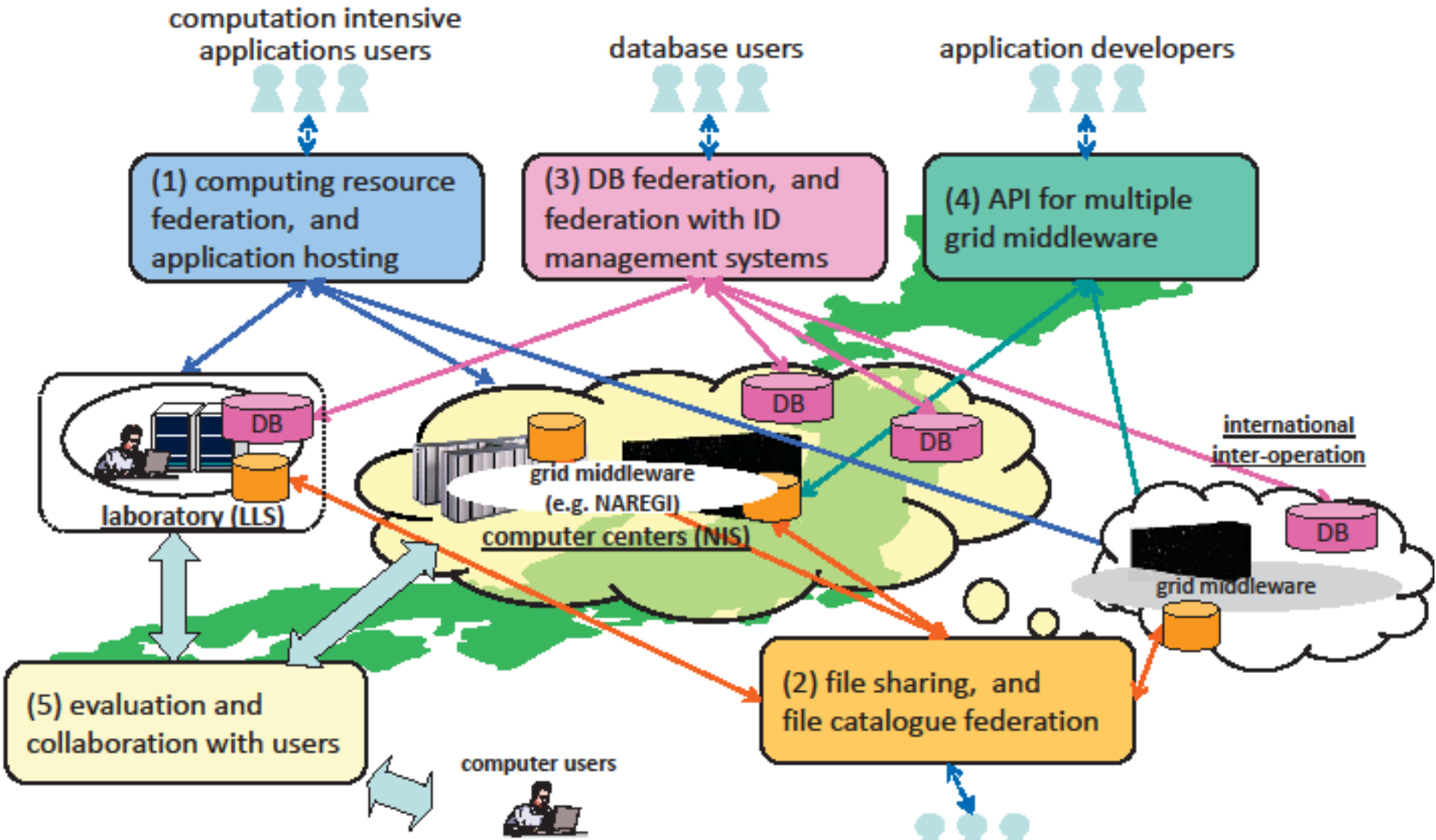
Asia-PoP : A proposal of the high throughput data sharing testbed

