



# 100G Lambda networking in CERNET and IPv4/IPv6 traffic engineering

Xing Li

2017-09-27

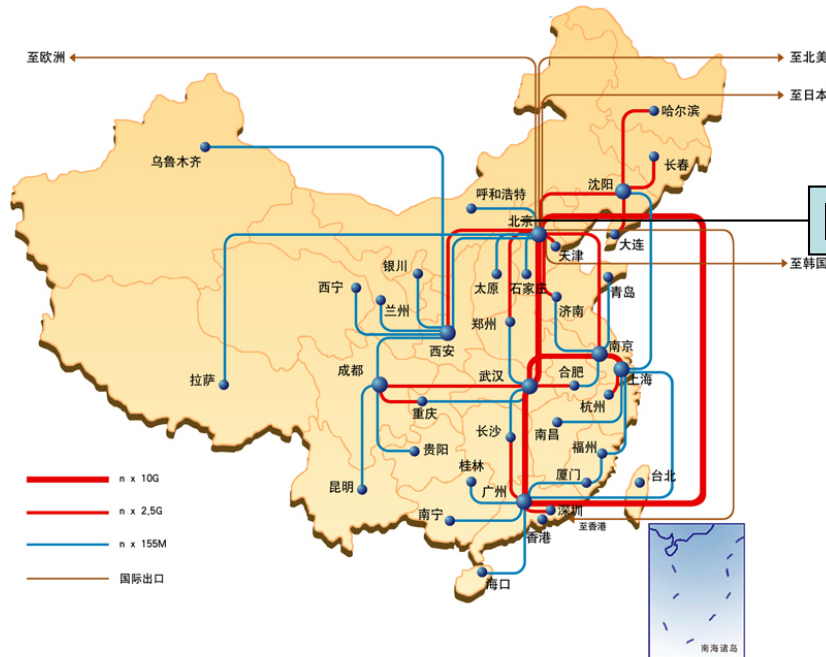
# Outline



- CERNET and CERNET2
- 100G experience
- Traffic engineering building blocks
  - Address switching
  - IVI
- Current projects
- Remarks

# CERNET Layer 3

## CERNET (IPv4)



## CERNET2 (IPv6)



- CERNET is the first (1994) nation wide Internet backbone in China.
- CERNET ranks 23 in global CIDR report.
- Over 2,000 universities on CERNET with about 20M subscribers.

- Built in 2004, with national coverage
- CERNET2 is the largest IPv6 backbone in China.
- About 300 universities connected to CERNET2 with about 3M subscribers.

