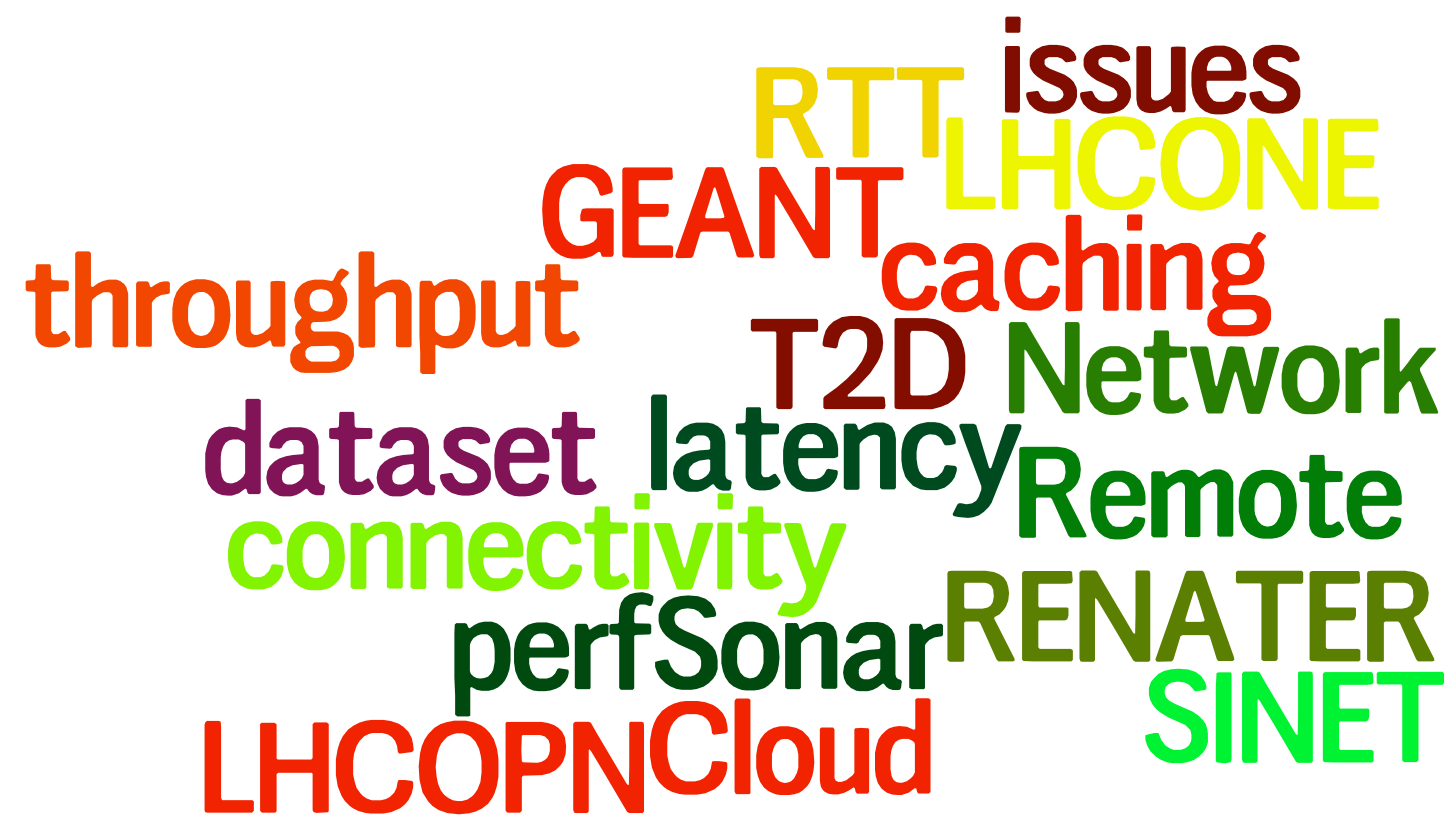


Network issues on FR cloud



A word cloud of network-related terms. The words are arranged in a cluster, with some overlapping. The colors of the words are: 'RTT' (yellow), 'issues' (dark red), 'LHCONE' (yellow), 'GEANT' (red), 'caching' (red), 'throughput' (orange), 'T2D' (dark red), 'Network' (green), 'dataset' (purple), 'latency' (dark green), 'Remote' (green), 'connectivity' (light green), 'perf' (dark green), 'Sonar' (dark green), 'RENATER' (green), 'LHCOPN' (red), 'Cloud' (red), and 'SINET' (green).

RTT issues
LHCONE
GEANT caching
throughput T2D Network
dataset latency Remote
connectivity
perf Sonar RENATER
LHCOPN Cloud SINET

ATLAS sites and connectivity

- ATLAS computing model is evolving
 - Experience of one year of data taking
 - Tools and monitoring are getting more mature
- New category of sites : Direct T2s (**T2Ds**)
 - primary hosts for datasets (analysis) and for group analysis
 - get and send data from different sites
- Connectivity is important for T2Ds (not only...)

Aim :
all FR-cloud T2s
should be T2D

validated T2Ds

- CA : CA-SCINET-T2, CA-VICTORIA-WESTGRID-T2, SFU-LCG2_DATADISK, CA-MCGILL
- DE : DESY-HH, DESY-ZN, MPPMU, LRZ-LMU, CSCS-LCG2, GOEGRID, UNI-FREIBURG,
- ES : IFIC-LCG2, IFAE, UAM-LCG2, LIP-LISBON, NCG-INGRID-PT
- FR : GRIF-LPNHE, GRIF-LAL, IN2P3-LAPP, IN2P3-LPC, IN2P3-LPSC, BEIJING-LCG2
- IT : INFN-NAPOLI-ATLAS, INFN-MILANO-ATLASC, INFN-ROMA1
- UK : UKI-LT2-QMUL, UKI-NORTHGRID-MAN-HEP, UKI-SCOTGRID-GLASGOW, UKI-NO
- US : AGLT2, MWT2_UC, NET2, SLACXRD, SWT2 CPB

Network performance monitoring

- Monitoring tools developed in 2011
 - ATLAS 'sonar' : 'calibrated' file transfers by the ATLAS Data Distribution system, from storage to storage
 - perfSONAR : network performance (throughput, latency, traceroute, ...)
- Non stable network performances : also side effects on storage system at CCIN2P3

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

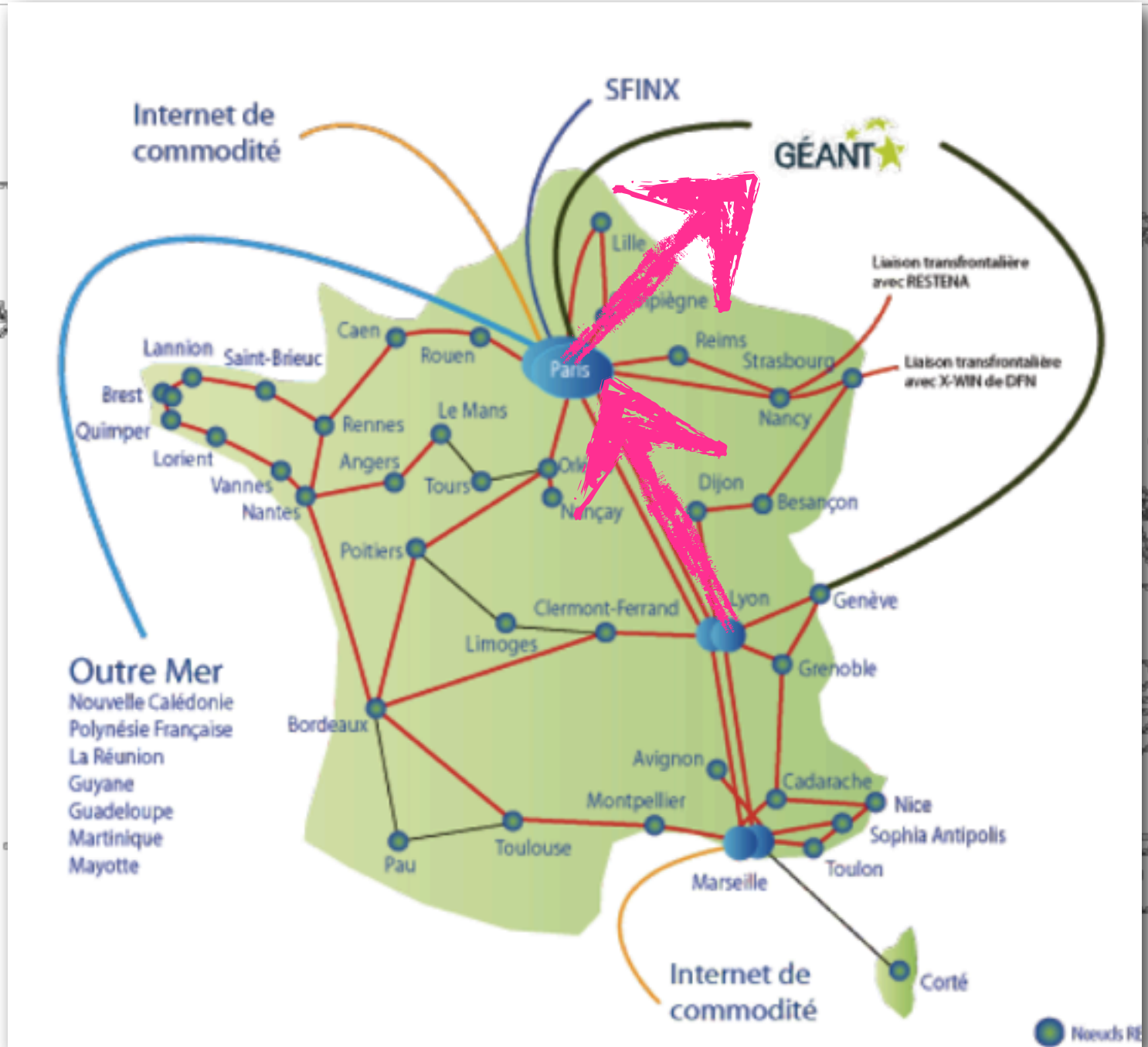
Data are transferred from site to site through a lot of
networks (multi-hop) and software layers

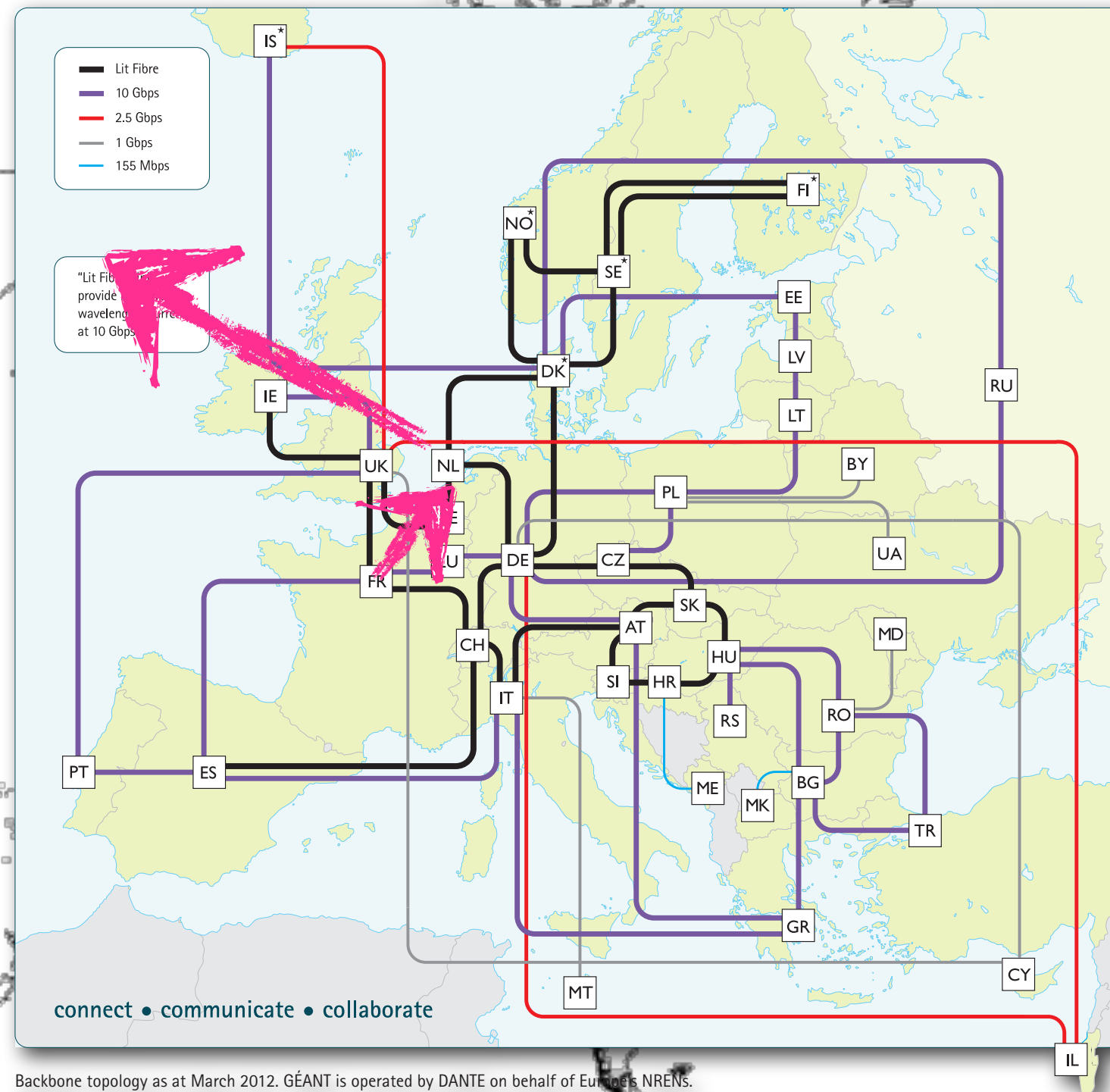
ICEPP
素粒子物理国際研究センター
International Center for Elementary Particle Physics