Go Iwai, KEK/CRC

With particular acknowledgements to Shin'ichiro Takizawa, TITECH

Courtesy K. Aida

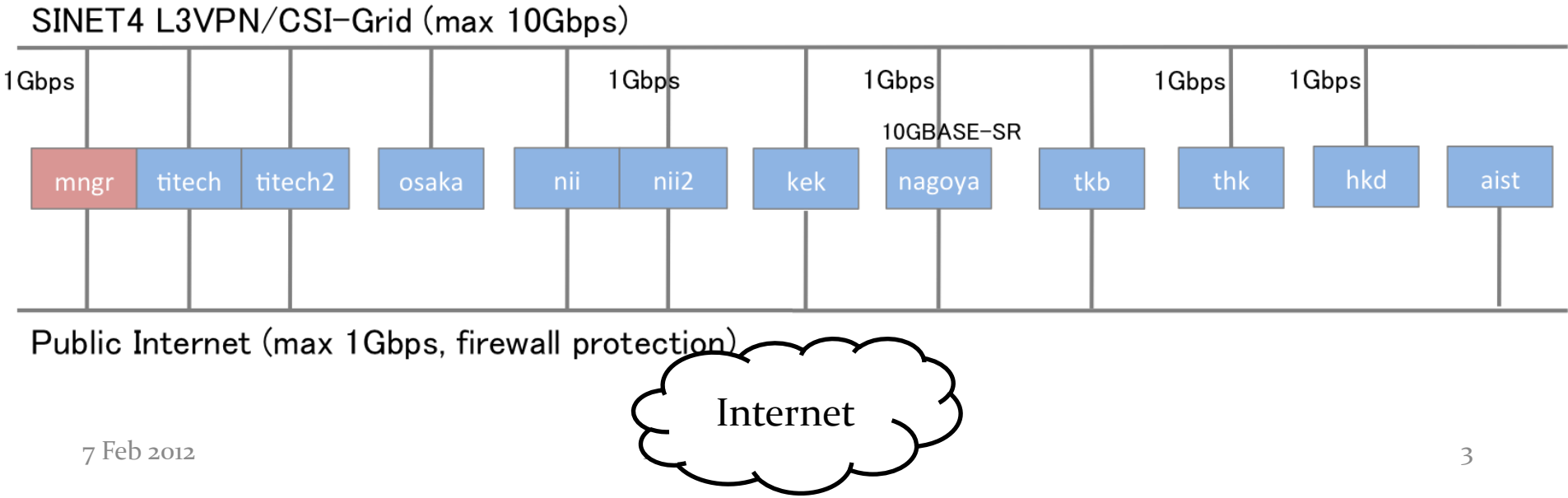
REsources liNKage for E-sclence (RENKEI)