# Peering

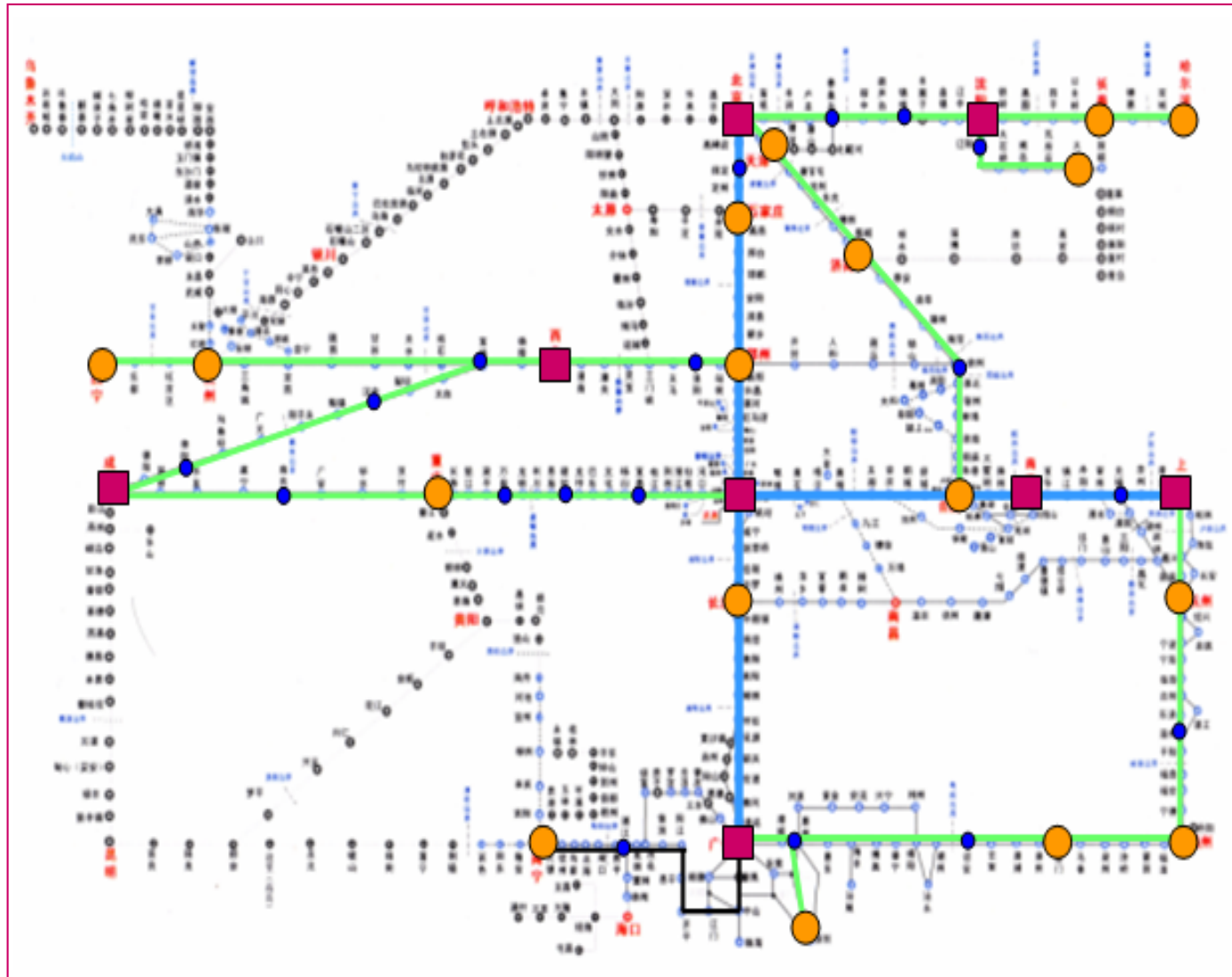


- Academic
  - Beijing  $\leftrightarrow$  Los Angeles 10G
  - Beijing  $\leftrightarrow$  London 10G
  - Beijing  $\leftrightarrow$  Hong Kong 10G (100G)

# CERNET fiber



# CERNET DWDM

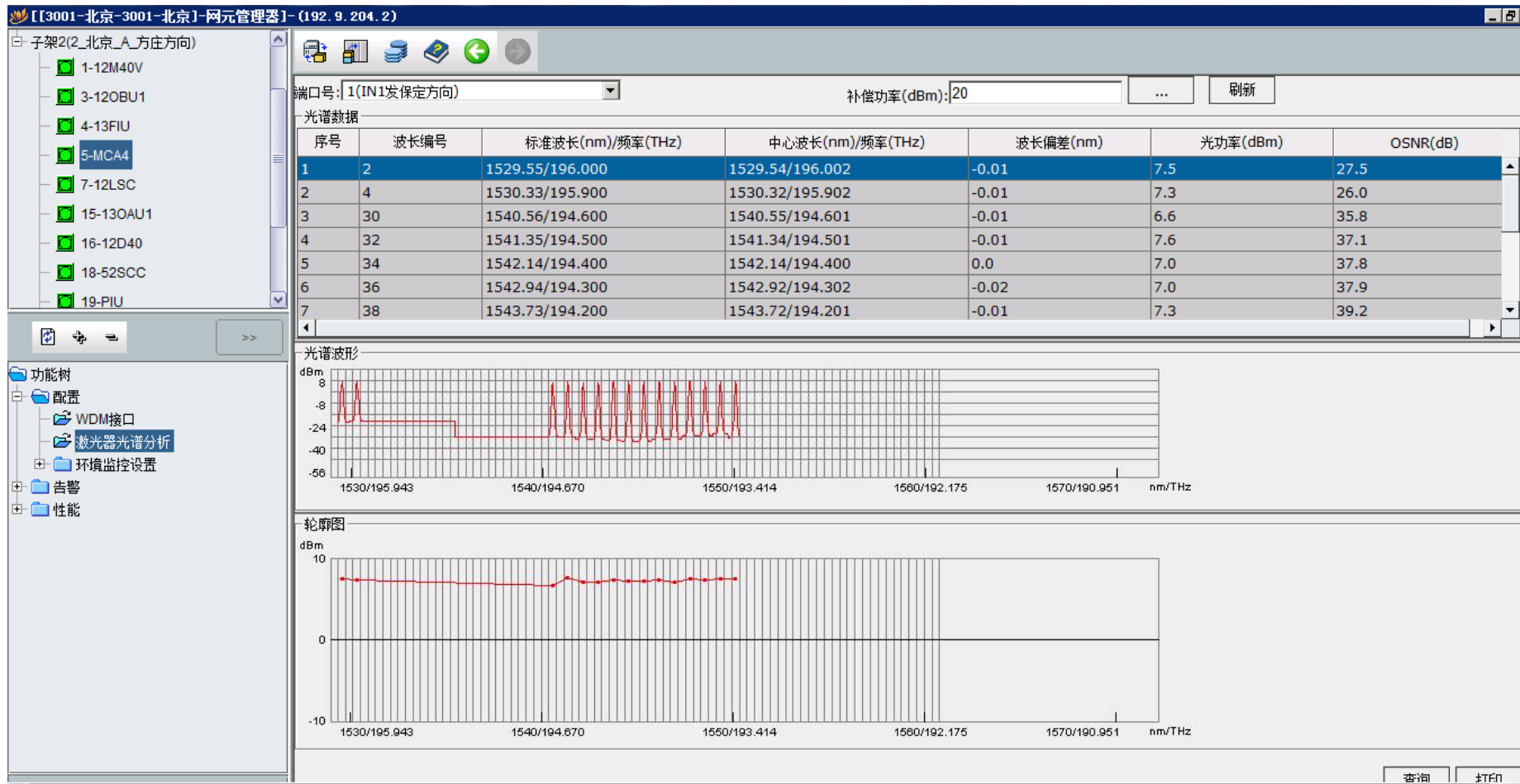


# DWDM design considerations



- 100G transport network
  - Point-to-point link
  - No OTN cross-connects
- 100G & 10G mixed system
  - Conservative considerations
    - Not sure 100G is stable at the design time (2012)
  - Economical considerations
    - 100G → TMUX multiple 10Gs
    - Or mixed 100G/10G

