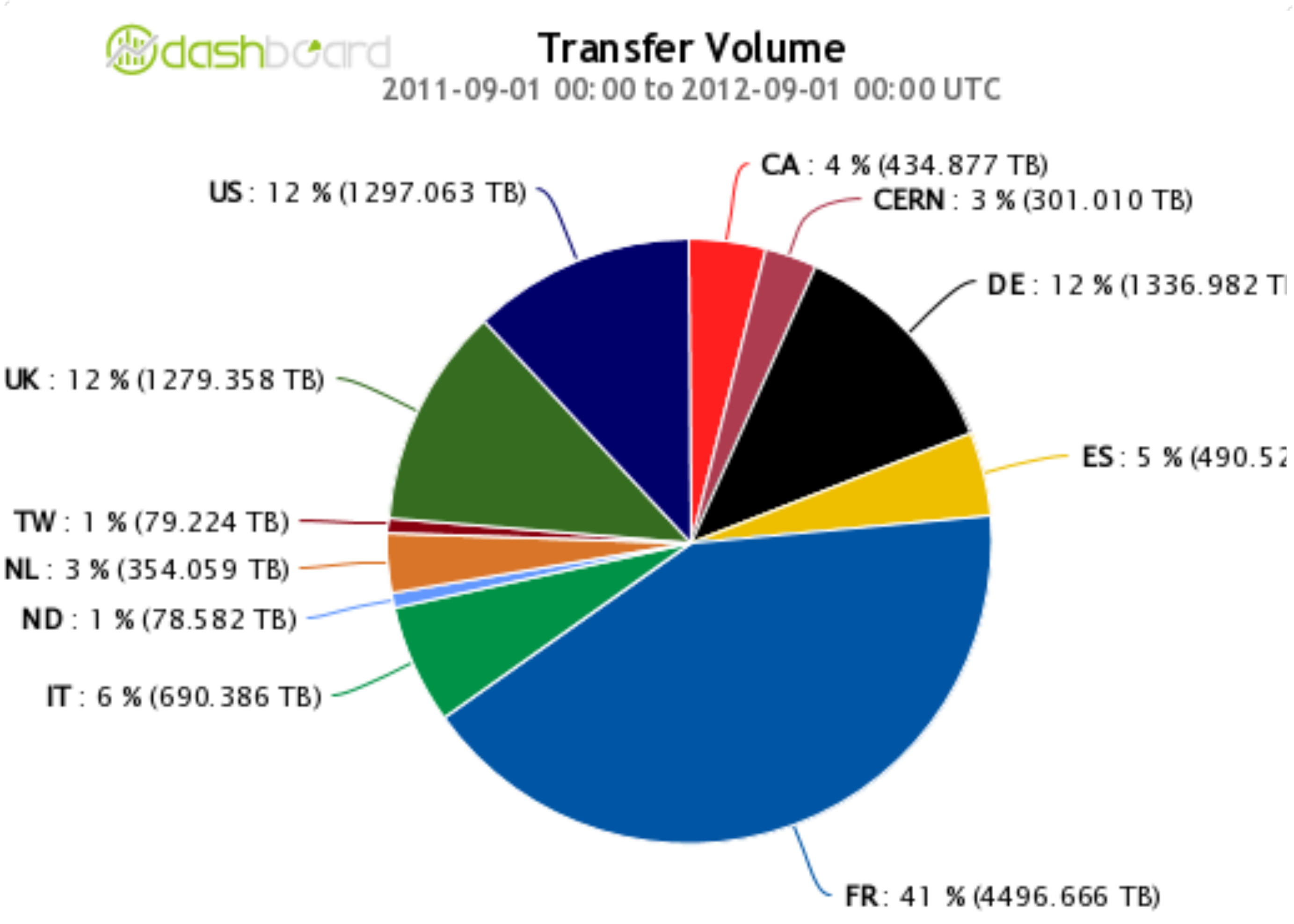
Last 3 months

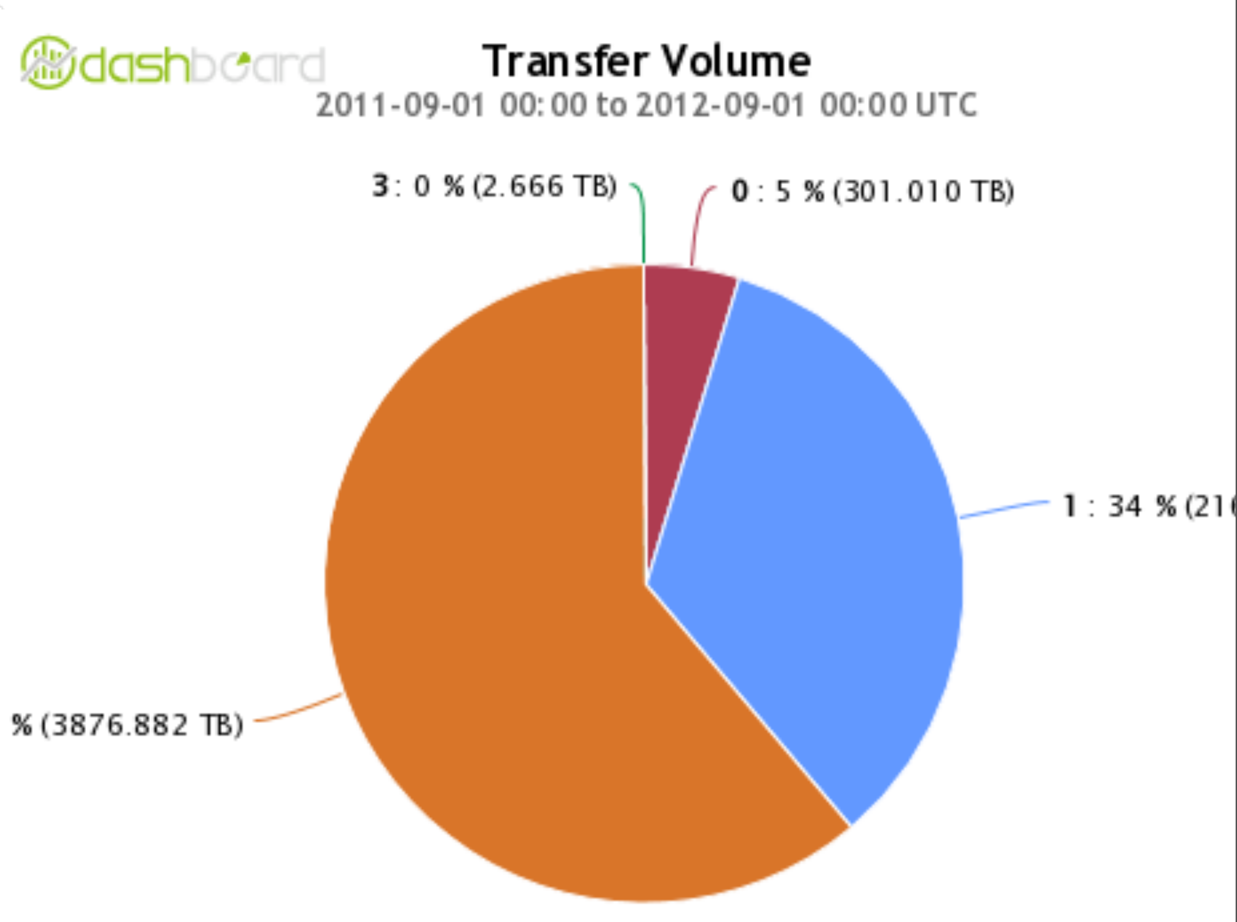
No so stable

better by ~5% for CCIN2P3 →Tokyo

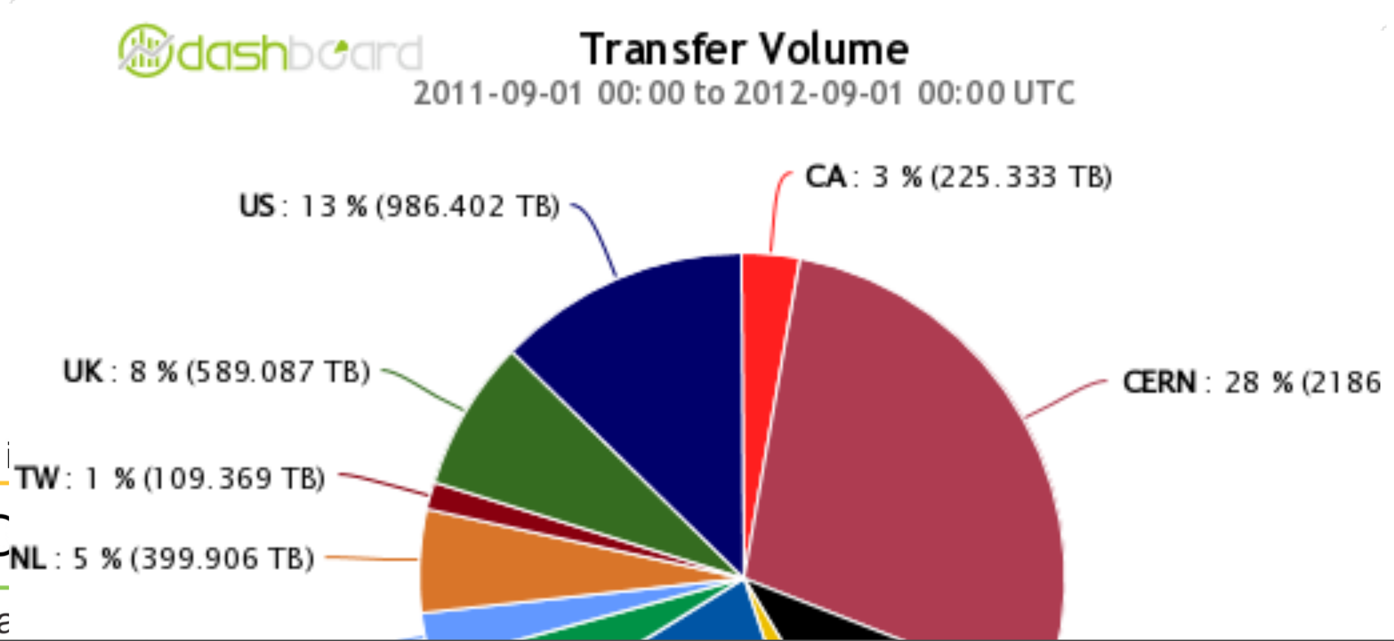
http://psps.perfsonar.net/

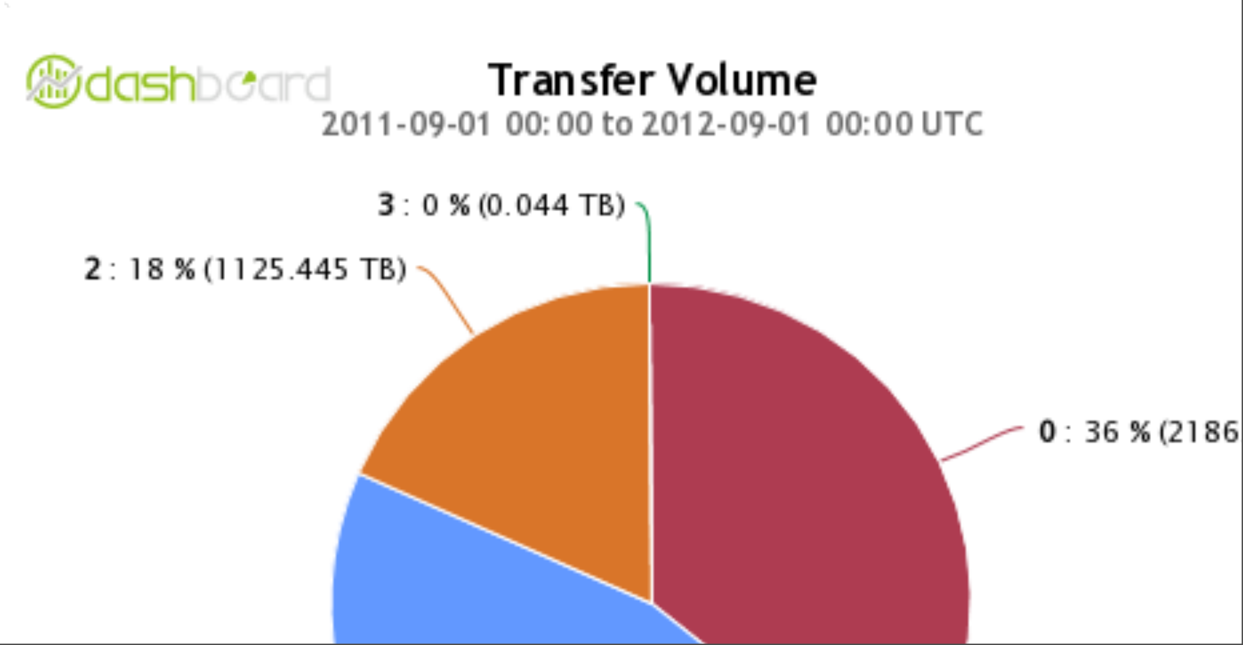
CCIN2P3 Exports



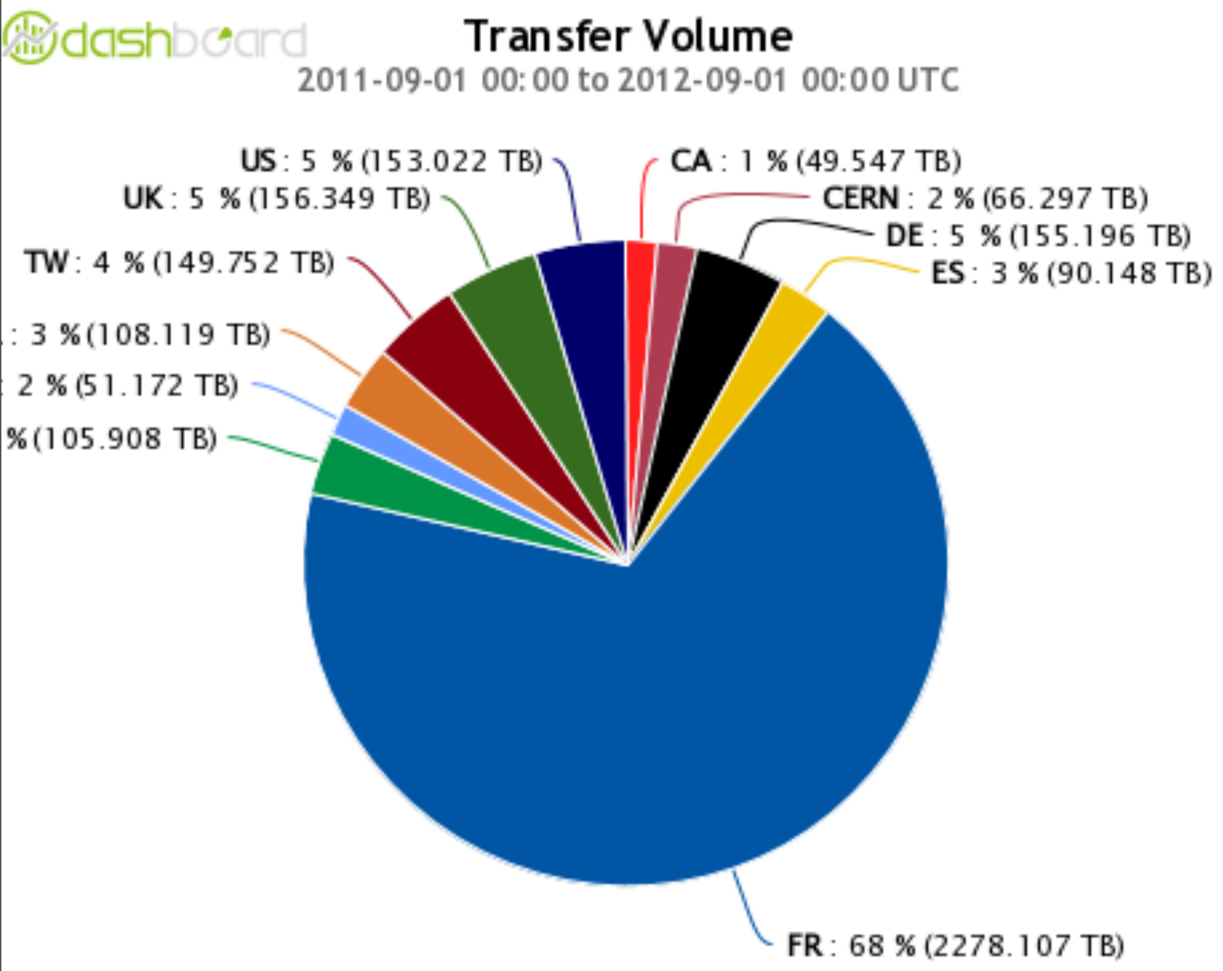


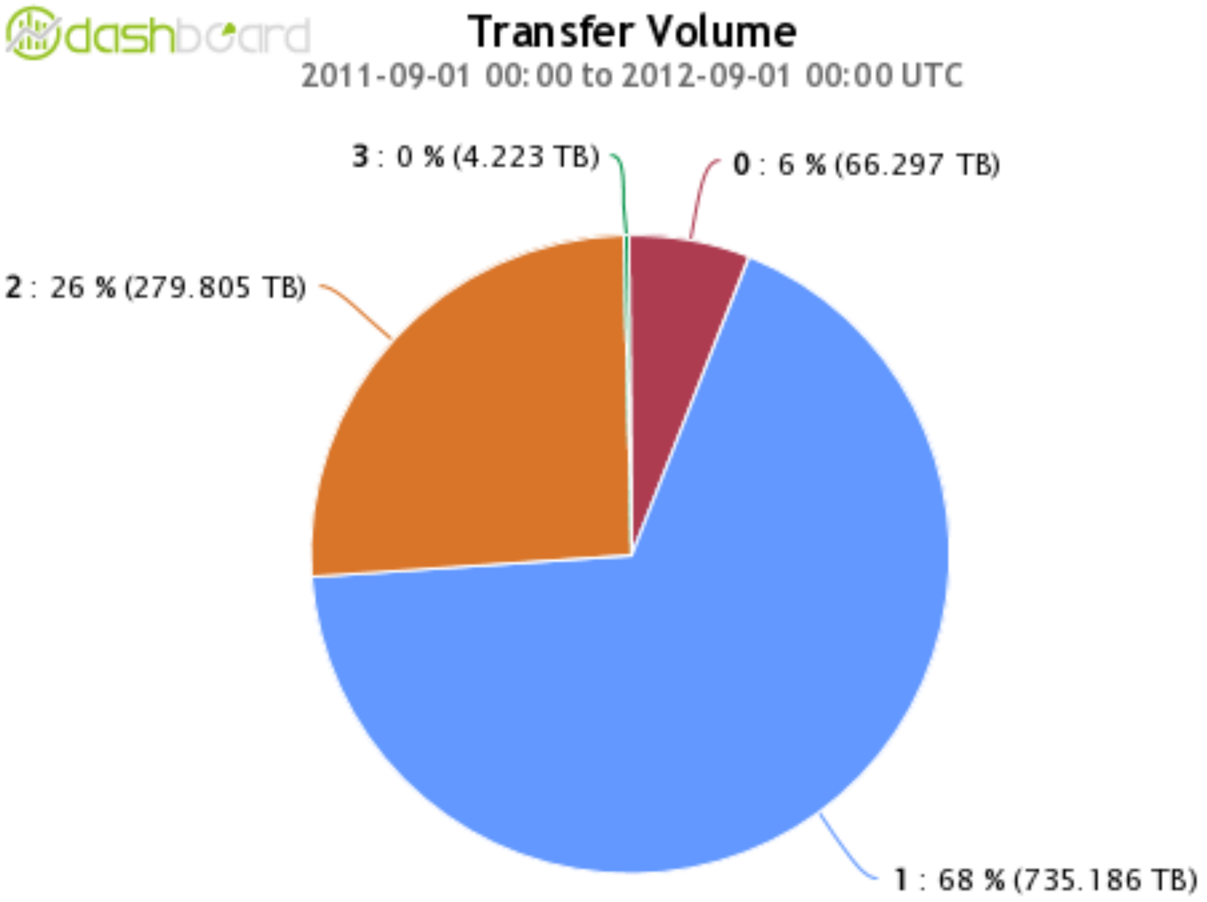