RENKEI-PoP is a products of RENKEI collaboration

What's RENKEI-PoP

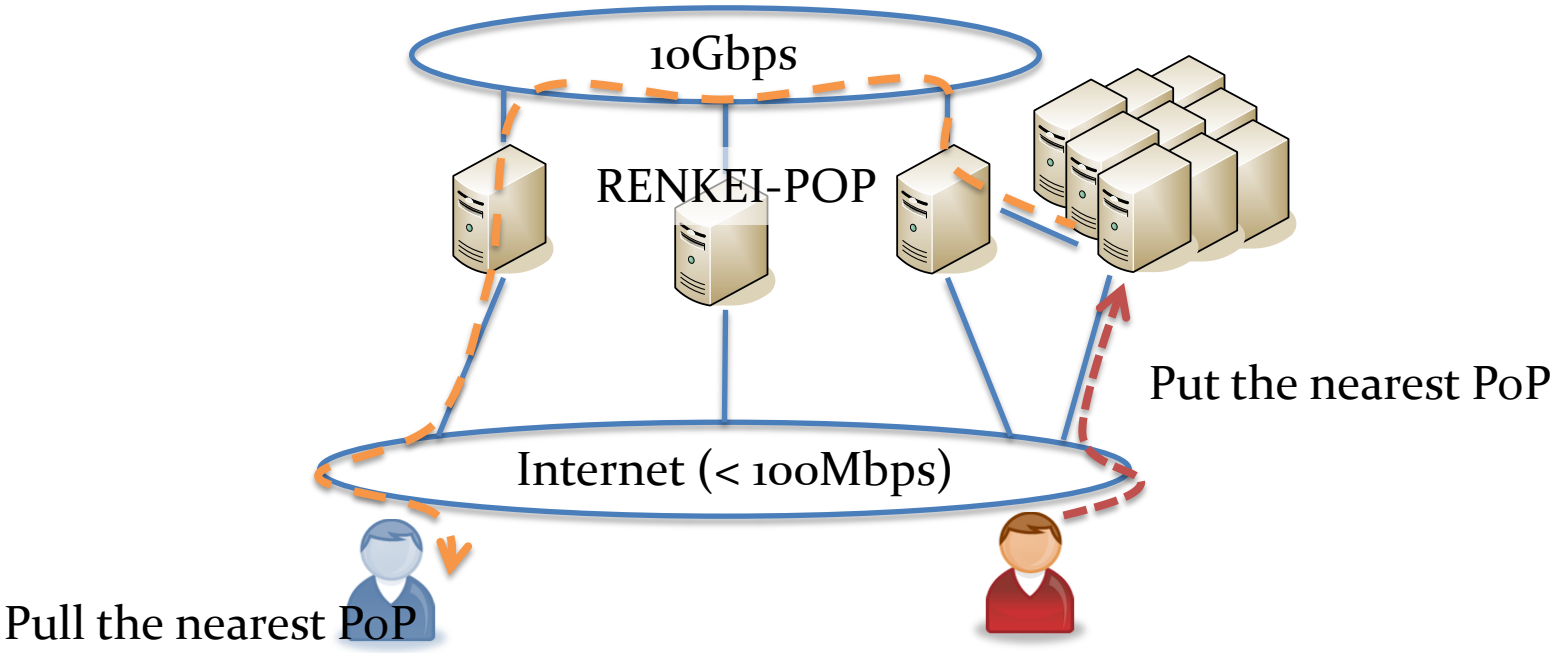
- An appliance for e-Science Data Federation
- Installed in computer centers as one of gateway servers and connected to a 10 Gbps high speed R&E network (SINET)