CCIN2P3

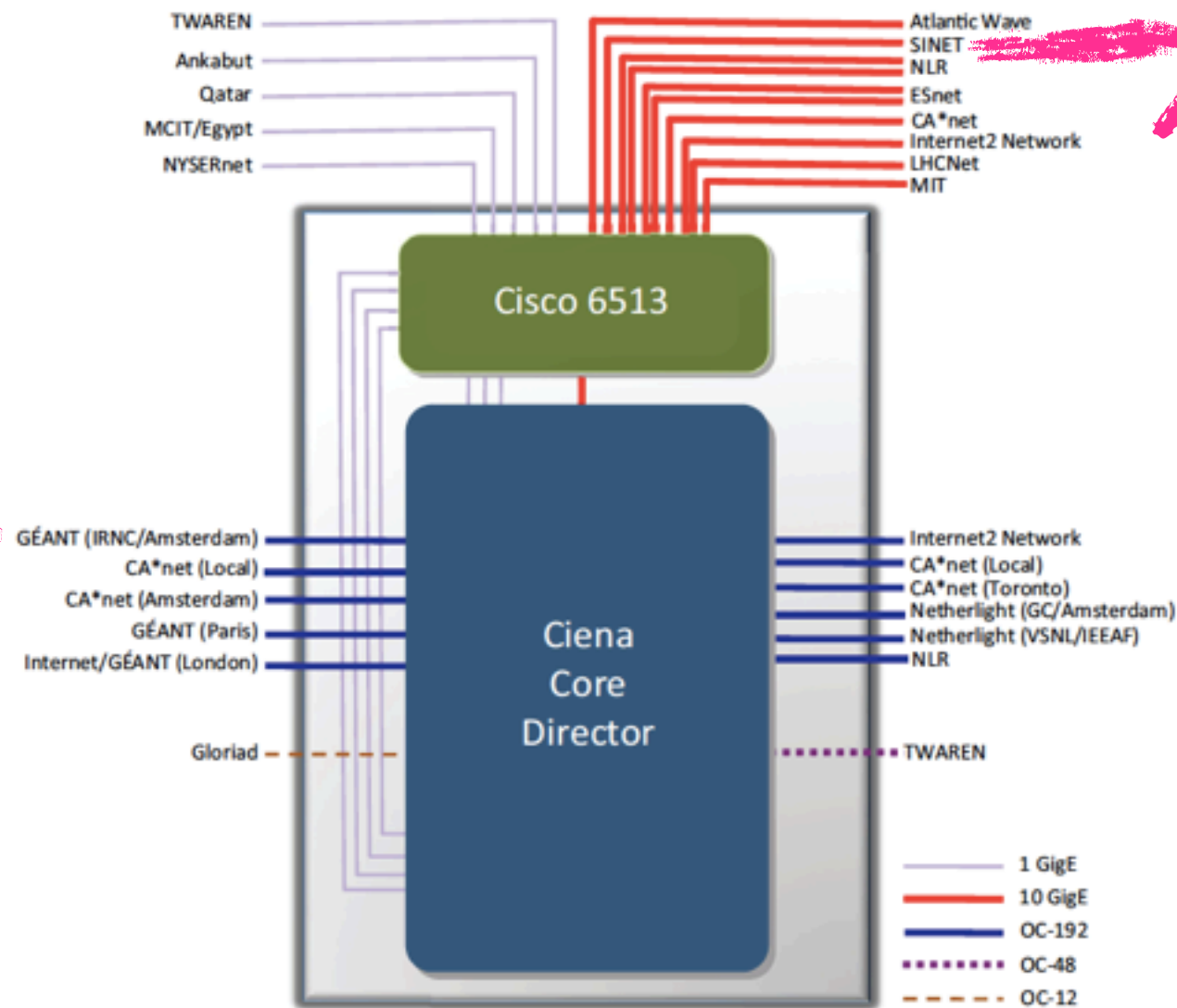
ideal view

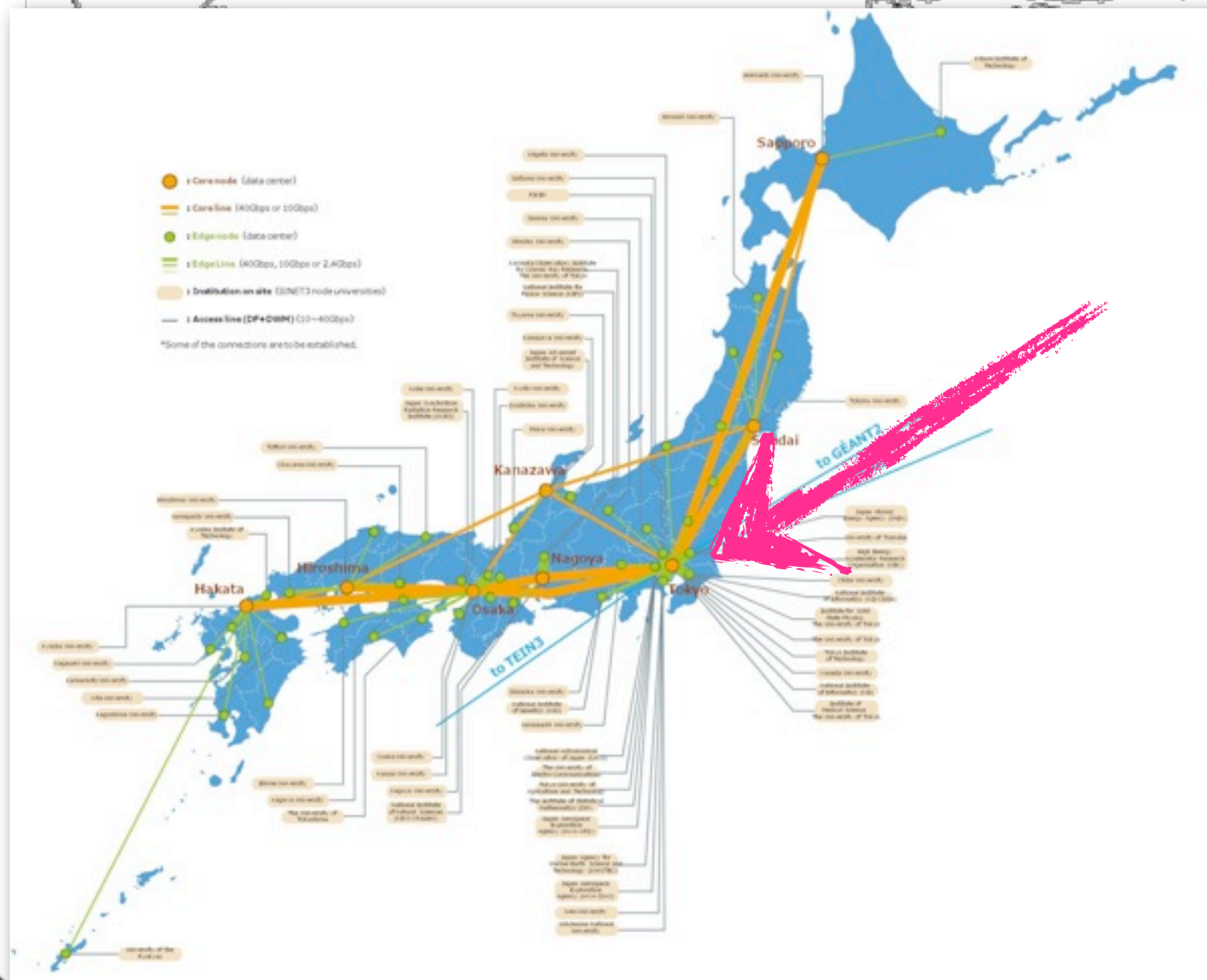
The reality ➡





MAN LAN TOPOLOGY

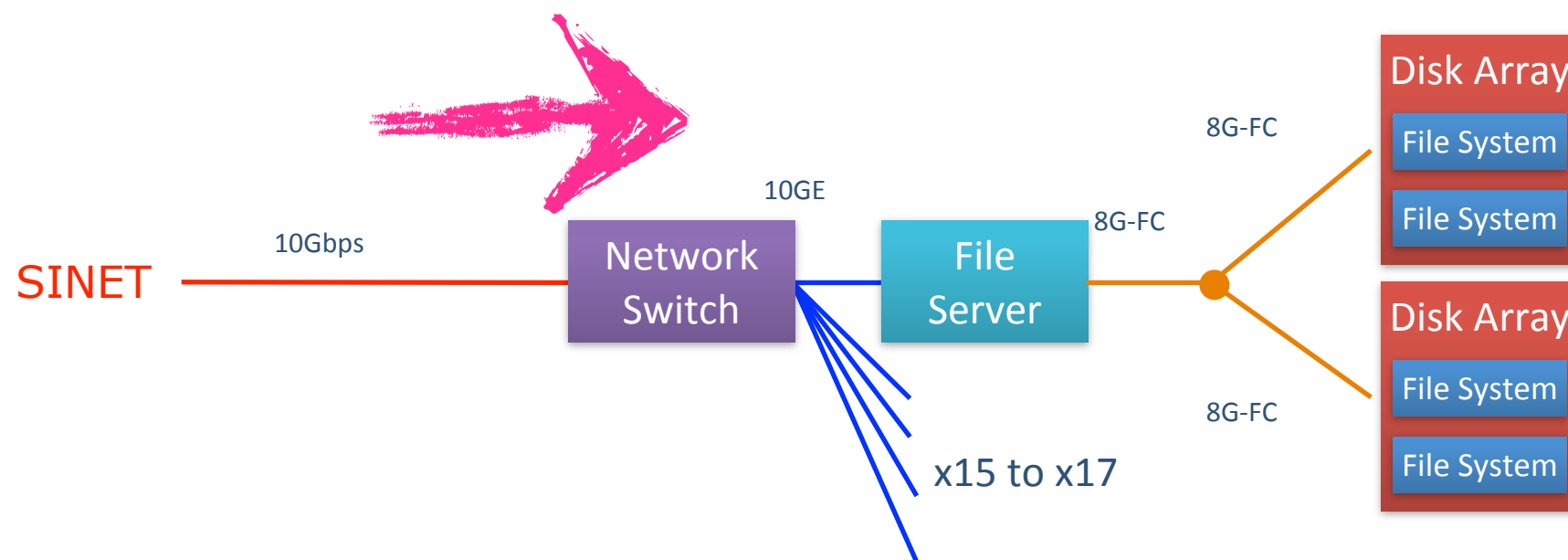




ICEPP

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

TOKYO-LCG2



- T2Ds: Tier2s directly connected to Tier1s of different clouds
 - “directly” in the DDM topology*ATLAS transfers from T2D to any T1*
- T2Ds should
 - demonstrate “good” connectivity from/to every T1/T2D
 - provide a certain level of commitment*Good connectivity*
- T2Ds could be de-commissioned if they degrade
 - Performances monitored*
- There is no “maximum number” of T2Ds
 - We should not create too many channels (FTS performance)

T2D: revising the criteria

Current criteria

- All transfers from the candidate T2D to 10/12 T1s for big files ('L') must be above 5 MB/s during the last week and during 3 out of the 4 last weeks.
- All transfers from 10/12 T1s to the candidate T2D for big files must be above 5 MB/s during the last week and during 3 out of the 4 last weeks

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>

- New T2D candidates
 - Sites/clouds should contact Central Operations if they want to be a candidate T2D
- Site will start receiving more Sonar Tests (large files from/to all T1s and T2Ds)
 - If the performance is acceptable, results are reported to the ADC Weekly Meeting and the site is declared a T2D
 - Need for FTS configuration of channels at T1s and this is done approx every 3 months
- So far, T2Ds candidates are monitored by squads
 - Cloud and sites themselves should take the necessary actions (monitoring, improving performance, reporting results to Central Operations)

perfSONAR(-PS) and ATLAS

Needed on every FR-cloud site

- Being deployed on T2D sites
- To measure latency and throughput (2 machines) between sites (matrix)
- Details in Shawn McKee's talk GDB
- Matrix monitoring available

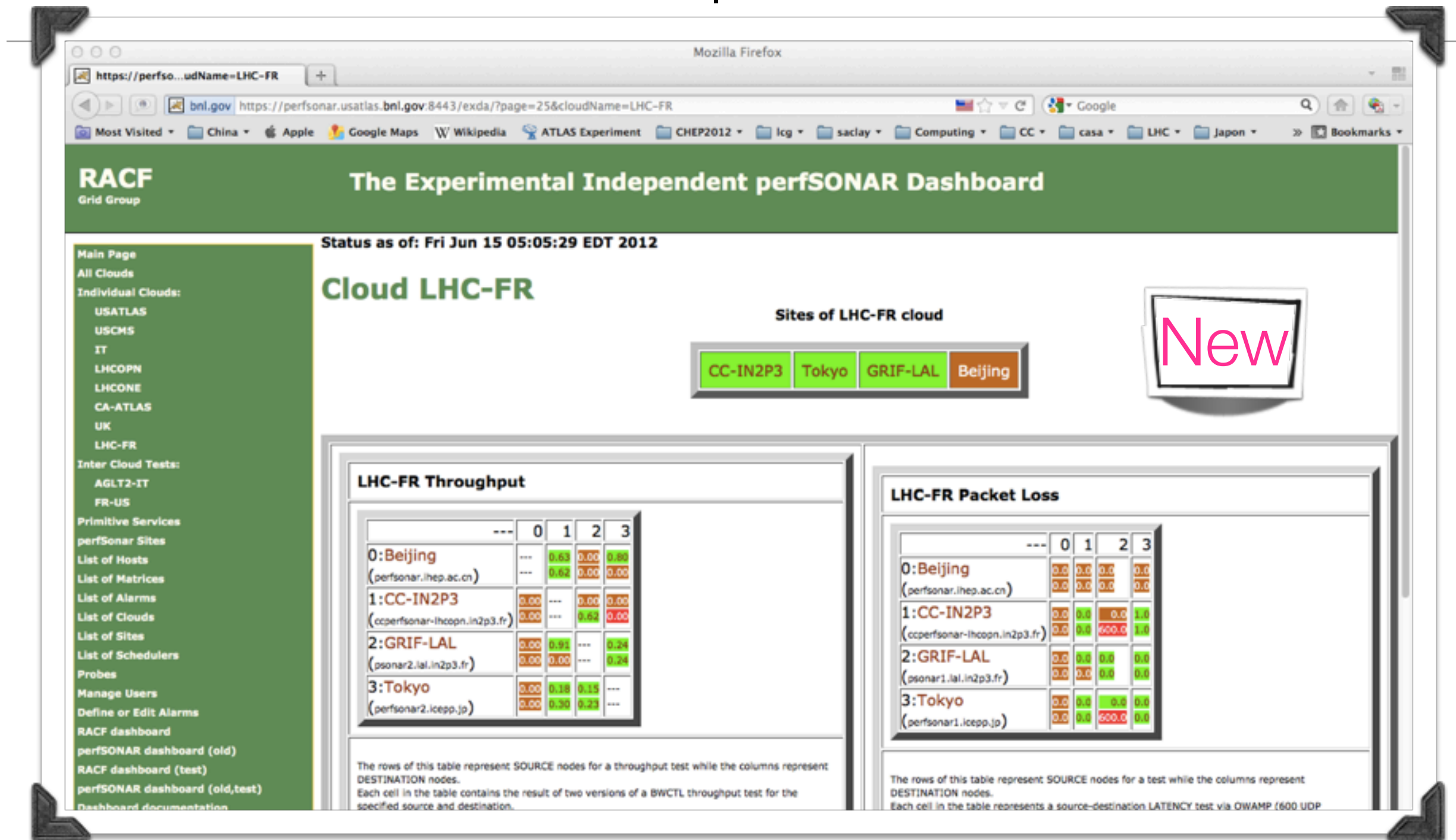


<http://psps.perfsonar.net/>

Summary for LHCONE

- ❄ Our specific goal in setting up perfSONAR-PS for LHCONE is to acquire before and after network measurements for the selected early adopter sites. This is **not** the long-term network monitoring setup for LHCONE...that is **TBD**
- ❄ Details of which sites and how sites should setup the perfSONAR-PS installations is documented on the Twiki at: <https://twiki.cern.ch/twiki/bin/view/LHCONE/SiteList>
- ❄ In the next few slides I will highlight some of the relevant details