To Lyon



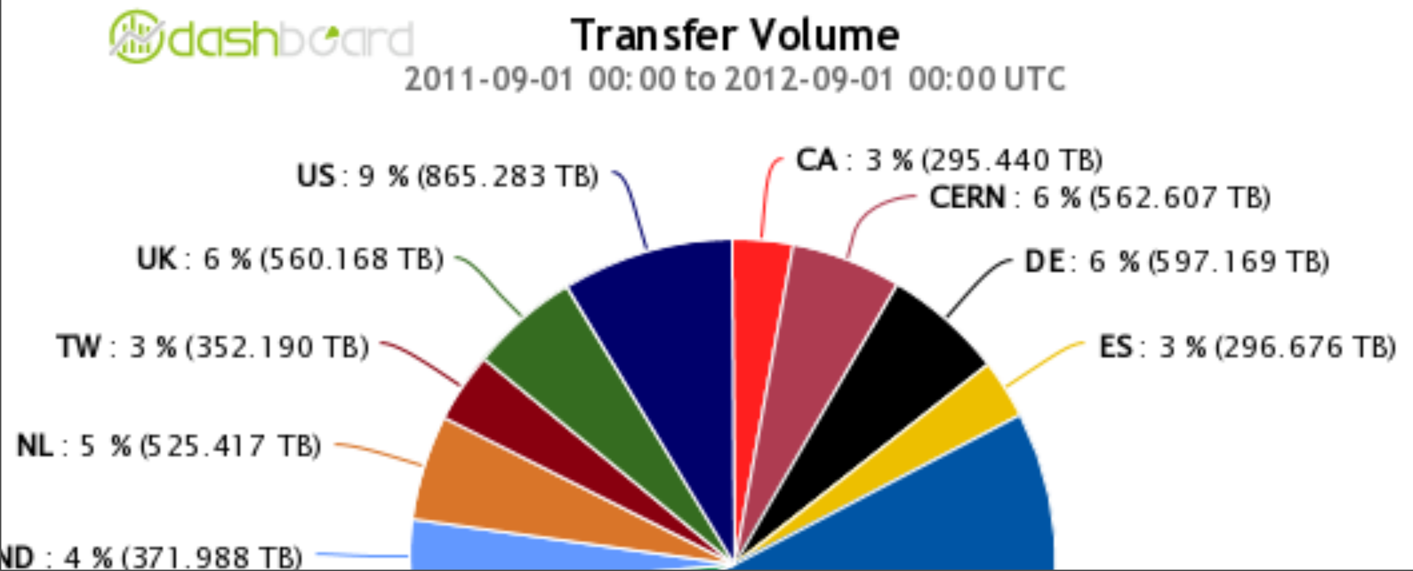


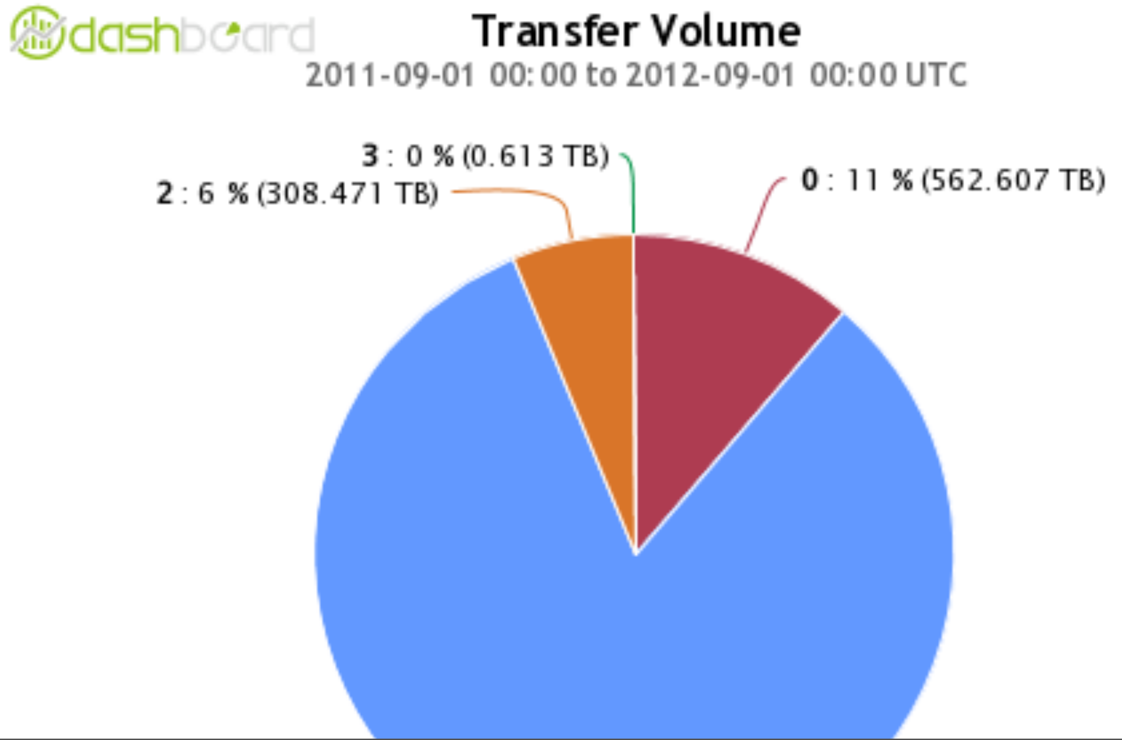
T2s Exports



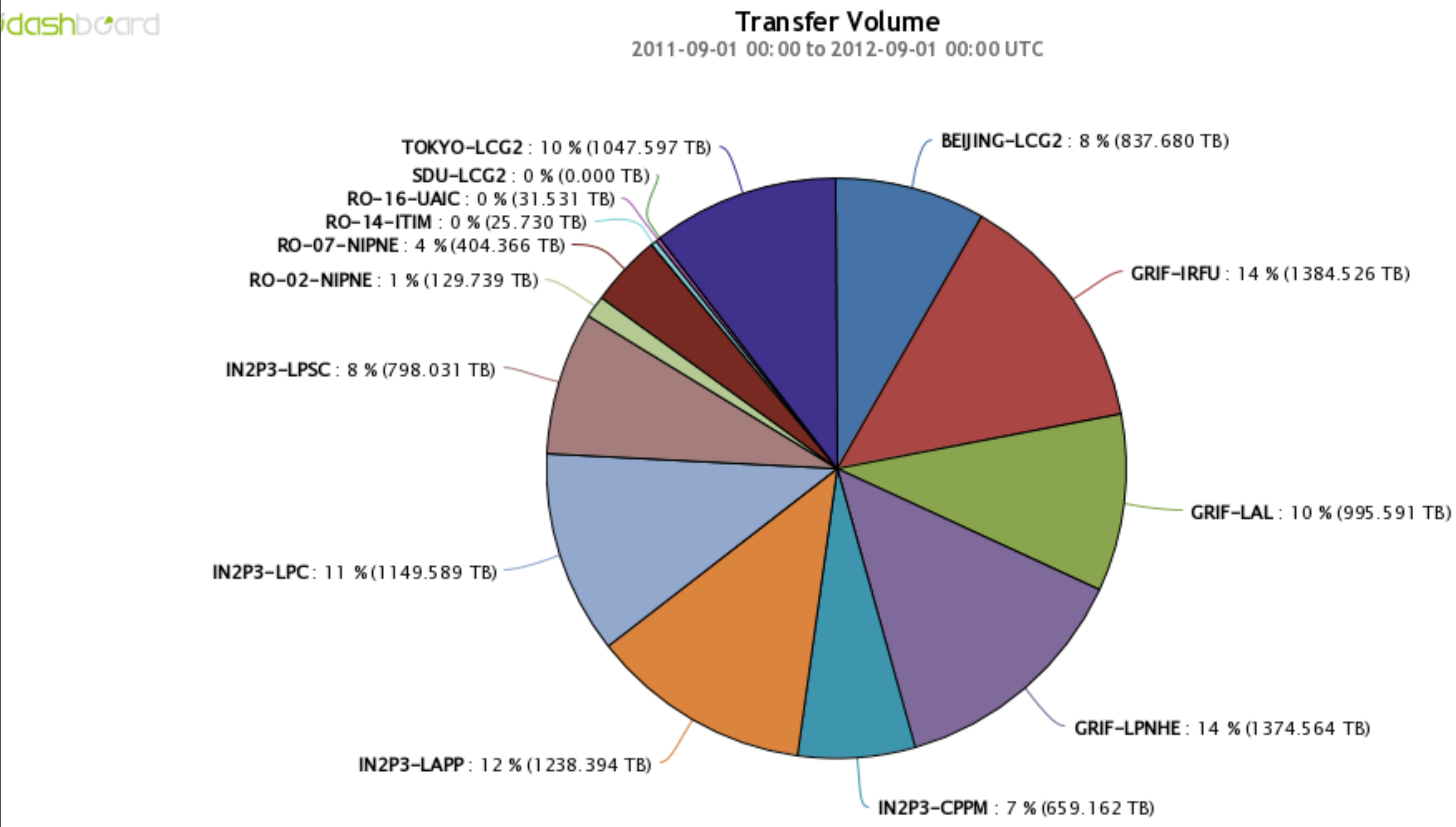


T2s Import





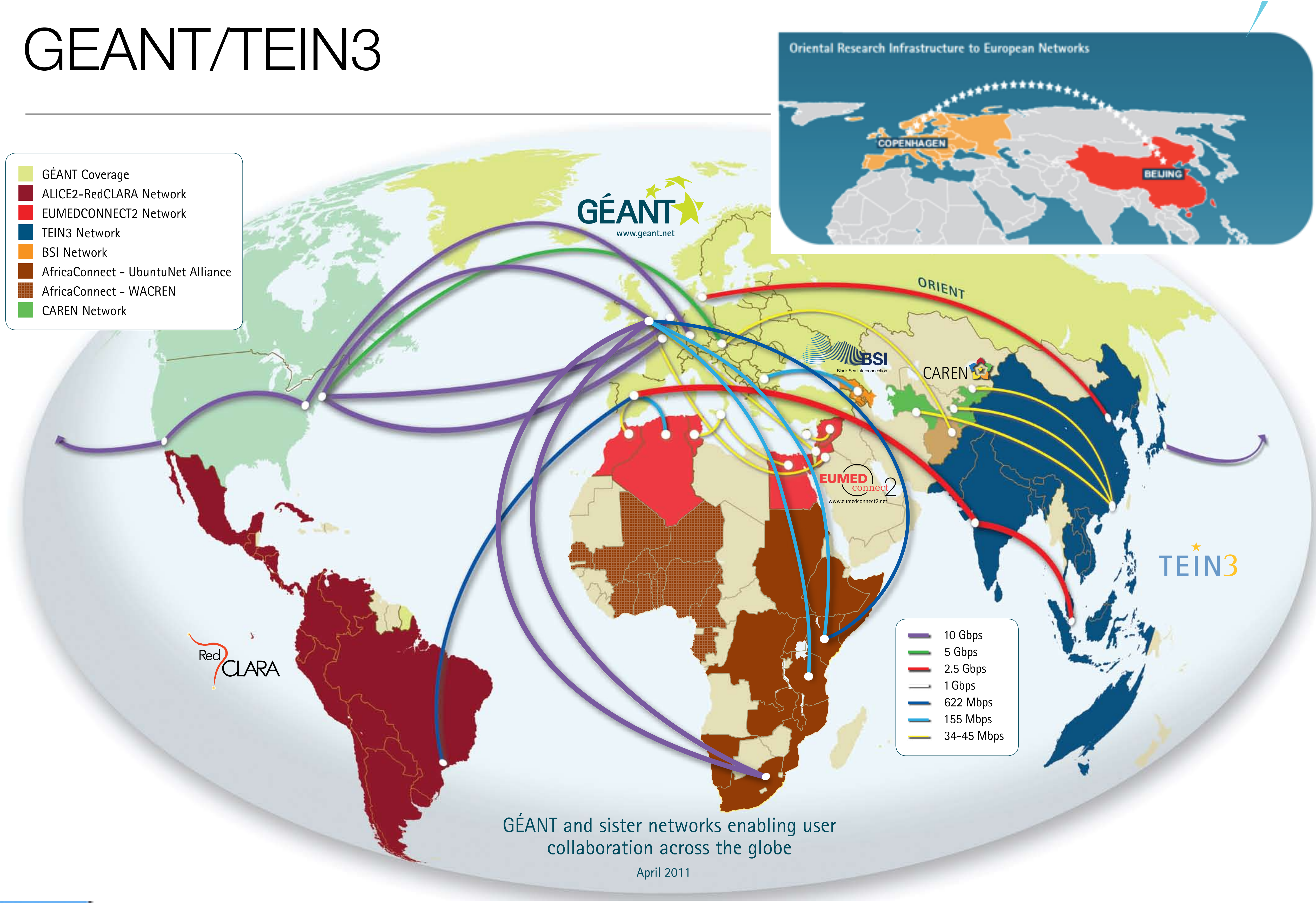
destination on FR cloud



trans-Siberian route

ORIENT

GEANT/TEIN3

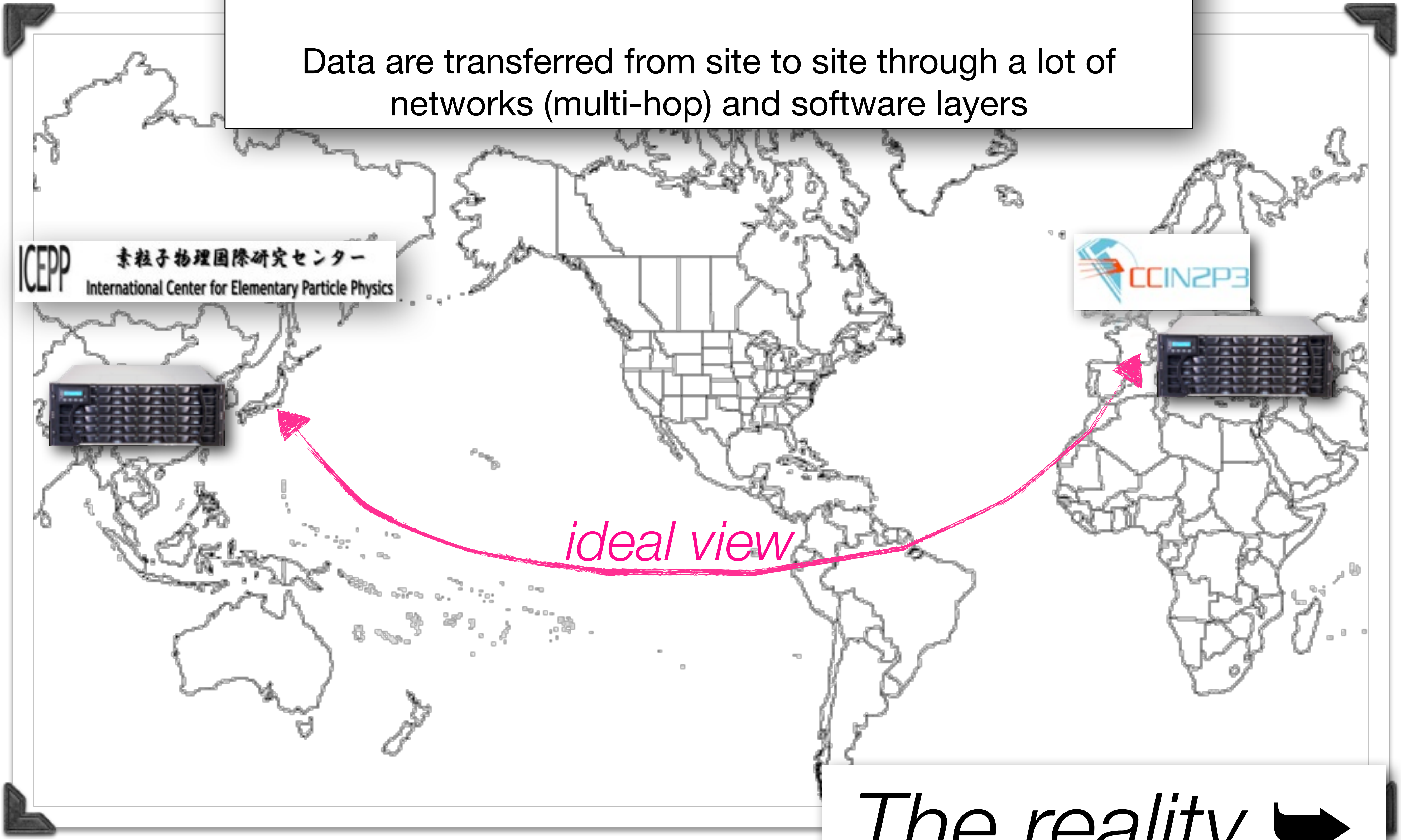


