

A short history of High Energy Physics activity between Japan and Korea

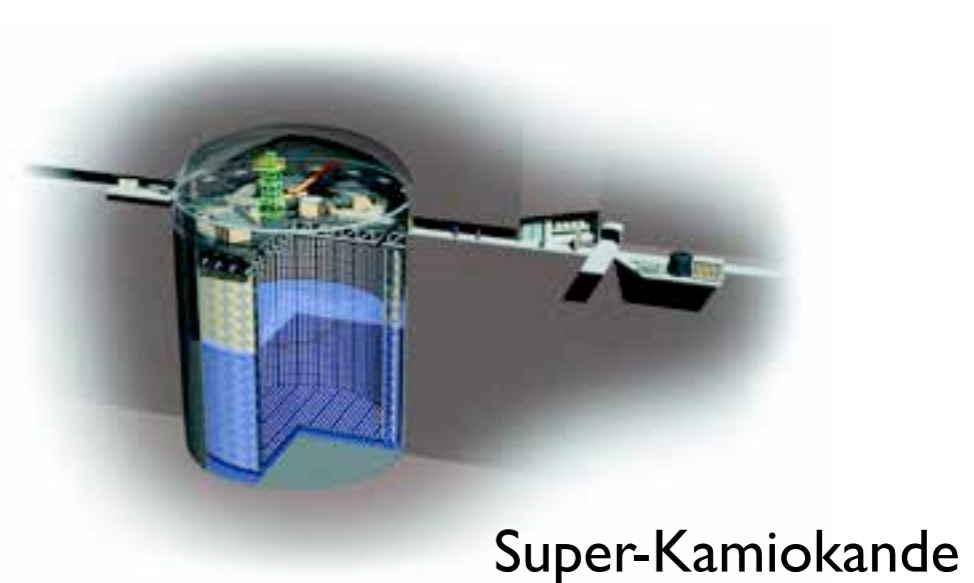
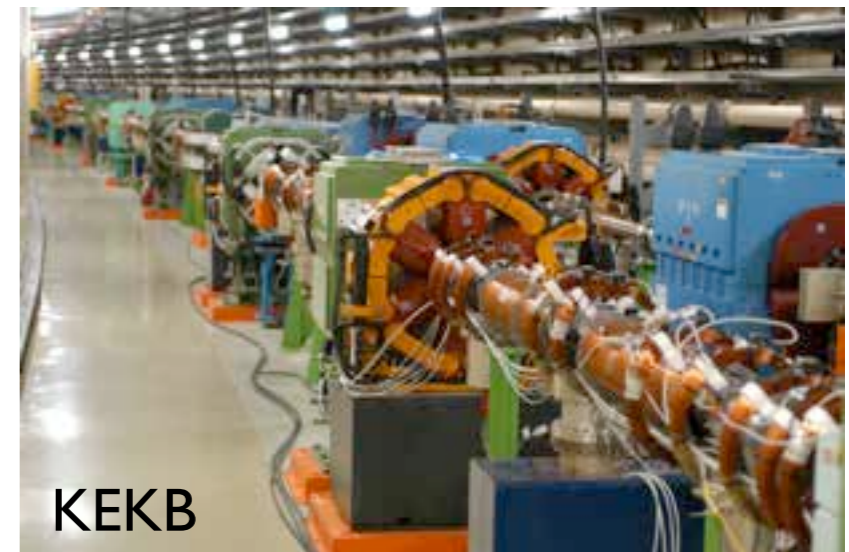
Soh Y. Suzuki (soh.suzuki@kek.jp)

Computing Research Center

High Energy Accelerator Research Organization (KEK)

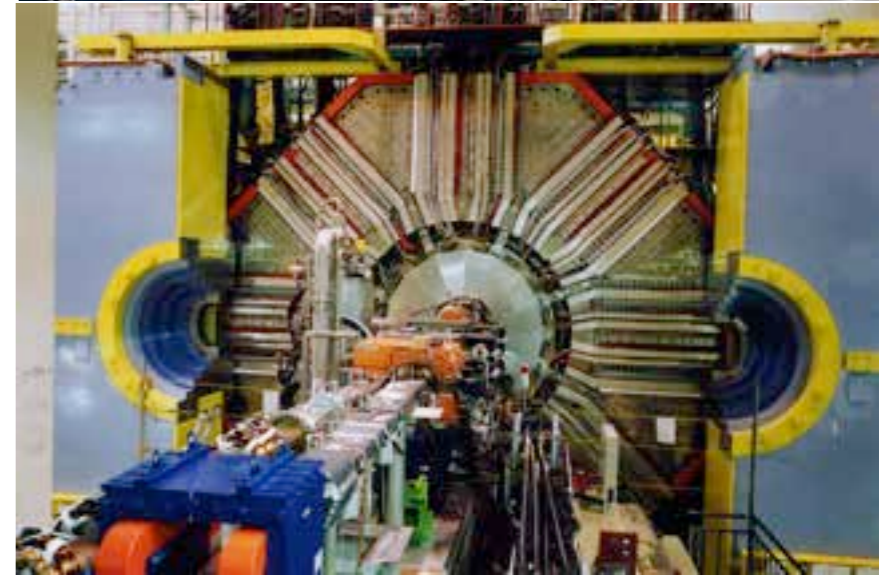
High Energy Physics

- is study of sub-atomic particles to understand the Universe
- Category of experiments
 - ▶ Accelerator
 - ▶▶ Accelerate protons or electrons to produce sub-atomic particles
 - ▶▶ accelerated bunch of particles is called as "beam"
 - ▶ Non-Accelerator
 - ▶▶ e.g. Cosmic-Ray observation
- Most of experiments are driven by each international collaboration

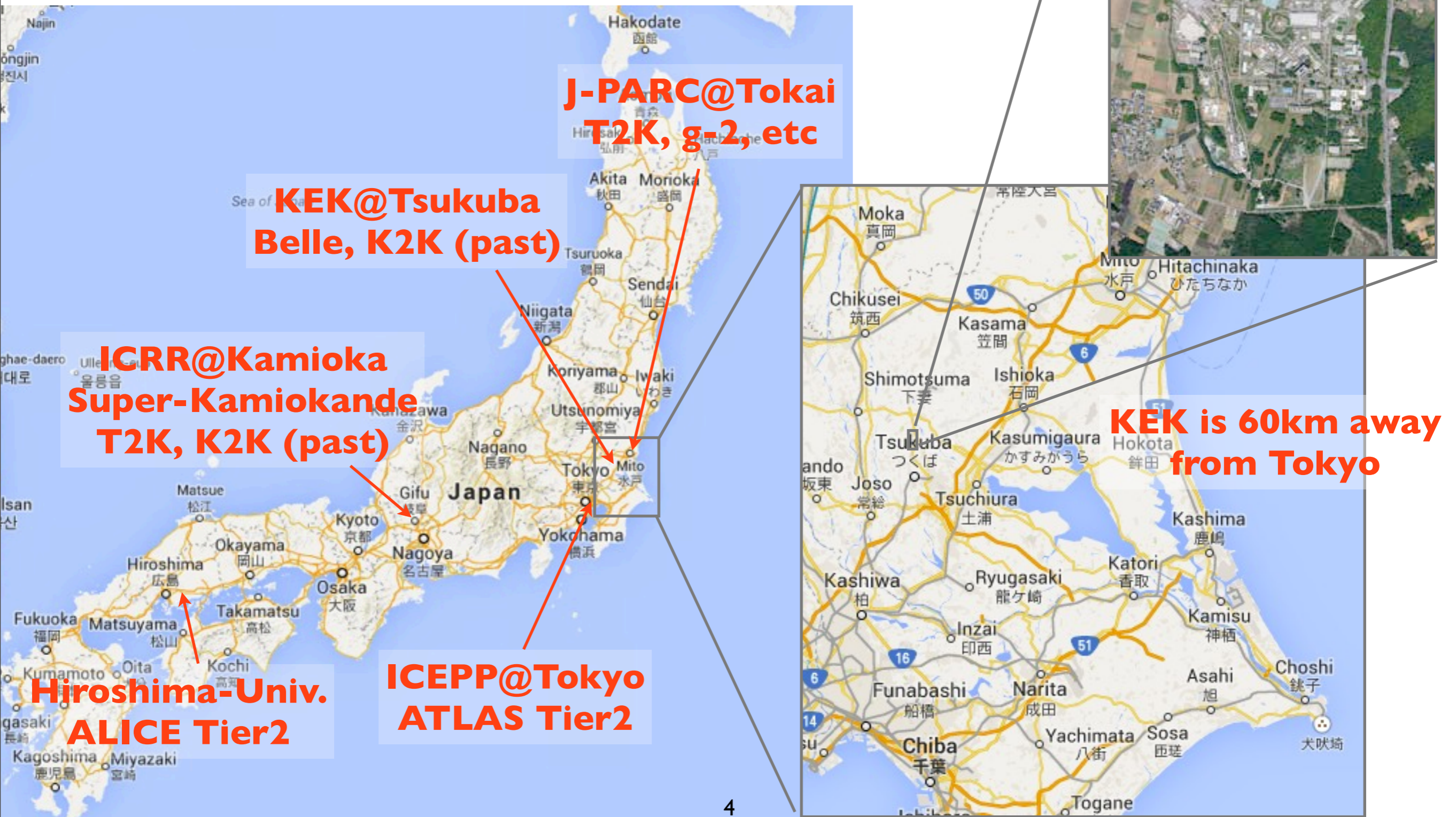


Facilities for experiments

- Accelerator to provide beam
- Detector to take data using the beam
- Computer to analysis the data
 - ▶ Huge data from accelerator experiments need computer for analysis
 - ▶ Once accelerator starts the operation, computer must not stop recording data by their problem



