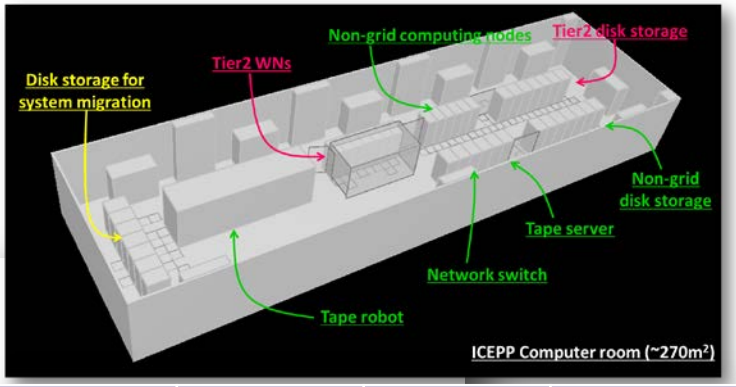
Report from the Tokyo Tier-2

Tomoaki Nakamura
KEK-CRC / ICEPP UTokyo

ICEPP regional analysis center



Resource overview

Support only ATLAS VO in WLCG as Tier2.
Provide ATLAS-Japan dedicated resource for analysis.
The first production system for WLCG was deployed in 2007.
Almost of hardware are prepared by three years lease.
System have been upgraded in every three years.
Current system is the 3rd generation system.
~10,000 CPU cores and 6.7PB disk storage (T2 + T3).

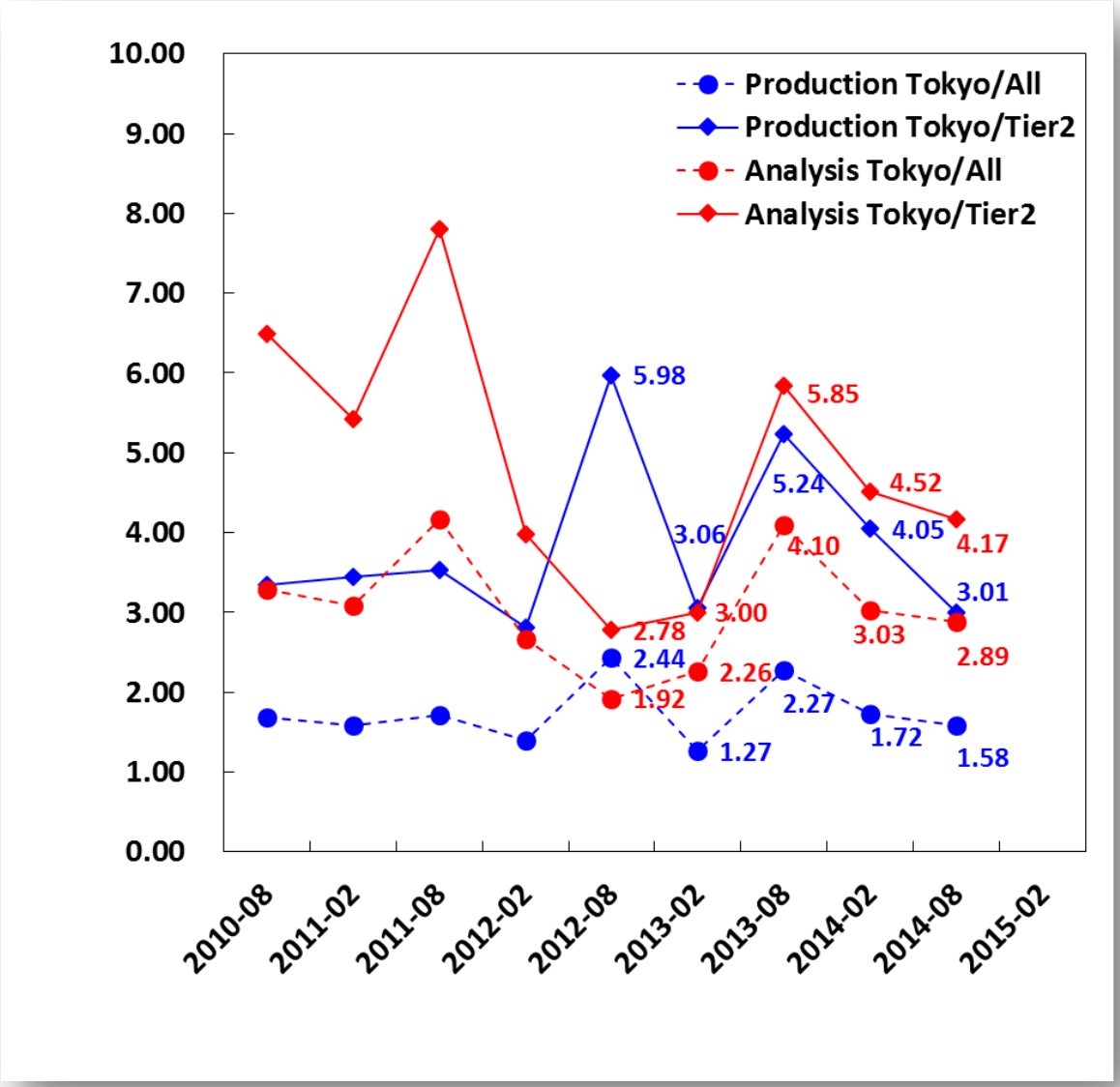
Single VO and Simple and Uniform architecture