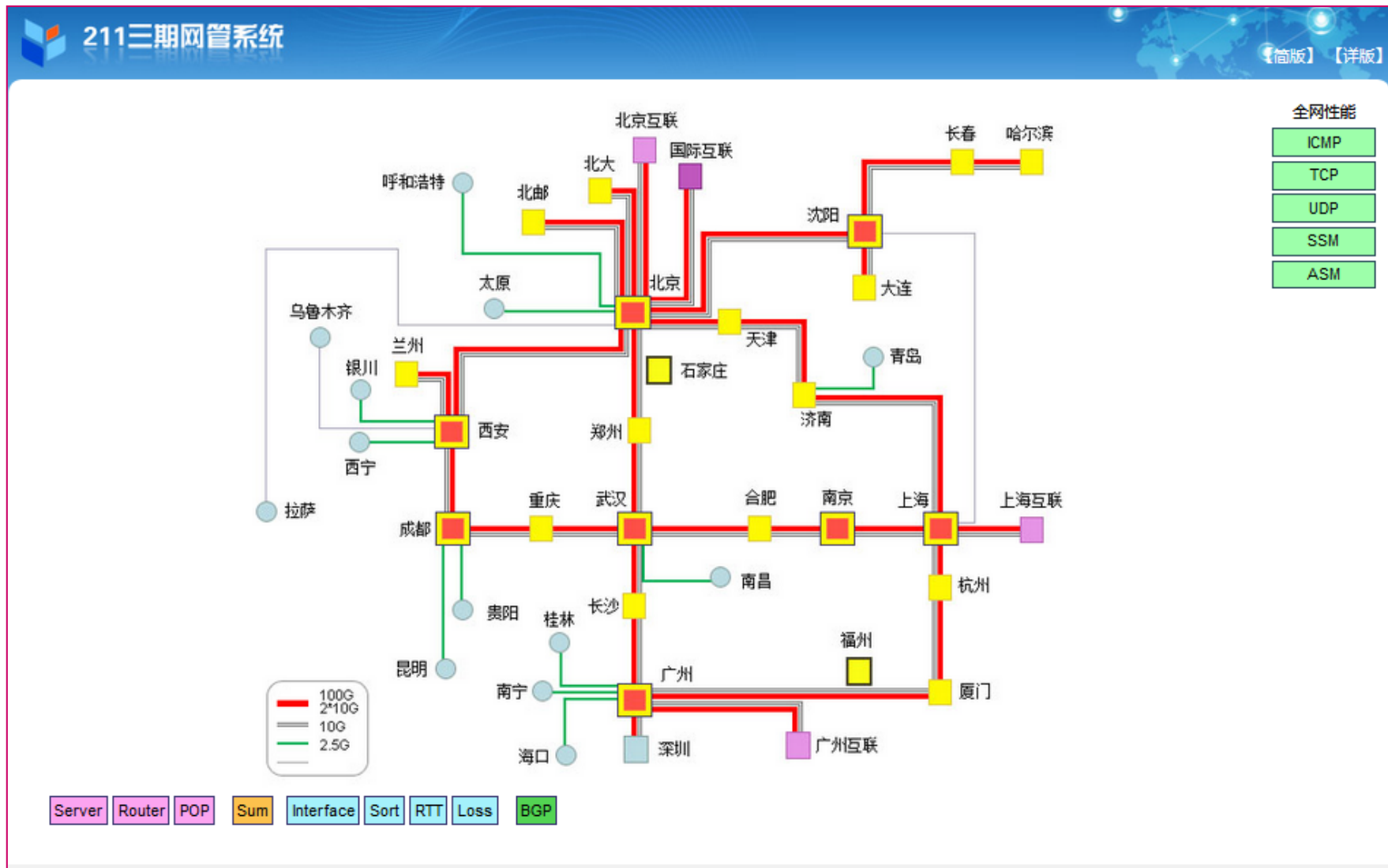
# DWDM deployment considerations



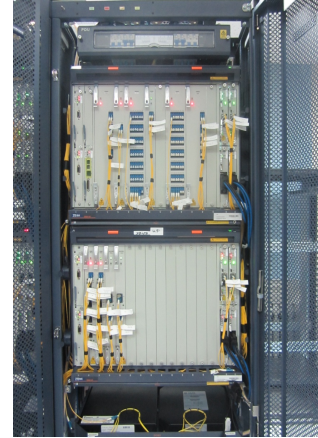
- 10G is sensitive to dispersion
- 100G is sensitive to attenuation