Major international collaboration in Japan



AS2505 HEPnet-J

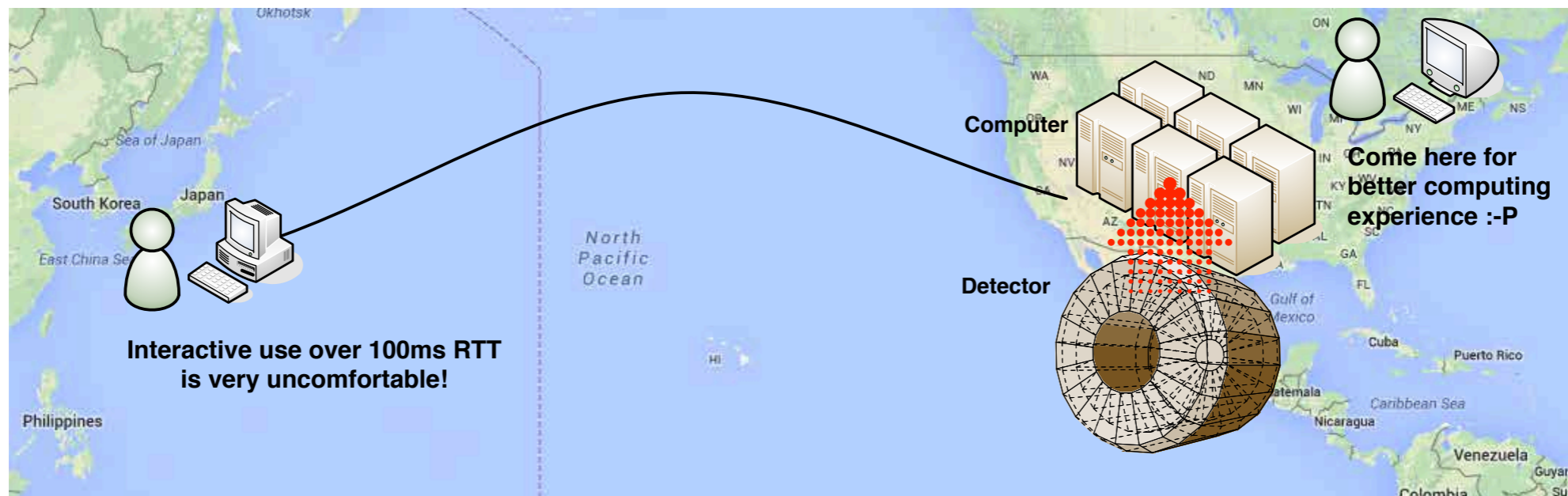
- HEPnet-J connects many HEP groups of universities in Japan
- Their IP prefixes are independent from their campus network.
 - ▶ but the primary network of their campus network is SINET
- Gate to Internet is placed in KEK
 - ▶ operated by KEK staff
 - ▶ Formerly, it has several international links. (to US, TW, CN, SU), but most of them are already shutdowned or replaced by VCs by NREN.

Really brief history

- ~2000: KR-JP link of IMnet through SINET
 - ▶ Priority is keeping reachability **Dedicated Link age**
 - ▶ but frequently the route from SNU to KEK went to US
- 2000~2003: ATM VC between APAN-JP and KEK
 - ▶ priority is keeping better path for short RTT **ATM age**
- 2003~2005: routed as ordinary Internet **Unfortunate age**
- 2005~2014: VC by JGN-X between APAN-JP and KEK
- 2014~: VC by SINET between APAN-JP and KEK

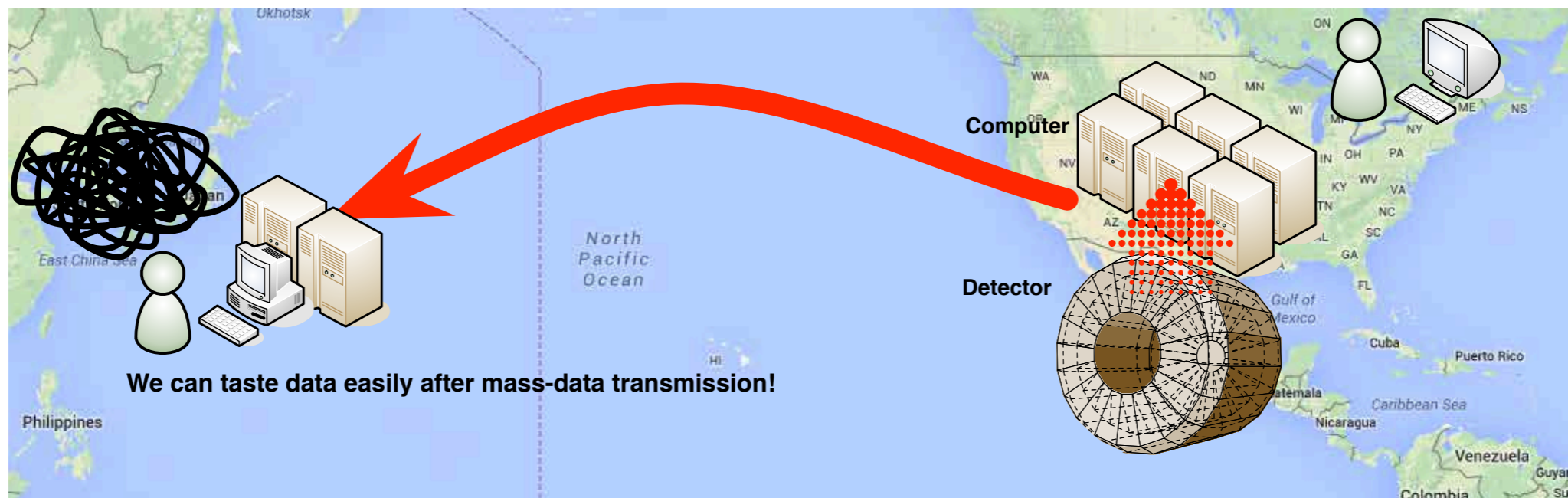
Formerly, it was centralized

- The experiment site should provide all of them.
 - ▶ All of the data are stored and preserved at there.
 - ▶ Collaborators may use computer remotely from their home institute,
 - ▶▶ Requires world wide network
 - ▶▶ Protocol was DECnet (~1990s) + Internet (~ now)
- Or copy skimmed (most important) part of data to their institute.
 - ▶ by tape cartridges
- Network was mainly for communication of collaboration and remote access.



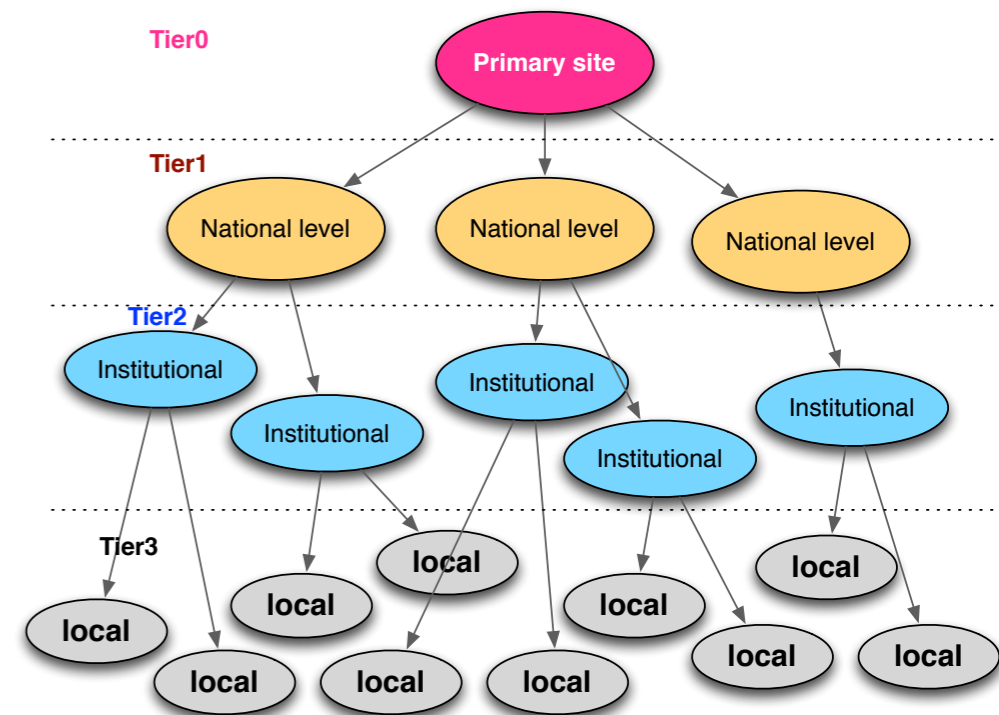
Wide-band network has changed it

- Skimmed data can be copied via network
 - ▶ Collaborator can timely and quickly analyze using the skimmed data at their institute.
- Abroad collaborators predicted they can analyze by computers nearby them, no need to wait for the long batch queue on the host site!
 - ▶ Soon after that, the long latency of trans-pacific networks destructed the beautiful dream.
 - ▶ There was no handy application to transmit data over shared networks whose RTT is longer than 300 msec.
 - ▶ It was solved by Grid systems such as LHC Computing Grid or Open Science Grid.