FR-cloud view in ATLAS perfSonar dashboard



<https://perfsonar.usatlas.bnl.gov:8443/exda/?page=25&cloudName=LHC-FR>

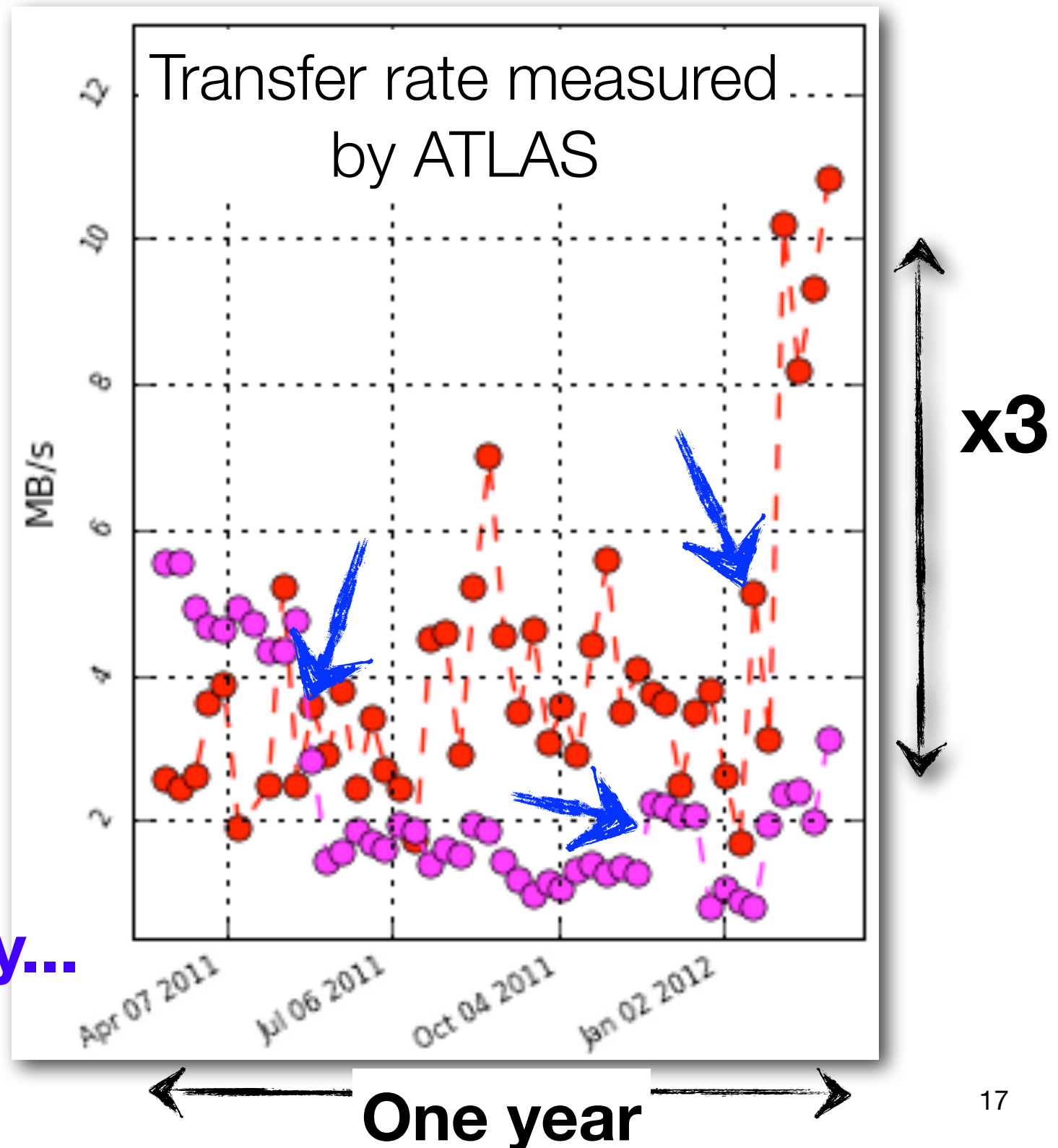
ATLAS transfers to/from Beijing over one year

Beijing → CCIN2P3

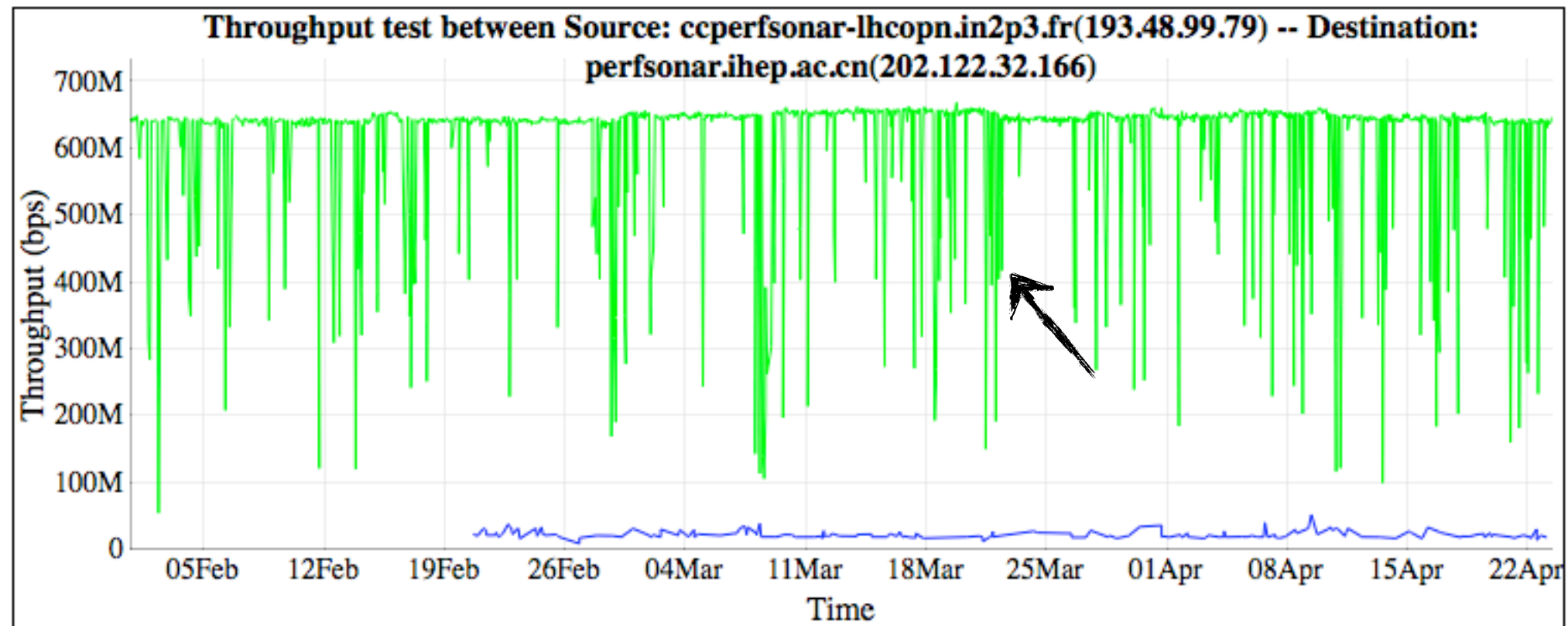
CCIN2P3 → Beijing

*asymmetry (why?) in
transfer rate
performance reversed
over last year*

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



Network throughput measured with perfsonar



<- 1 month

1 month ->

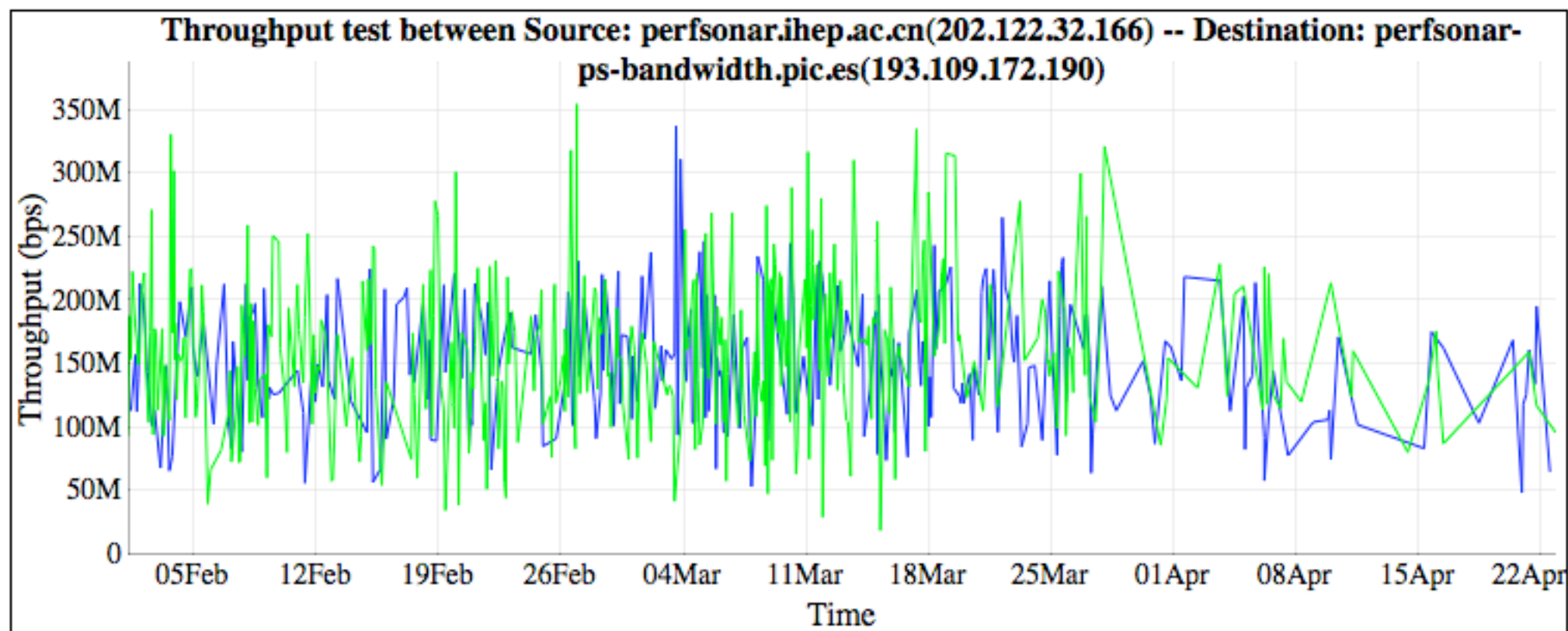
Timezone: GMT+0200 (CEST)

CCIN2P3 → Beijing

Beijing → CCIN2P3

Direction	Max throughput(bps)	Mean throughput(bps)	Min throughput(bps)
Src-Dst	52.76M	23.19M	9.49M
Dst-Src	669.33M	616.84M	55.11M