# Layer 3 topology



# Equipment



Juniper T4000



Cisco CRS-3

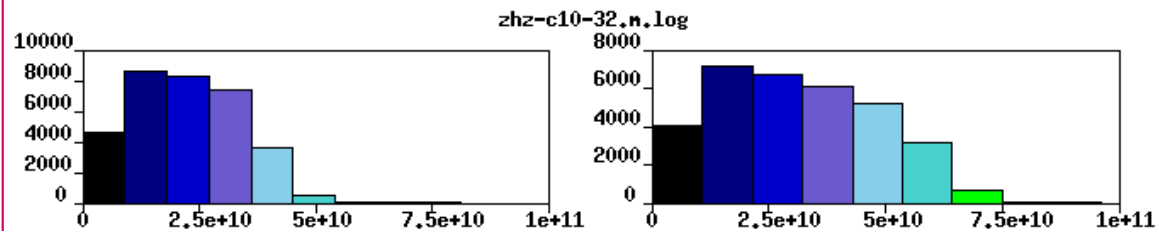
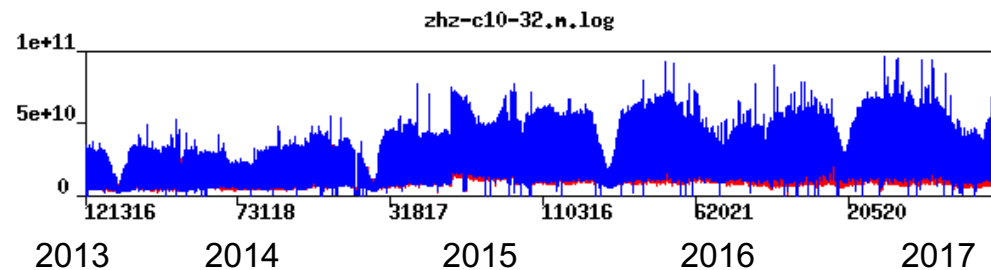


Huawei NE40

# IP Traffic examples

111 - 2017092316

```
zhz-c10-32.m.log in. act= 41.0123G (31458/33114) [max= 81.1666G med= 21.7031G min= 0 ] {ave= 22.2291G  
zhz-c10-32.m.log out act= 59.8893G (31458/33114) [max= 96.3341G med= 30.2292G min= 0 ] {ave= 31.3001G  
data
```



# 100G to supercomputing center in Wuxi

国家超级计算无锡中心  
National Supercomputing Center in Wuxi

About Us | News

INNOVATION COOPERATION SHARING EXCELLENCE

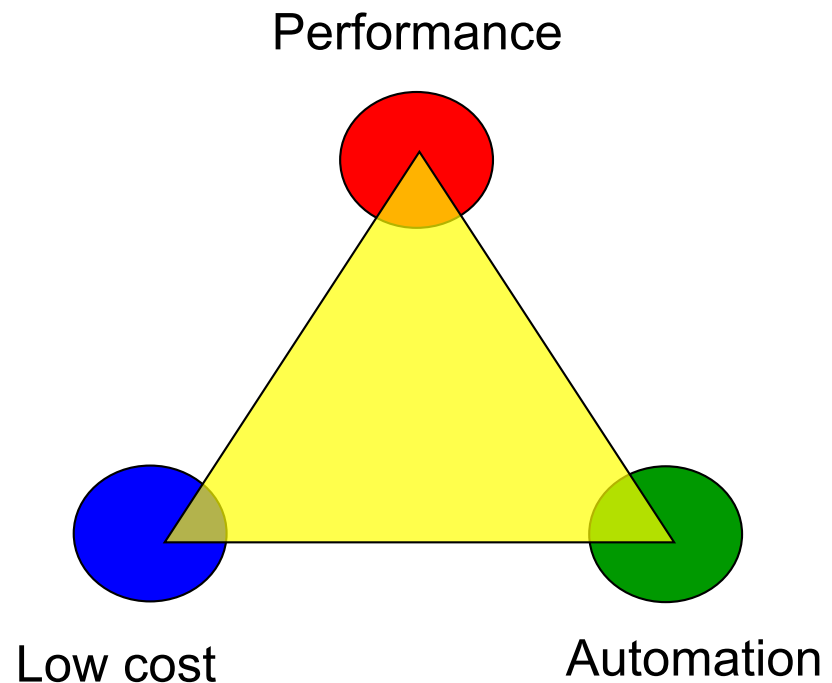
神威  
太湖之光

THE SUNWAY TAIHULIGHT SYSTEM IS THE WORLD'S FIRST SUPERCOMPUTER WITH PEAK PERFORMANCE OVER 100PFLOPS.  
THE ENTIRE COMPUTING SYSTEM IS BASED ON THE SW26010 MANY-CORE PROCESSOR.