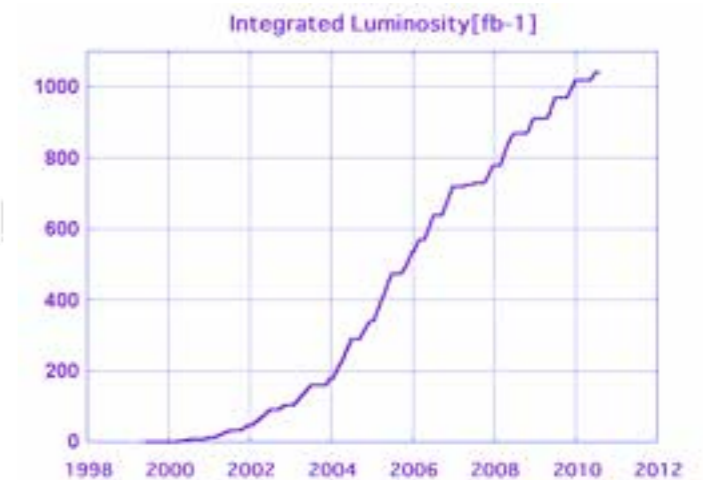
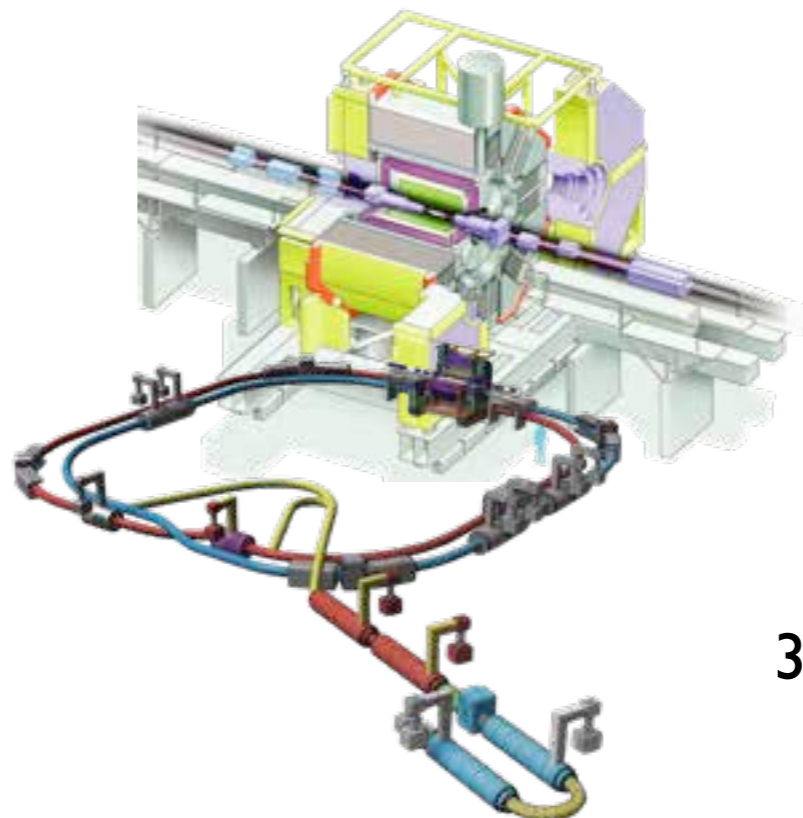
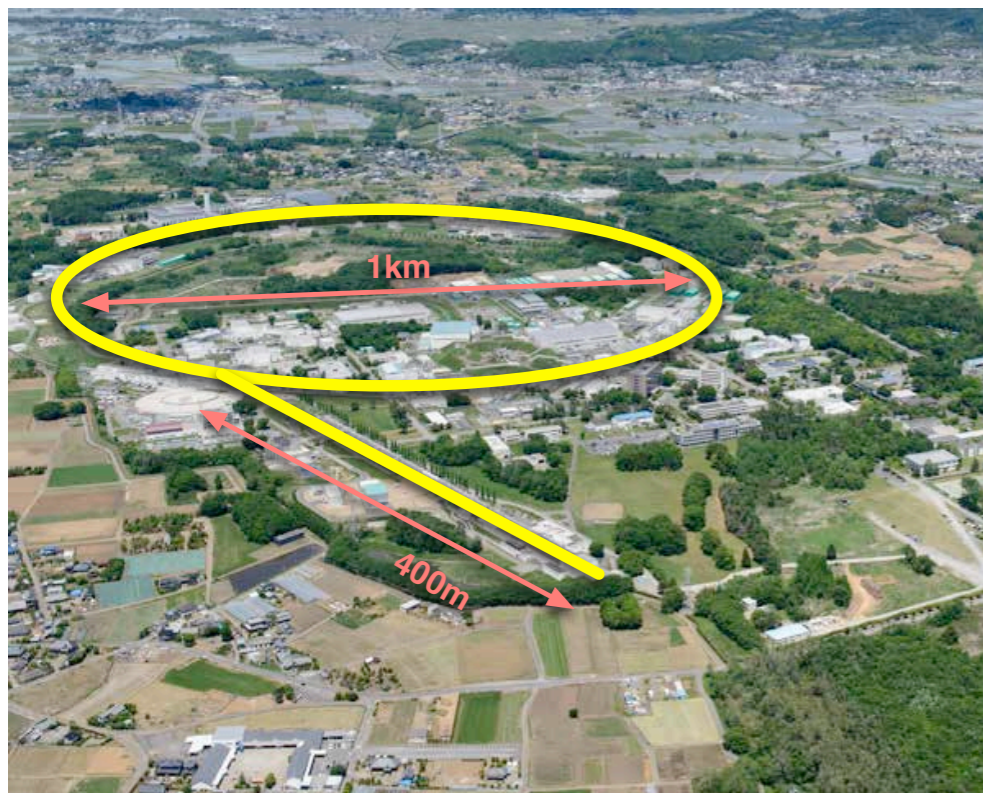
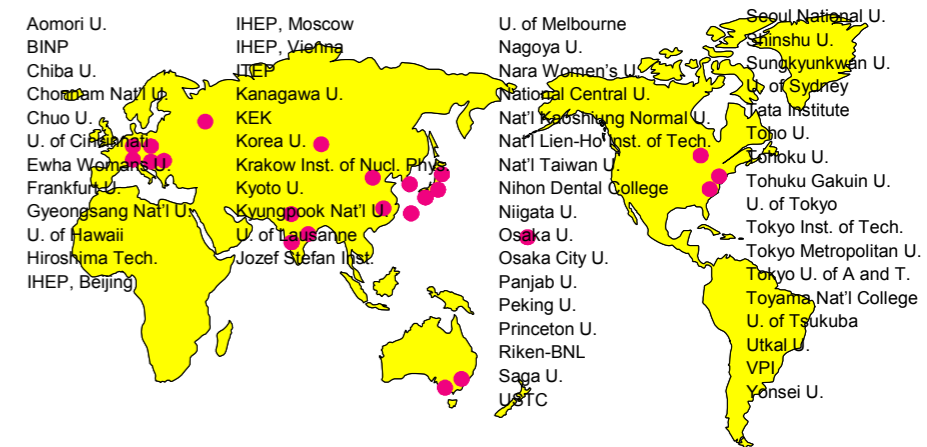
De-centralization

- The scale of experiments becomes larger and larger
 - ▶ Needs of computing resources also
- Single institute can't provide sufficient computing resources only by itself.
 - ▶ Data will be scattered and preserved to many collaborator sites.
 - ▶▶▶ CPU power + Large storage
 - ▶▶▶ **High speed network connection to Internet**
 - ▶▶▶ Operational human power
- By the LHC experiment which is a largest collaboration in HEP, it's tier-structure becomes popular.
- KISTI has a Tier-1 for ALICE of LHC experiment



Belle Experiment in KEK

- Research on violation of the symmetry between particles and anti-particles using large number of B-mesons. (1999~2010)
- 13 countries, 57 institutes, 400 collaborators
- Centralized computing model
- Data rate was 15MB/s ~ 30MB/s in average