Beijing <-> PIC



<- 1 month

1 month ->

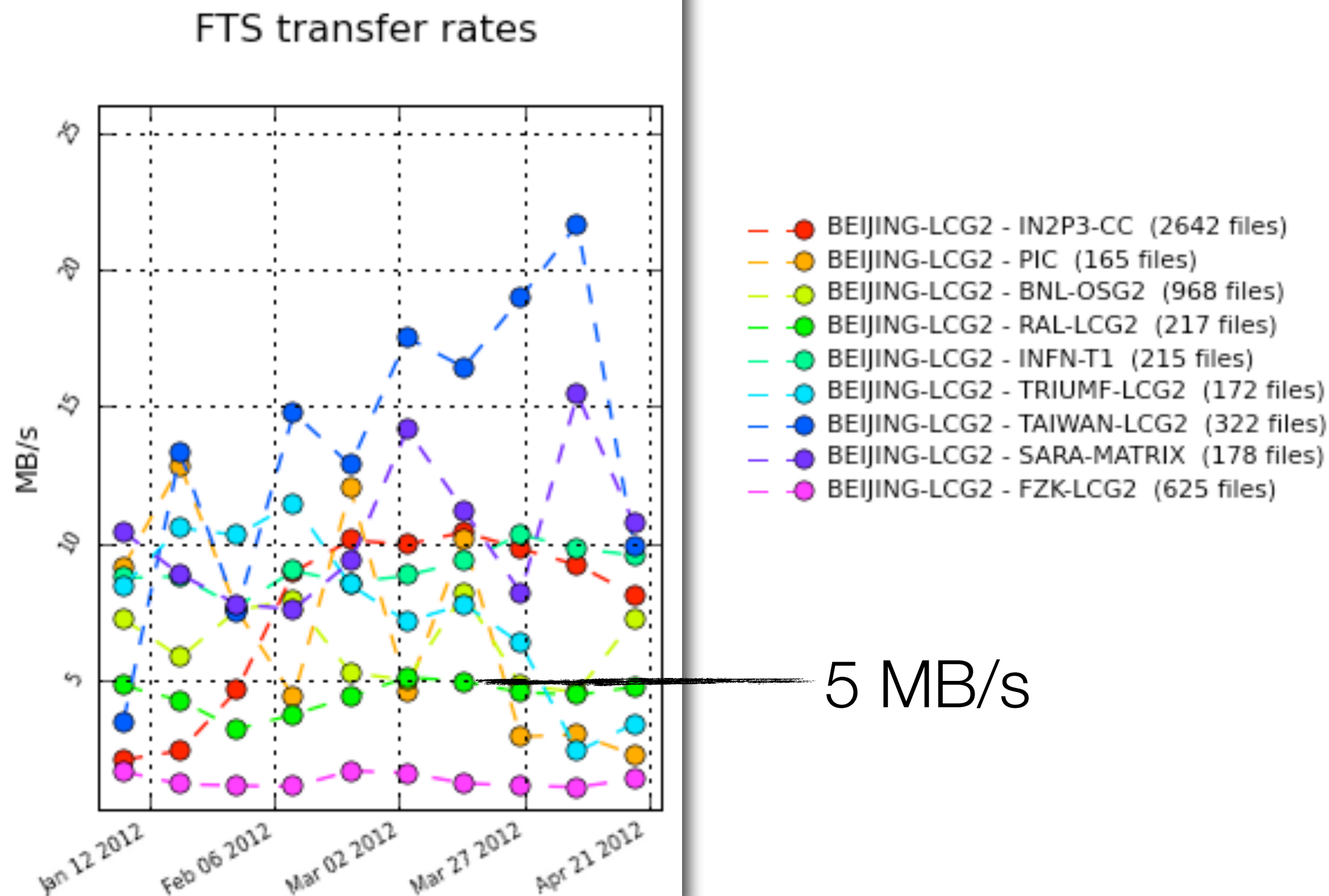
Timezone: GMT+0200 (CEST)

Beijing → PIC

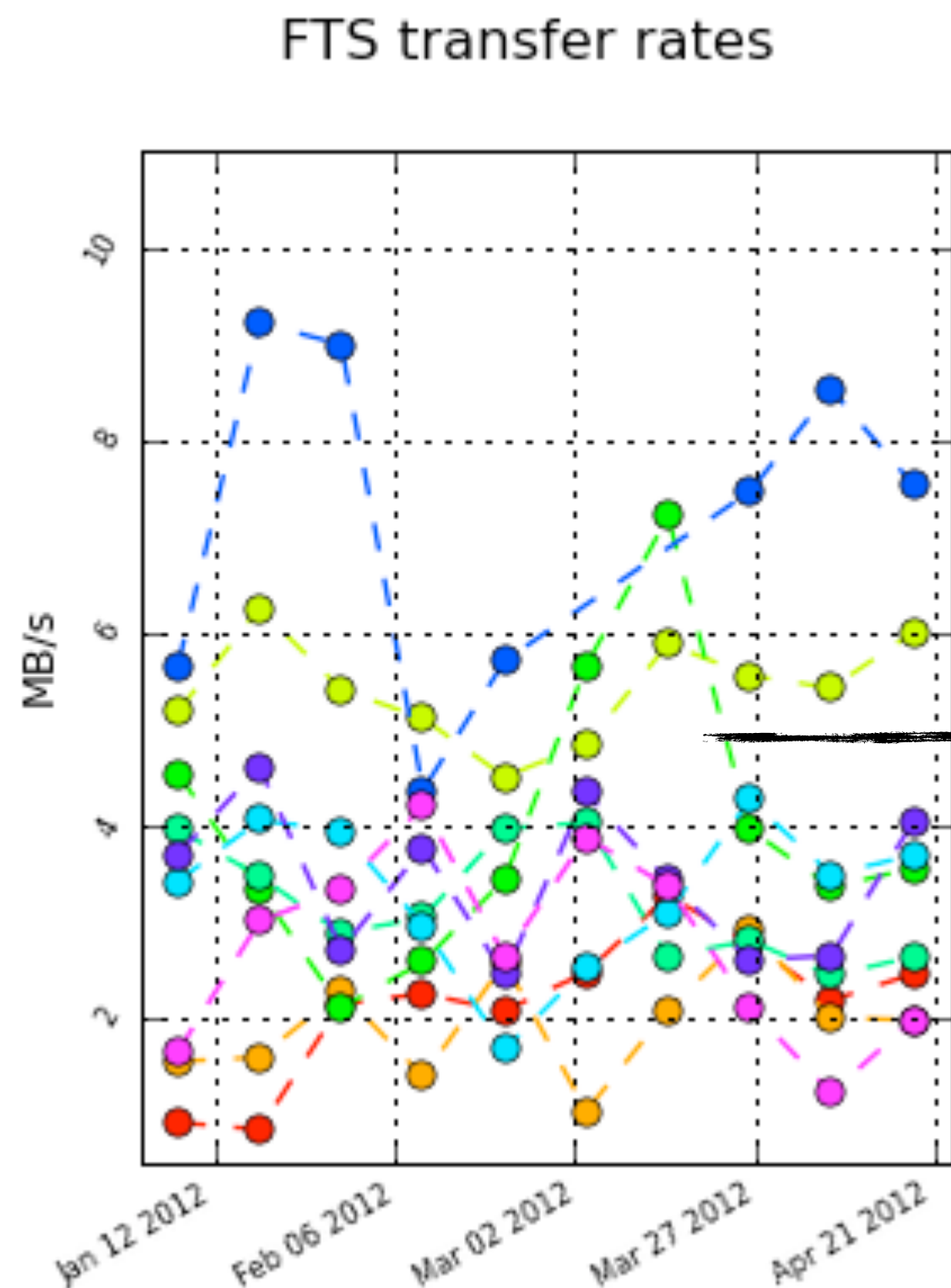
PIC → Beijing

Direction	Max throughput(bps)	Mean throughput(bps)	Min throughput(bps)
Src-Dst	337.62M	152.54M	49.1M
Dst-Src	354.95M	159.85M	19.23M

Beijing to T1s



T1s to Beijing



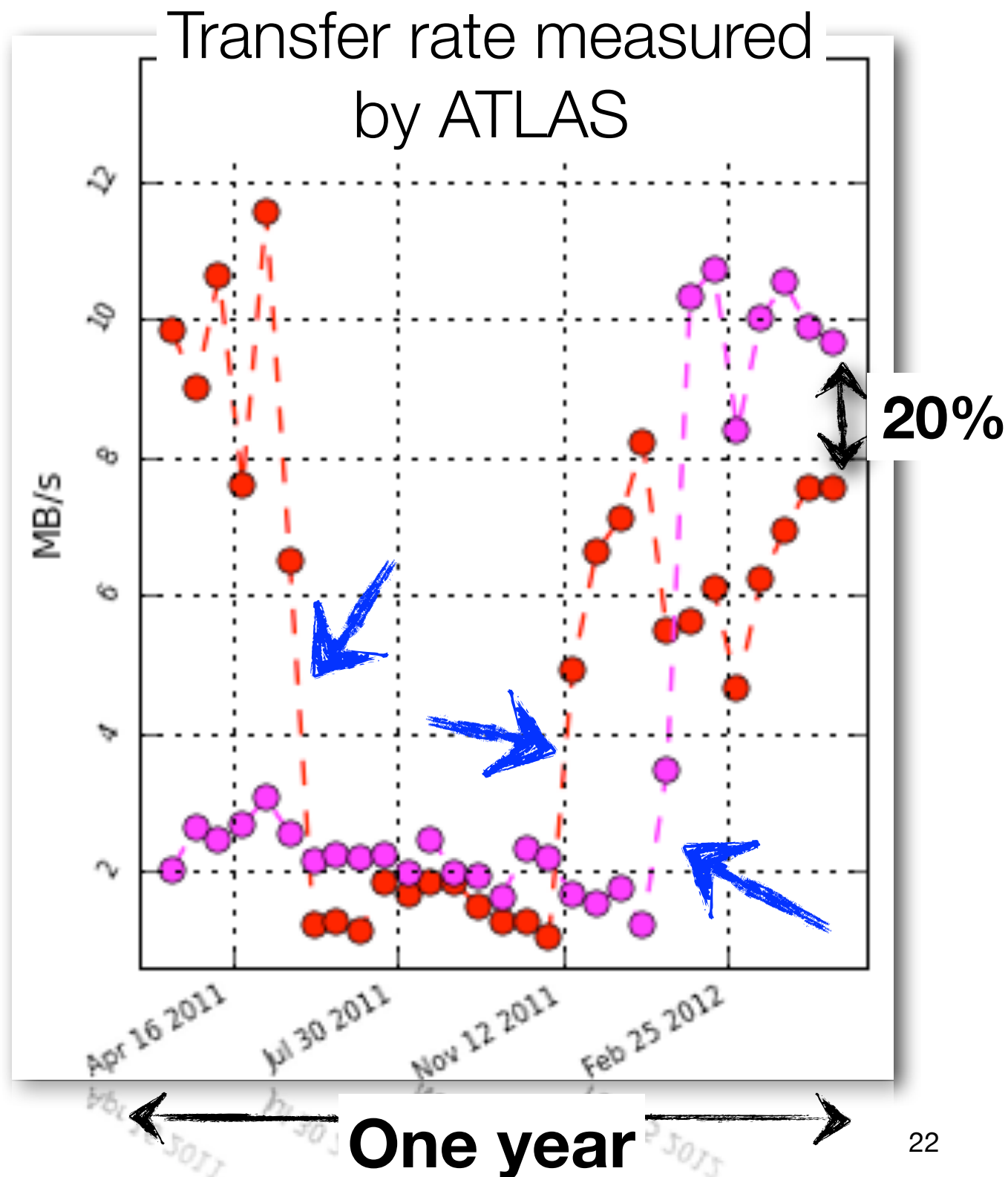
- IN2P3-CC - BEIJING-LCG2 (20899 files)
- PIC - BEIJING-LCG2 (1612 files)
- BNL-OSG2 - BEIJING-LCG2 (6152 files)
- RAL-LCG2 - BEIJING-LCG2 (4580 files)
- INFN-T1 - BEIJING-LCG2 (2619 files)
- TRIUMF-LCG2 - BEIJING-LCG2 (4510 files)
- TAIWAN-LCG2 - BEIJING-LCG2 (3193 files)
- SARA-MATRIX - BEIJING-LCG2 (2656 files)
- FZK-LCG2 - BEIJING-LCG2 (2790 files)