GÉANT and sister networks enabling user collaboration across the globe

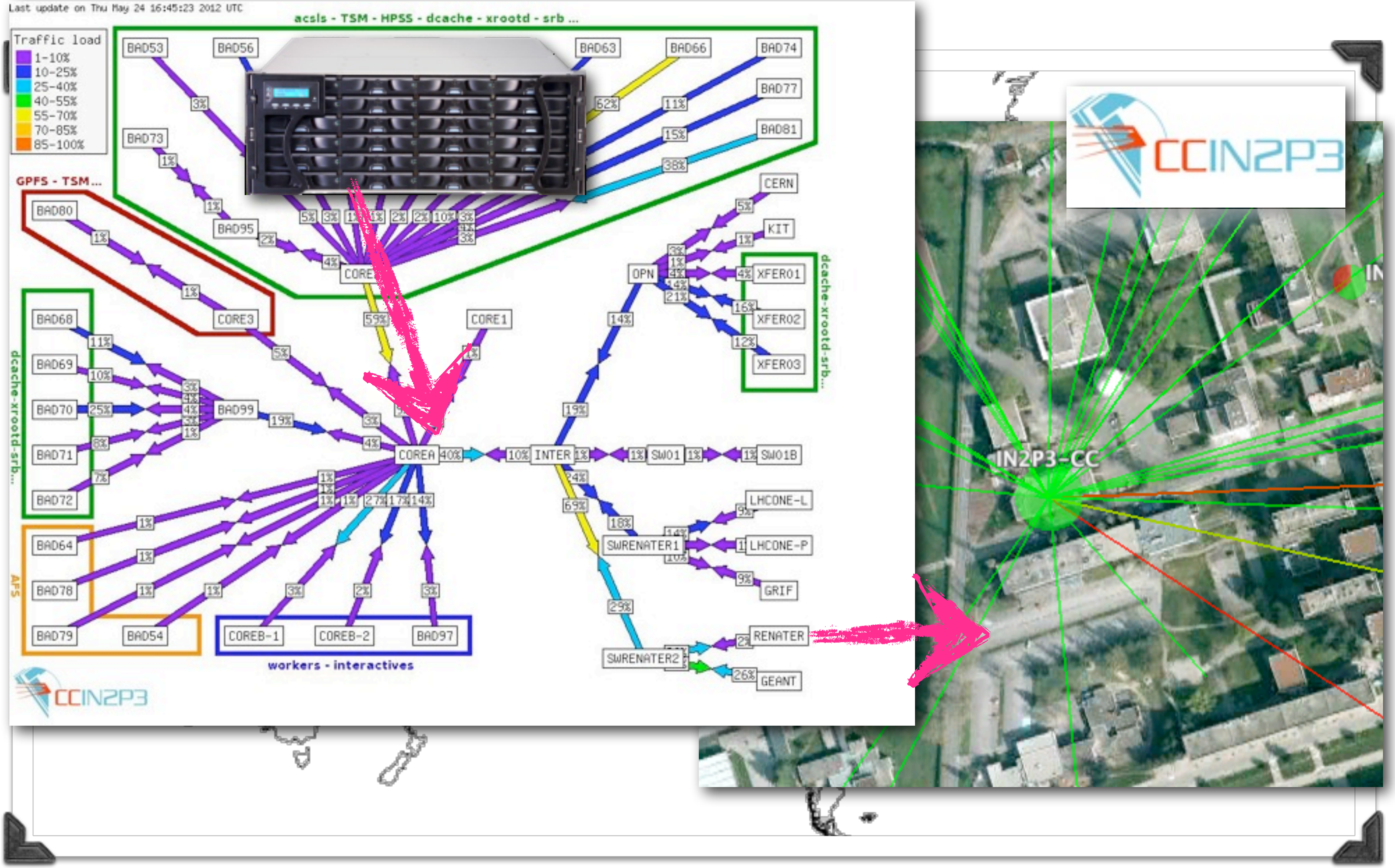
April 2011

Tokyo is far from CCIN2P3 : ~300 ms RTT (Round Trip Time)
Throughput ~ 1 / RTT

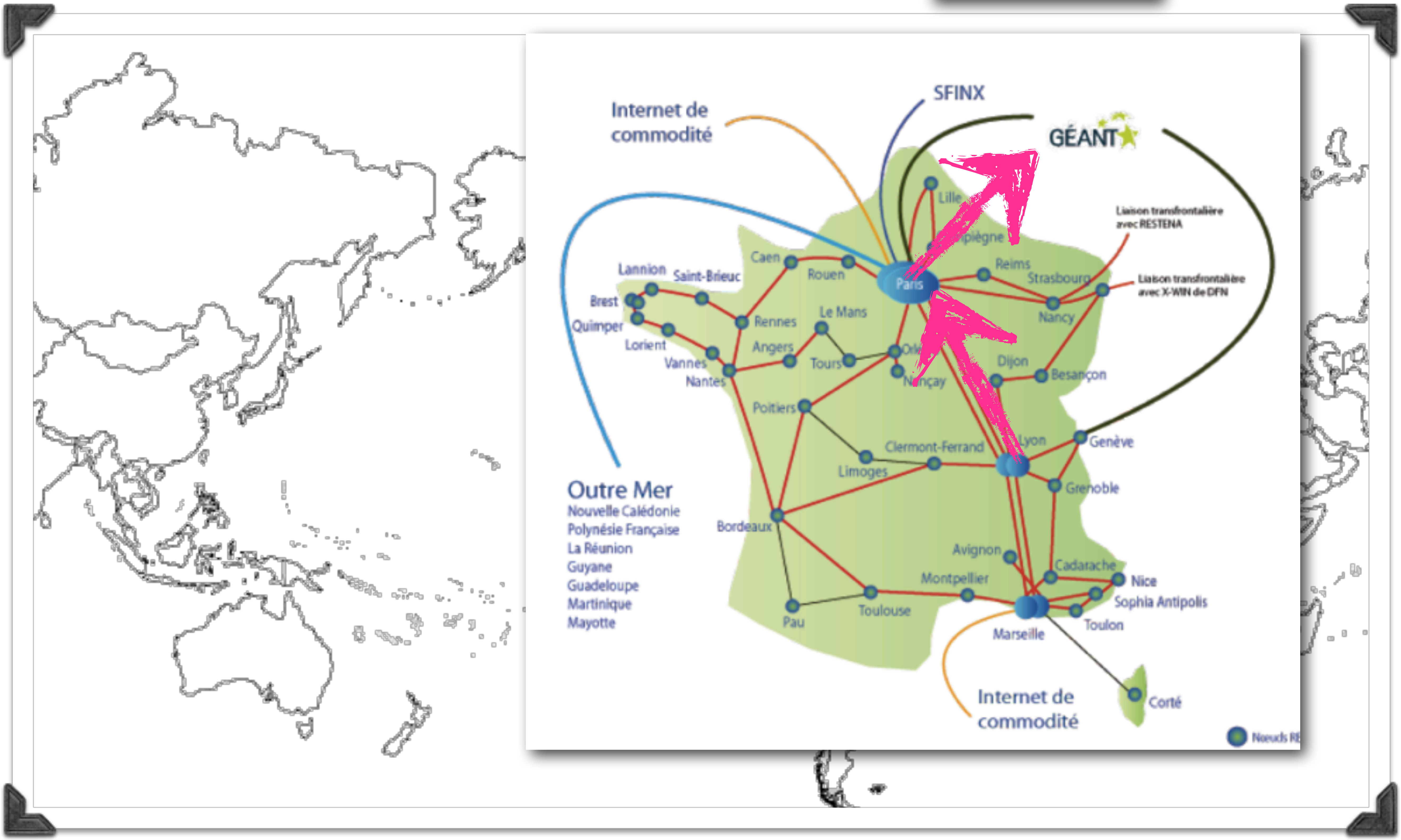
Data are transferred from site to site through a lot of networks (multi-hop) and software layers