Deployed PoPs

Examples of PoP specification will be shown later

Simple Scenario of RENKEI-PoP

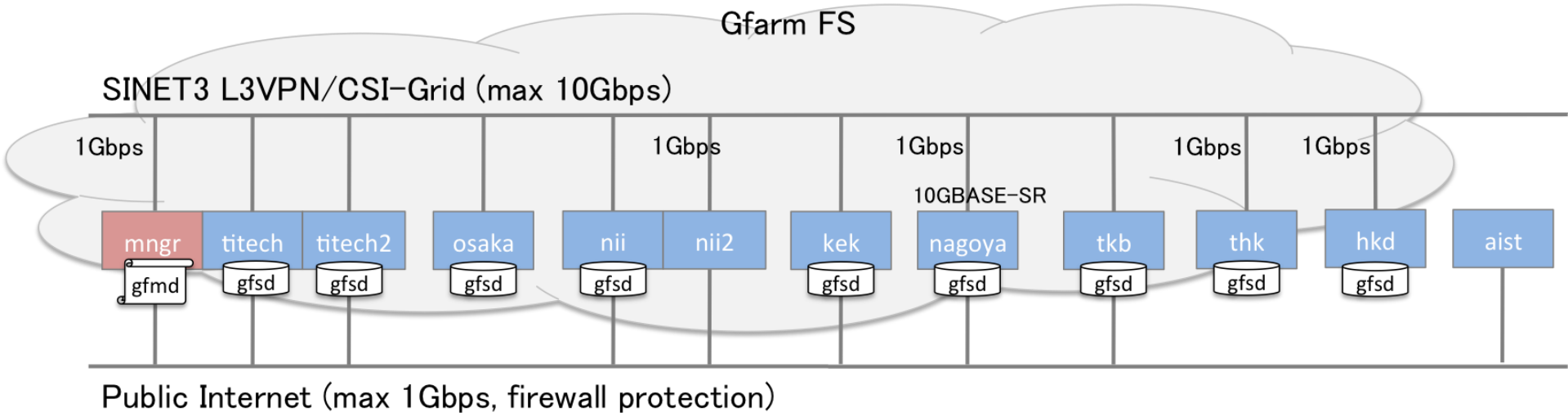


- Users write data in the nearest PoP
- Users read data on the nearest PoP

RENKEI-PoP Provides…

- file transfer/sharing between PoPs and between PoP and User's work machines
  - scp, GridFTP, gsi-ssh, Gfarm
- VM-based hosting service
  - A cloud-like hosting service based on OpenNebula

Gfarm – A middleware for network shared file system

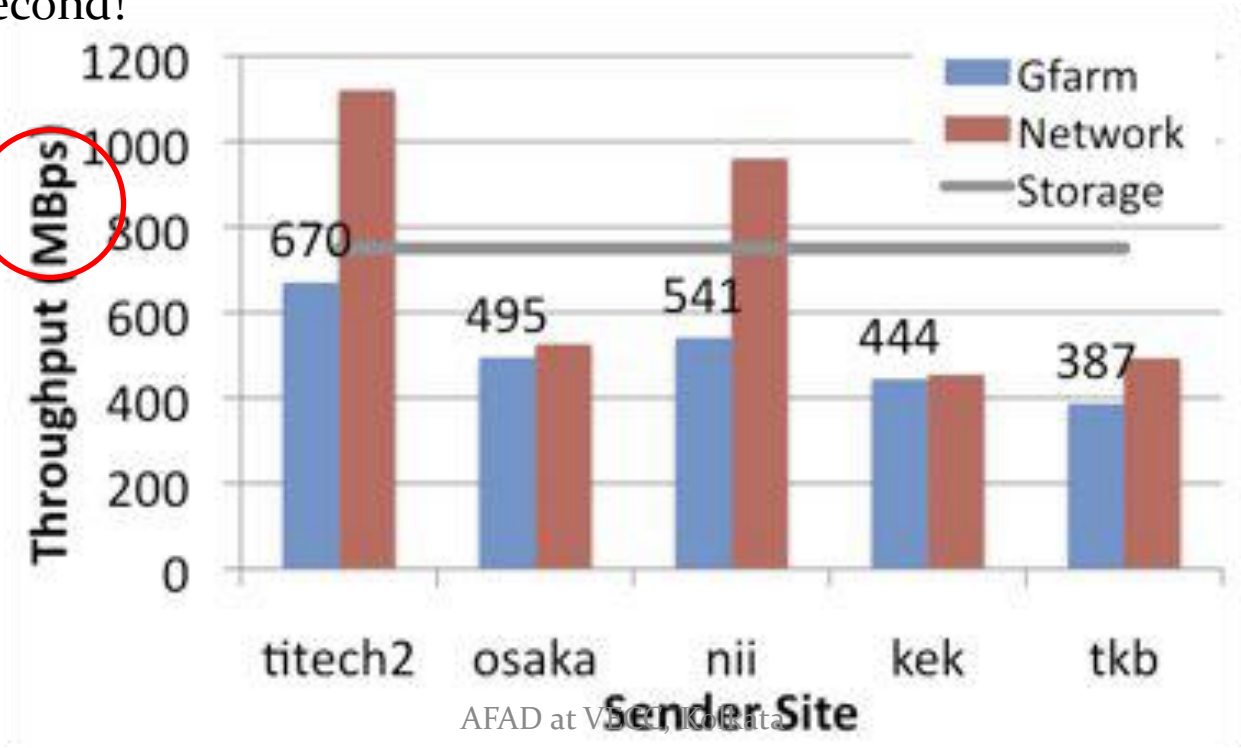



```
[iwai@rpop-kek ~]$ gfdf -h
1K-blocks    Used   Avail Use% Host
    27.2T  18.5G   27.2T   0% rpop-hkd.r.cc.titech.ac.jp
     222G  97.7G    124G  44% rpop-kek.cc.kek.jp
    27.2T   1.1T   26.1T   4% rpop-nagoya.cc.nagoya-u.ac.jp
    27.2T   378G   26.9T   1% rpop-nii.naregi.org
    27.2T  22.9G   27.2T   0% rpop-osaka.hpc.cmc.osaka-u.ac.jp
    27.2T  22.6G   27.2T   0% rpop-thk.r.cc.titech.ac.jp
    27.2T  85.0G   27.2T   0% rpop-titech.cc.titech.ac.jp
    27.2T  17.2G   27.2T   0% rpop-tkb.ccs.tsukuba.ac.jp
-------------------------------
     191T   1.8T    189T   1%
```