5 MB/s

each site is different

ATLAS transfers to/from Tokyo over one year

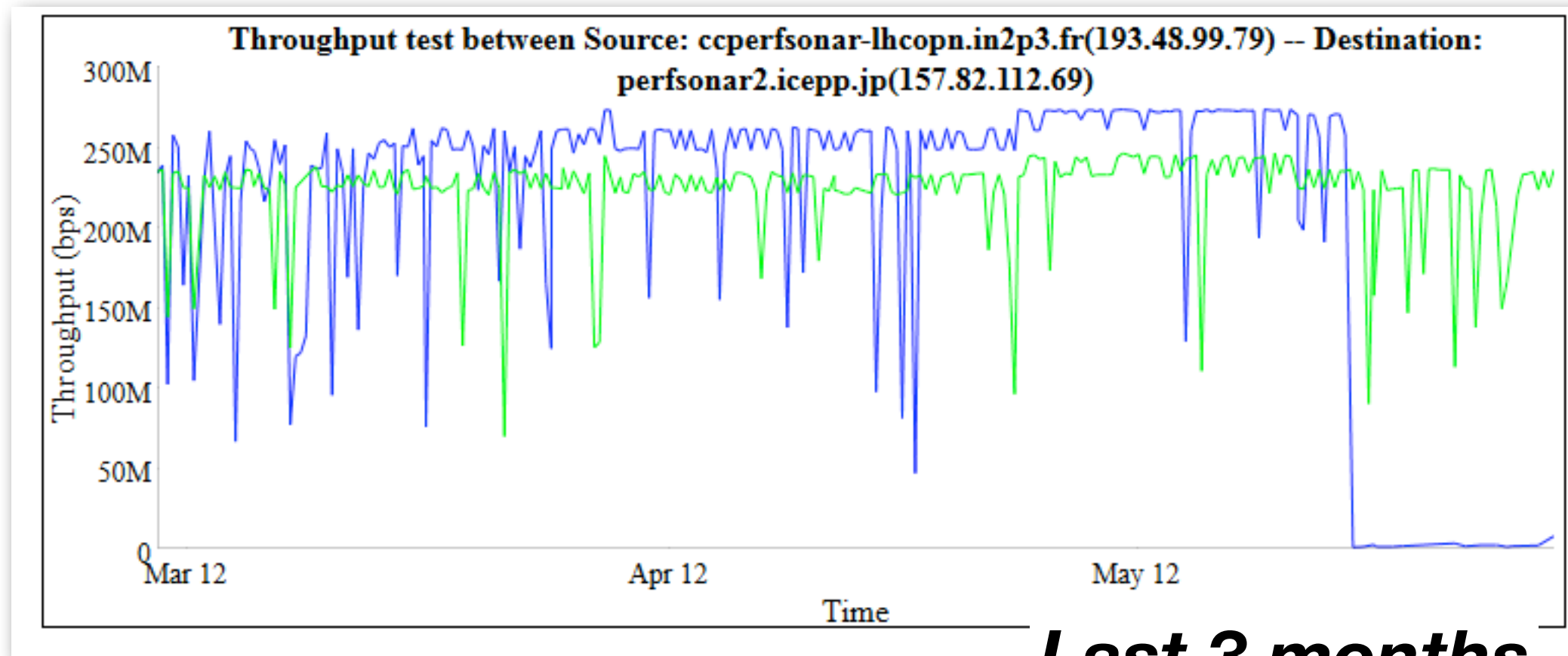
CCIN2P3 → Tokyo
Tokyo → CCIN2P3



Network throughput measured with perfSONAR

CCIN2P3 → Tokyo

Tokyo → CCIN2P3

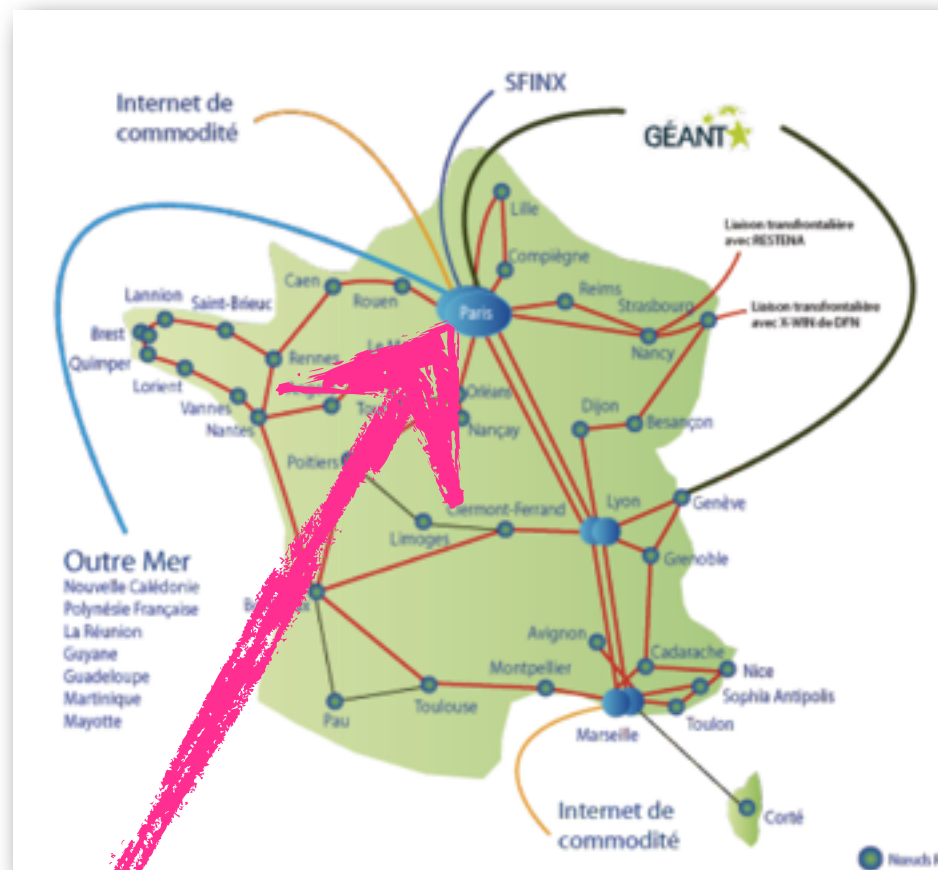


Last 3 months

*No so stable
better by ~5% for CCIN2P3 → Tokyo*

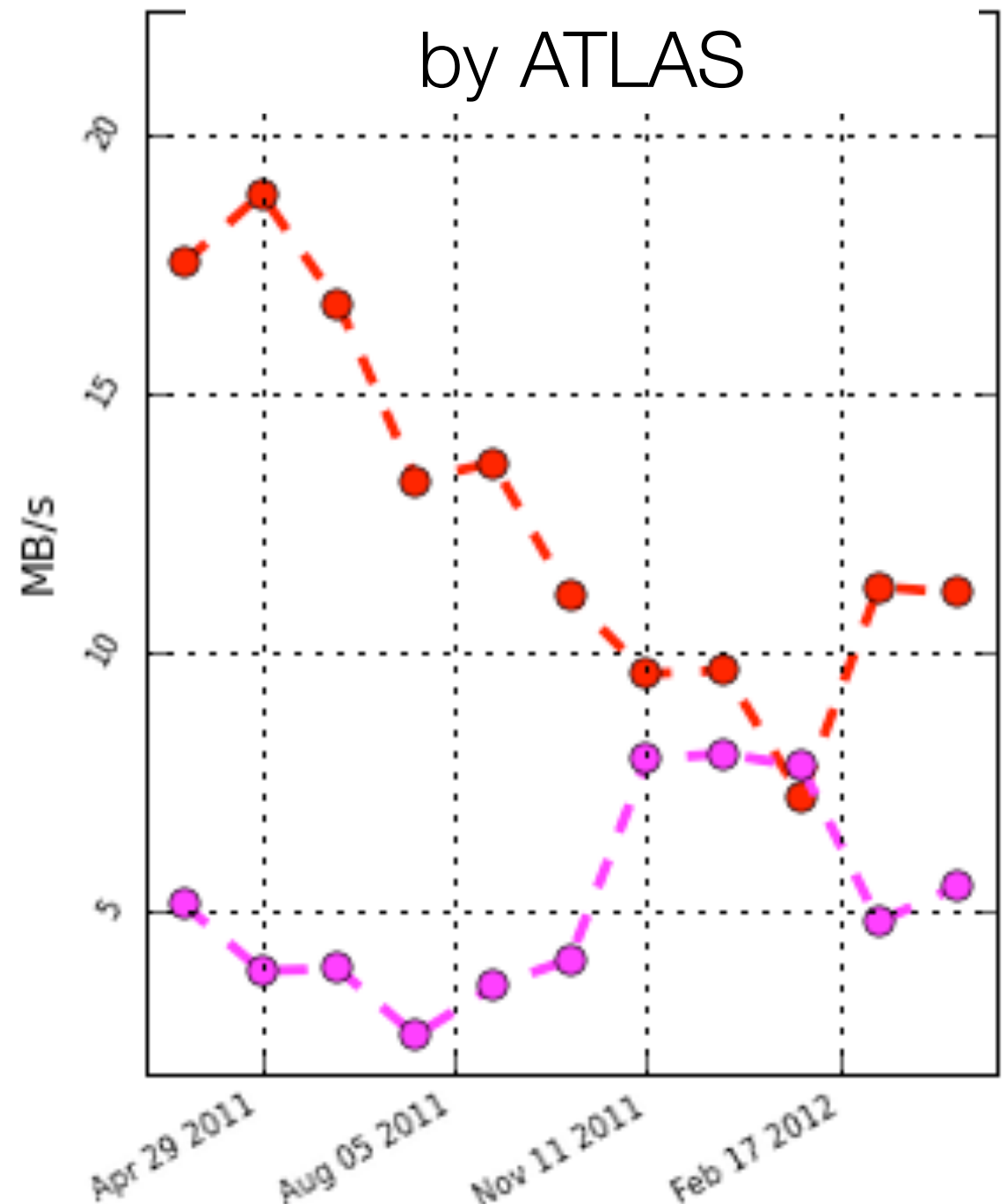
Different pattern for another French site

GRIF-LPNHE → Tokyo
Tokyo → GRIF-LPNHE

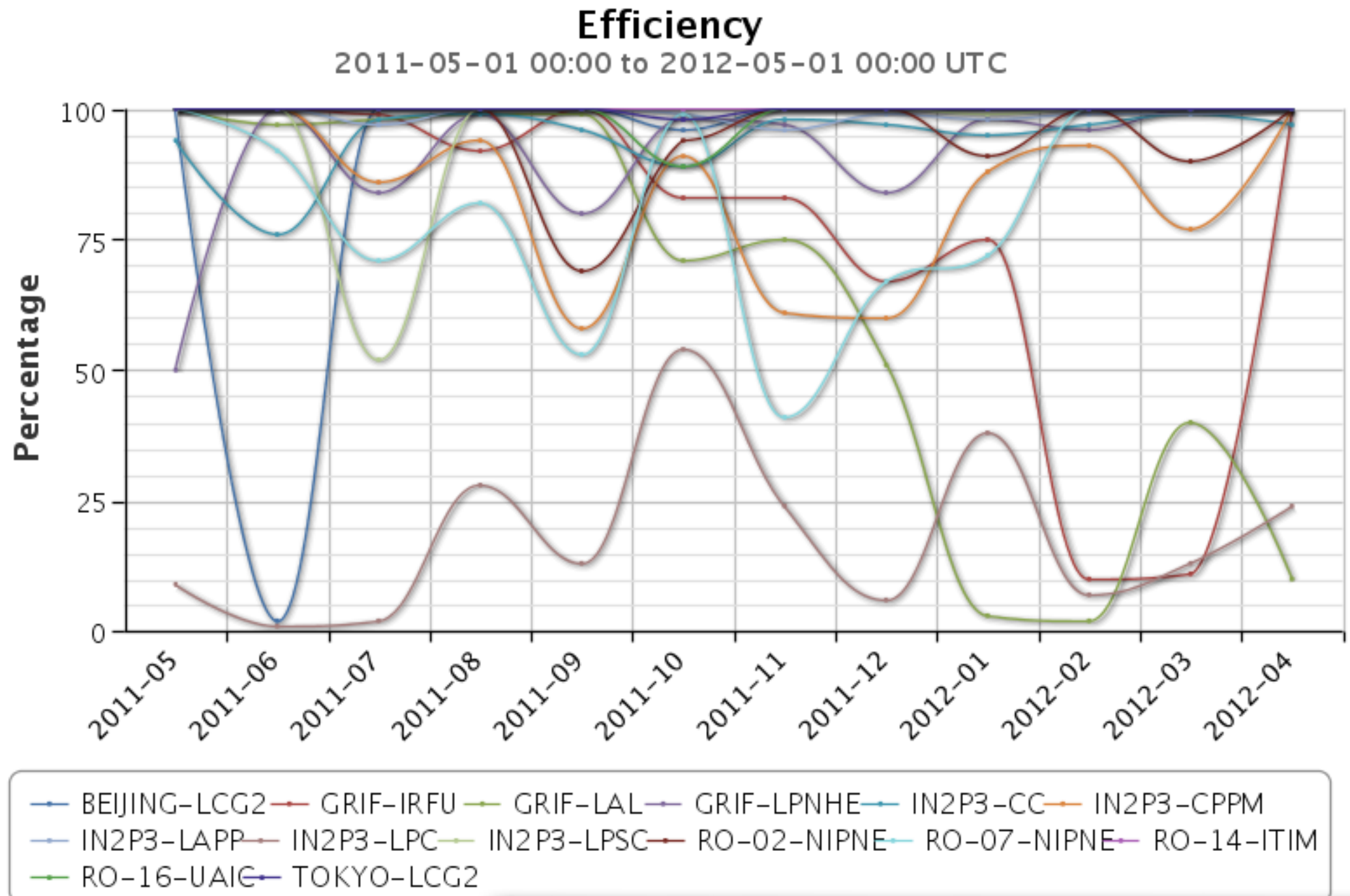


GRIF-LPNHE = Paris

Transfer rate measured
by ATLAS

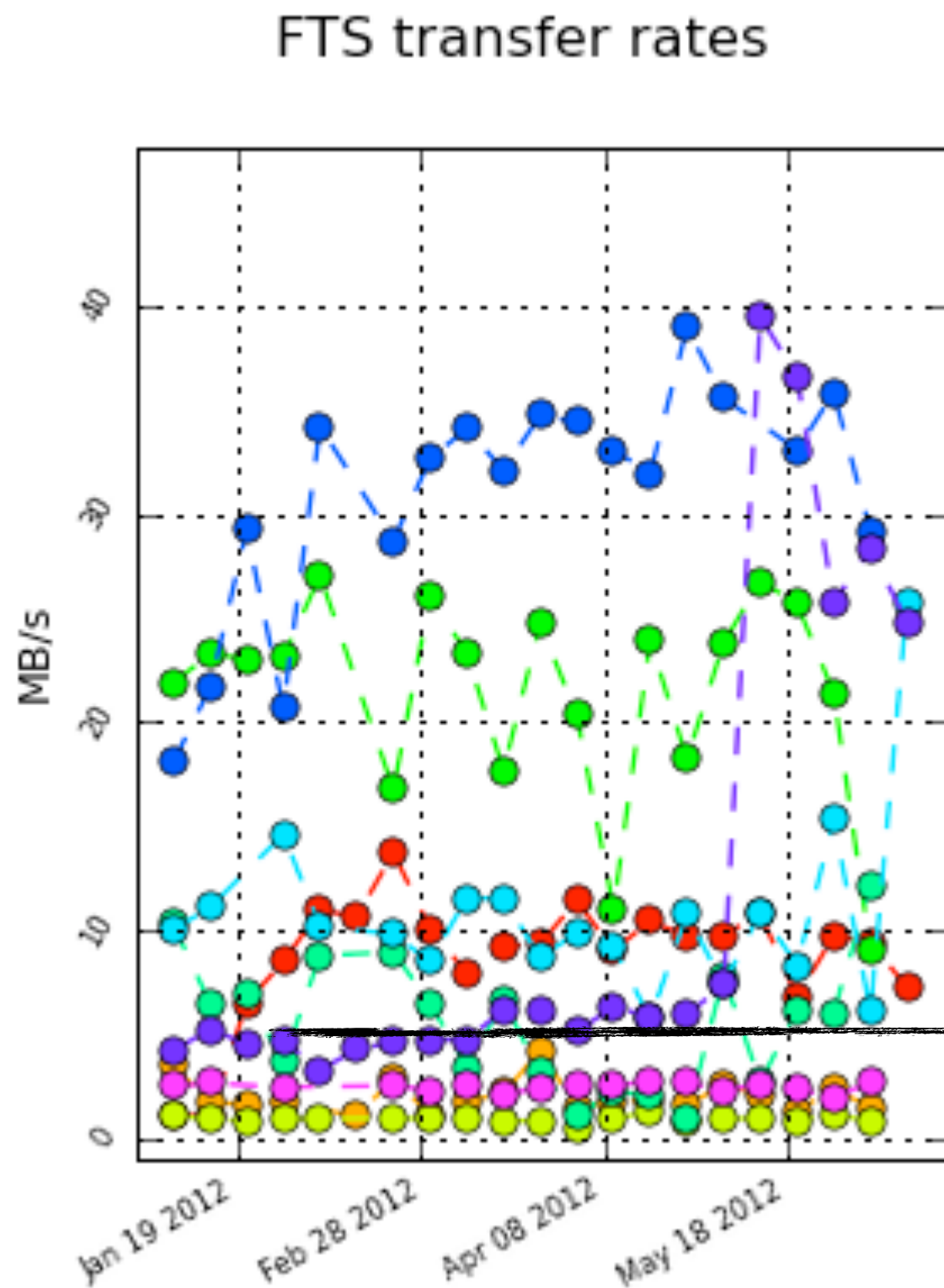


Efficiency of data transfer to Tokyo from FR-cloud sites



Each site is different, room for improvement everywhere...