ICEPP
素粒子物理国際研究センター
International Center for Elementary Particle Physics
CCIN2P3
ideal view

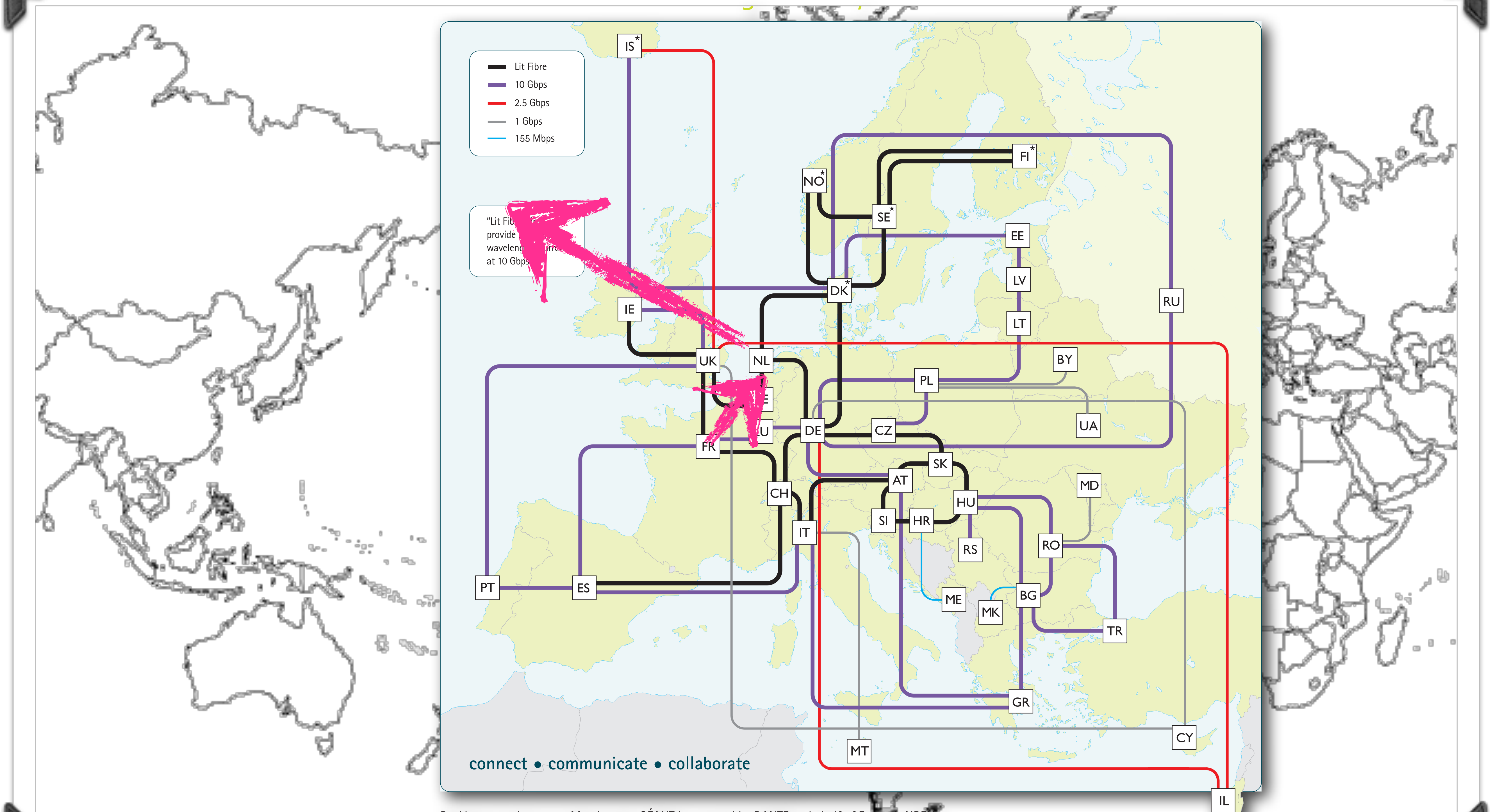
The reality ➥


Last update on Thu May 24 16:45:23 2012 UTC
Traffic load
1-10%
10-25%
25-40%
40-55%
55-70%
70-85%
85-100%
acsls - TSM - HPSS - dcache - xrootd - srb ...
GPFS - TSM ...
dcache-xrootd-srb...
AFS
workers - interactives
dcache-xrootd-srb...
BAD53
BAD56
BAD63
BAD66
BAD74
BAD77
BAD81
BAD73
BAD95
BAD80
CORE3
CORE1
CORE
BAD68
BAD69
BAD70
BAD71
BAD72
BAD99
COREA
INTER
SW01
SW01B
OPN
CERN
KIT
XFER01
XFER02
XFER03
SWRENATER1
SWRENATER2
LHCONE-L
LHCONE-P
GRIF
RENATER
GEANT
BAD64
BAD78
BAD79
BAD54
COREB-1
COREB-2
BAD97
CCIN2P3
IN2P3-CC