Tel : 0510-8519 5508 | Add: 1 Yinbai road, Binhu district, Wuxi, Jiangsu province, China ©Copyright©National Supercomputing Center in Wuxi

# CERNET user requirements (1)

- Networks
  - QoS Layer 3 path
  - Bandwidth guaranteed Layer 2 Links
  - Measurement
  - Trouble shooting
  - Security
  - Low cost
- Research groups
  - Performance
  - Automation
  - Low cost



# CERNET user requirements (2)

---

- Layer 3 services
  - IPv4 global, commodity (best effort)
  - IPv4 academic (bandwidth on demand)
  - IPv6 global
  - IPv4 and IPv6 translation
- Layer 3 VPN services
- Layer 2 Point-to-point
- SDN service



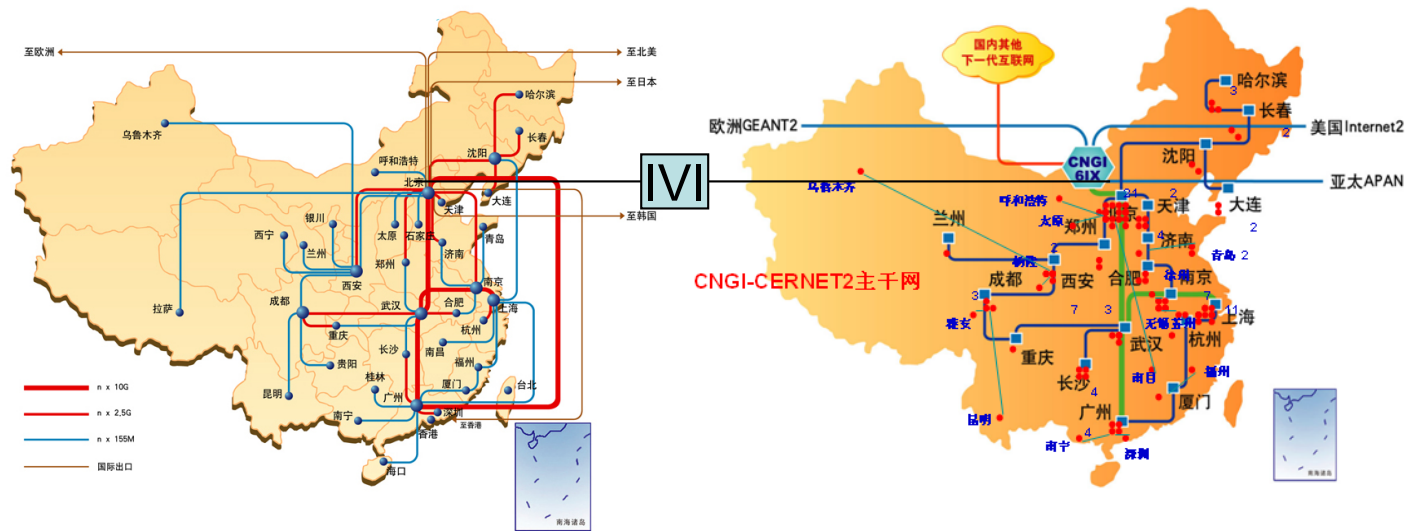
# Building blocks

---

- Routing
  - IPv4 BGP
  - IPv6 BGP
  - IPv6 multicast SSM
- Layer 2
  - point-to-point
- Layer 3 VPN
  - BGP VRF
- SDN
- SRv6

- Address switching
- IPv4/IPv6 translation (IVI)
- Scalable Application Specific Measurement (SASM)