How Good?

- Move 14GB astronomy data from RENKEI-PoPs in other sites to one in Tokyo Tech
- Applied network tuning techniques are
  - Tune kernel TCP and flow control parameters
  - Configure some device specific properties, such as TSO, interrupt interval, etc.

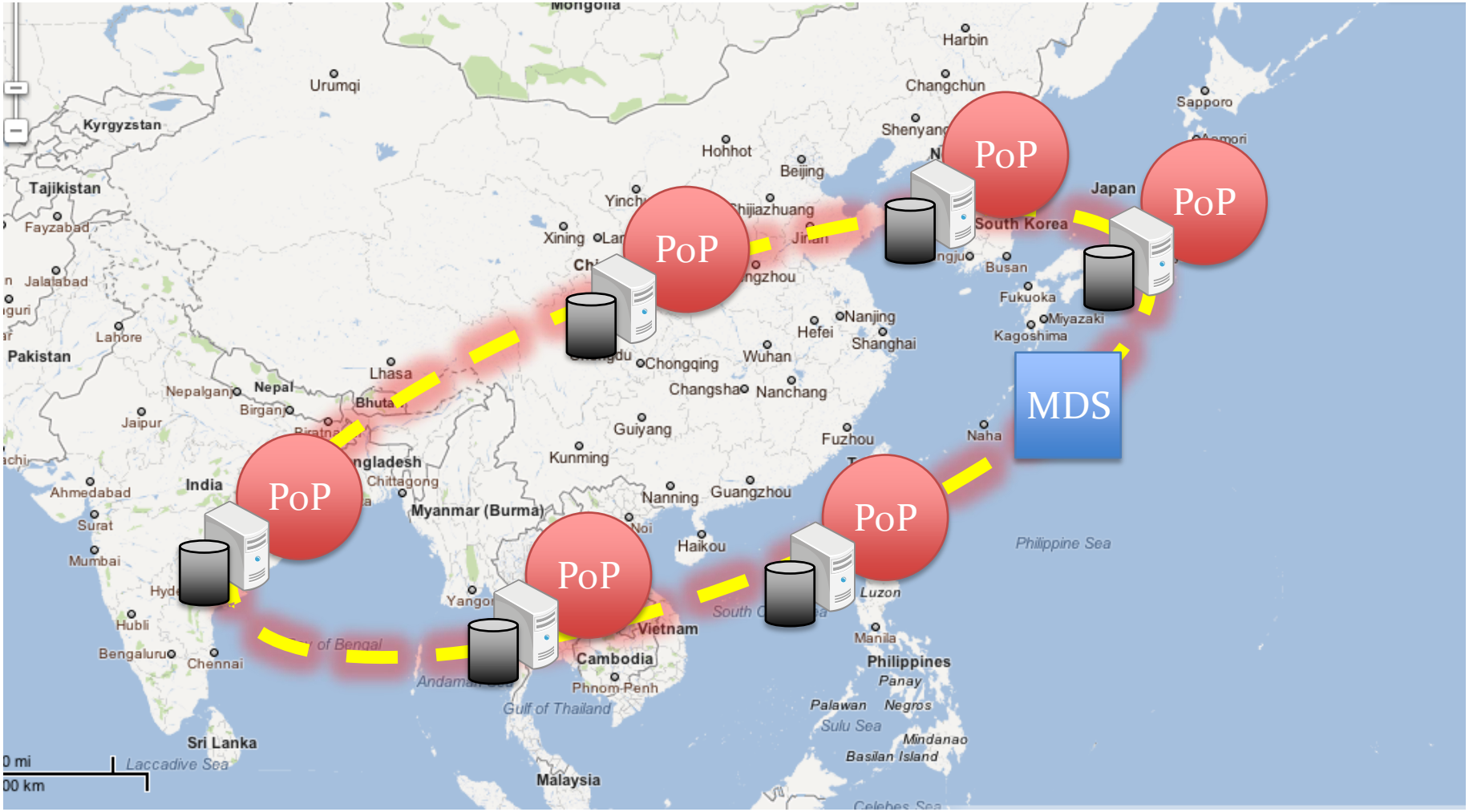
Mega Byte / second!