Tokyo to T1s

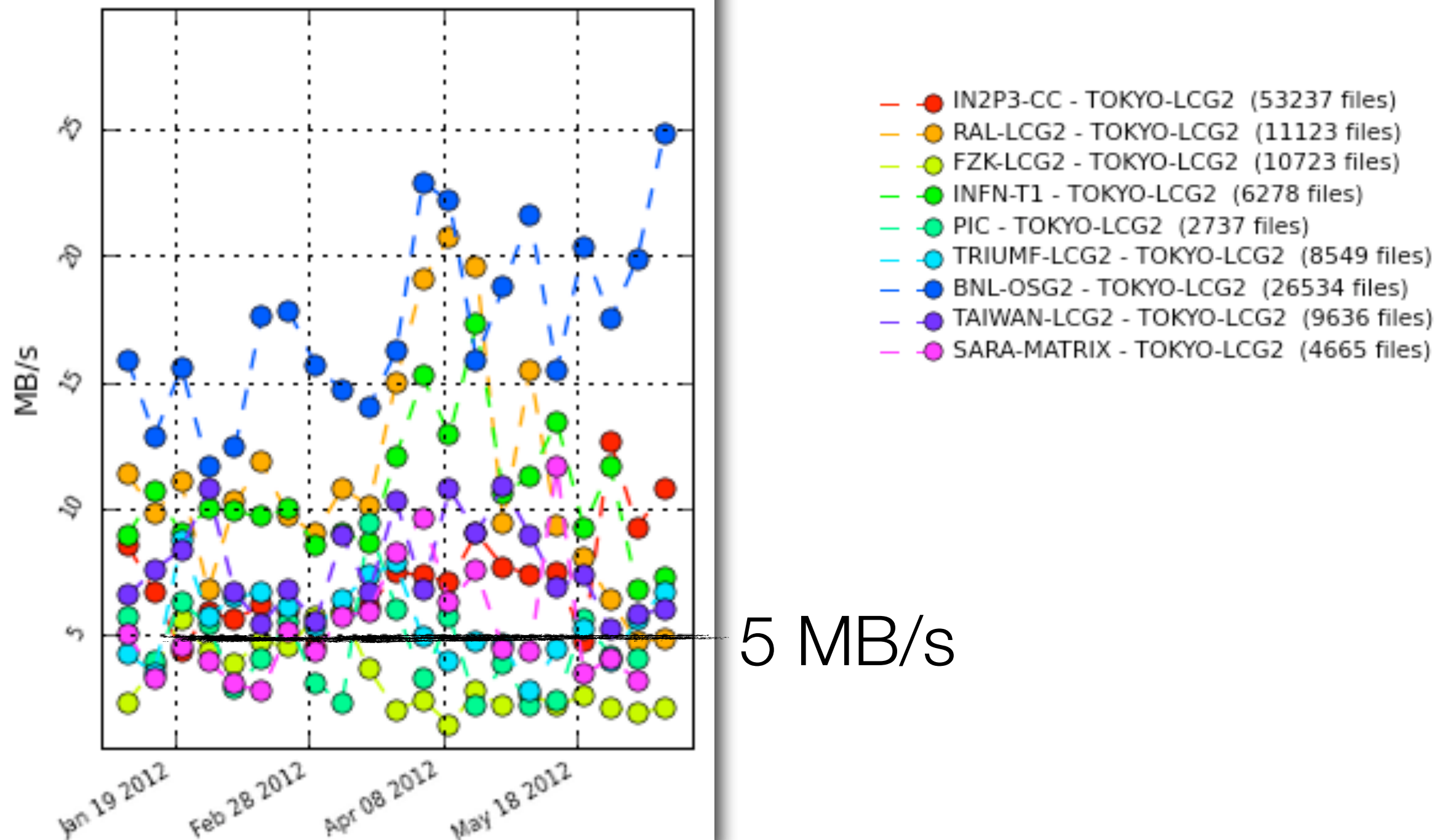


- TOKYO-LCG2 - IN2P3-CC (22368 files)
- TOKYO-LCG2 - RAL-LCG2 (107 files)
- TOKYO-LCG2 - FZK-LCG2 (110 files)
- TOKYO-LCG2 - INFN-T1 (107 files)
- TOKYO-LCG2 - PIC (110 files)
- TOKYO-LCG2 - TRIUMF-LCG2 (105 files)
- TOKYO-LCG2 - BNL-OSG2 (116 files)
- TOKYO-LCG2 - TAIWAN-LCG2 (4099 files)
- TOKYO-LCG2 - SARA-MATRIX (100 files)

5 MB/s

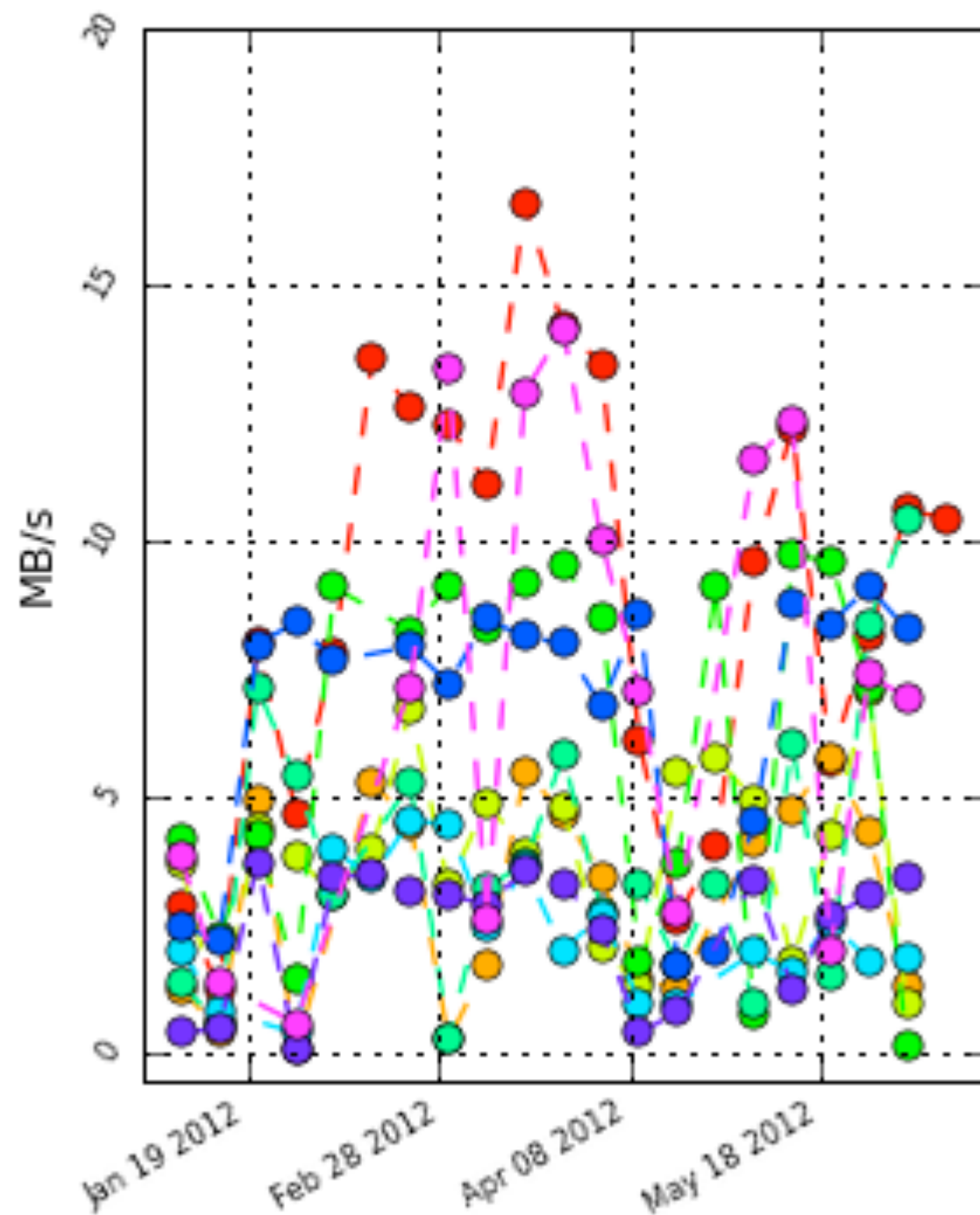
T1s to Tokyo

FTS transfer rates



RO-02 to T1s

FTS transfer rates

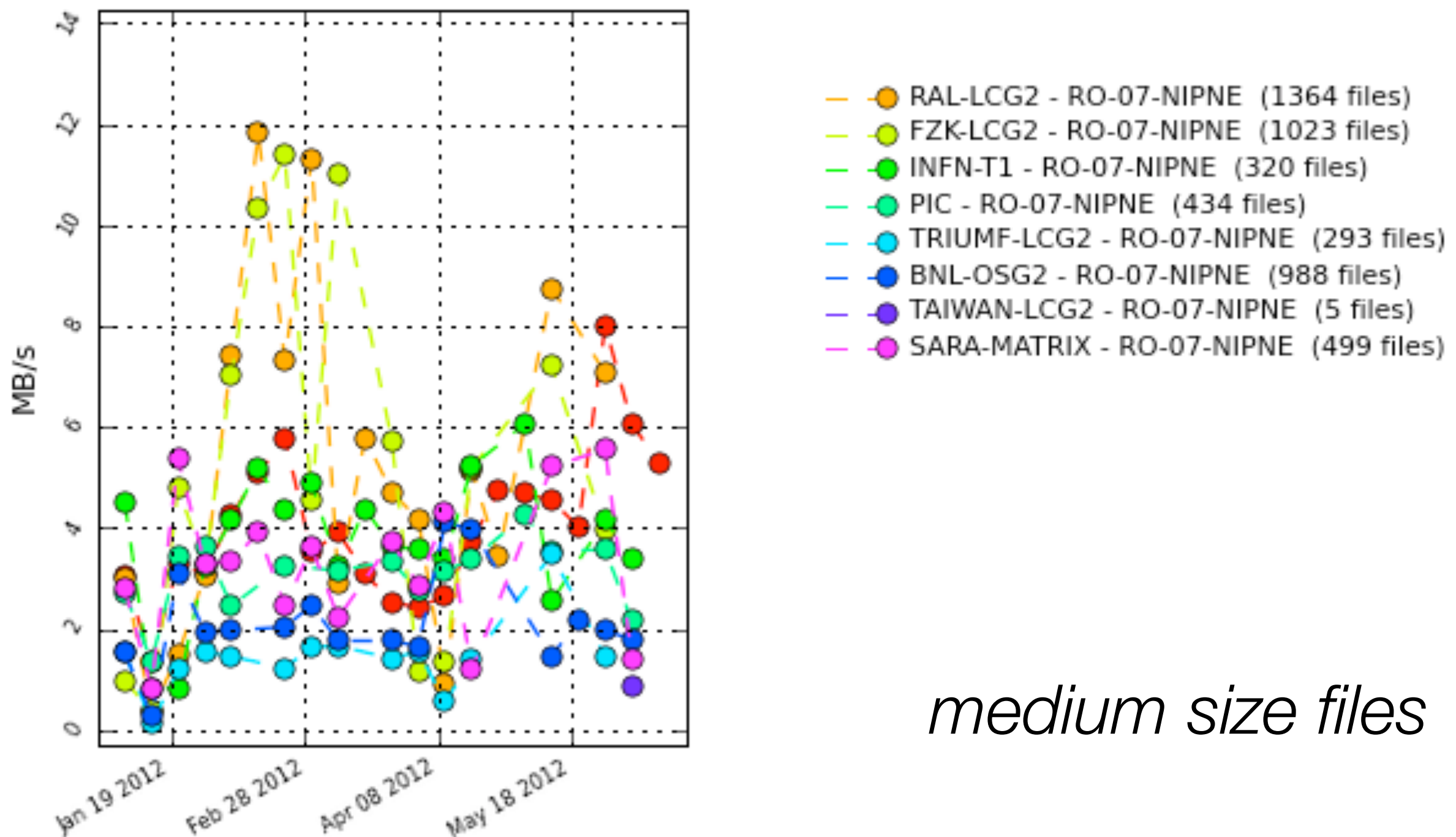


- RO-07-NIPNE - IN2P3-CC (43338 files)
- RO-07-NIPNE - RAL-LCG2 (100 files)
- RO-07-NIPNE - FZK-LCG2 (210 files)
- RO-07-NIPNE - INFN-T1 (110 files)
- RO-07-NIPNE - PIC (105 files)
- RO-07-NIPNE - TRIUMF-LCG2 (100 files)
- RO-07-NIPNE - BNL-OSG2 (110 files)
- RO-07-NIPNE - TAIWAN-LCG2 (252 files)
- RO-07-NIPNE - SARA-MATRIX (95 files)