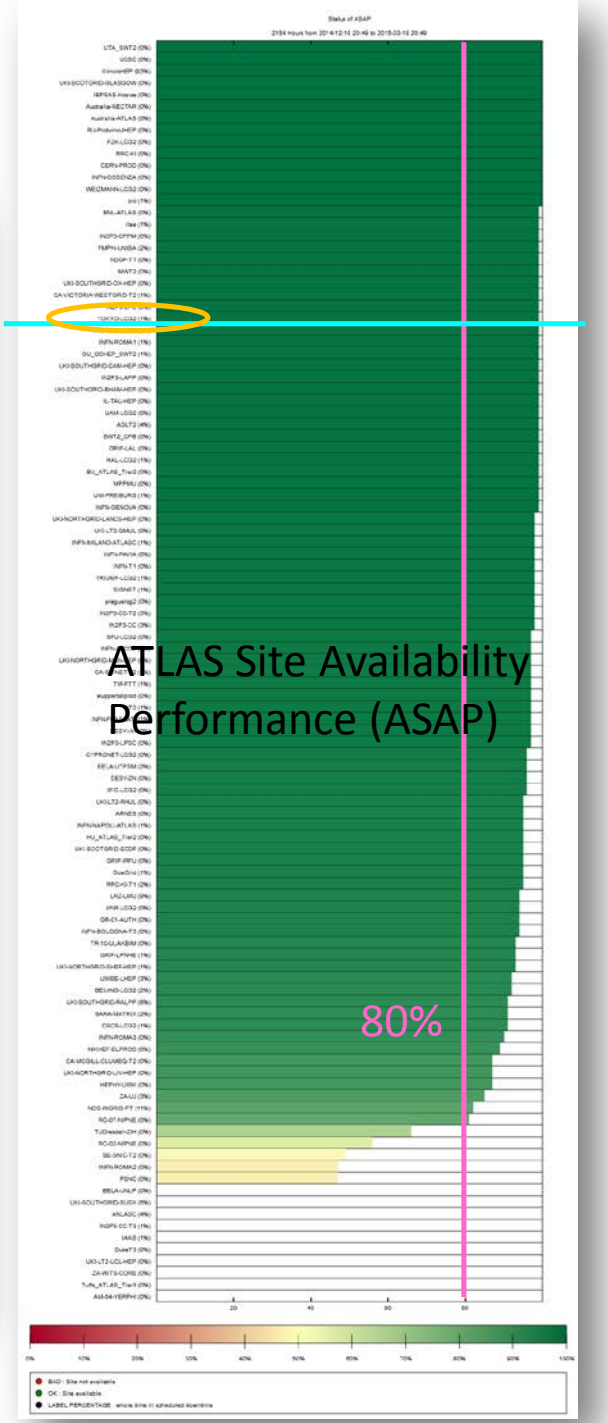
	2013	2014	2015
CPU pledge	16000 [HS06]	20000 [HS06]	24000 [HS06]
CPU deployed	43673.6 [HS06-SL5] (2560core)	46156.8 [HS06-SL6] (2560core)	46156.8 [HS06-SL6] (2560core)
Disk pledge	1600 [TB]	2000 [TB]	2400 [TB]
Disk deployed	2000 [TB]	2000 [TB]	2400[TB]

Dedicated staff

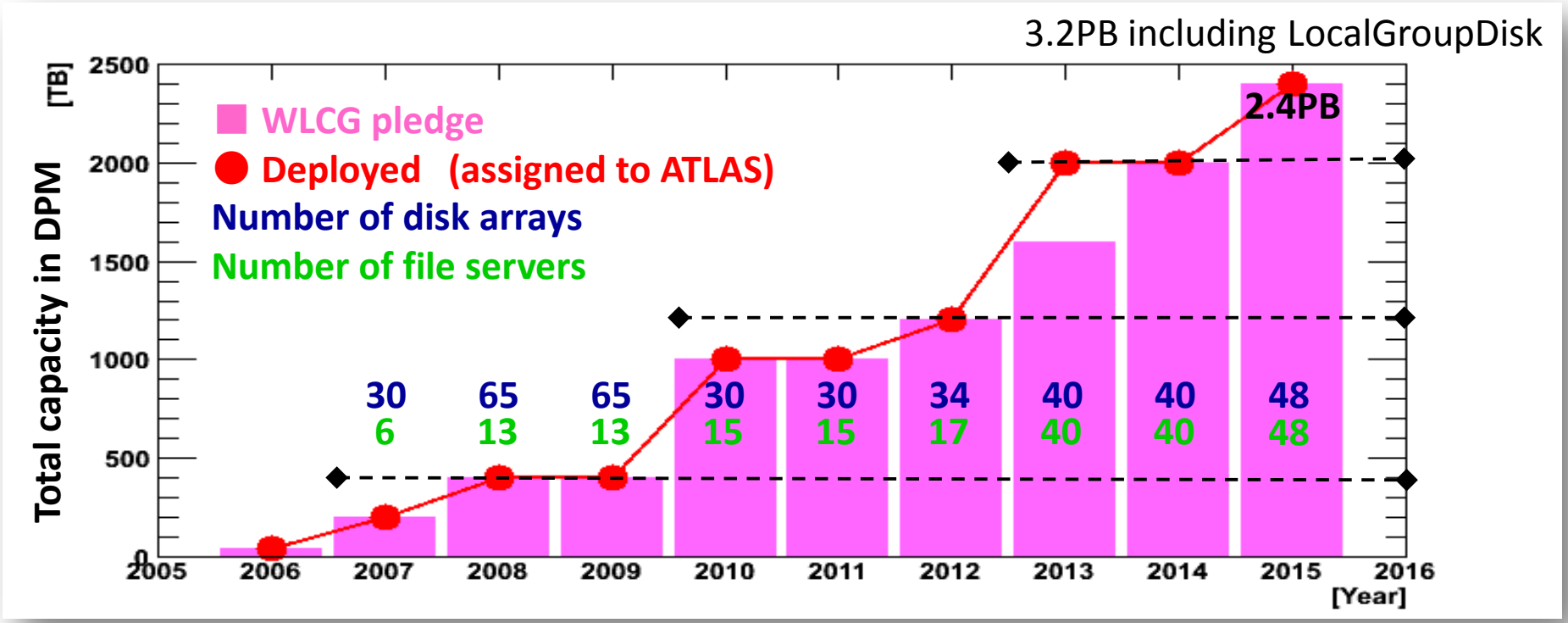
Tetsuro Mashimo:	fabric operation, procurement, Tier3 support
Nagataka Matsui:	fabric operation, Tier3 support
Tomoaki Nakamura (KEK):	Tier2 operation, Tier3 analysis environment
Hiroshi Sakamoto:	site representative, coordination, ADCoS
Ikuo Ueda:	ADC coordinator, site contact with ADC
System engineer from company (2FTE):	fabric maintenance, system setup