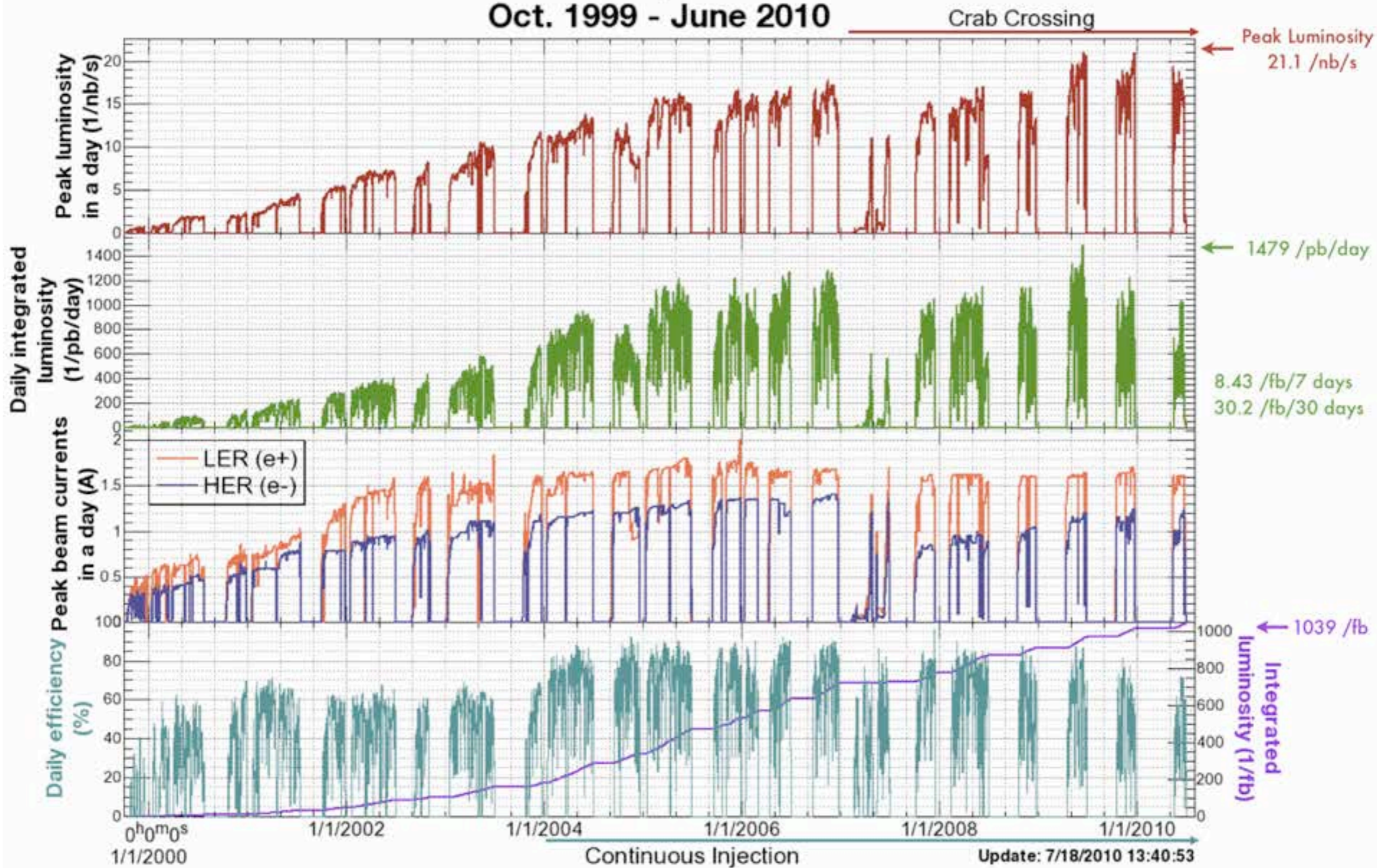


3.5PB tapes were consumed at the exp. end

History of computing facility for Belle

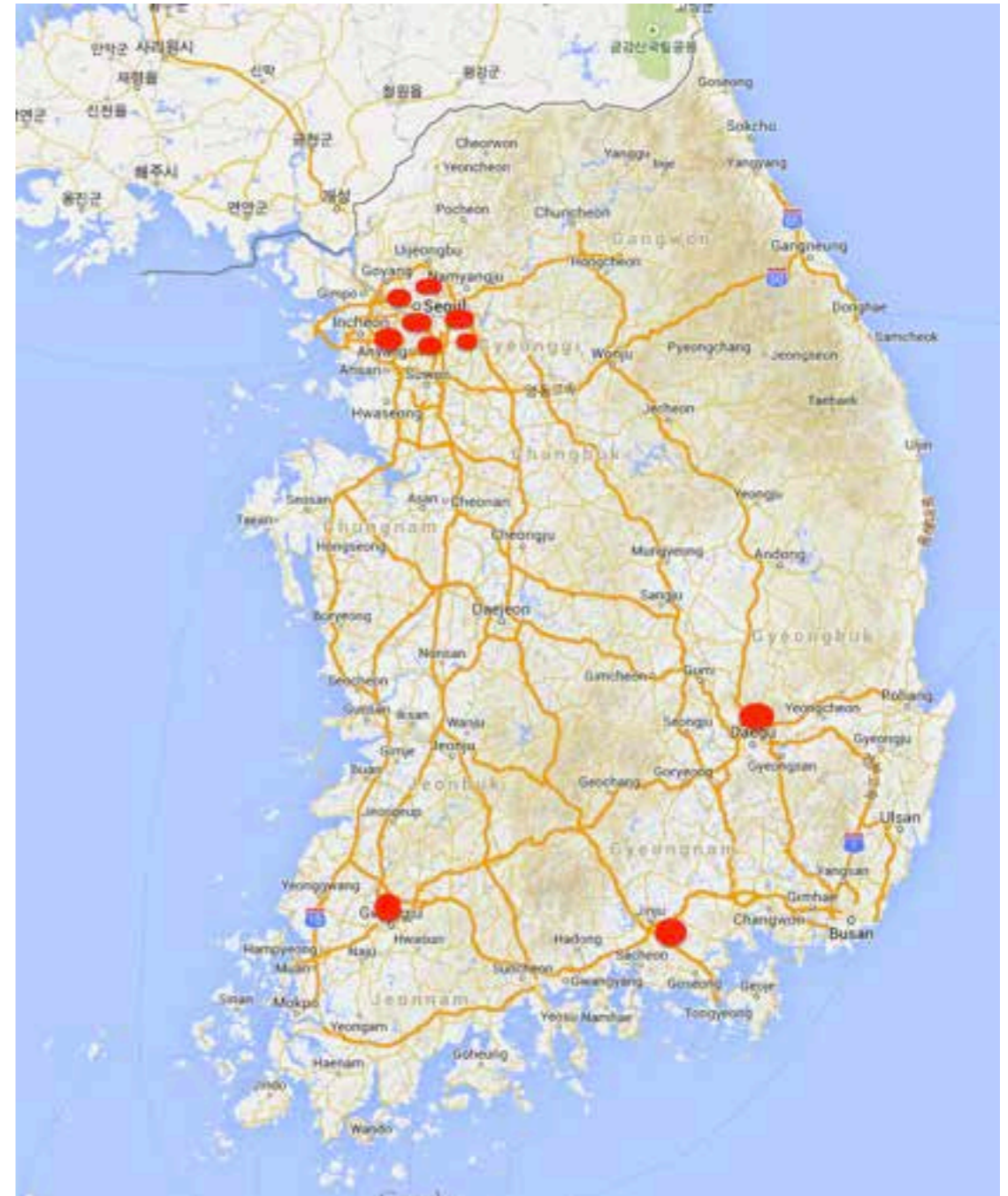
	1999	2001	2006	2009
CPU (spec cint2000)	~100 (SPARC WS)	~1200 (WS+PC)	~42500 (PC)	~115200 (PC)
Disk (TB)	4	9	1000	1500
Tape (TB)	160	620	3500	3500
# of Servers	4	11	96	96
Servers for Interactive use	28	23	128	128

Luminosity of KEKB Oct. 1999 - June 2010



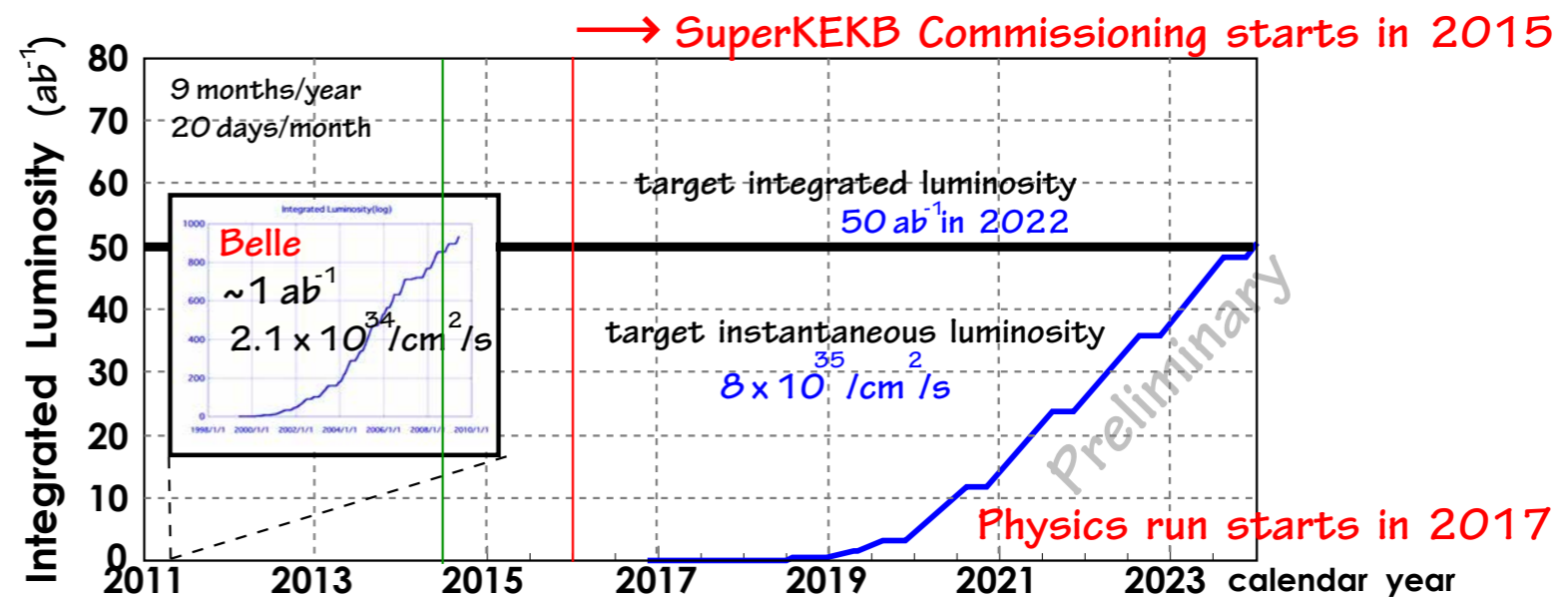
Belle collaborators from Korea

- Chonnam National Univ.
- Gyeongsang National Univ.
- Hanyang Univ.
- KISTI
- Korea Univ.
- Kyungpook National Univ.
- Seoul National Univ.
- Soongsil Univ
- Sungkyunkwan Univ.
- Yonsei Univ.