# Routing



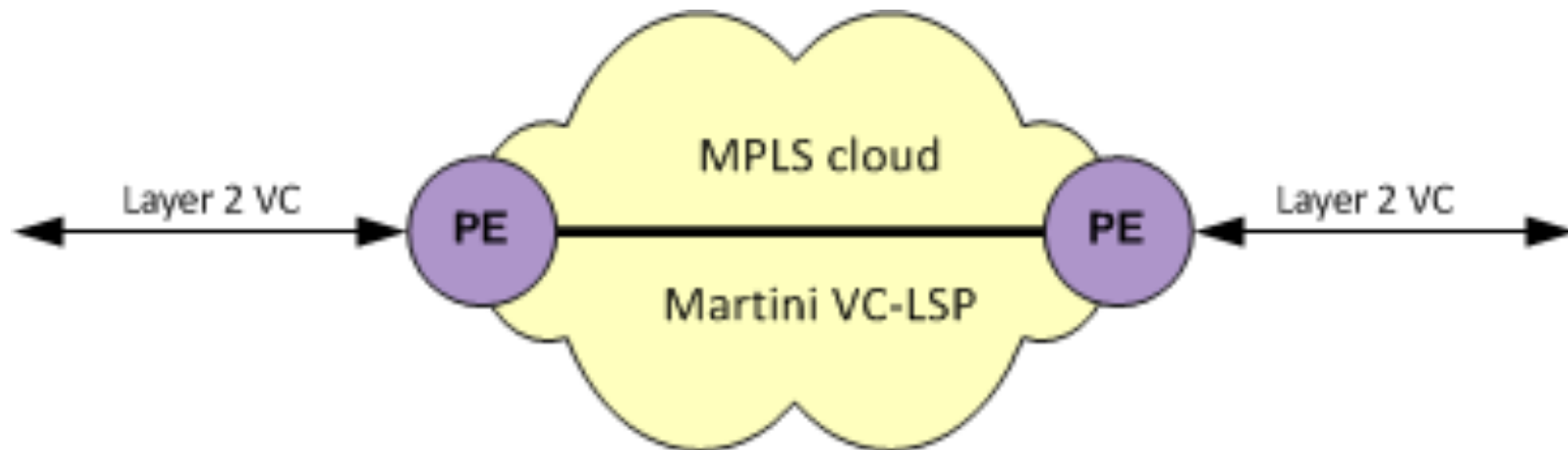
IPv4

IPv6

- Providing services already
- Asymmetric routing
- Best effort

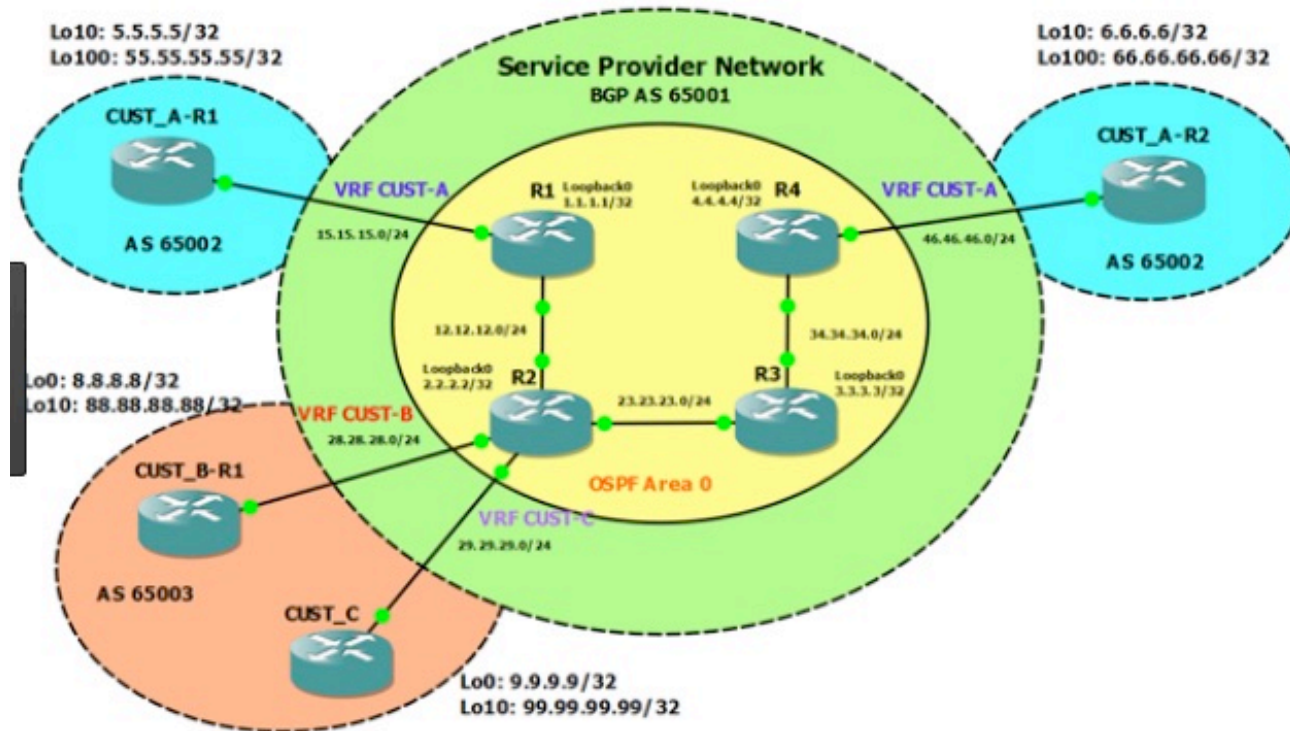


# Layer 2 point-to-point



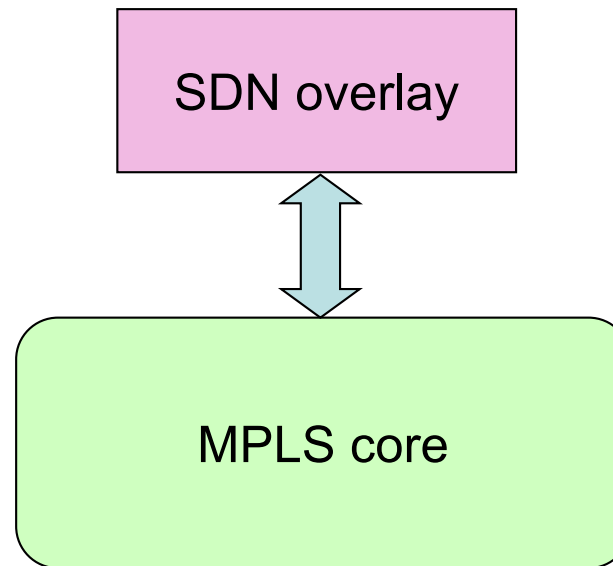
- Providing services already
- Scalability
- Inter domain