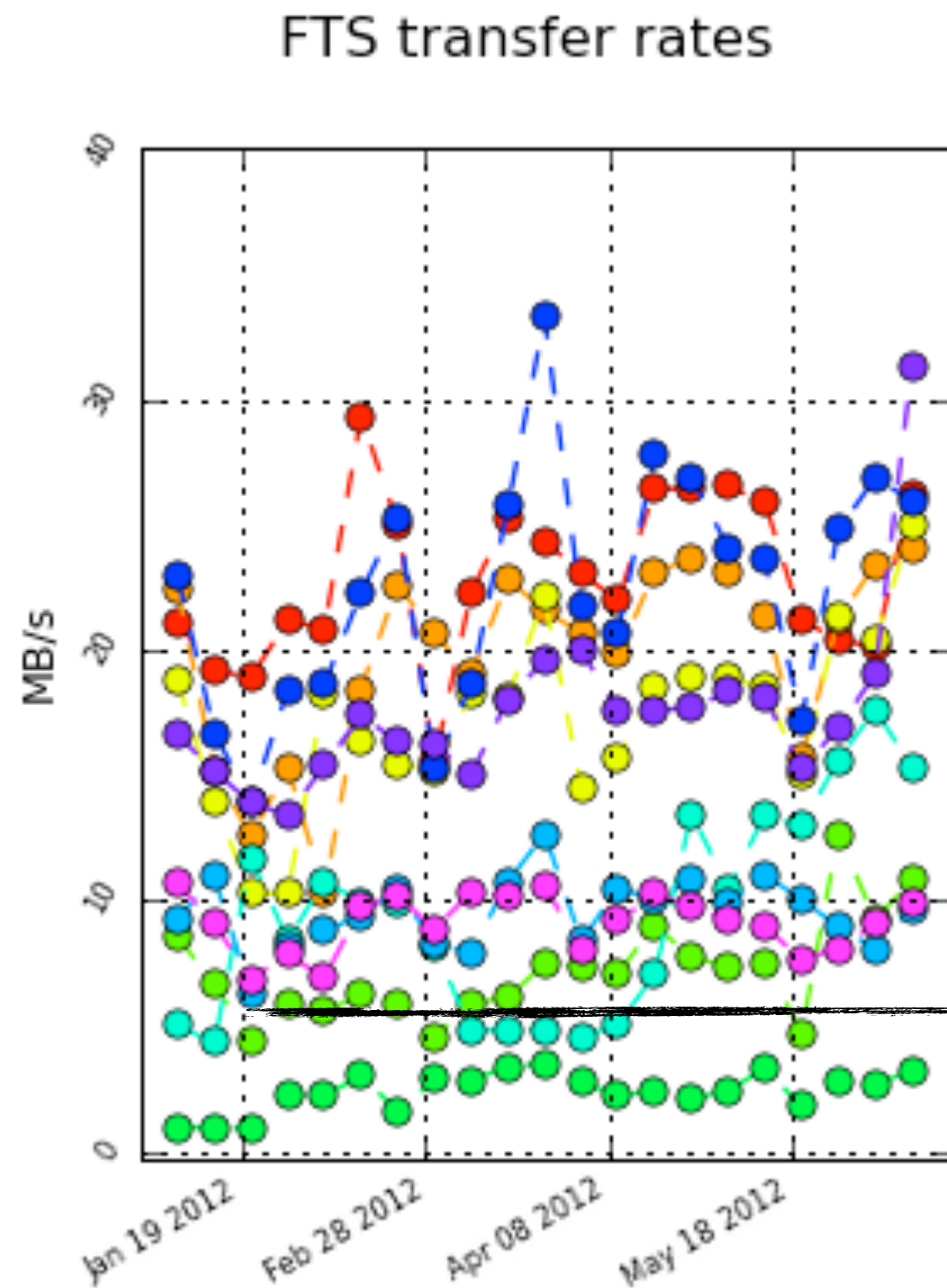
medium size files

T1s to RO-02

FTS transfer rates



CCIN2P3 to FR-cloud T2s

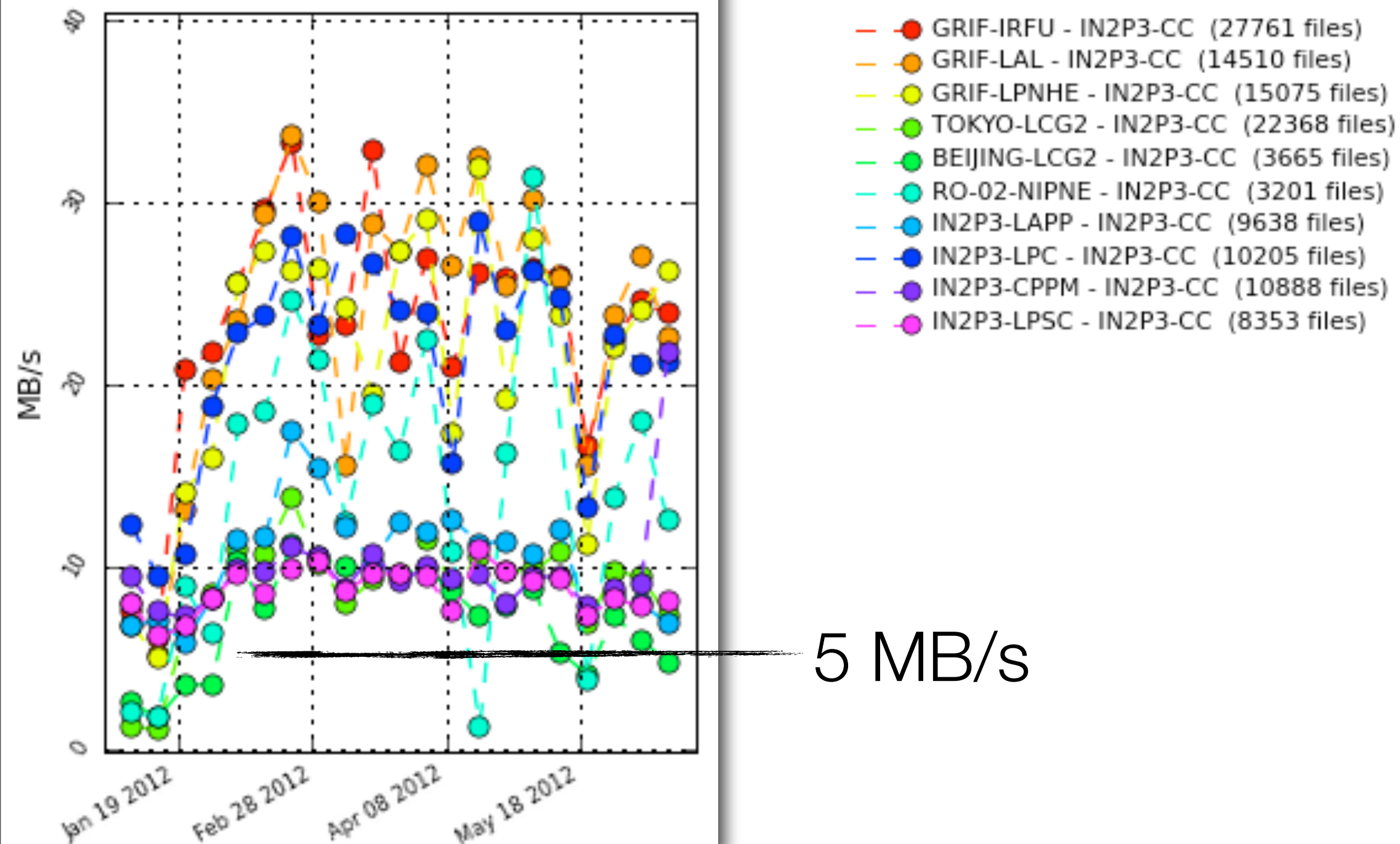


- IN2P3-CC - GRIF-IRFU (175713 files)
- IN2P3-CC - GRIF-LAL (38934 files)
- IN2P3-CC - GRIF-LPNHE (45210 files)
- IN2P3-CC - TOKYO-LCG2 (53237 files)
- IN2P3-CC - BEIJING-LCG2 (33933 files)
- IN2P3-CC - RO-02-NIPNE (19737 files)
- IN2P3-CC - IN2P3-LAPP (36288 files)
- IN2P3-CC - IN2P3-LPC (33911 files)
- IN2P3-CC - IN2P3-CPPM (84984 files)
- IN2P3-CC - IN2P3-LPSC (44030 files)

5 MB/s

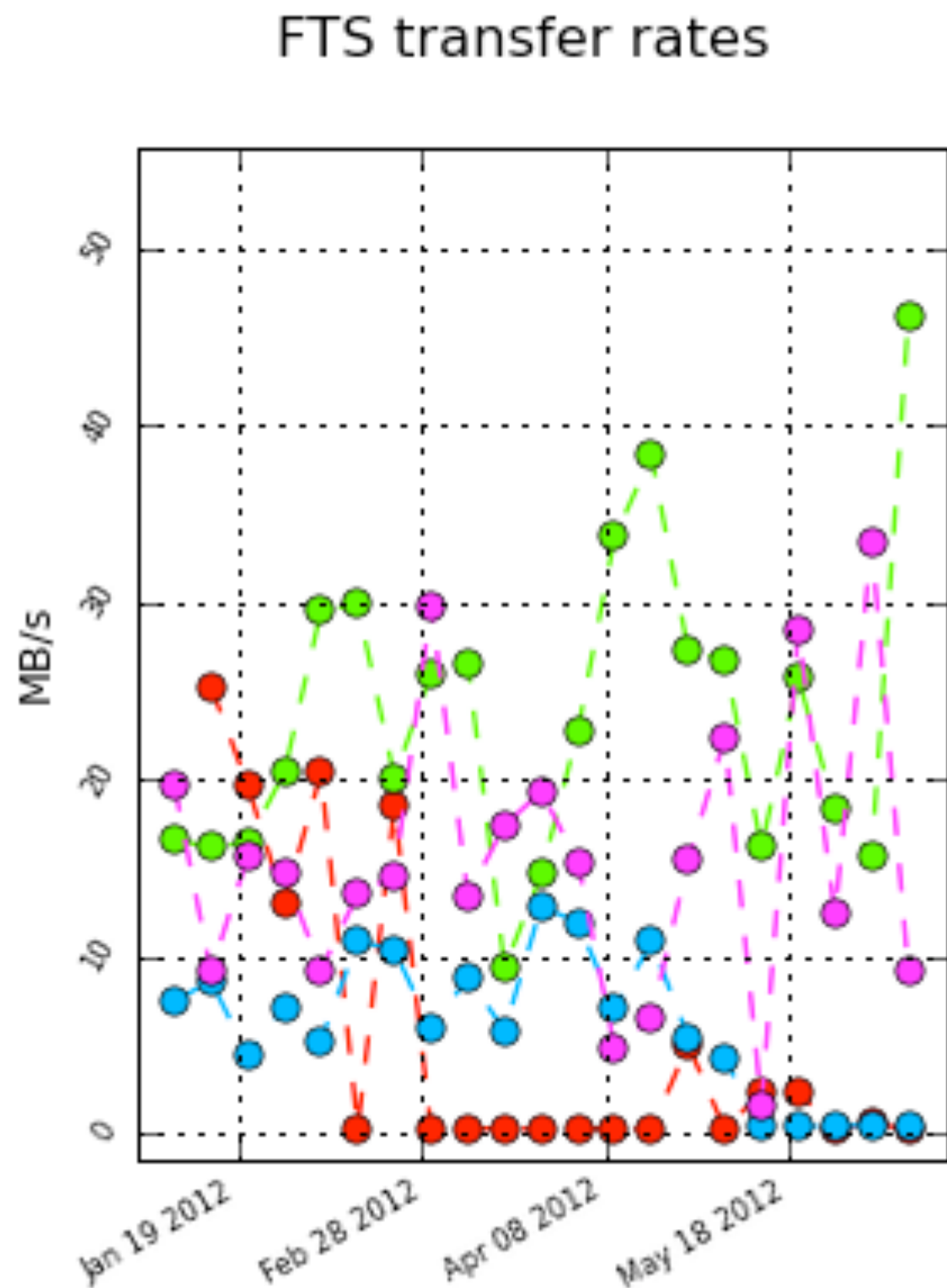
FR-cloud T2s to CCIN2P3

FTS transfer rates



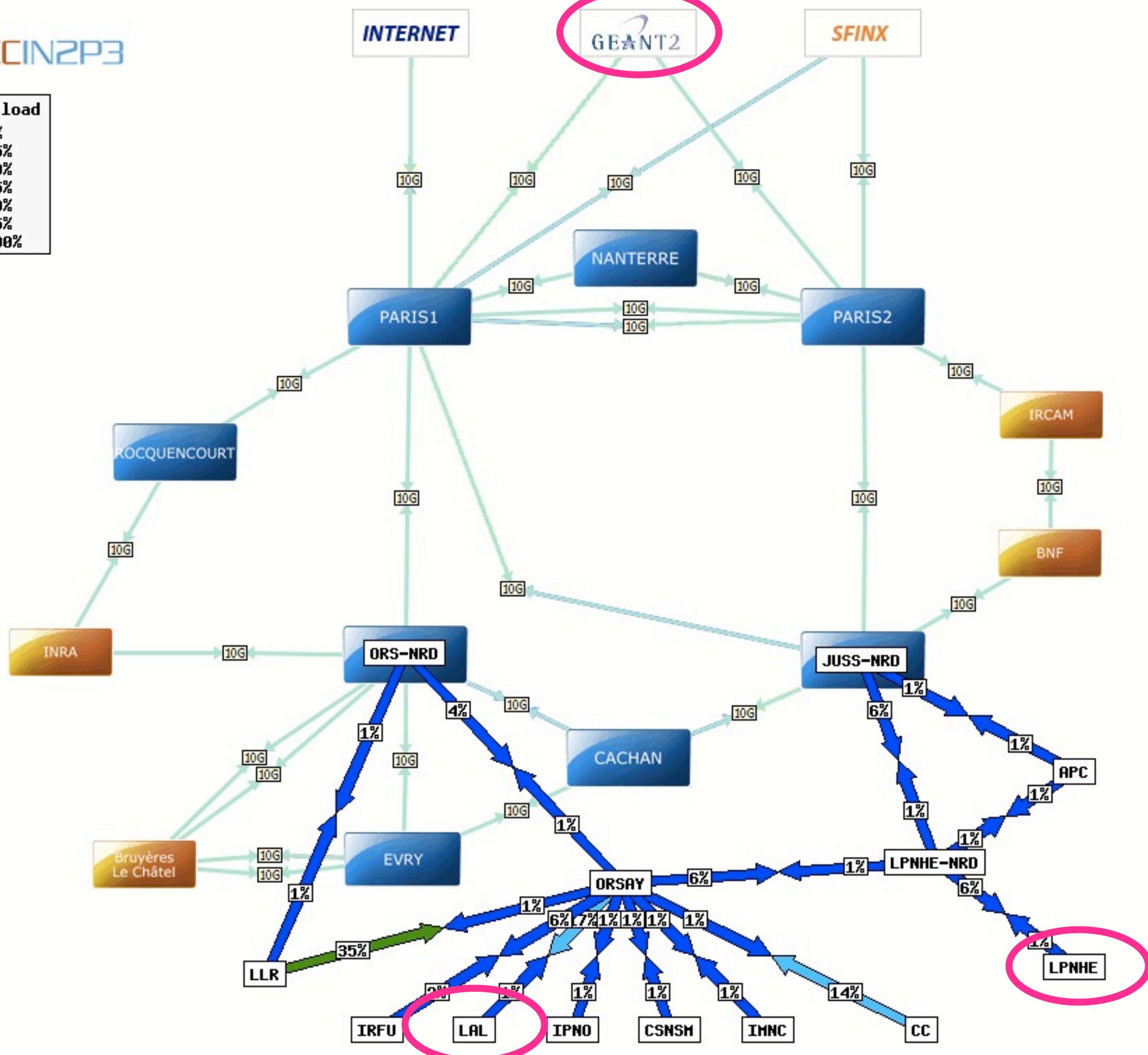
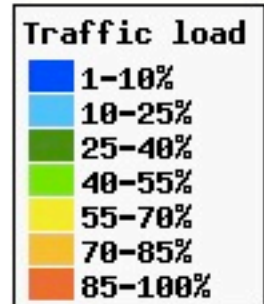
5 MB/s