Belle II experiment

- Accelerator will be upgraded to provide the 40 times higher luminosity.
- Impossible to provide sufficient resources only by KEK,
 - Currently, existing CPU power is almost exhausted.
- institutes will store a certain amount of data.

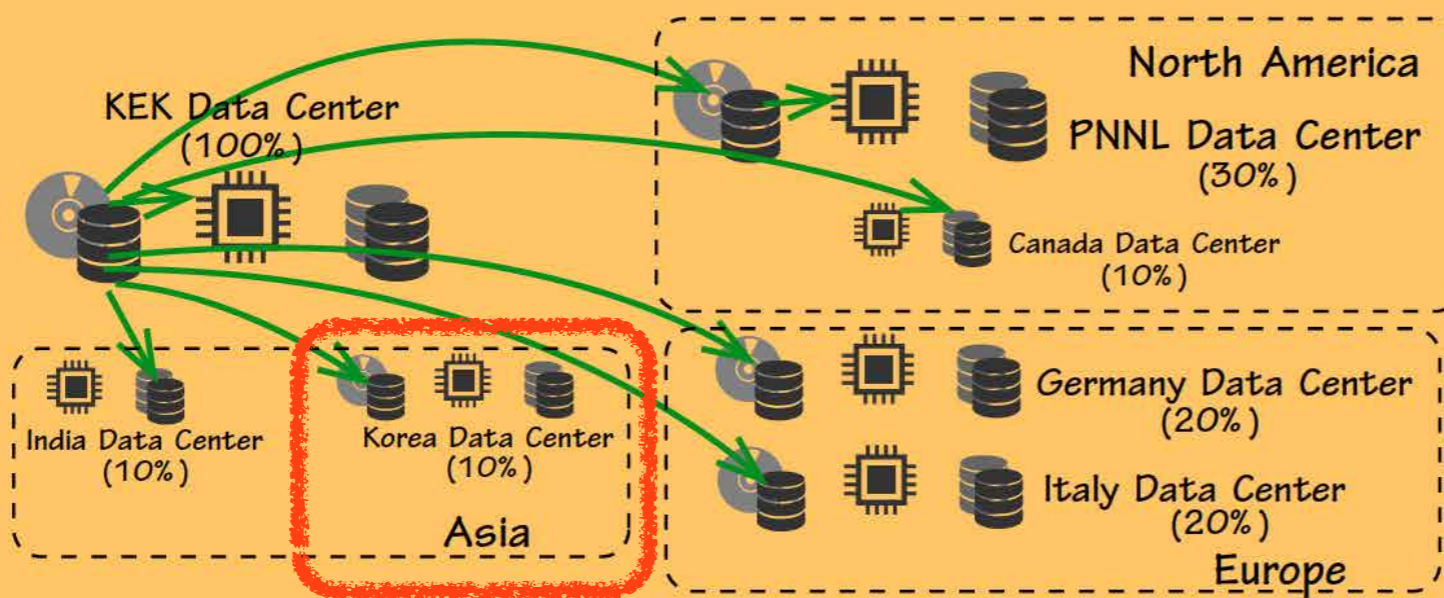


Raw data distribution

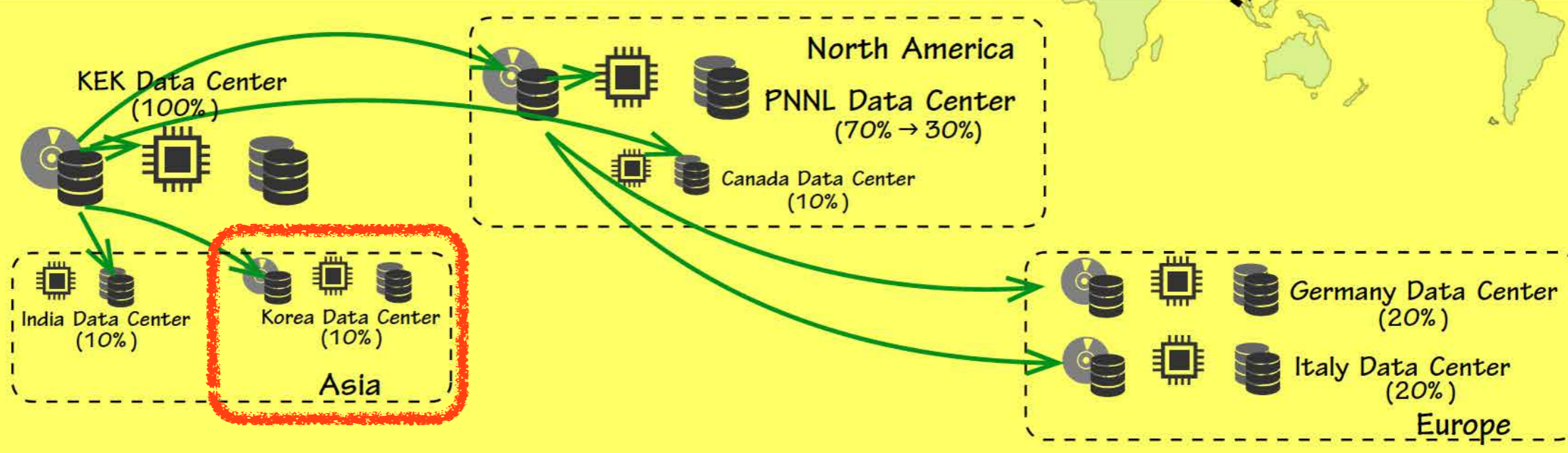
until Year 3



Scenario 1
(copy from KEK)



Scenario 2
(2step copy, KEK → PNNL → Europe)



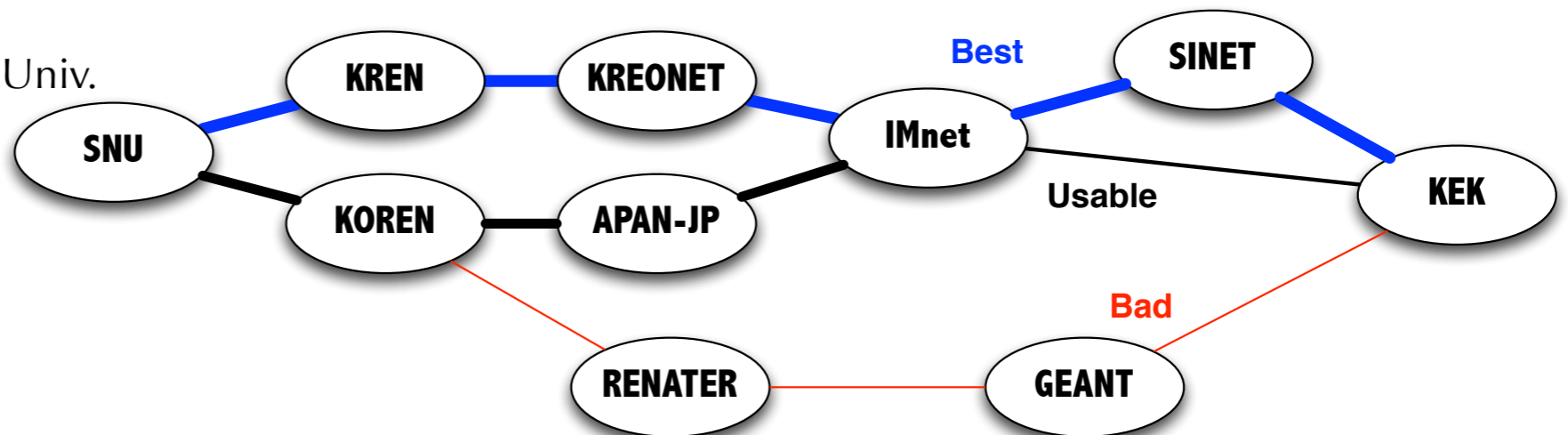
A few PB storage area is needed for 10% of rawdata

Interactive use and file transfer

- In the case of JP-KR, the problem is not RTT, but bandwidth.
- In the former model, the activity between JP-KR was just interactive use of computer facility and most of the problem was just reachability.

▶ Sometimes route from KEK is mis-filtered and inaccessible from Korea

▶ e.g. Seoul National Univ.