# Layer 3 VPN



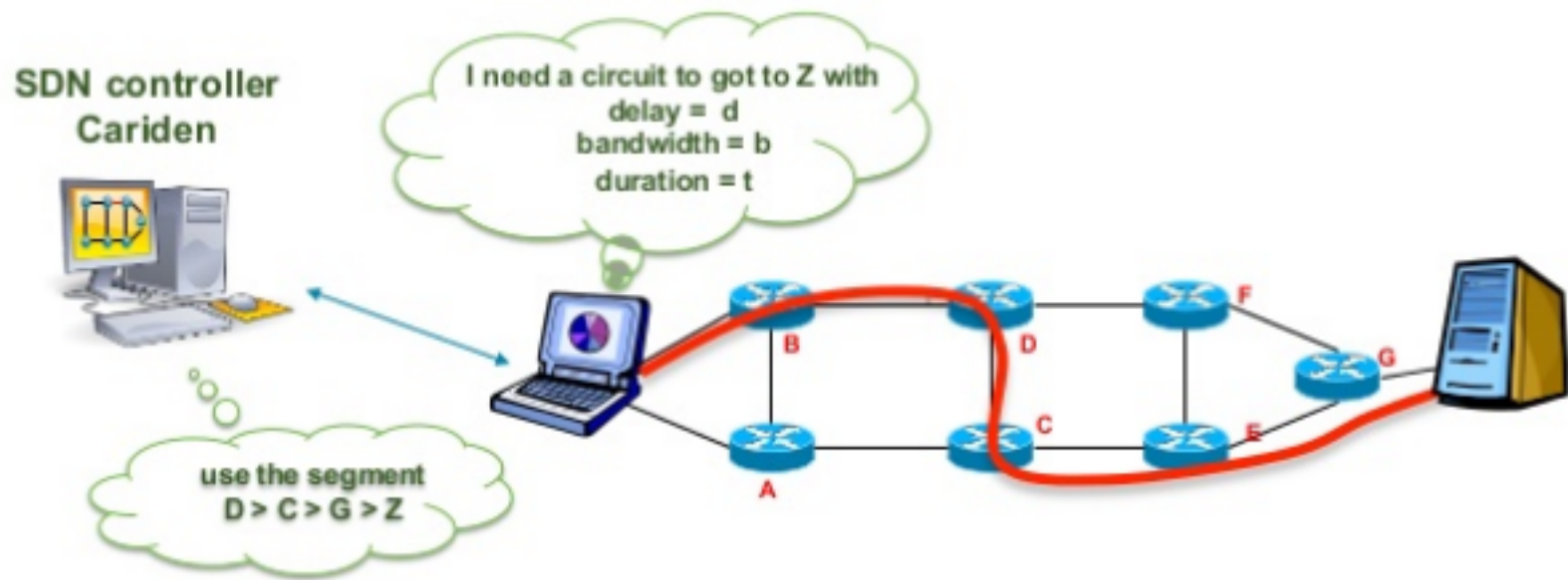
- Provide services soon
- Scalability
- Inter domain

# SDN



- Provide services soon (overlay model)
- Production service?

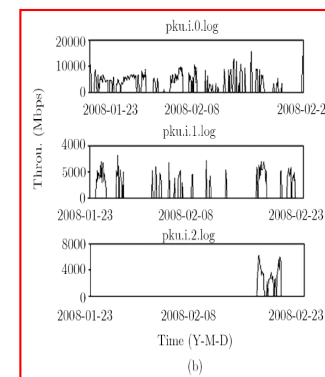
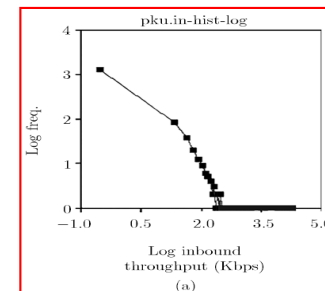
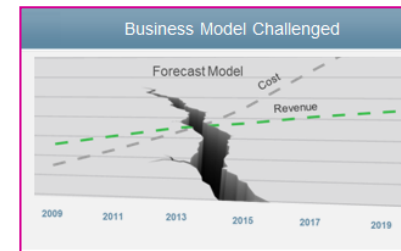
# SRv6



