

Upgrade of network connection between KEK and SINET

Soh Y. Suzuki, Tomoaki Nakamura, Go Iwai,
Fukuko Yuasa, Takashi Sasaki, and Toshiaki Kaneko
Computing Research Center
High Energy Accelerator Research Organization (KEK)

What is KEK? - Our Mission

- Scientists at KEK use accelerators and perform research in high-energy physics to answer the most basic questions about the universe as a whole, and the matter and the life it contains. (<http://www.kek.jp/en/About/>)
- 60km away from Tokyo
 - Most of the research network projects are NOT directly accessible.
 - Accessibility by NREN is very important

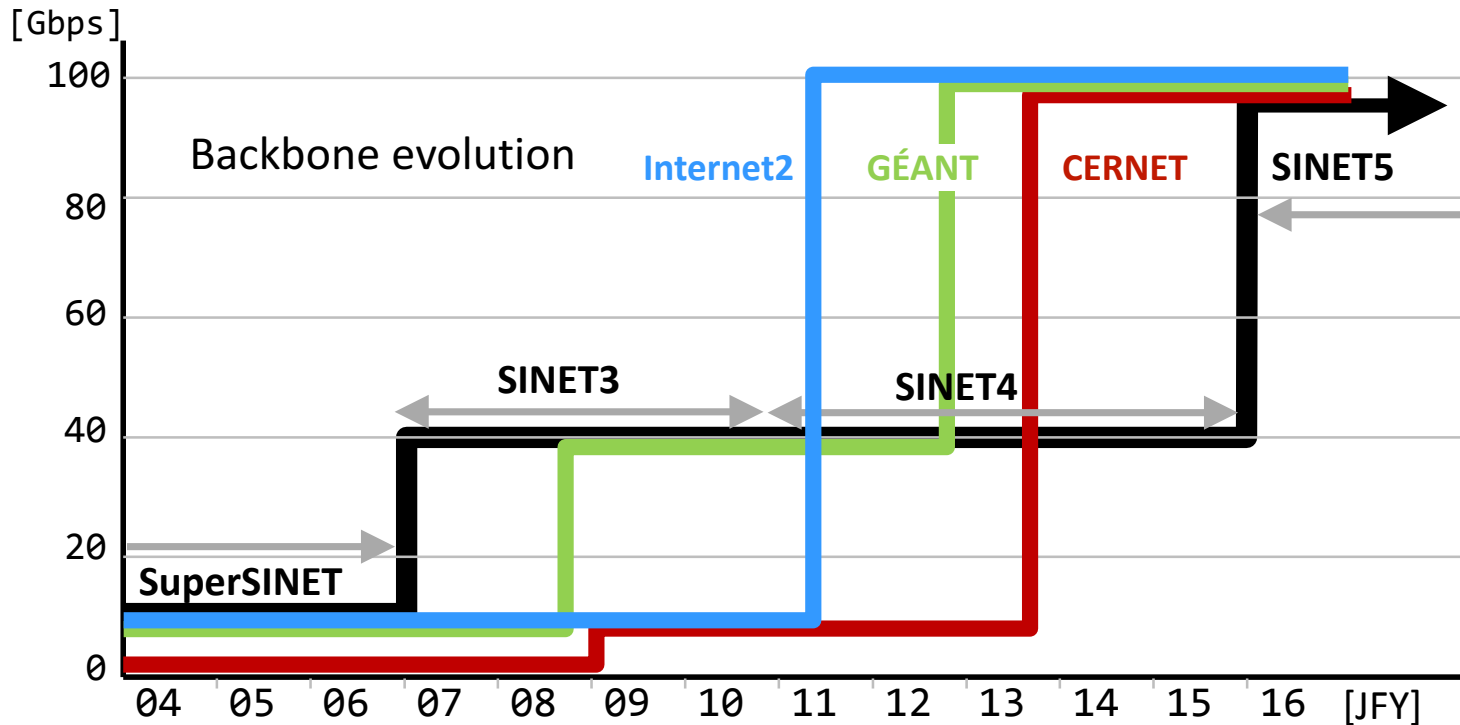
SINET is a NREN in Japan

- Most of the connectivity for HEP researchers in Japan is provided by SINET
 - Mainly for researches using network
 - SINET4 placed nodes at all prefectures