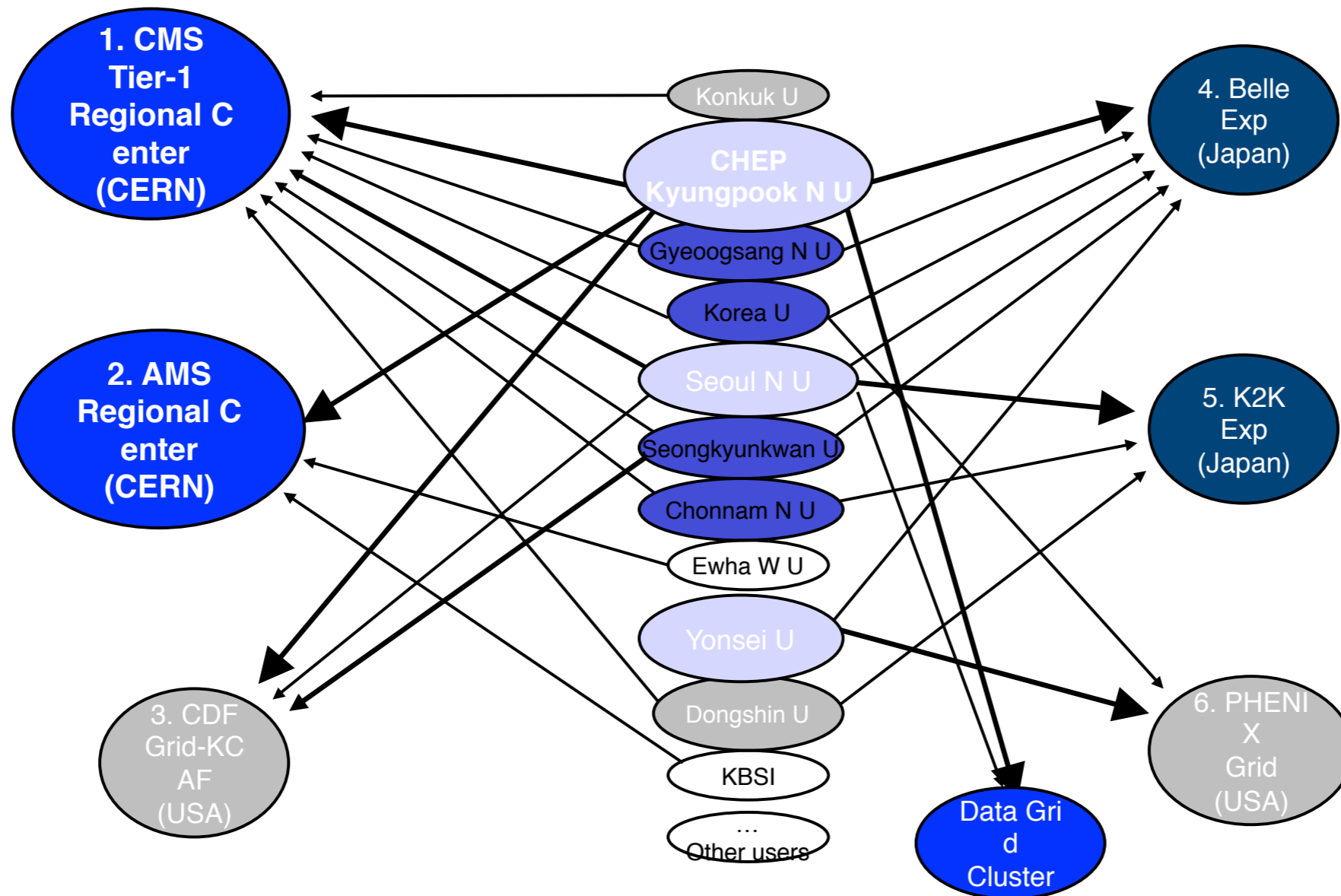


- Busan and Fukuoka 1 Gbps line opened at 2004, the situation was dramatically changed.
 - ▶ Center of HEP (CHEP) in KNU tasted the performance of file transfer.
 - ▶ CHEP - APAN-KR - QGPOP - SINET - KEK

Relation reported at APAN2003




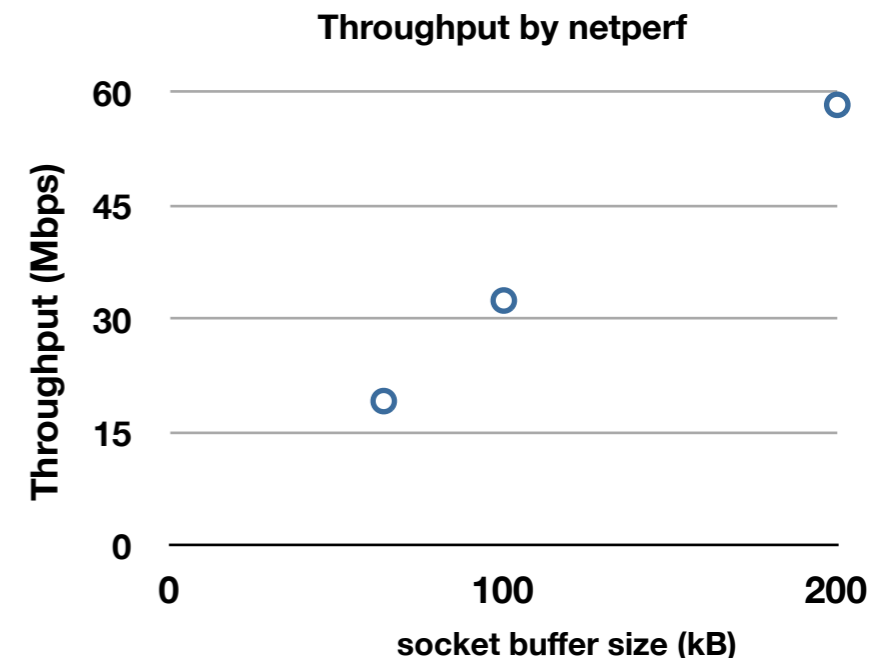
Participation of Institutions in the HEP Data Grid Project



By Youngdo Oh,
at APAN2003

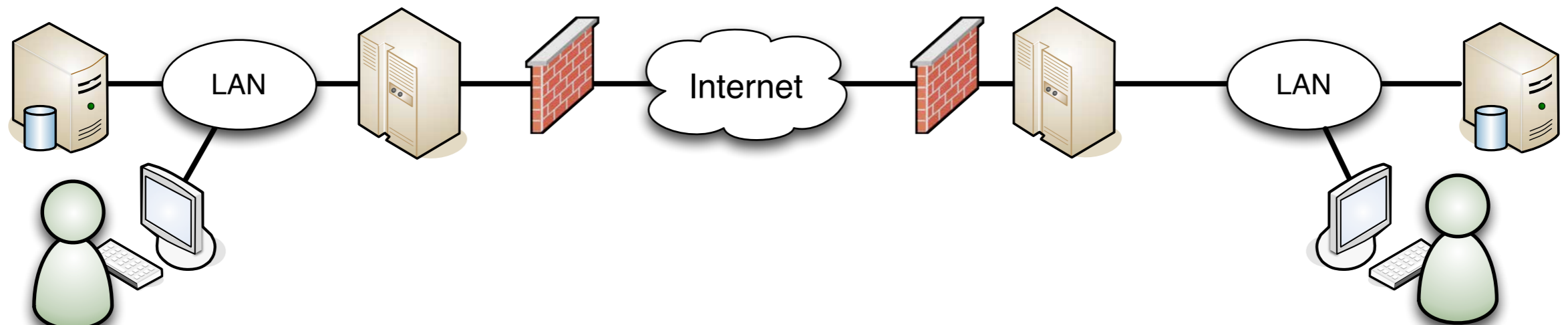
CHEP(KNU) - KEK performance

- RTT: 
- Netperf showed the necessity of window size extension
- Multi-stream application achieved better speed.
 - ▶ bbcp achieved 110Mbps by 10 streams although it wasn't popular.
 - ▶ HPN patch for SSH didn't yet appear.
 - ▶ Manual and occasional file transfer
- Already major labs in EU,US was moving to use grid-ftp.



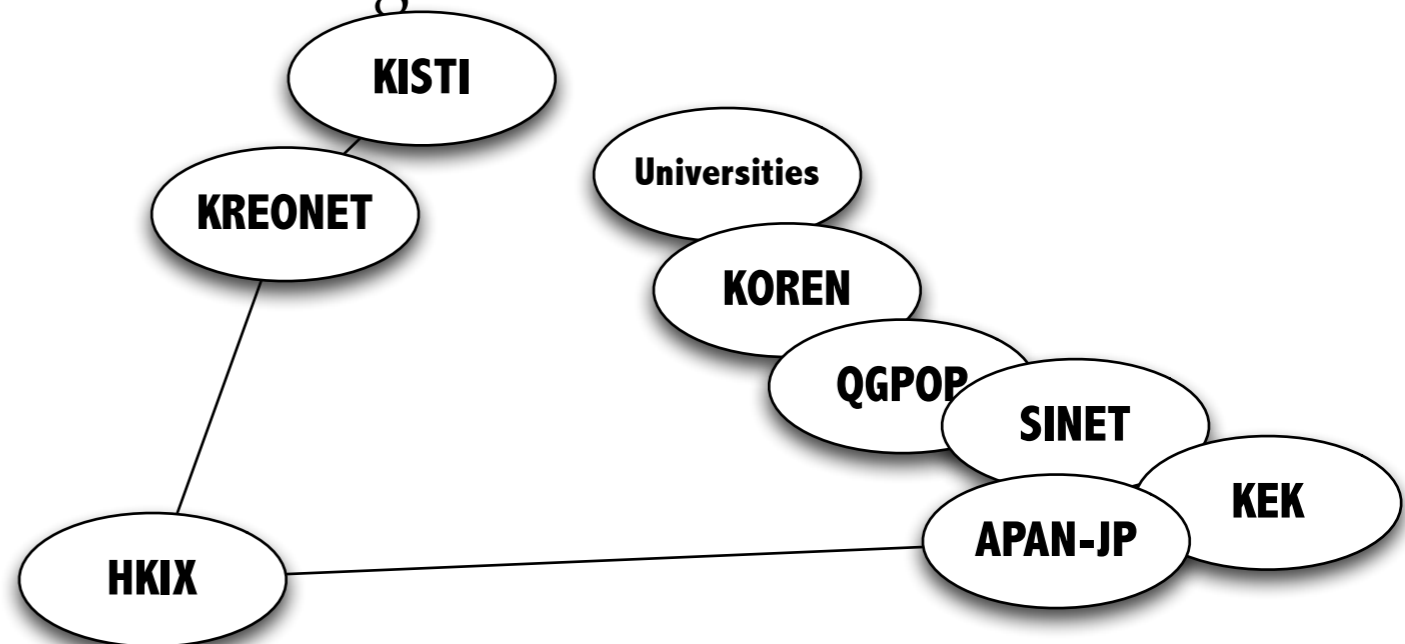
scp is not so fast, but...