Status in ATLAS





Disk Storage



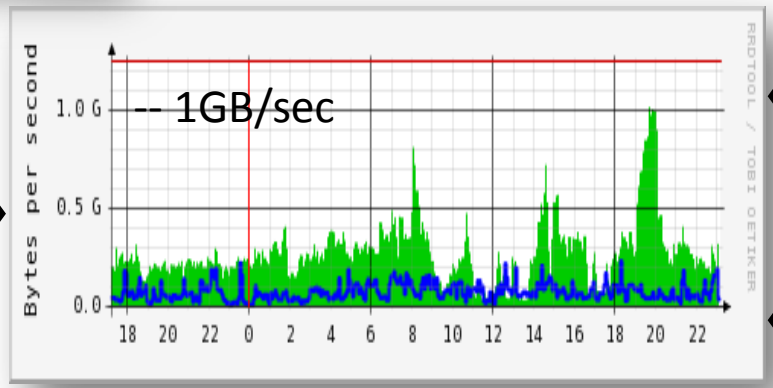
Pilot system for R&D	1st system 2007 - 2009	2nd system 2010 - 2012	3rd system 2013 - 2015	4th system 2016 - 2019
	16x500GB HDD / array **5disk arrays / server** **XFS on RAID6** **4G-FC via FC switch** **10GE NIC**	**24x2TB HDD / array** **2disk arrays / server** **XFS on RAID6** **8G-FC via FC switch** **10GE NIC**	**24x3TB HDD / array** **1disk array / server** **XFS on RAID6** **8G-FC without FC switch** **10GE NIC**	

Performance of internal network

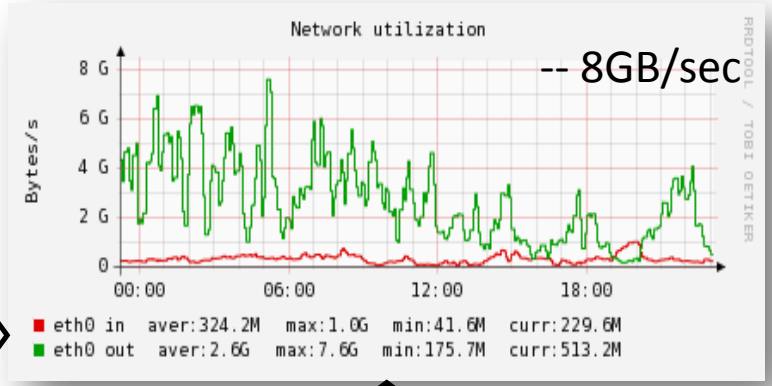
(a)

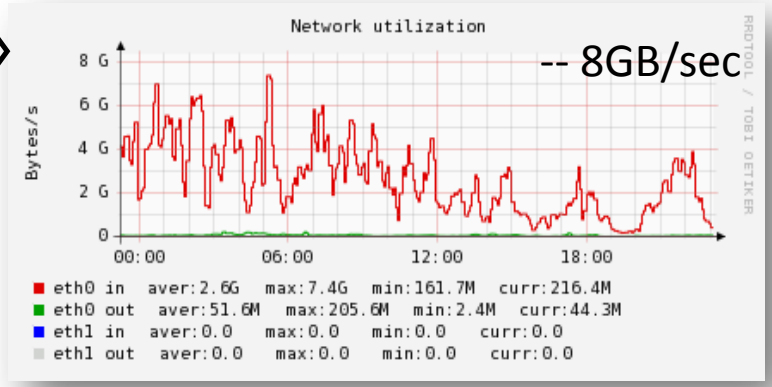
(b)

File servers
10GE x 40 nodes
8G-FC with disk array (66TB)



WAN
10Gbps





Center network switch
Brocade MLXe-32
non-blocking 10Gbps

10Gbps x 176 x 2 port

(c)

Worker nodes
10GE x 160 nodes
(minimum 5Gbps/node)

Data transfer throughput (2013)

Sustained transfer rate

Incoming data: ~100MB/sec in one day average

Outgoing data: ~50MB/sec in one day average

300~400TB of data in Tokyo storage is replaced within one month!

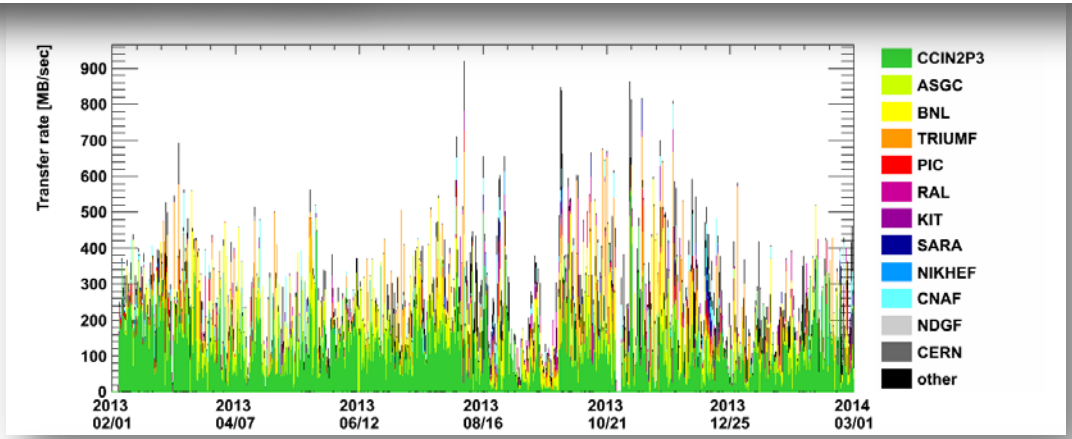
Peak transfer rate

Almost reached to 10Gbps

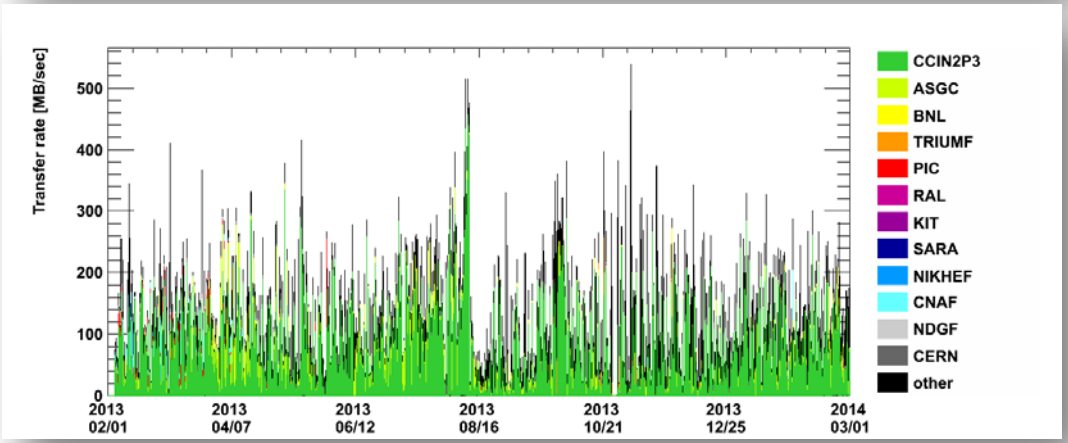
Need to increase bandwidth and stability!