SINET4 to SINET5

- SINET is a project of NII
- SINET5 is the 5th term



Changes from 4 to 5

- Backbone upgrade from multi-10G or 40G to multi-100G
- JP-US upgrade from 3x10G to 100G+10G
- JP-EU 2x10G New
- Subscribers can newly use 100G-LR4, 40G-LR4 at all nodes
 - 10G-LR, 1000LX, 1000T also OK, but no SR,SX even from nearby collocation space.
- Shutdown of all former-nodes in universities

100G access line for KEK

- Some optical fibers around KEK are terribly old. Our experience on another link was,
 - Power loss > 10dB in 10km
 - 10G-LR: warns low opt-power, but no packet loss
 - 40G-LR4: achieved 40Gbps, but 1 loss per 1 min.
 - 10G-ER: No warn, no loss.
- Initially we hoped 100G-LR4 is directly reachable via dark fibers, but lot of fear.

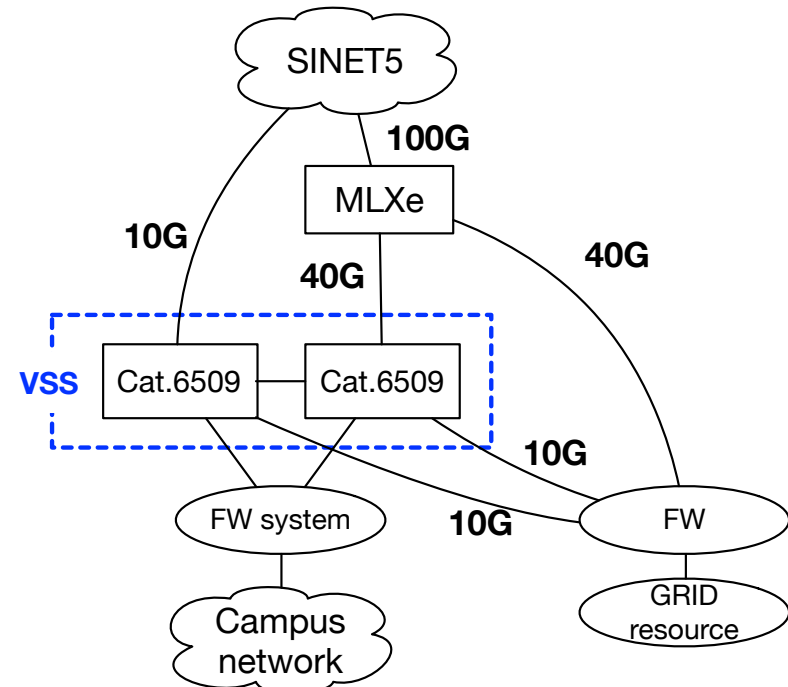
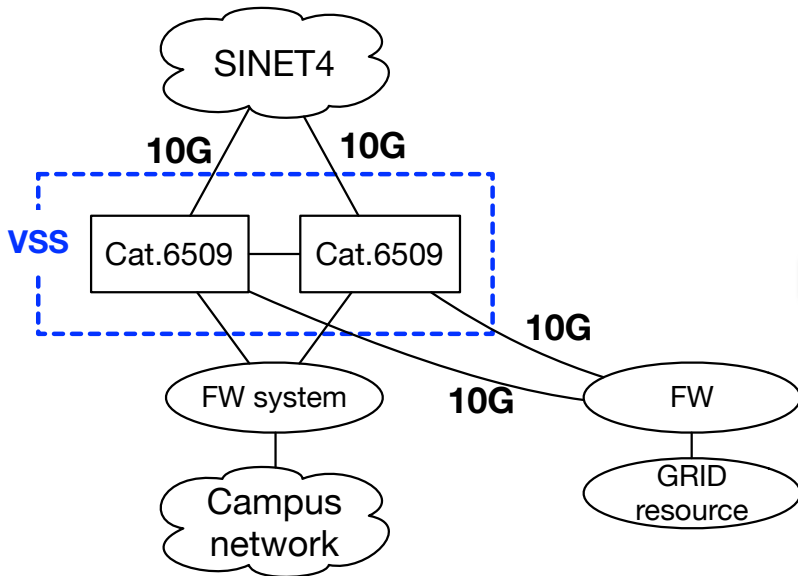
For reliability and stability,