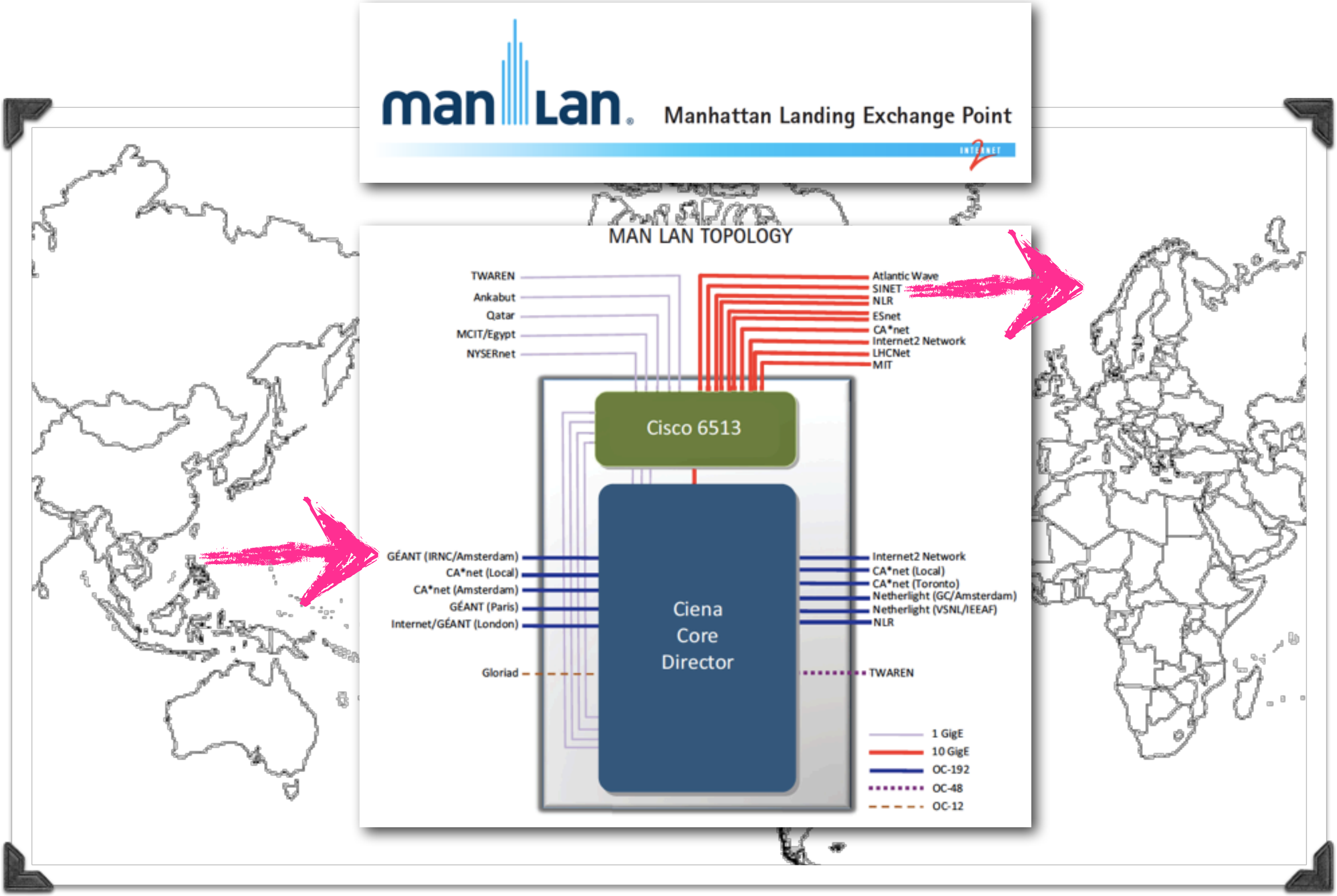
Renater


SFINX
Internet de commodité
GÉANT
Liaison transfrontalière avec RESTENA
Liaison transfrontalière avec X-WIN de DFN
Lille
piègne
Caen
Rouen
Paris
Reims
Strasbourg
Nancy
Lannion
Saint-Brieuc
Brest
Quimper
Lorient
Vannes
Nantes
Rennes
Le Mans
Angers
Tours
Orl
Nançay
Dijon
Besançon
Poitiers
Clermont-Ferrand
Lyon
Genève
Limoges
Grenoble
Bordeaux
Avignon
Cadarache
Nice
Sophia Antipolis
Montpellier
Toulouse
Pau
Marseille
Toulon
Corté
Internet de commodité
Outre Mer
Nouvelle Calédonie
Polynésie Française
La Réunion
Guyane
Guadeloupe
Martinique
Mayotte
Noeuds RE

GEANT2



Backbone topology as at March 2012. GÉANT is operated by DANTE on behalf of Europe's NRENs.



manLan Manhattan Landing Exchange Point

INTERNET2

MAN LAN TOPOLOGY

TWAREN
Ankabut
Qatar
MCIT/Egypt
NYSERnet

Atlantic Wave
SINET
NLR
ESnet
CA*net
Internet2 Network
LHCNet
MIT

Cisco 6513

Ciena Core Director

GÉANT (IRNC/Amsterdam)
CA*net (Local)
CA*net (Amsterdam)
GÉANT (Paris)
Internet/GÉANT (London)

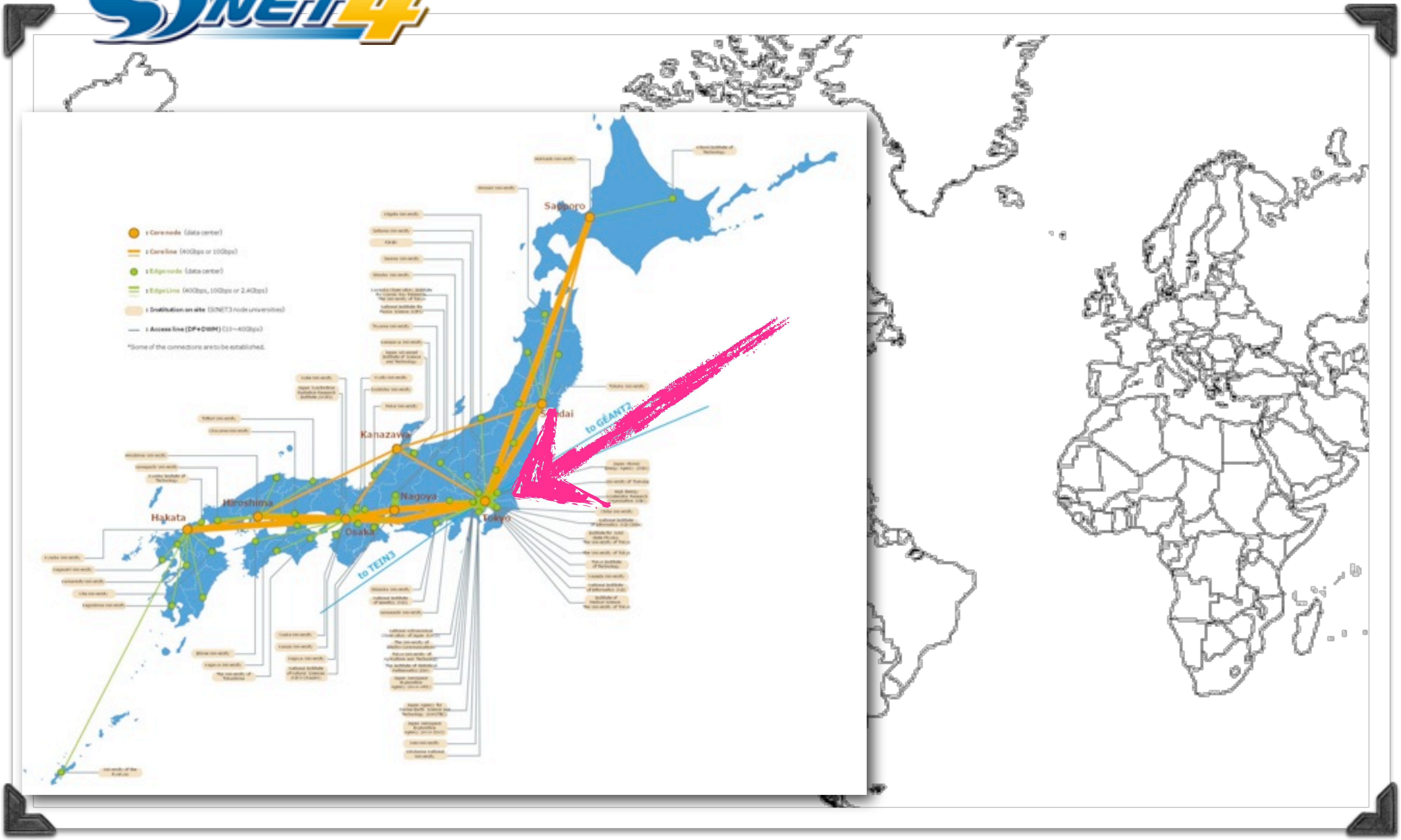
Gloriad

Internet2 Network
CA*net (Local)
CA*net (Toronto)
Netherlight (GC/Amsterdam)
Netherlight (VSNL/IEEAF)
NLR

TWAREN

─	1 GigE
━	10 GigE
━	OC-192
·······	OC-48
- - - -	OC-12

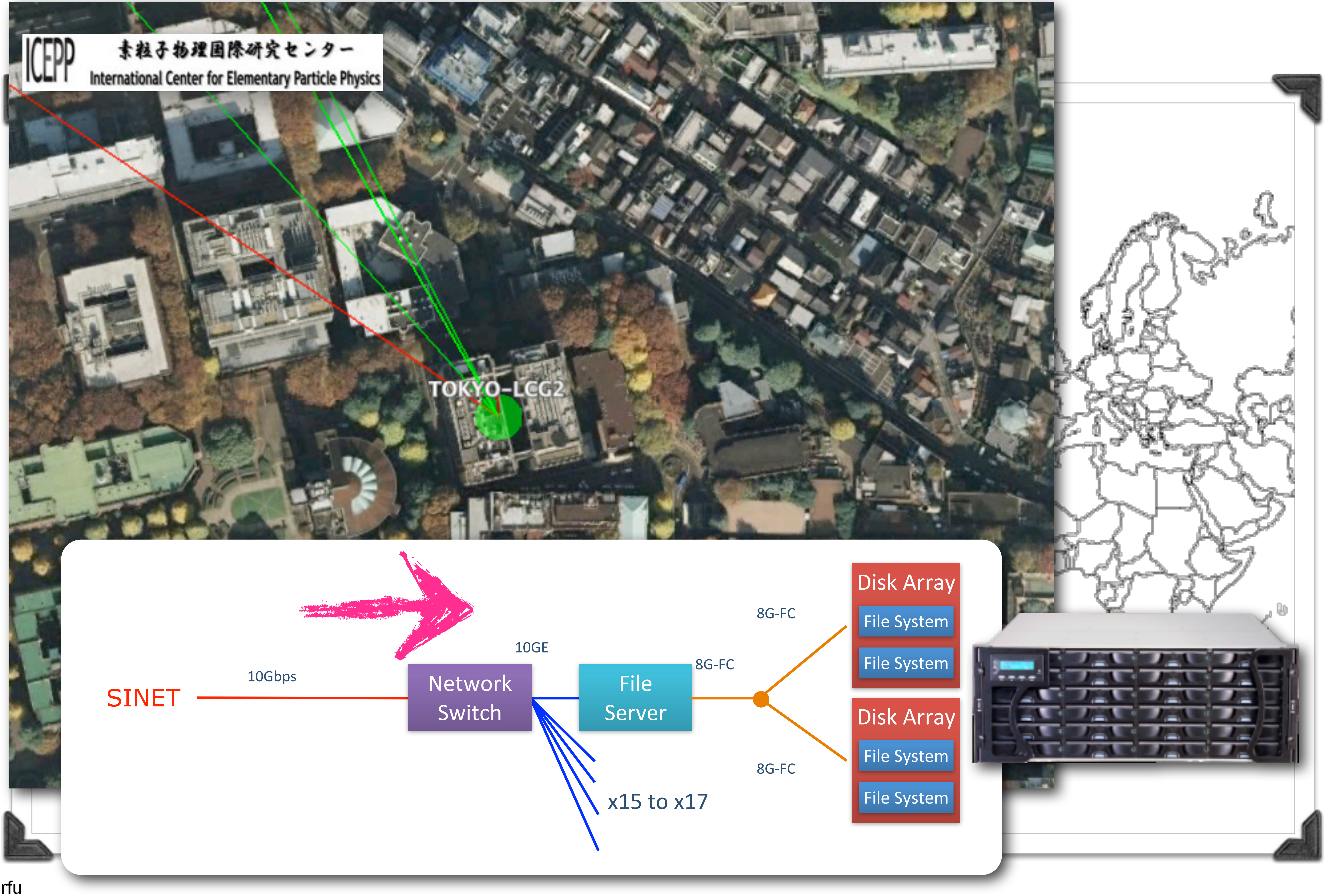
SINET4


: Core node (data center)
: Core line (40Gbps or 10Gbps)
: Edge node (data center)
: Edge Line (40Gbps, 10Gbps or 2.4Gbps)
: Institution on site (SINET3 node universities)
: Access line (DF+DWM) (10~40Gbps)
*Some of the connections are to be established.
Sapporo
Sendai
to GÉANT2
Kanazawa
Nagoya
Tokyo
Hiroshima
Hakata
Osaka
to TEIN3

irfu

cea

saclay


ICEPP
素粒子物理国際研究センター
International Center for Elementary Particle Physics
TOKYO-LCG2