2013 (10 min. ave.)

Monitored by file servers (extract from grid FTP logs)

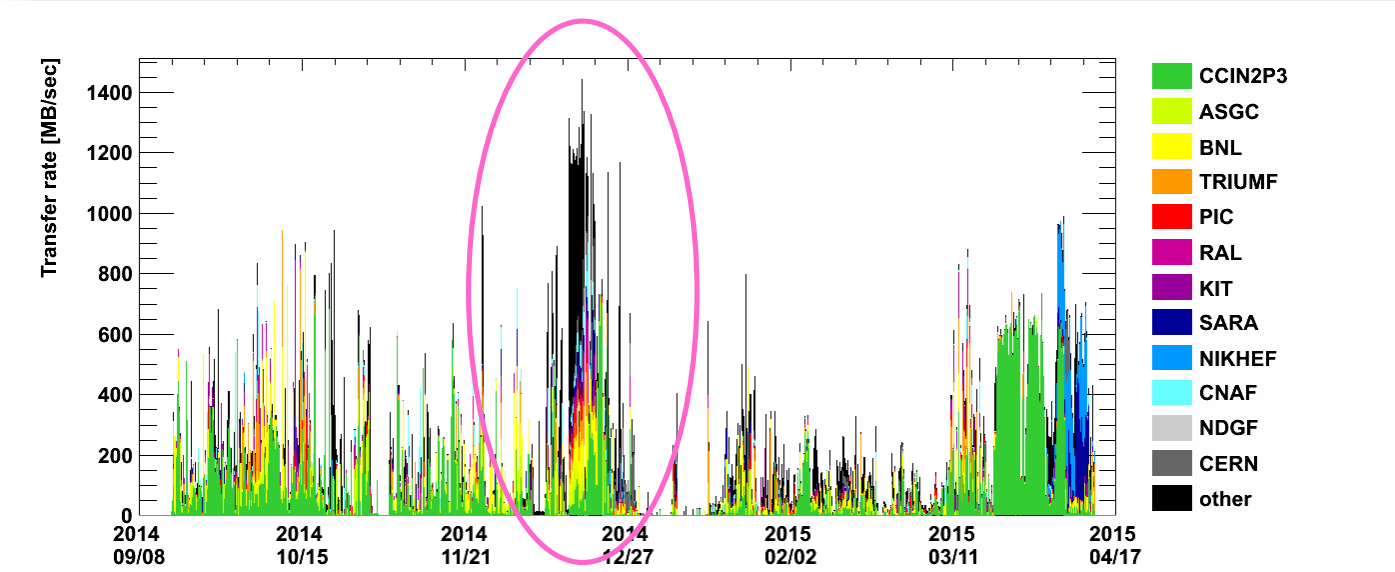


Incoming data

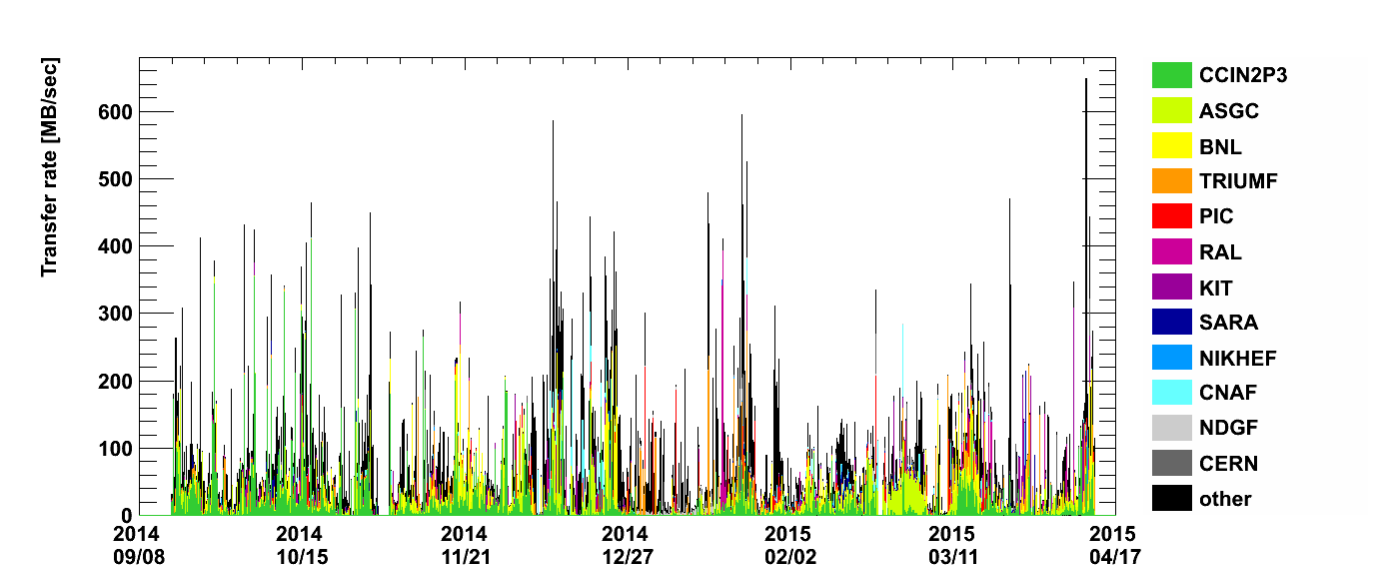


Outgoing data

Latest 3 months

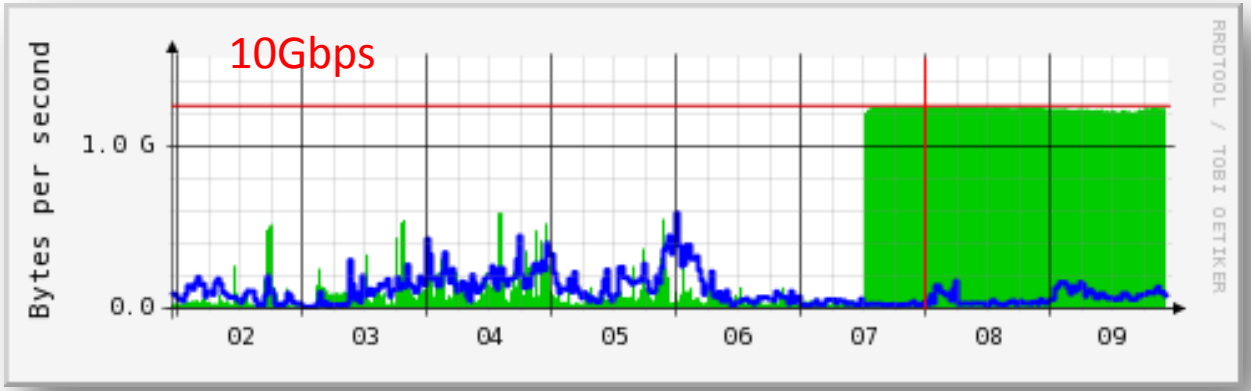