- We chose 100G WDM for longer distance
 - 10G(internet) + 10G(inter-campus for J-PARC) + 100G
- But we bumped against a 100G-LR4 CFP2 with a initial malfunction (bad production?).
 - link flapping at a few Hz
 - replaced by newer one and problem had gone (hopefully)



Border switch for SINET5

- We used a pair of Catalyst 6509 to handle 10G links for for SINET4
 - No 100G line card is available
 - 40G line card provides only 10-12Gbps per stream
- Added Brocade MLXe4 with 100G+40G
 - Now used as L2 switch



Monitoring and Filtering

- KEK was connected to SINET via 2x10G
 - Monitoring by optical taps on most outer links, tap aggregation switches, and IDS
 - Filtering by full 10G Firewalls
- For 10G+100G
 - Monitoring ports on older border switch instead of optical taps.
 - Filtering by Firewalls and ACLs on 40G switches

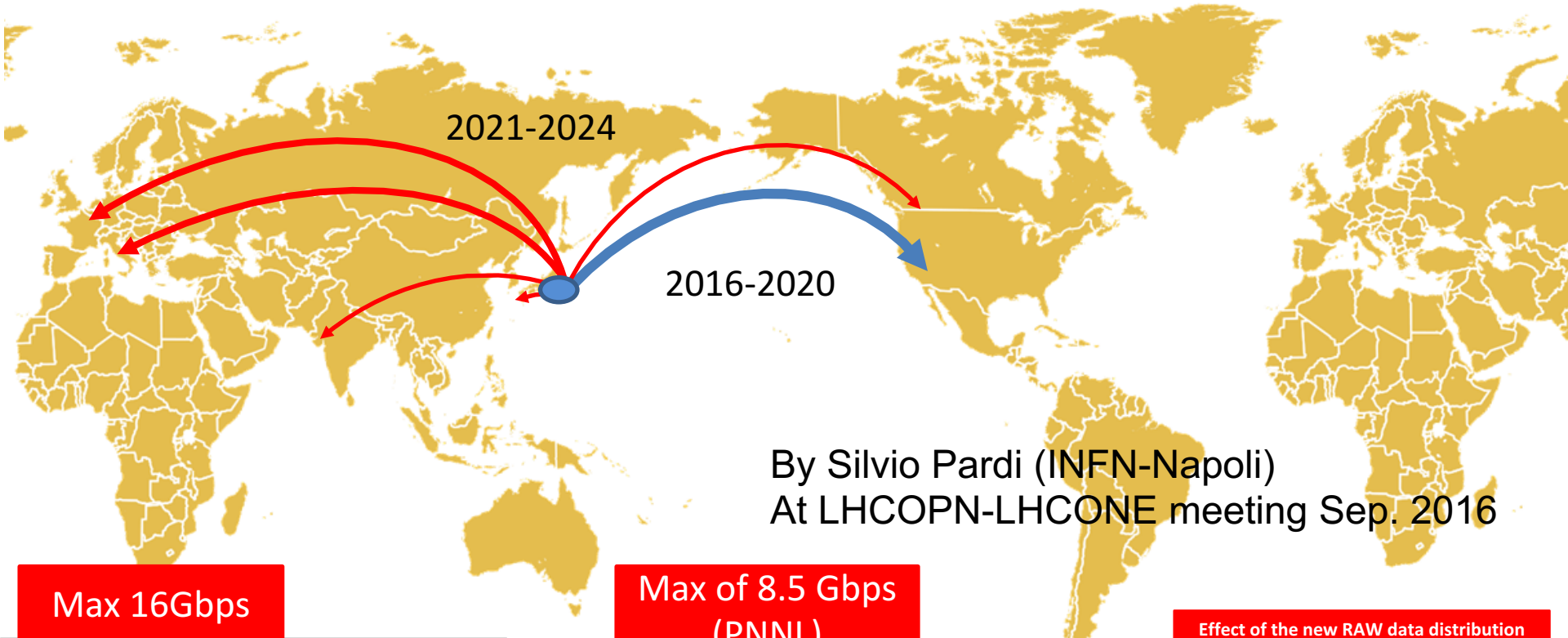
KEKCC and Belle II

- the KEK Central Computing system (KEKCC)
 - Provides GRID resources
 - Collaborators of KEK research projects can use it from remote.
 - The policy of the firewall differs from that for the KEK Campus network for KEK staffs and researchers geographically in KEK.
- Belle II collaboration is one of the largest user group

Connecting it to LHCONE

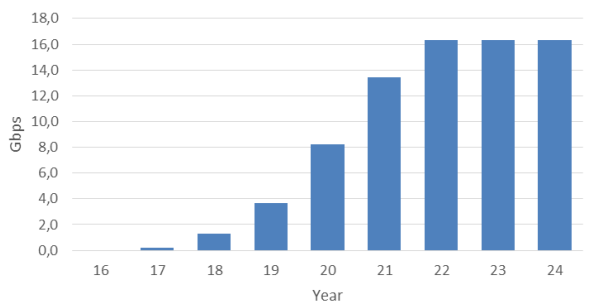
- Belle II needs huge computing resources
 - Exceeds the volume of present KEKCC
- World wide co-operation with foreign HEP sites is essential, like LHC
- Nowadays most of active HEP computing sites are in WLCG
 - LHCONE is the best way for the mass data transfer among them

Belle II experiment RAW Data Strategy



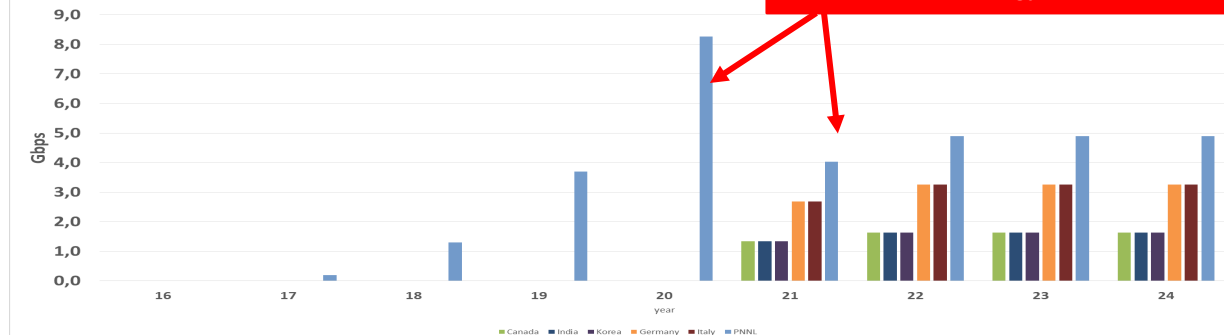
Max 16Gbps

RAW-DATA KEK OutBand



Max of 8.5 Gbps
(PNNL)