SINET

10Gbps

Network Switch

10GE

File Server

x15 to x17

8G-FC

8G-FC

8G-FC

Disk Array

File System

File System

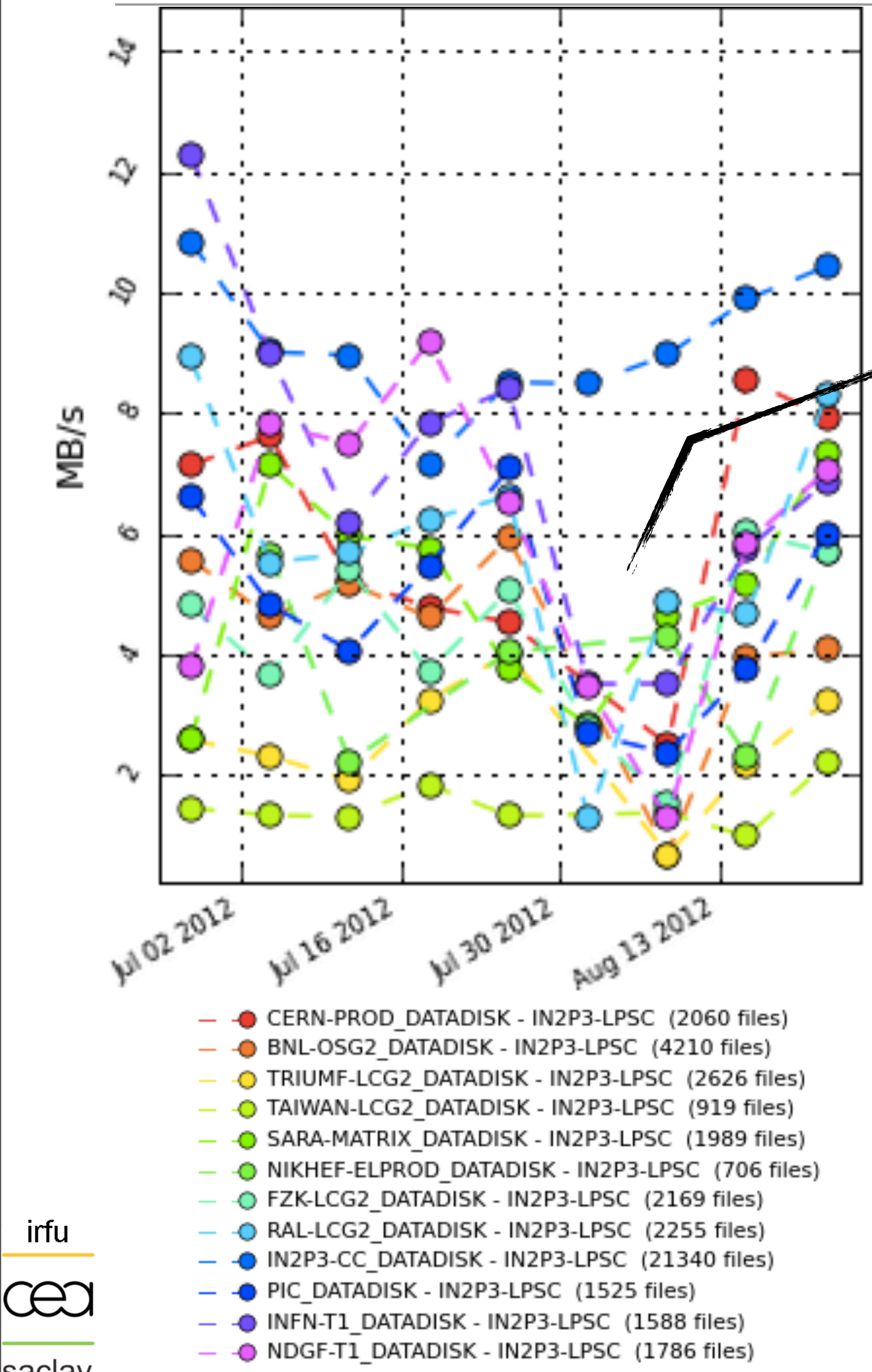
Disk Array

File System

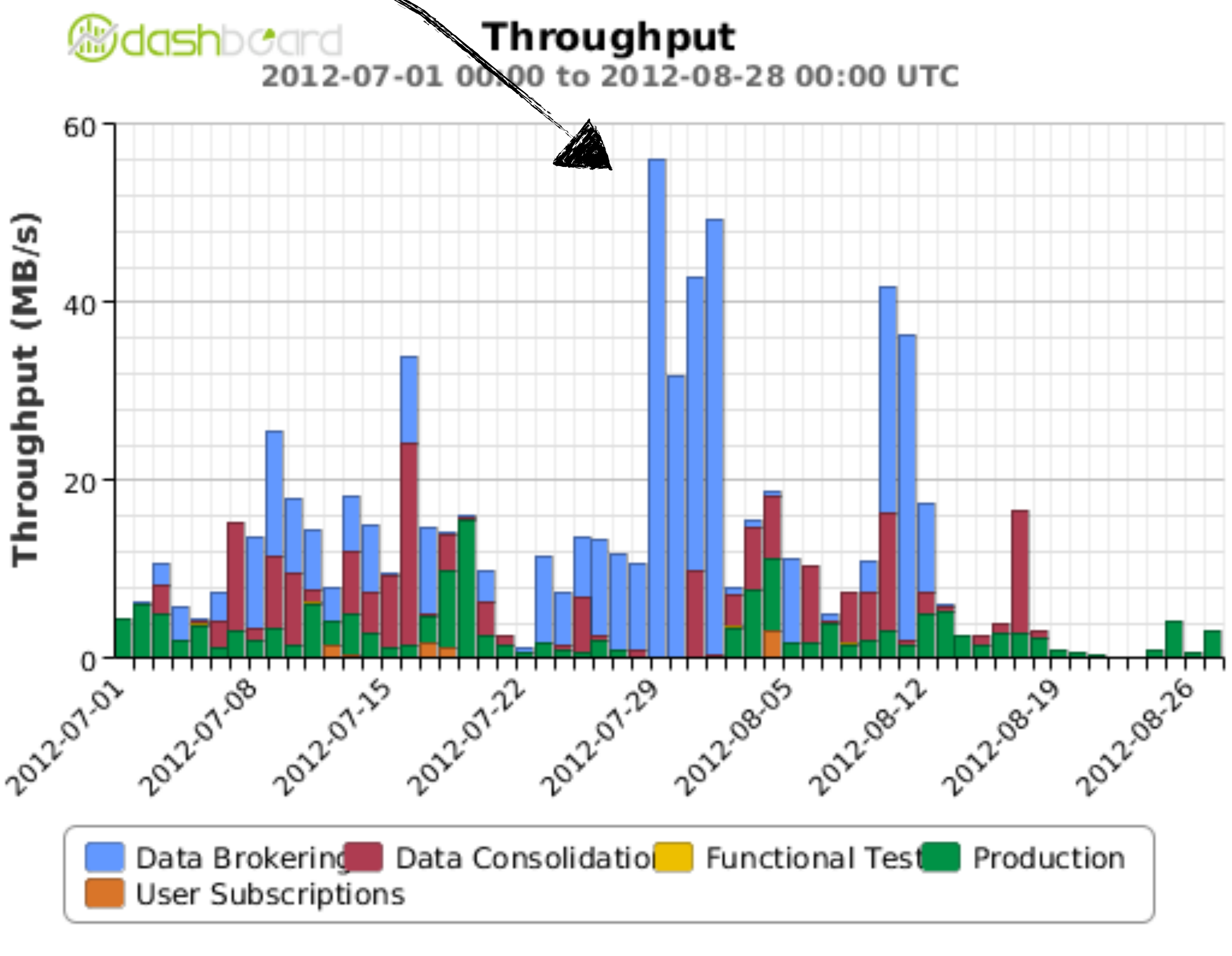
File System

T1s -> IN2P3-LPSC

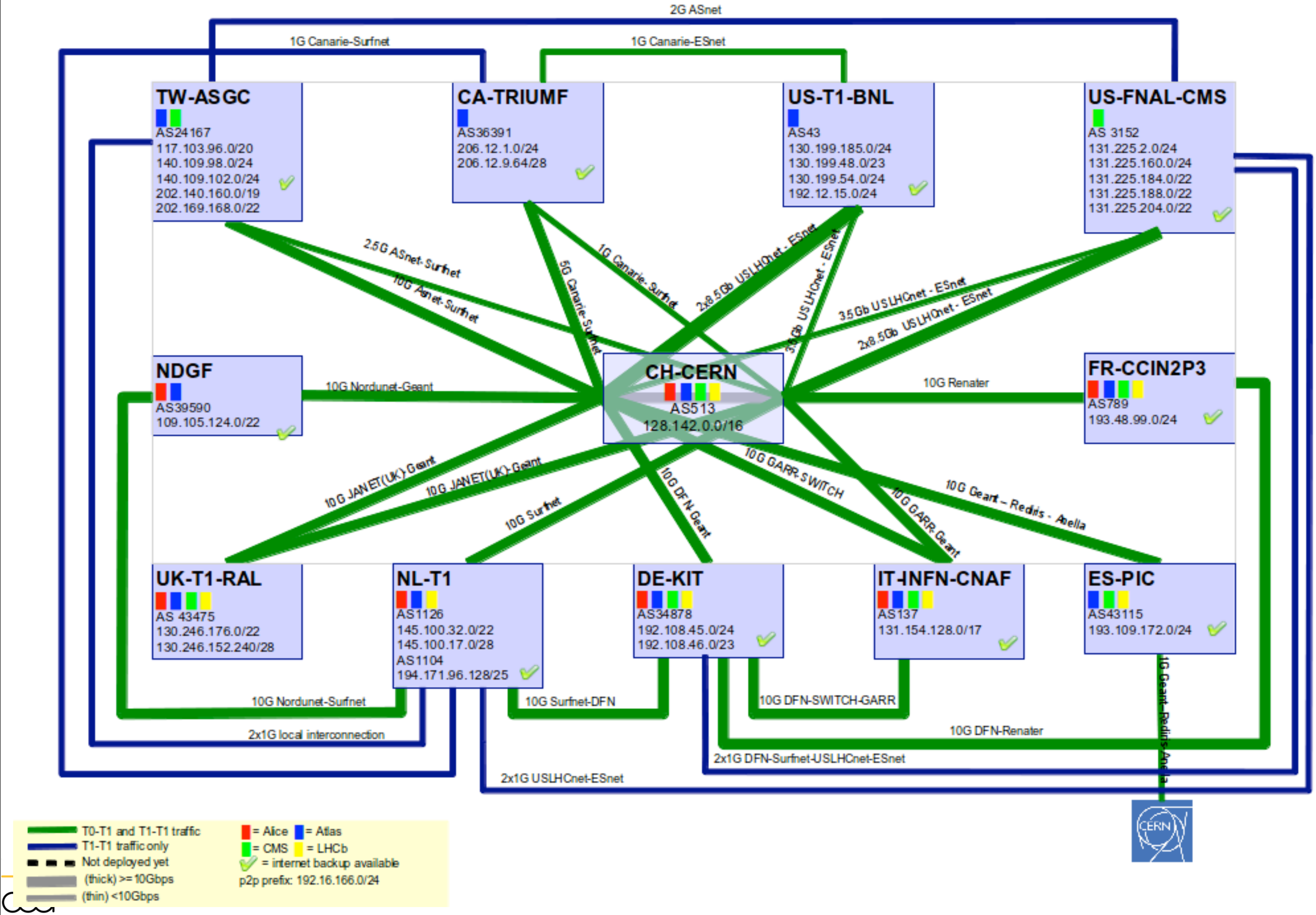
Heavy transfers from IN2P3-CC (T1) interference with FTS monitoring