- Still scp and rsync are popular for manual and occasional file transfer, but they are slow over international links.
- Why?
 - ▶ Many users can't expose their own host to Internet, so unusual new application are not available at the login server.
 - ▶▶ filtered by firewall controlled by IT section in their institutes.
 - ▶ If both users have grid resources, grid-ftp is usable. But not always so.
- Few people know HPN patch is effective, but most of sites do NOT apply the patch to their production system to avoid security risks related to SSH.
- The Belle experiment **was NOT** in the grid world at that time.



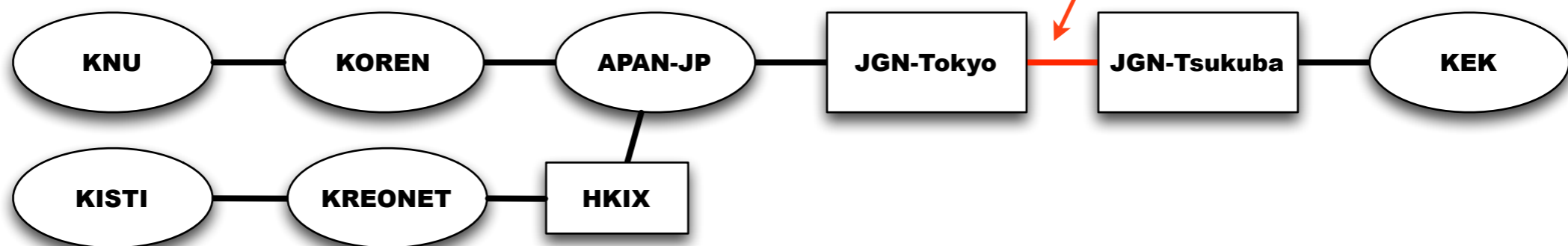
KISTI

- A HEP group born in KISTI and joined to Belle in 2008
- Belle group planned a data replication
- KISTI is in KREONET and route from KEK went to HKIX
 - ▶ relatively farther than sites in KOREN
- Route is KISTI - KREONET - APAN-JP - KEK
- HPNed SCP achieved ~ 420Mbps
- Never reached to 1 Gbps, so precise QoS was indispensable. It is quite un-welcome for the server managers.
- We observed severe packet loss higher than that, but **where is the bottle neck?**

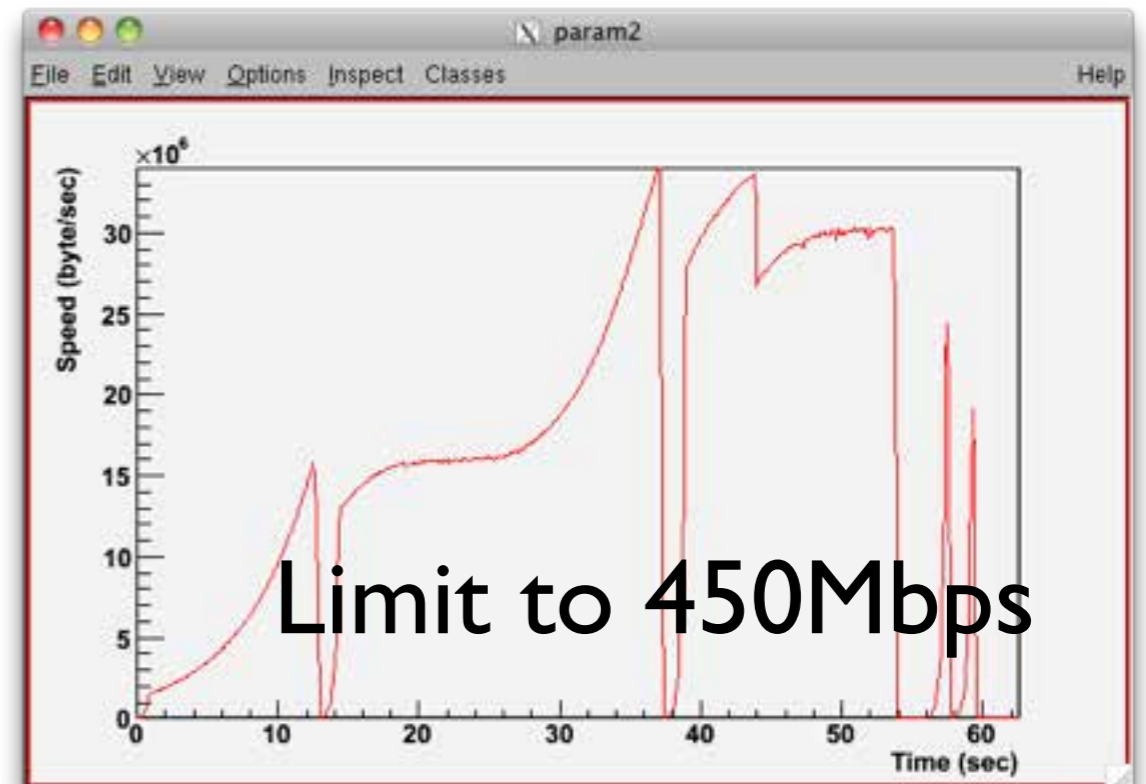
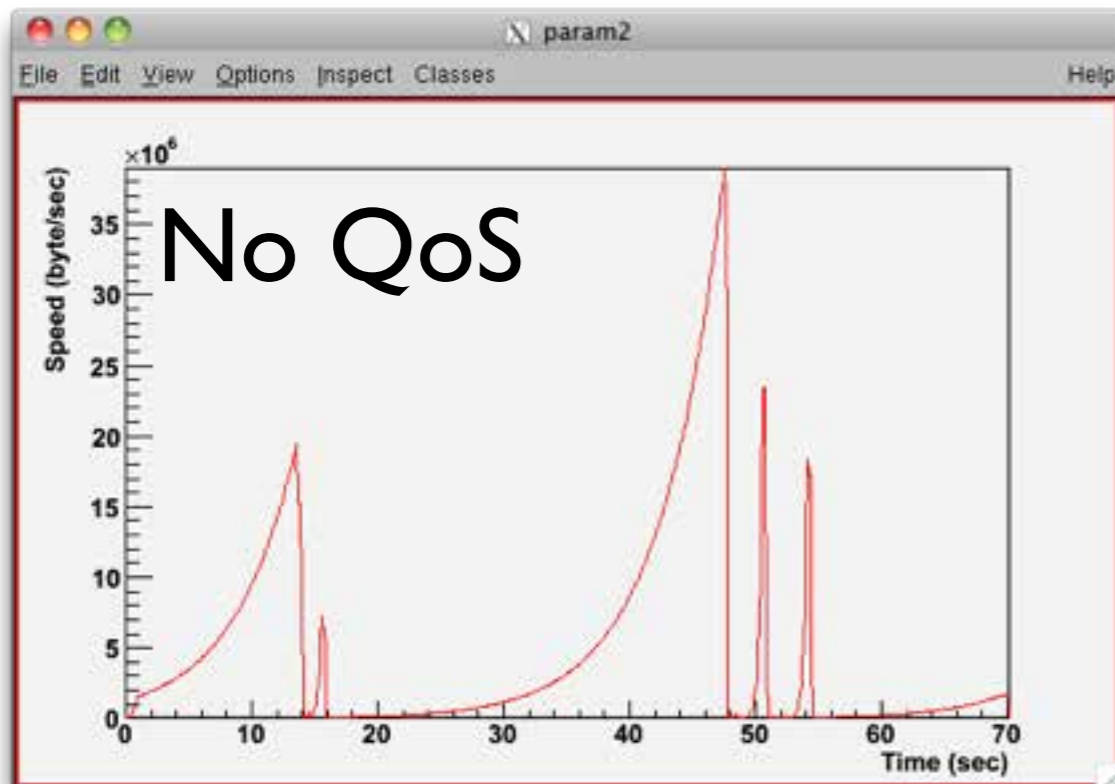
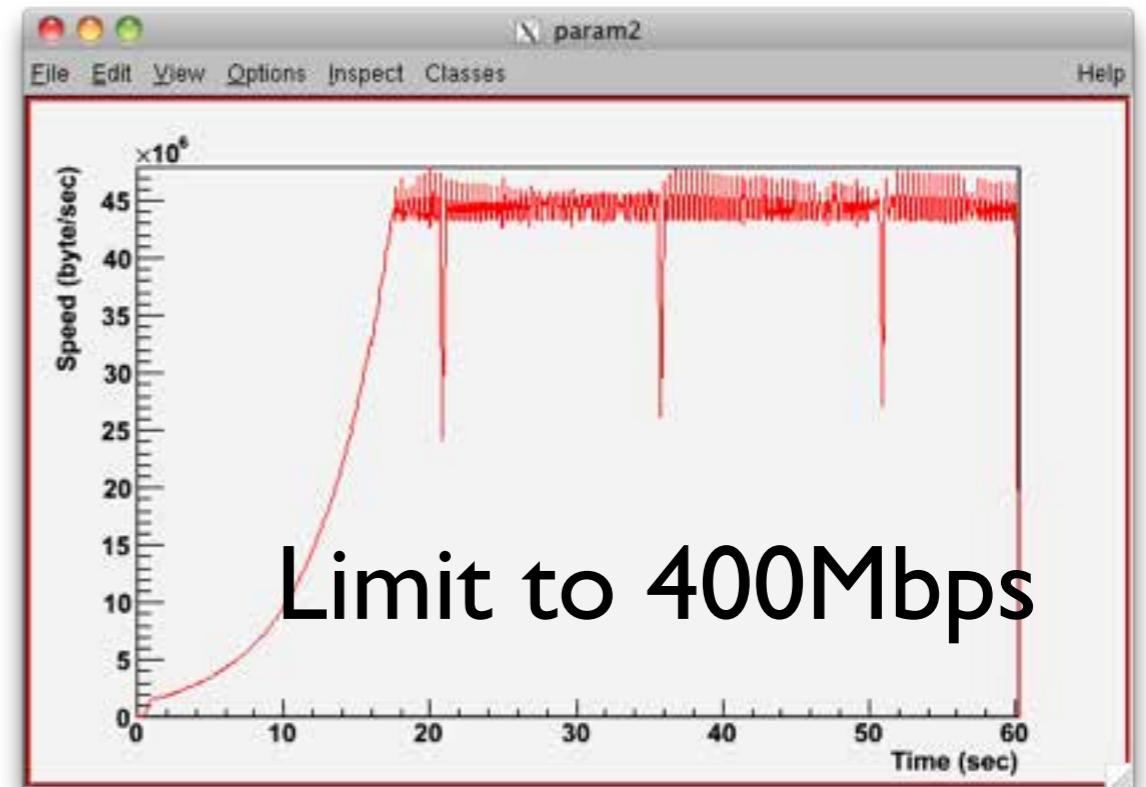