RAW Data InBand

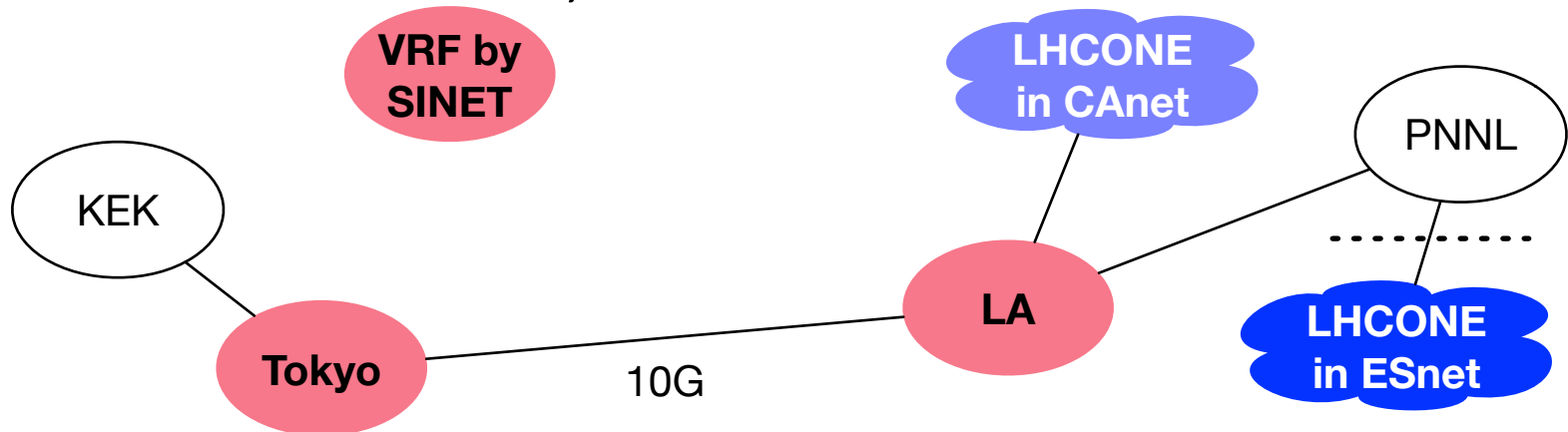


KEK is not in WLCG

- Major computing sites related Belle II are in WLCG and LHCONE
 - Only KEK and PNNL are not so.
- We asked the LHCONE community to connect us for the Belle II experiment
 - they accepted our request
- Then ESnet immediately connected PNNL to LHCONE
- KEK didn't immediately, because KEKCC was unable to handle policy routing for LHCONE

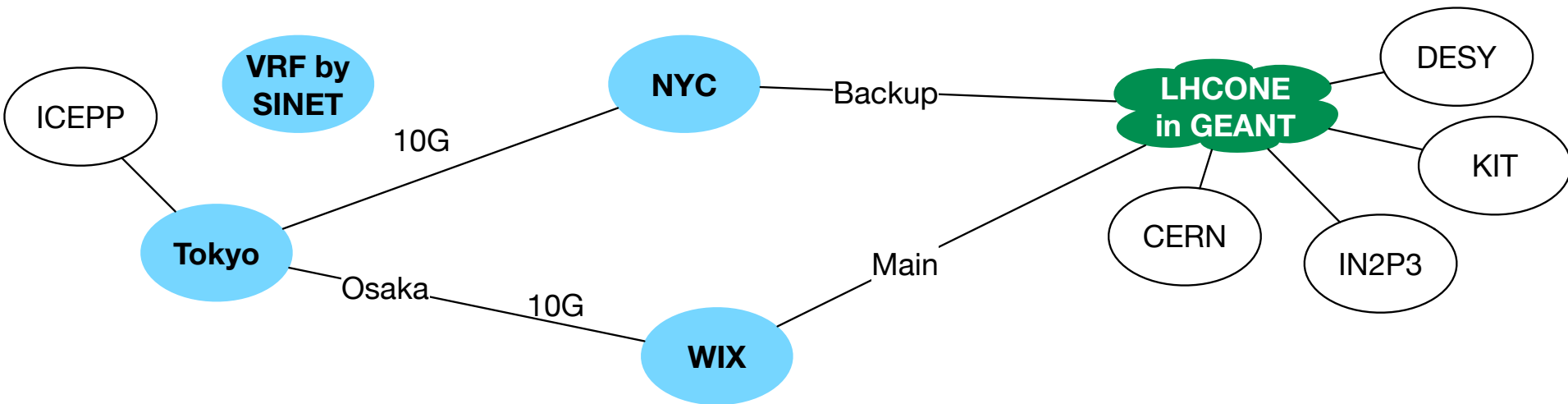
Our temporal solution

- Manual maintenance of static routes on grid-ftp servers in KEKCC (about 15 prefixes).
- Closed VPN by SINET and ESnet
 - This VPN is similar to LHCONE, but not LHCONE
- Then,
 - CANARIE joined their LHCONE VRF to this VPN for HEPnet CANADA, without transit