GRIF-LAL vs GRIF-LPNHE

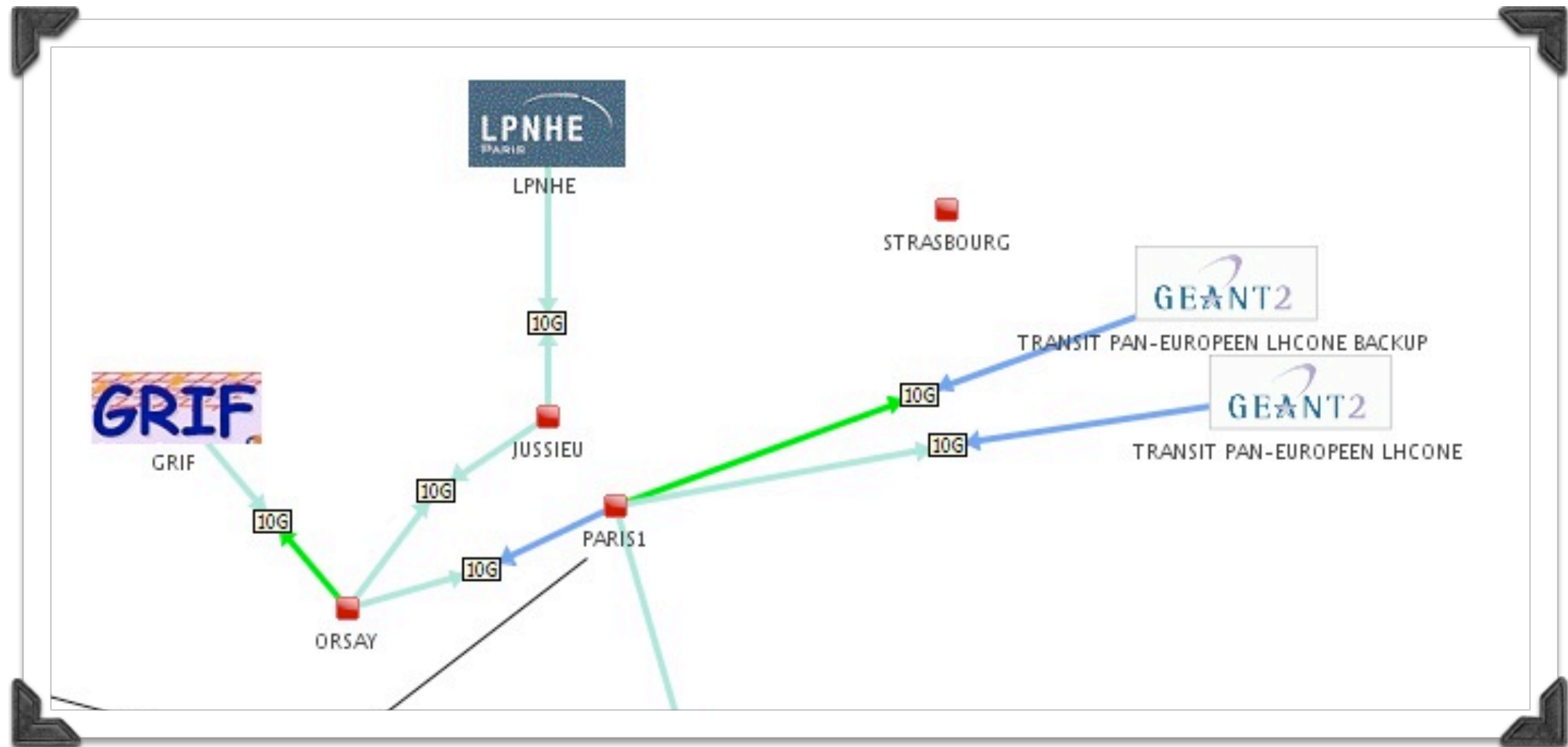


- GRIF-LAL - BNL-OSG2 (748 files)
- GRIF-LPNHE - BNL-OSG2 (7745 files)
- GRIF-LAL - TAIWAN-LCG2 (629 files)
- GRIF-LPNHE - TAIWAN-LCG2 (1339 files)

LAL connectivity issues

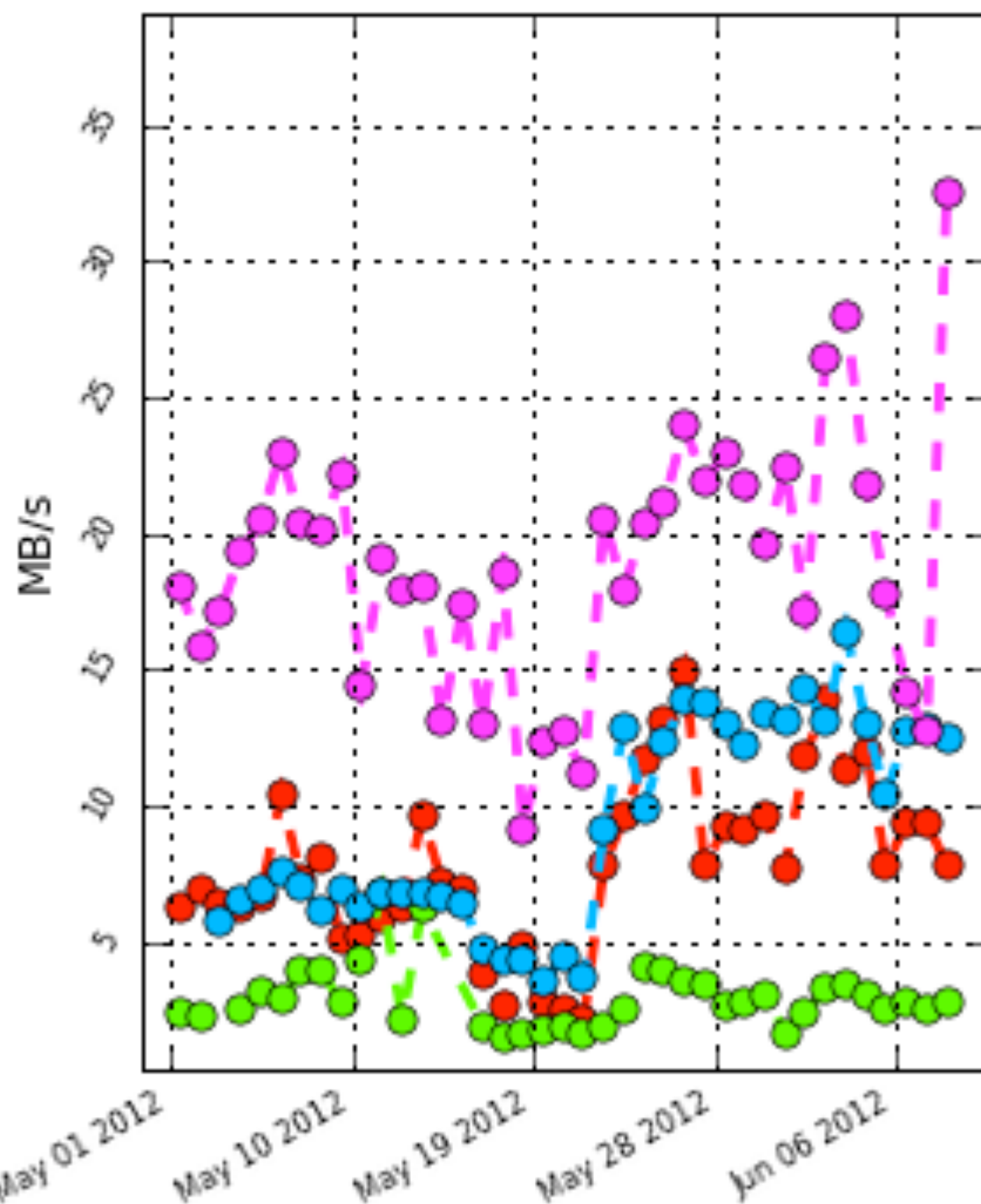


is this LHCONE map correct?



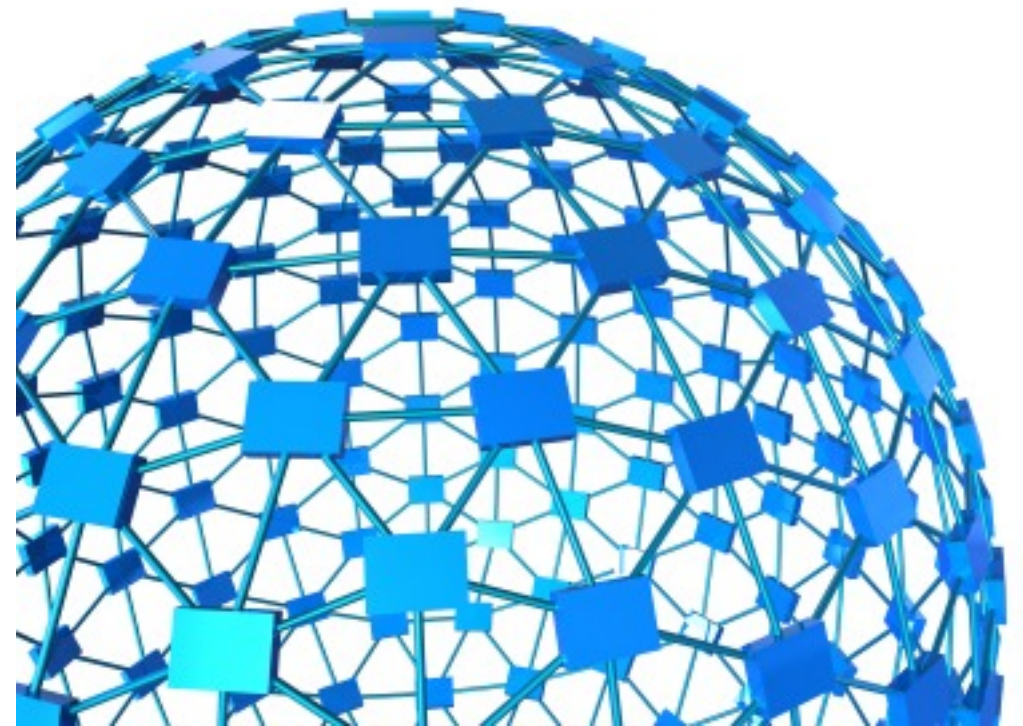
May 23 effect

FTS transfer rates

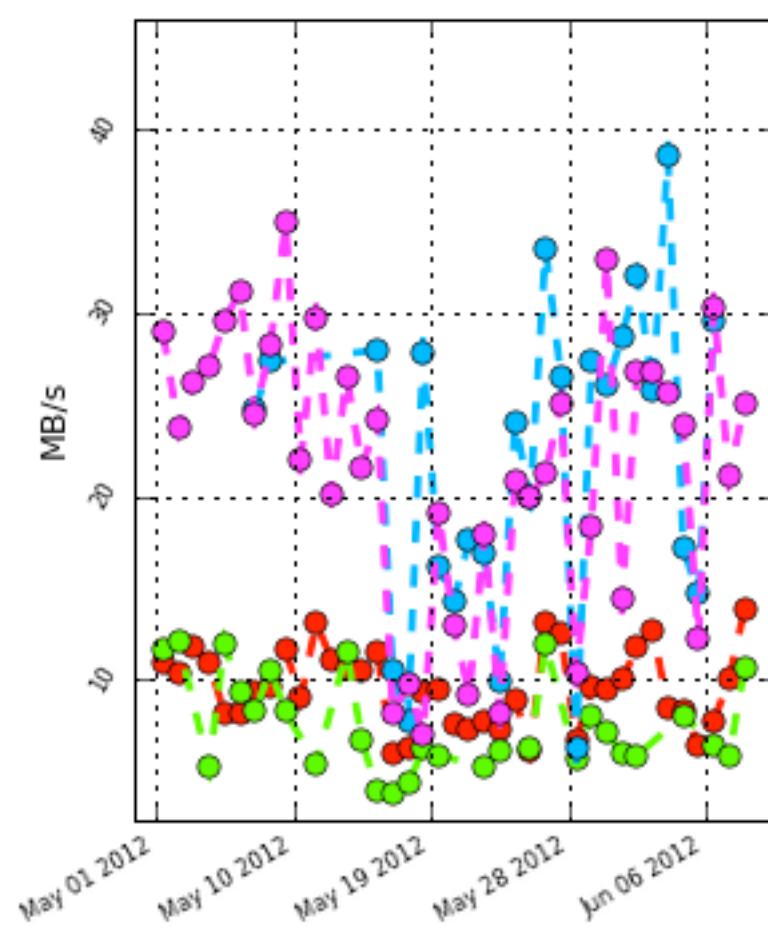


- IN2P3-CC - TOKYO-LCG2 (18218 files)
- IN2P3-CC - BEIJING-LCG2 (8263 files)
- IN2P3-CC - RO-07-NIPNE (20831 files)
- IN2P3-CC - GRIF-LPNHE (10619 files)

The end....



FTS transfer rates



- TOKYO-LCG2 - IN2P3-CC (3978 files)
- BEIJING-LCG2 - IN2P3-CC (542 files)
- RO-07-NIPNE - IN2P3-CC (821 files)
- GRIF-LPNHE - IN2P3-CC (4999 files)