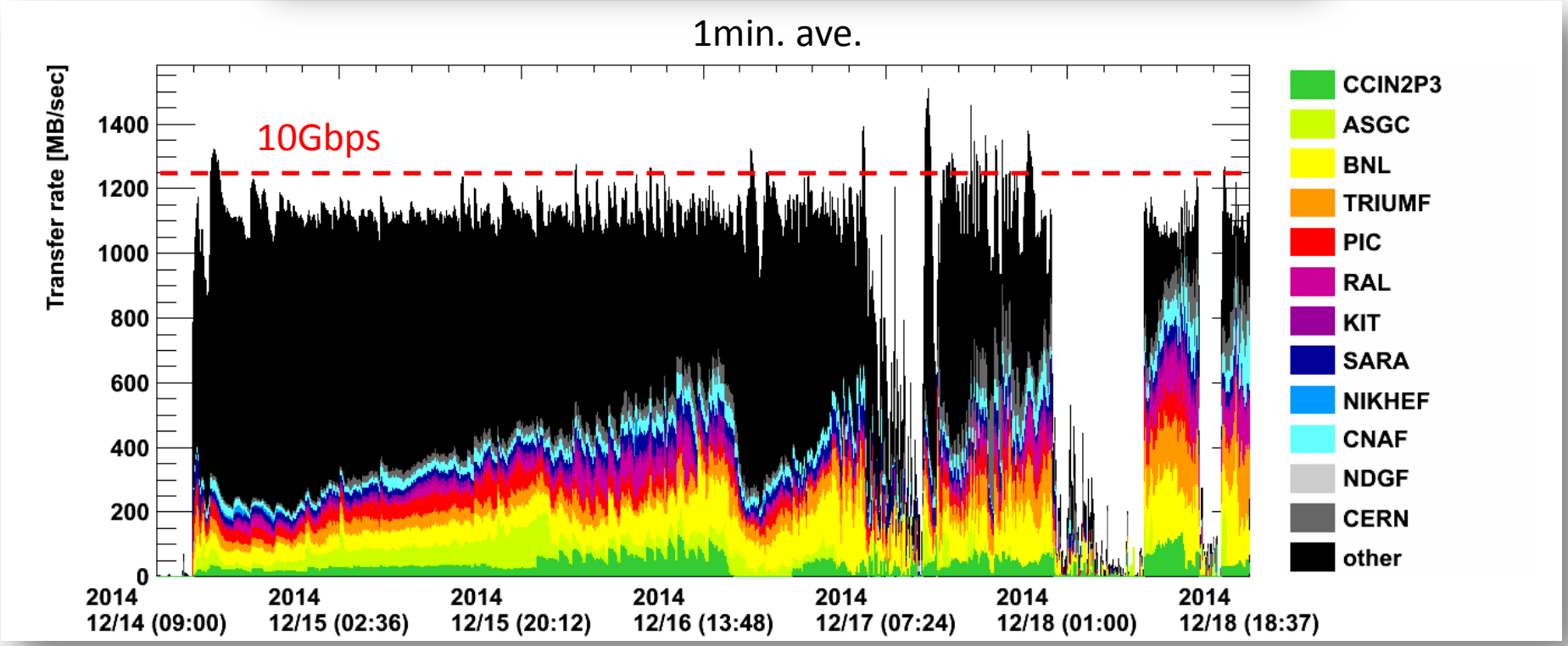
Incoming data



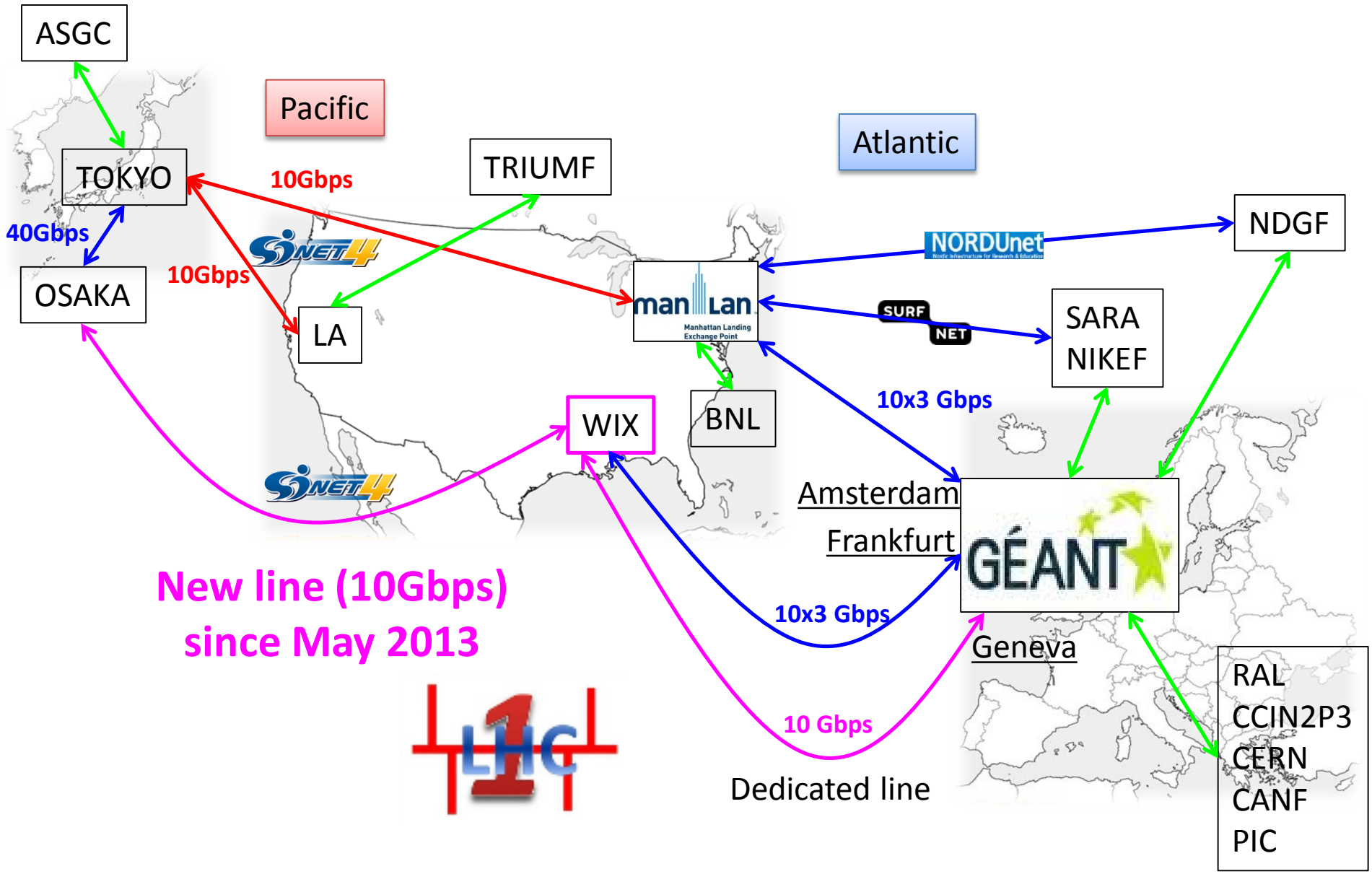
Outgoing data

Dec. 2014

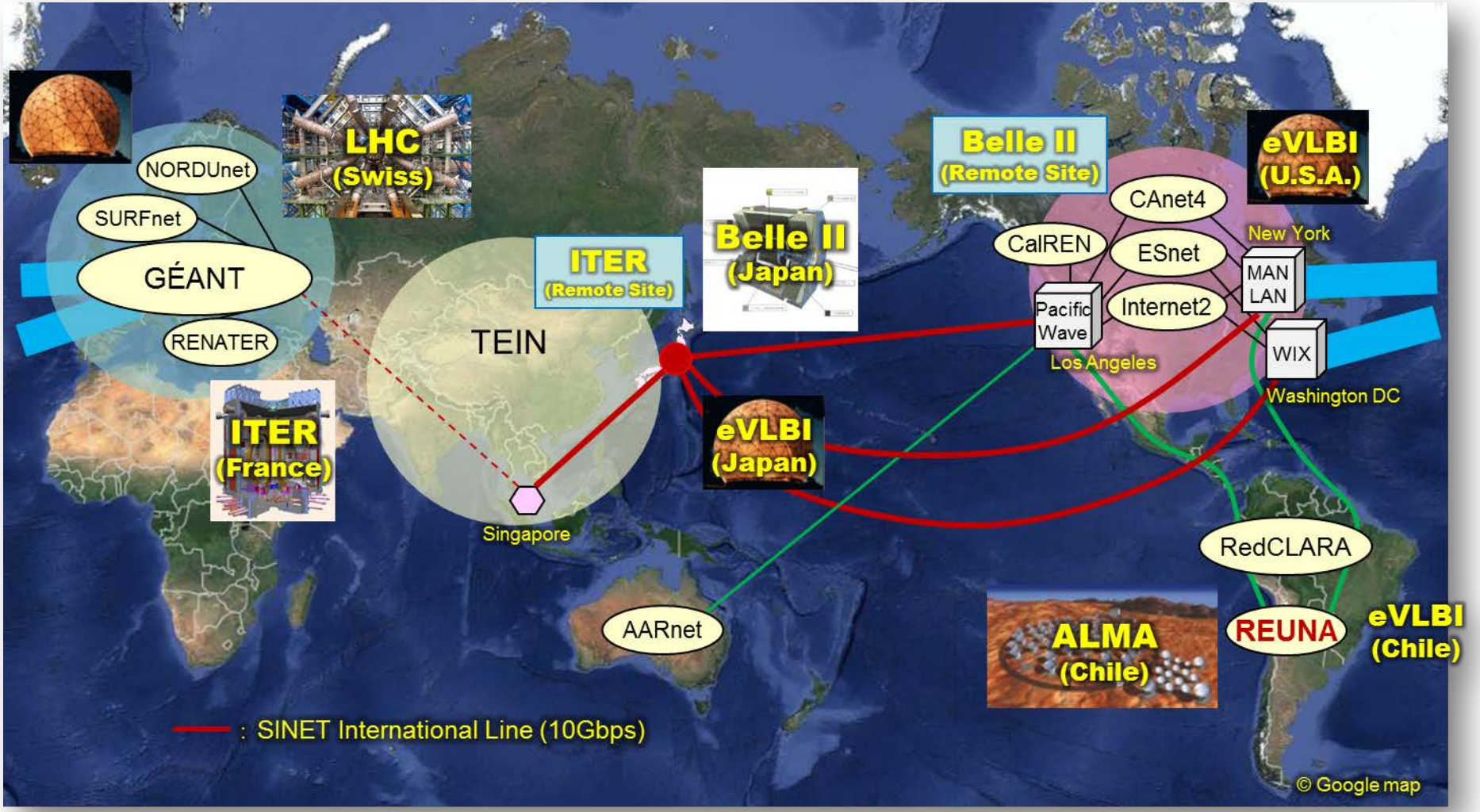

10Gbps
Bytes per second
1.0 G
0.0
02
03
04
05
06
07
08
09
RRDTOOL / TOBI OETIKER


1min. ave.
10Gbps
Transfer rate [MB/sec]
1400
1200
1000
800
600
400
200
0
2014 12/14 (09:00)
2014 12/15 (02:36)
2014 12/15 (20:12)
2014 12/16 (13:48)
2014 12/17 (07:24)
2014 12/18 (01:00)
2014 12/18 (18:37)
CCIN2P3
ASGC
BNL
TRIUMF
PIC
RAL
KIT
SARA
NIKHEF
CNAF
NDGF
CERN
other