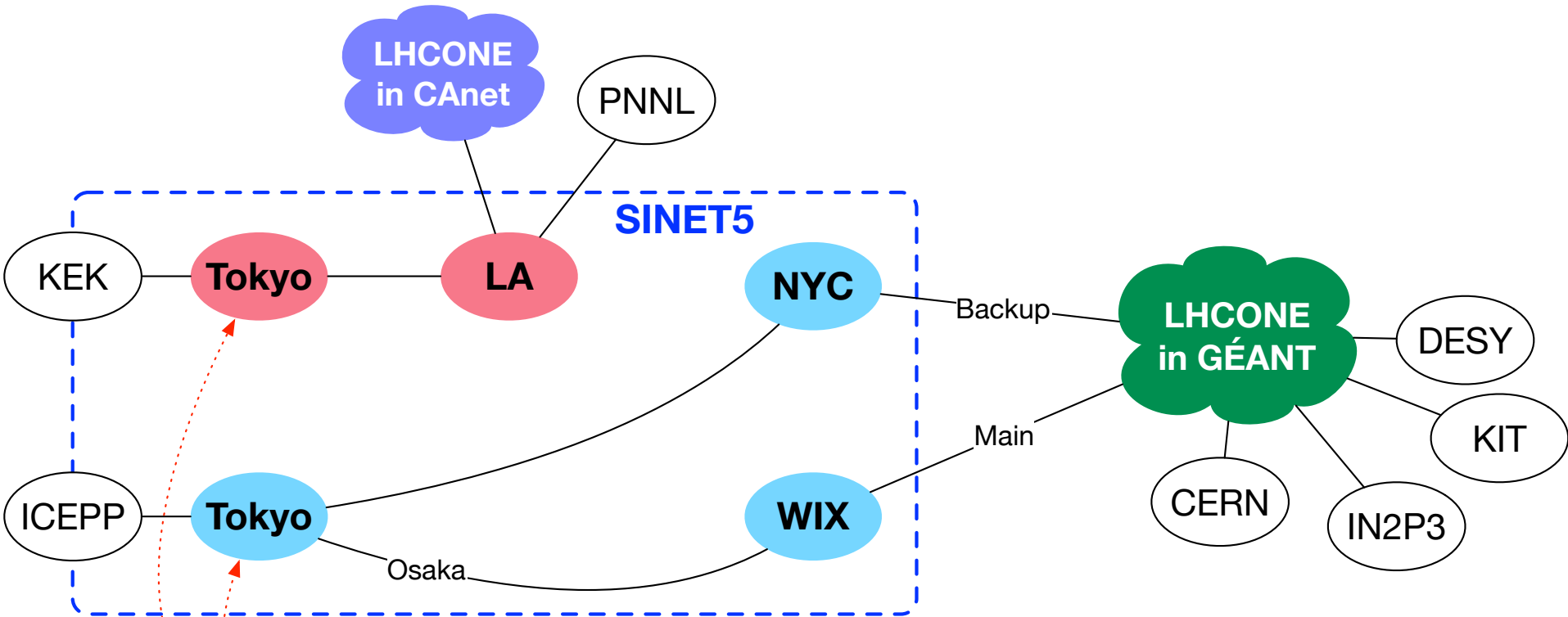


Real LHCONE connection from Japan

- ATLAS Tier2 at ICEPP in the University of Tokyo
- Only to European sites via GEANT
 - Not connected to LHCONE sites in U.S.