Asia-PoP: Storage Provisioning Service over the Asia-WAN

- Asia-PoP might be helpful for people who need:
  - An extra resource capability e.g. at system migration
  - A place for the important data to be replicated e.g. at the natural disaster
  - A wide-distributed testbed for data sharing but easy to control
- We don't aim to use this as a production service.

Rough Sketch of Asia-PoP



Software to be Installed

- Networked File System
  - pNFS: being more popular
  - Gfarm
- OpenSSH: Not fast but easy accessing
- GridFTP: Widely utilized

Specification Examples (Quoted from RENKEI-PoP)

- The peak performance for inter-site disk-to-disk transfer is 1 GBps (Giga Bytes per second)
  - Each RENKEI-PoP has a high throughput storage device that uses HBA to connect 8 to 16 SSDs or HDDs to achieve 1GBps IO throughput and has 10GbEthernet NIC for the remote site connection.
  - Inter-site transfer is supported by SINET L3VPN/CSI-Grid 10Gbps (Giga bits per second) high speed network.

Price Examples (Quoted from RENKEI-PoP)

Generation	Gen 1	Gen 2	Gen 3
CPU	Intel Core i7 965 Extreme (3.2GHz)	Intel Core i7 975 Extreme (3.33GHz)	Intel Xeon W3565 (3.2GHz)
Memory	DDR3-1333 12GB (2GB x6)	DDR3-1333 12GB (2GB x6)	DDR3-1333 ECC 24GB (4GB x6)
NIC	Supermicro(NetXEN) or Myricom 10GbE	Mellanox, or Myricom 10GbE	Myricom 10GbE
HBA	Adaptec RAID 5085, or Adaptec RAID 5805	Adaptec RAID 5805	Adaptec RAID 5805
Disk	Intel SSD X25-E Extreme (SLC) 32GB x8	2TB HDD x16	2TB HDD x16
Price	**US$ 18,000**	**US$ 11,000**	**US$ 12,000**

We aim in Asia-PoP...

- Easy-to-use even without special clients
- Fast but not special setup in users
- Better transfer performance and relationship in Asian research institutes

What do you think?

- If you have interests please comment me what you think.

Many thanks for your attentions!

go.iwai@kek.jp

Or goiwai on the Twitter and Facebook ☺