IN2P3-CC -> IN2P3-LPSC



LHCOPN



2G ASnet

1G Canarie-Surfnet

1G Canarie-ESnet

TW-ASGC
🟦🟩
AS24167
117.103.96.0/20
140.109.98.0/24
140.109.102.0/24
202.140.160.0/19
202.169.168.0/22
✓

CA-TRIUMF
🟦
AS36391
206.12.1.0/24
206.12.9.64/28
✓

US-T1-BNL
🟦
AS43
130.199.185.0/24
130.199.48.0/23
130.199.54.0/24
192.12.15.0/24
✓

US-FNAL-CMS
🟩
AS 3152
131.225.2.0/24
131.225.160.0/24
131.225.184.0/22
131.225.188.0/22
131.225.204.0/22
✓

2.5G ASnet- Surfnet

10G Asnet- Surfnet

5G Canarie-Surfnet

1G Canarie- Surfnet

2x8.5Gb USLHCnet - ESnet

3.5Gb USLHCnet - ESnet

3.5Gb USLHCnet - ESnet

2x8.5Gb USLHCnet - ESnet

NDGF
🟥🟦
AS39590
109.105.124.0/22
✓

10G Nordunet-Geant

CH-CERN
🟥🟦🟩🟨
AS513
128.142.0.0/16

10G Renater

FR-CCIN2P3
🟥🟦🟩🟨
AS789
193.48.99.0/24
✓

10 G JANET(UK)-Geant

10 G JANET(UK)-Geant

10G Surfnet

10G DFN-Geant

10G GARR-SWITCH

10G GARR-Geant

10G Geant – Rediris - Anella

UK-T1-RAL
🟥🟦🟩🟨
AS 43475
130.246.176.0/22
130.246.152.240/28

NL-T1
🟥🟦🟨
AS1126
145.100.32.0/22
145.100.17.0/28
AS1104
194.171.96.128/25
✓

DE-KIT
🟥🟦🟩🟨
AS34878
192.108.45.0/24
192.108.46.0/23
✓

IT-INFN-CNAF
🟥🟦🟩🟨
AS137
131.154.128.0/17
✓

ES-PIC
🟦🟩🟨
AS43115
193.109.172.0/24
✓

1G Geant-Rediris-Anella

10G Nordunet-Surfnet

2x1G local interconnection

10G Surfnet-DFN

10G DFN-SWITCH-GARR

10G DFN-Renater

2x1G DFN-Surfnet-USLHCnet-ESnet

2x1G USLHCnet-ESnet

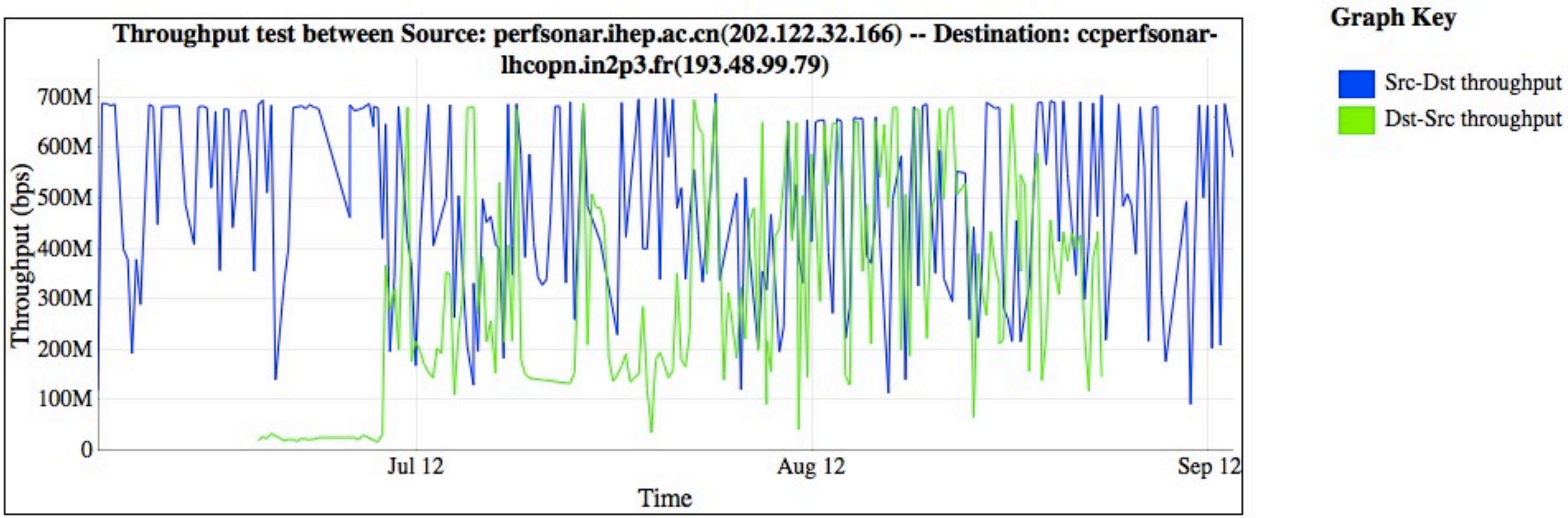
CERN

T0-T1 and T1-T1 traffic
T1-T1 traffic only
Not deployed yet
(thick) >= 10Gbps
(thin) <10Gbps

🟥 = Alice 🟦 = Atlas
🟩 = CMS 🟨 = LHCb
✓ = internet backup available
p2p prefix: 192.16.166.0/24

IN2P3-CC ↔ Beijing as seen by perfSonar

Beijing -> IN2P3-CC   IN2P3-CC -> Beijing



- Link unstable
- Asymmetry

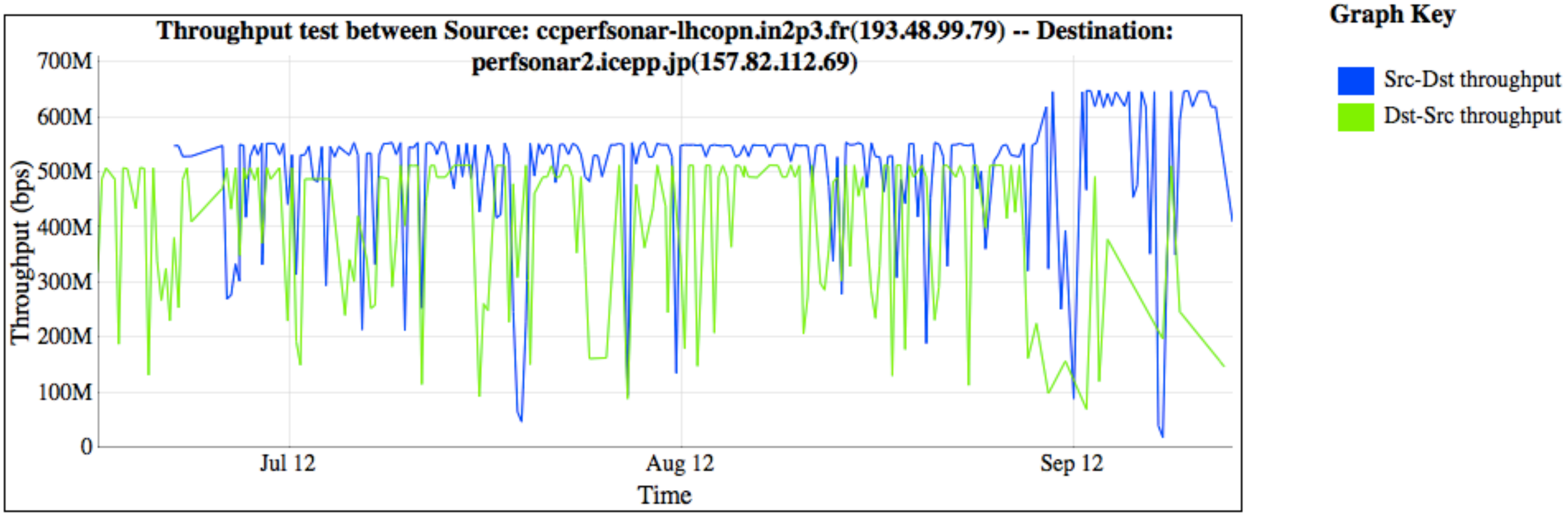
irfu

cea

saclay

IN2P3-CC ↔ Tokyo as seen by perfSonar

IN2P3-CC-> Tokyo Tokyo -> IN2P3-CC


Throughput test between Source: ccperfsonar-lhcopn.in2p3.fr(193.48.99.79) -- Destination: perfsonar2.icepp.jp(157.82.112.69)
Throughput (bps)
700M
600M
500M
400M
300M
200M
100M
Jul 12
Aug 12
Sep 12
Time
Graph Key
Src-Dst throughput
Dst-Src throughput