International connection to Tokyo

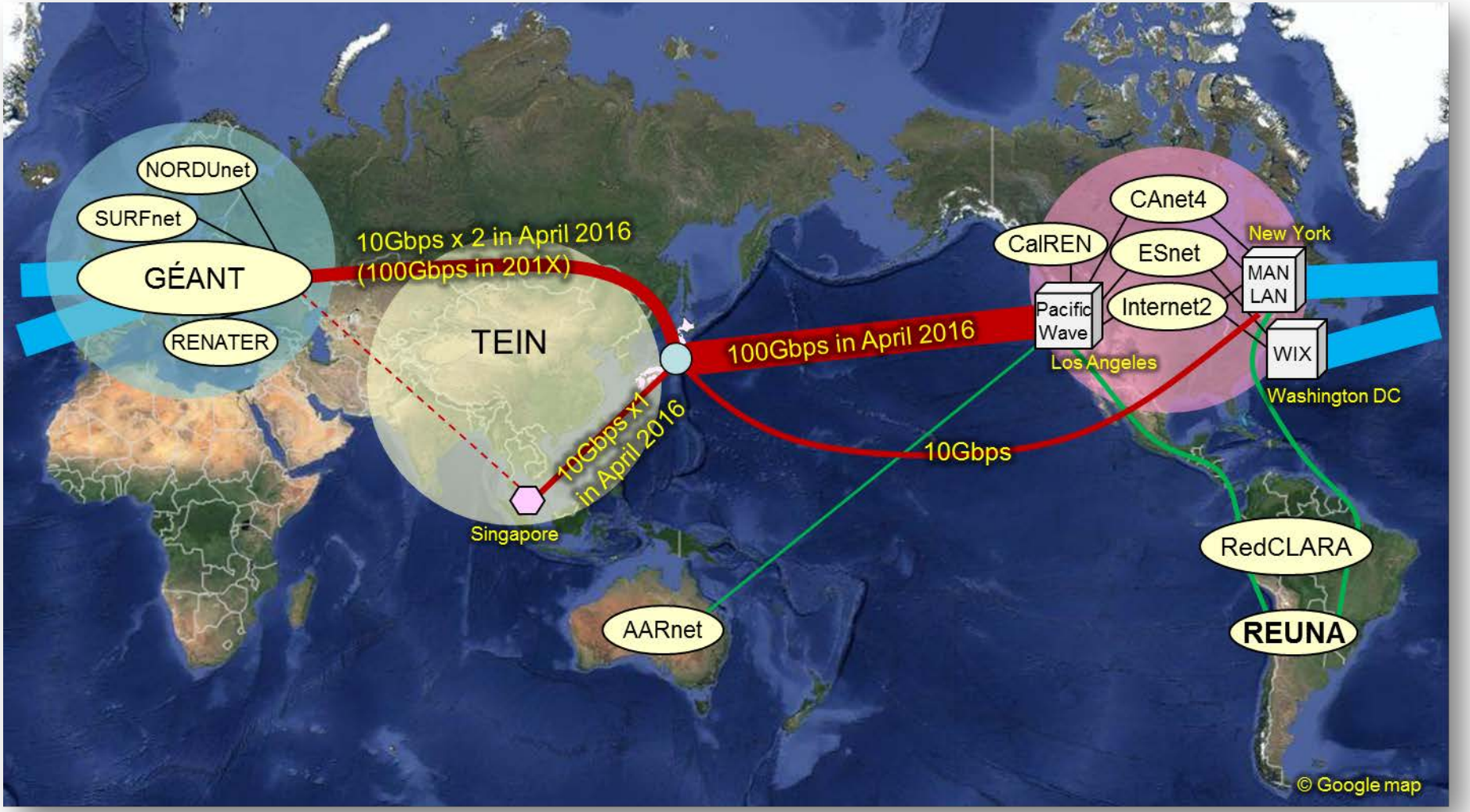


International network provided by NII



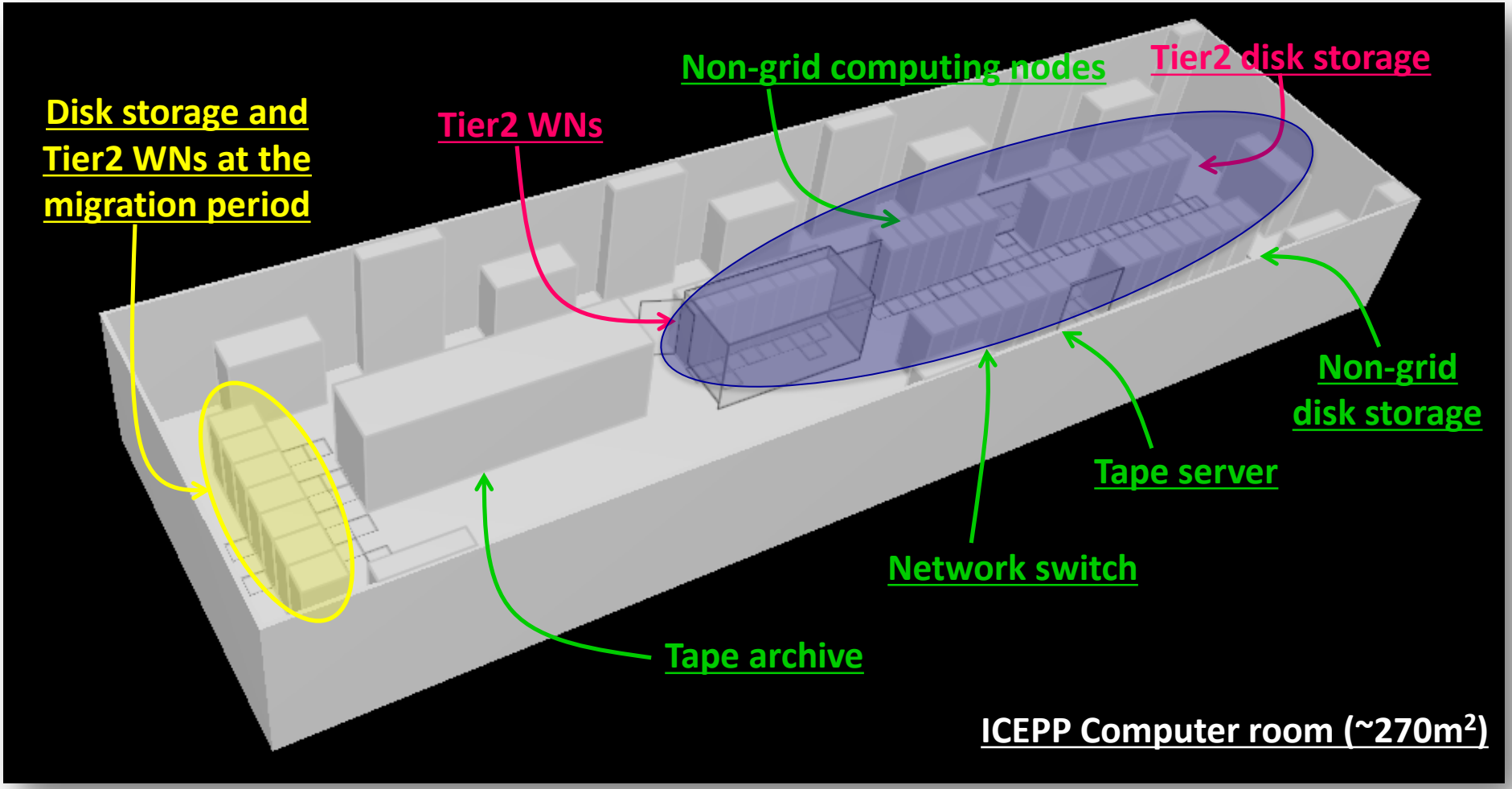
M. Nakamura (NII)