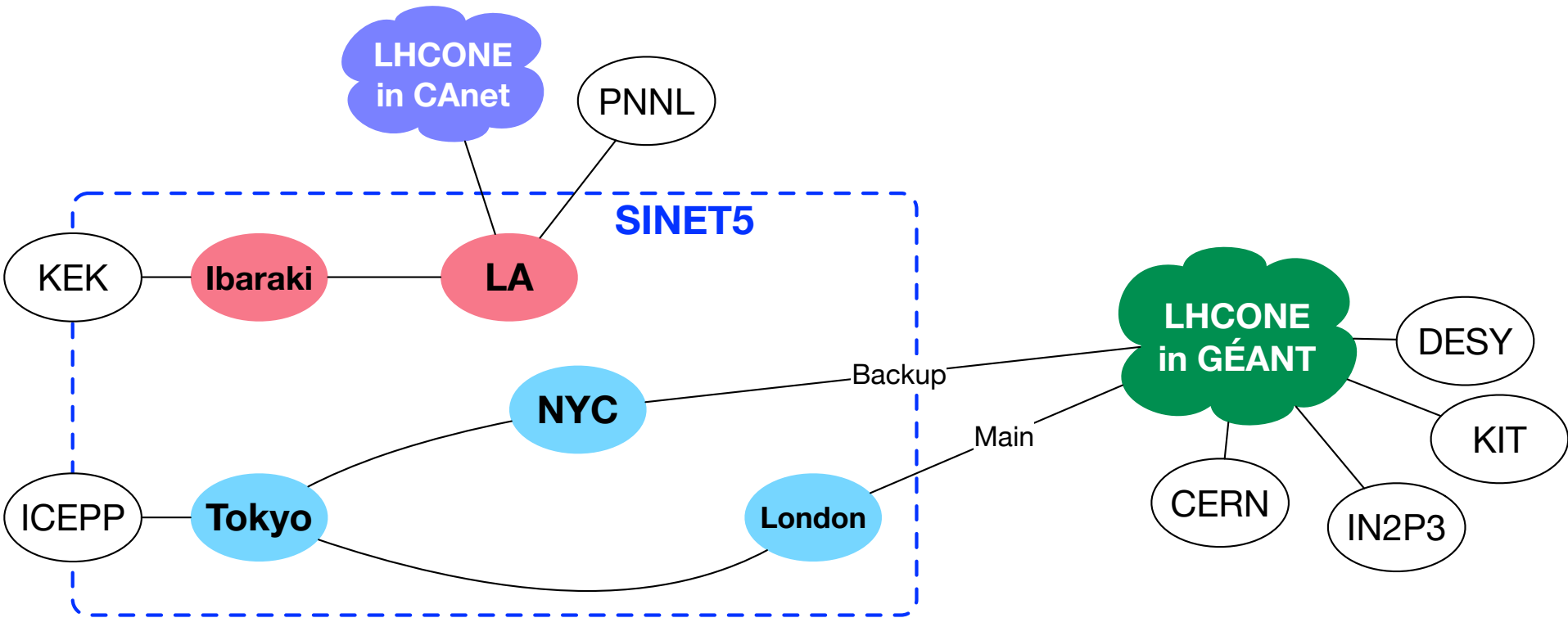


Until Feb. 2016 by SINET4



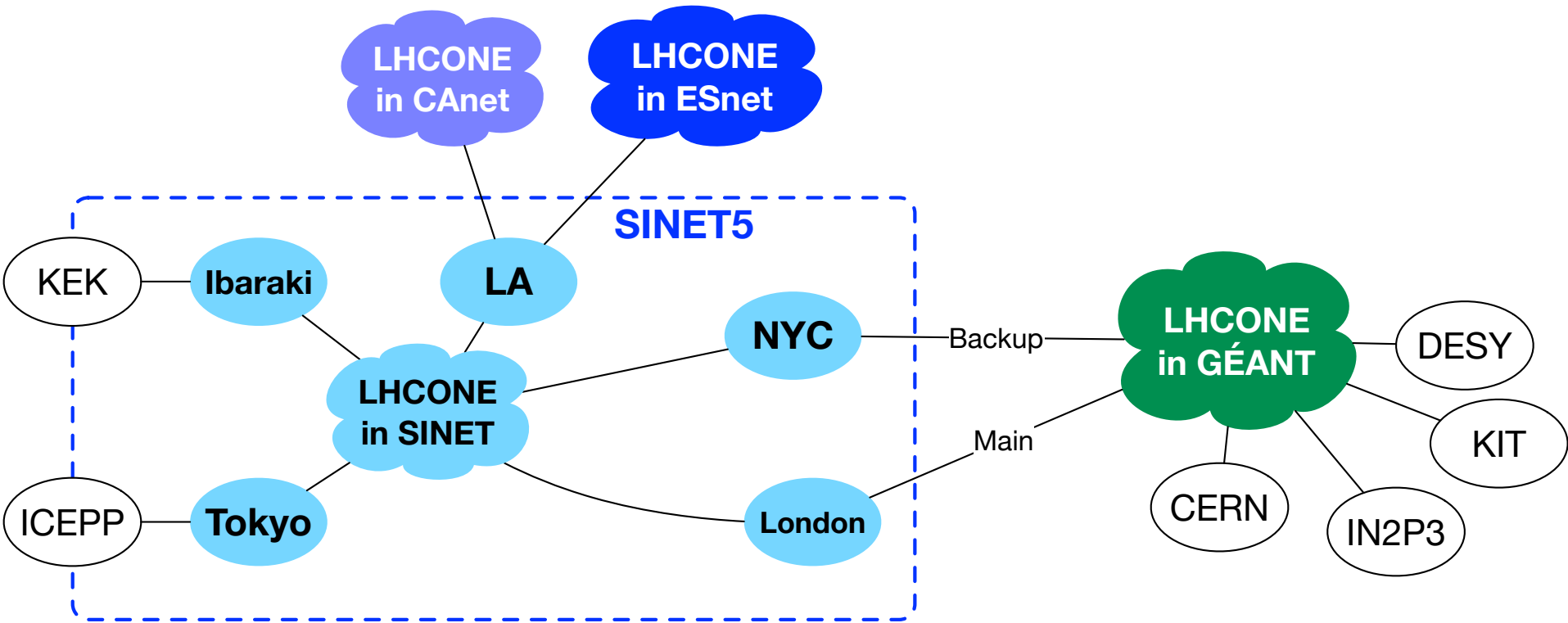
**Geographically same,
but different VRF**

Since Mar. 2016, by SINET5



- JP-LA is upgraded to 100G
- Moving from WIX to London reduced the RTT
- VRF speaking BGP with KEK is moved from Tokyo to Ibaraki

Since Sep. 2016 by SINET5



migration took about one week

- Plan
 - prepare new peer SINET-ESnet for LHCONE
 - shutown old peers SINET-ESnet for Belle II
 - stop route exchange on SINET-CAnet peer with keeping BGP connection
 - KEK-PNNL and KEK-HEPnet Canada use internet
 - SINET immigrate 2 VRFs for Belle II into LHCONE VRFs
 - check routes between KEK and LHCONE in GEANT
 - start new peer SINET-ESnet for LHCONE
 - check routes between ICEPP and LHCONE in ESnet
 - This took several retries and a retry needs a day.

What we learned

- Think about the filter at the neighbor side before advertising of new prefix(es)