It was in Japan

- The congested link is Tsukuba-Tokyo link.
- This link is used for the peer for KEK and APAN-JP.
- Traffic to Universities in KOREN have same problem.
- Situation becomes worse year by year, recently we could achieve only 400Mbps.
- Traffic to KOREN is also damaged

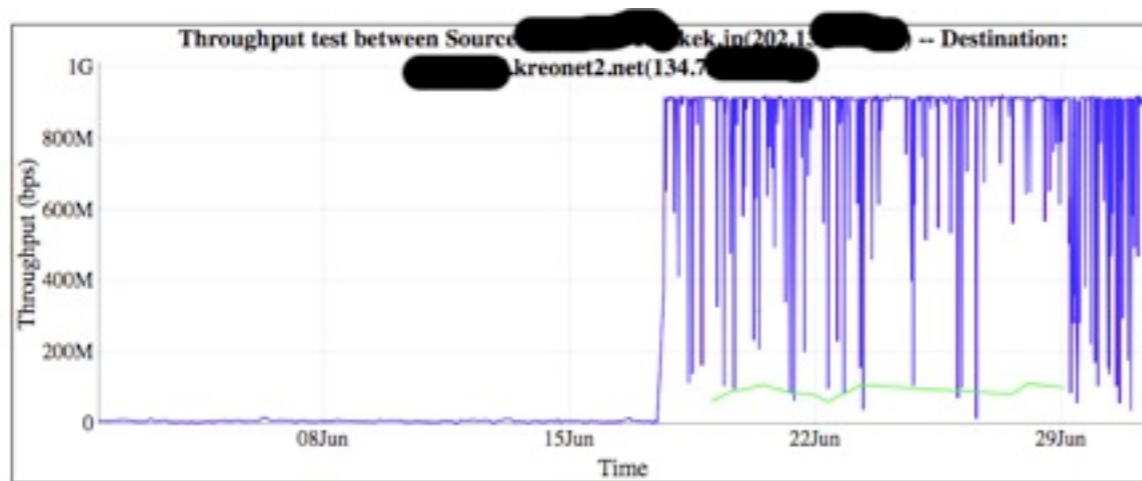


iperf test at 2010 from KEK to a node over KREONET2



Link replacement

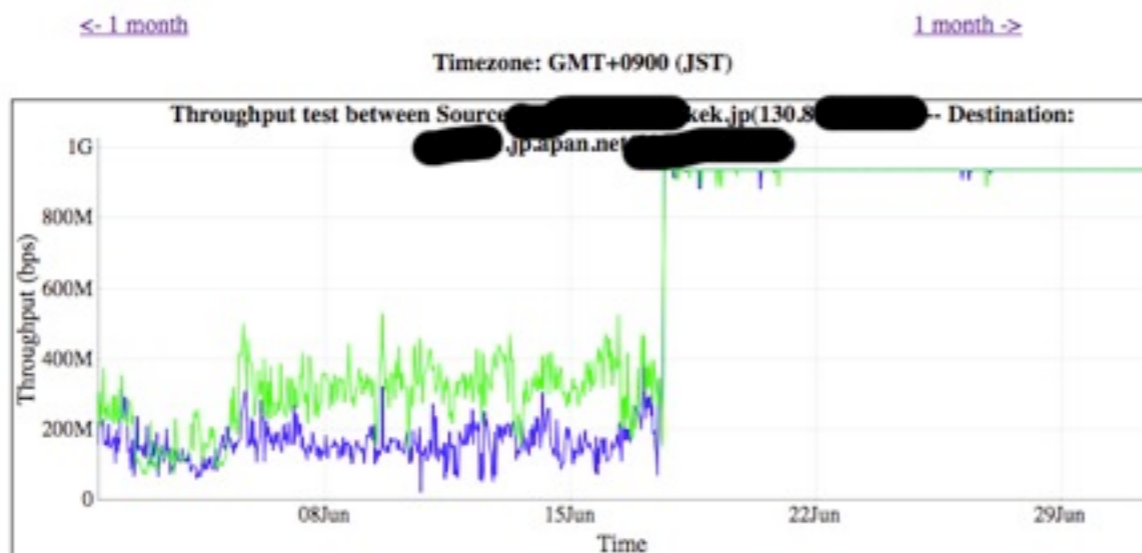
- At May 2014, We noticed KREONET will be reachable via JP-KR link.
 - ▶ Strong request to fix this jam condition
- The jammed link is superseded by a virtual circuit provided by SINET.
- The situation dramatically improved.



Graph Key

■ Src-Dst throughput
■ Dst-Src throughput

KEK - KREONET2

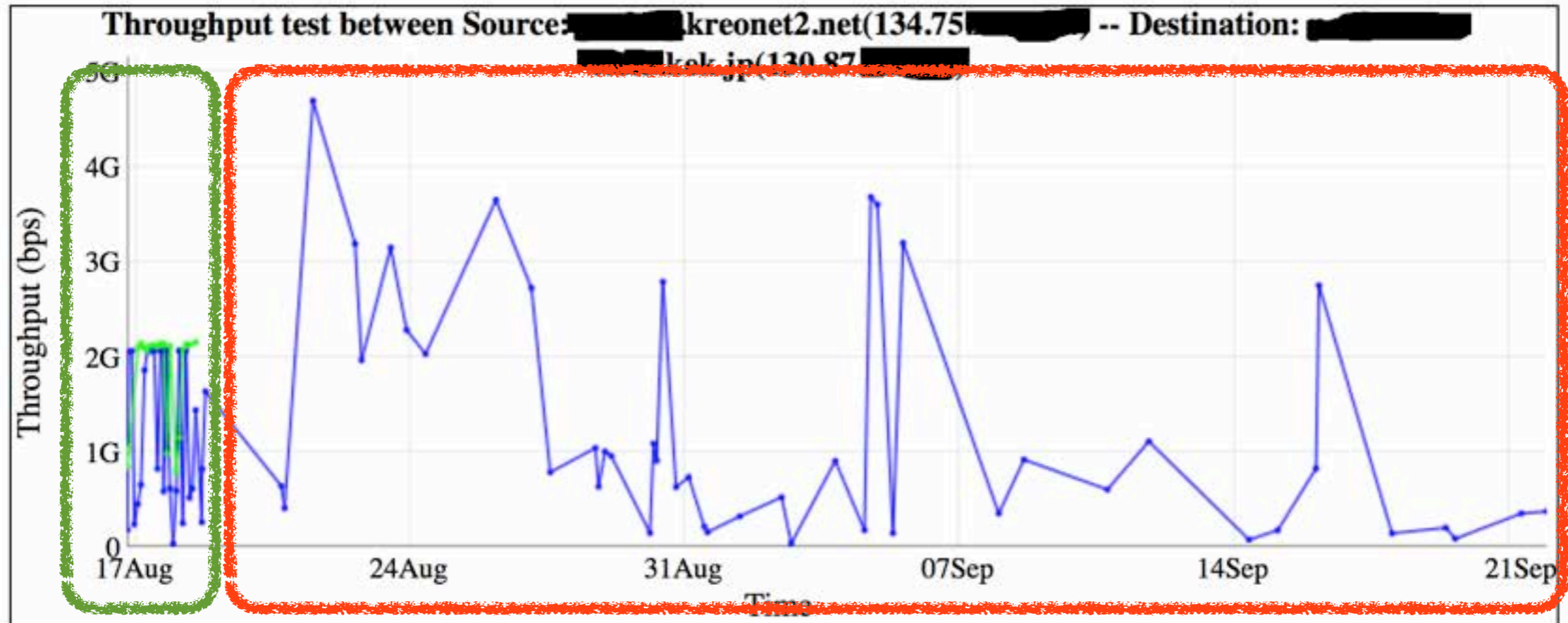


Graph Key

■ Src-Dst throughput
■ Dst-Src throughput

KEK - APAN-JP

Careless tuning is not good



Default installation of perfSONAR-PS

Window size extension for trans-pacific tests

Peak speed may be better, but average speed decreases.

Future prospects

- Belle II will start the operation and transmit rawdata to the data center in Korea.
- KISTI in KR and Hiroshima Univ. in JP are Tier-1 and Tier2 of the ALICE experiment in LHC respectively.
 - ▶ Data source site is automatically chosen by Grid middleware, files are transferred non-interactively.
 - ▶ Total bandwidth limited by facilities in Hiroshima University, upgrade is needed.
 - ▶ They are planning to join into LHCCONE to bypass firewall after few years.



Summary

- Current bandwidth between JP and KR is very helpful for our collaborations.
- Still main activity is interactive use and occasional file transfers, but several experiments are planning mass-data transfer.
- We deeply appreciate efforts for many years to improve the connection between Japan and Korea.