Upgrade from 2016 (international network)



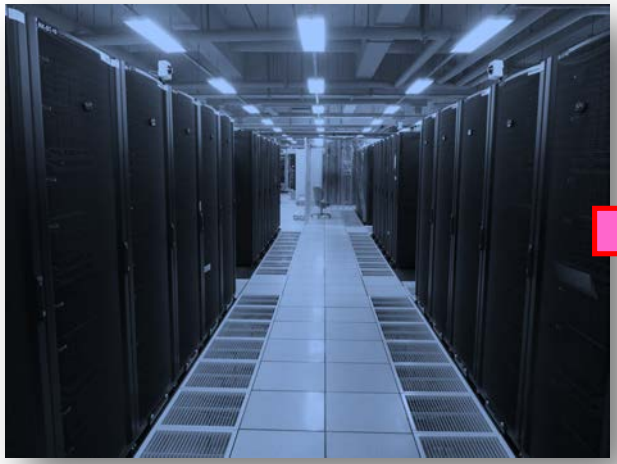
M. Nakamura (NII)

System upgrade (Dec. 2015)



System migration (Dec. 2015 – Jan. 2016)

3rd system



Clearance: 2 days

4th system

Construction: 1 week

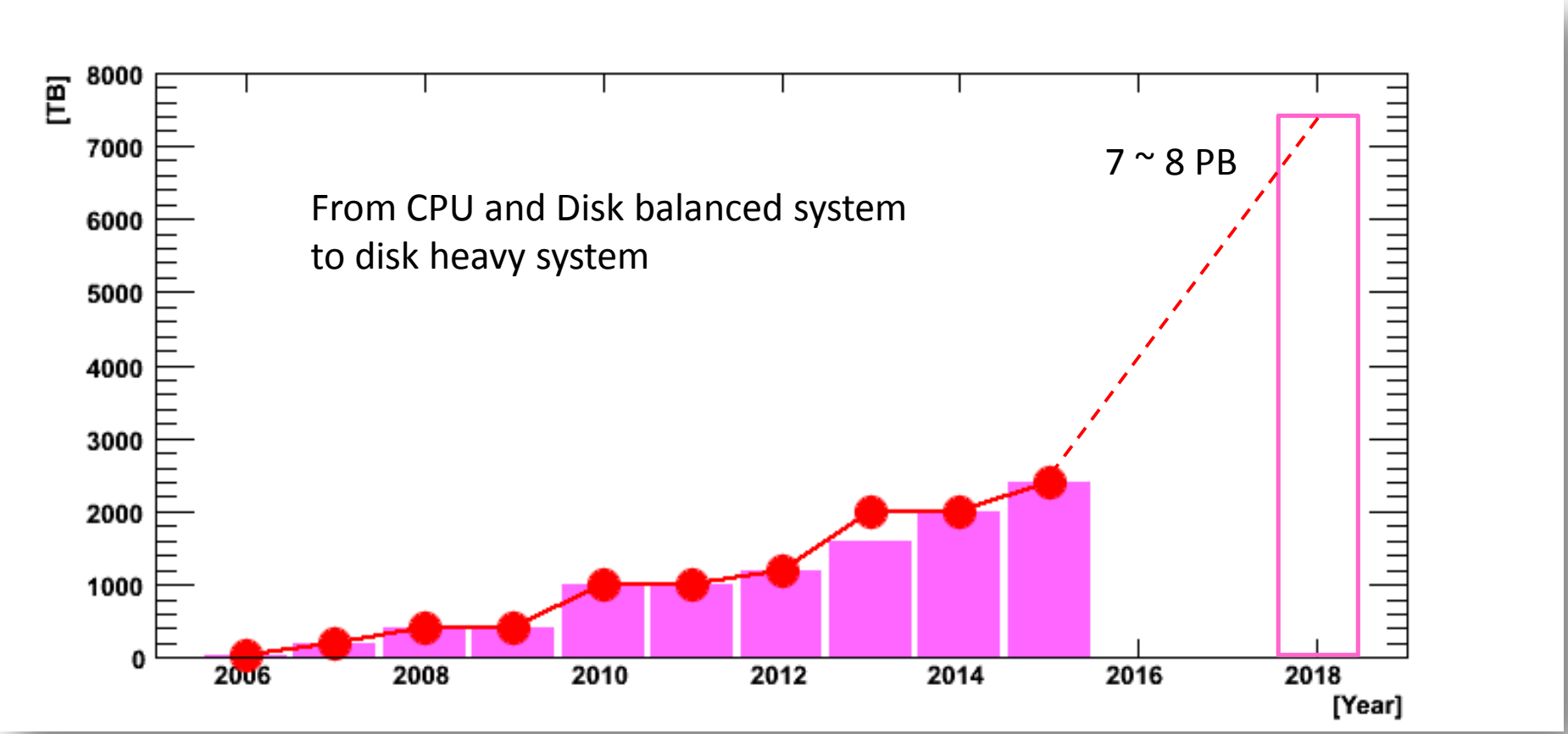
Data copy:
several week

Migration period

Copy back
several week

Running with reduced
number of WNs

Expected resources



4th system
(3 years)

Summary

Tokyo-LCG2 is running quite smoothly and providing several percent of ATLAS resources.

Current system will terminate in Dec. 2015. and will be migrated to the 4th generation system.

Concern: JPY is weak, 1$ = JPY80 at the previous migration, 1$ = JPY 120 now

Going to disk heavy system.

International network of Japan provided by Japanese NREN will be upgraded at April in 2016 (100Gbps + 10Gbps to US, 20Gbps to EU).

We can share the NRE network by more flexible and effective way after the upgrade among a lot of projects.

Wide area network bandwidth of TOKYO-LCG2 will be also upgraded from the current 10Gbps to 20Gbps (2016) to 40Gbps (future).