 - It may be a white list, not black list
 - It may be a prefix list, not AS list
- There is a direct peer between ESnet and CERN
 - SINET-ESnet-CERN was used instead of SINET-GEANT-CERN that we expected.
 - BGP community tuning is needed to use the London link of SINET

Then,

Summary

KEK Outgoing

Destination	SINET4 [Gbps]	SINET5 old KEKCC [Gbps]	SINET5 new KEKCC LHCONE [Gbps]	Increase over old KEKCC & LHCONE
PNNL	3.6	3.9	8.4	115%
DESY	3	3	-	-
KIT	3.5	3.2	-	-
CNAF	-	3.8	9.0	136%
NAPOLI	3	3	8.8	190%

KEK Incoming

Source	SINET4 [Gbps]	SINET5 old KEKCC [Gbps]	SINET5 new KEKCC LHCONE [Gbps]	Increase over old KEKCC & LHCONE
PNNL	4.6	6.3	-	-
DESY	4	8	-	-
KIT	5	7	-	-
CNAF	7	7	13.5	93%
NAPOLI	5.5	6.6	13	97%

Summary

- We migrated from 10Gx2 for SINET4 to 100G+10G for SINET5.
 - 10G is assigned ordinary Internet access from the campus network
 - 100G for inter-lab VPNs including LHCONE
- New KECC is now connected to LHCONE
 - Improved FTS3 throughput especially for EU sites