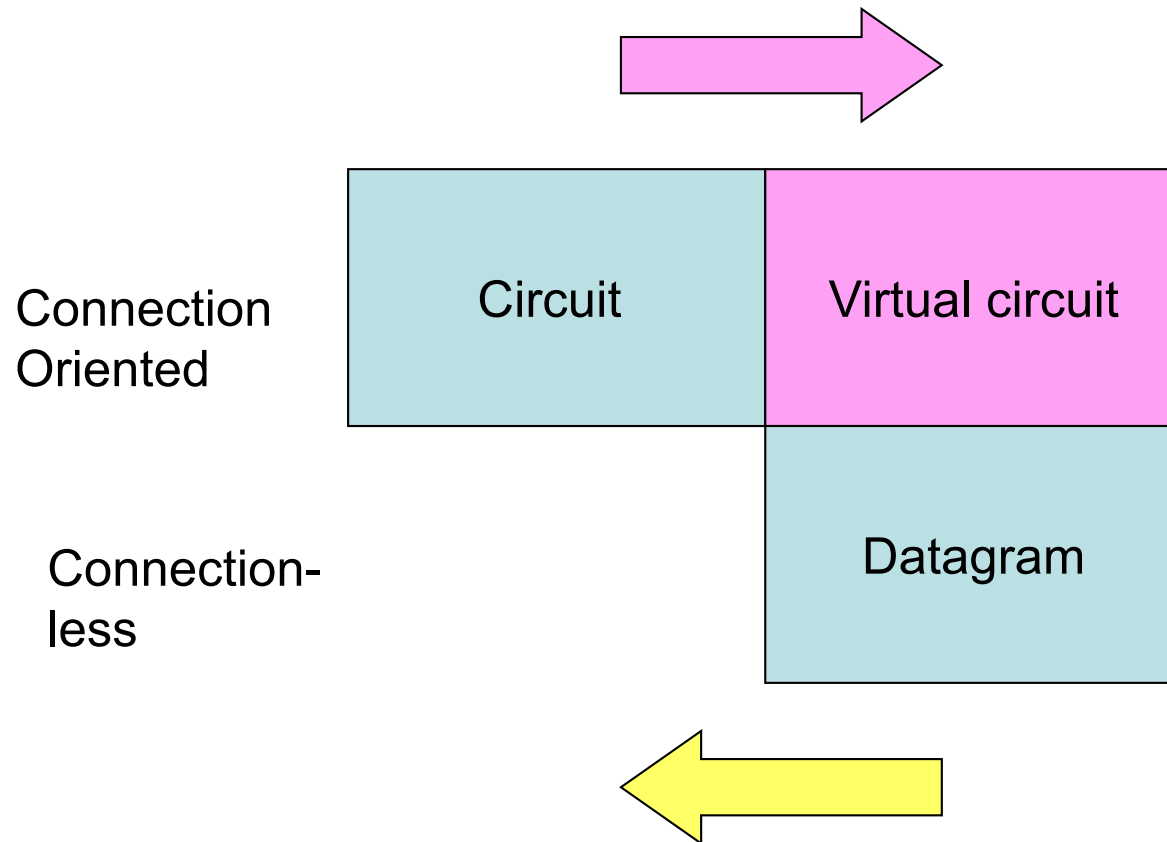
- Keep eyes open

# Natures of the Internet

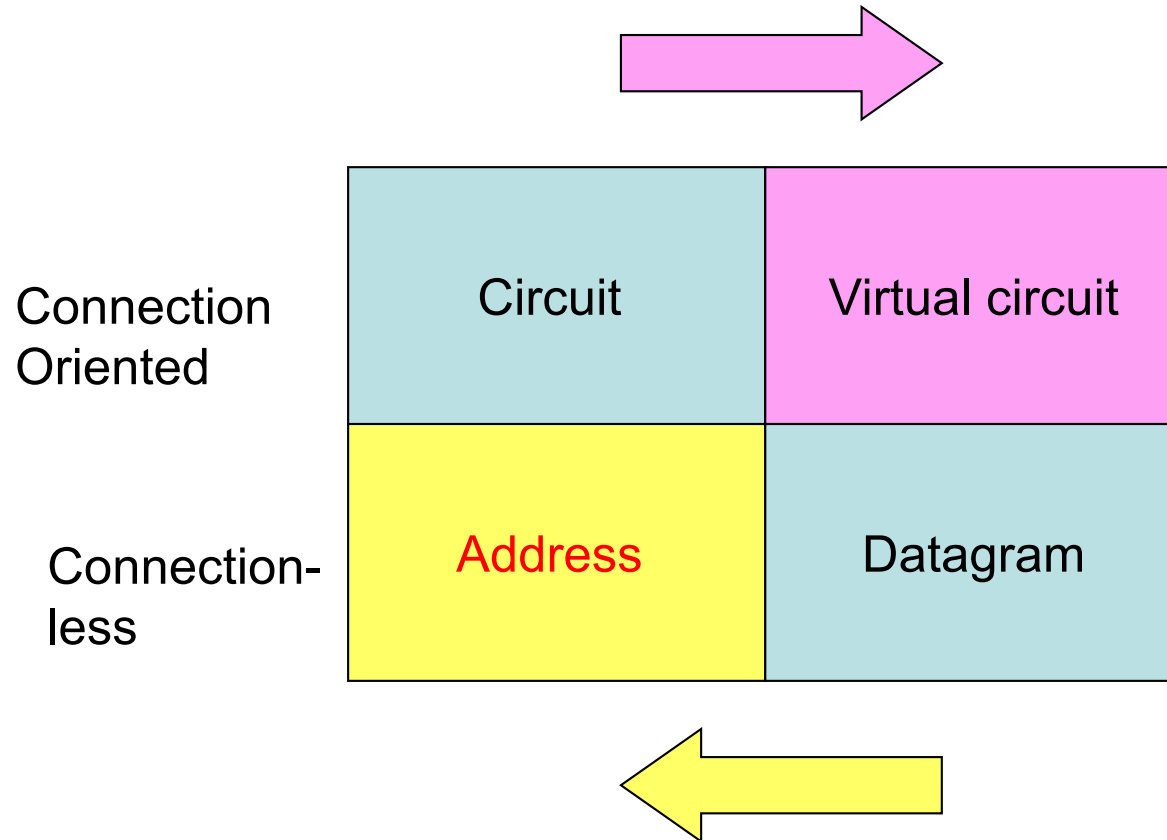
- Bandwidth is limited resource
- The user's behavior follows the power law
- The high performance applications are still Poisson distribution



# Switching technology

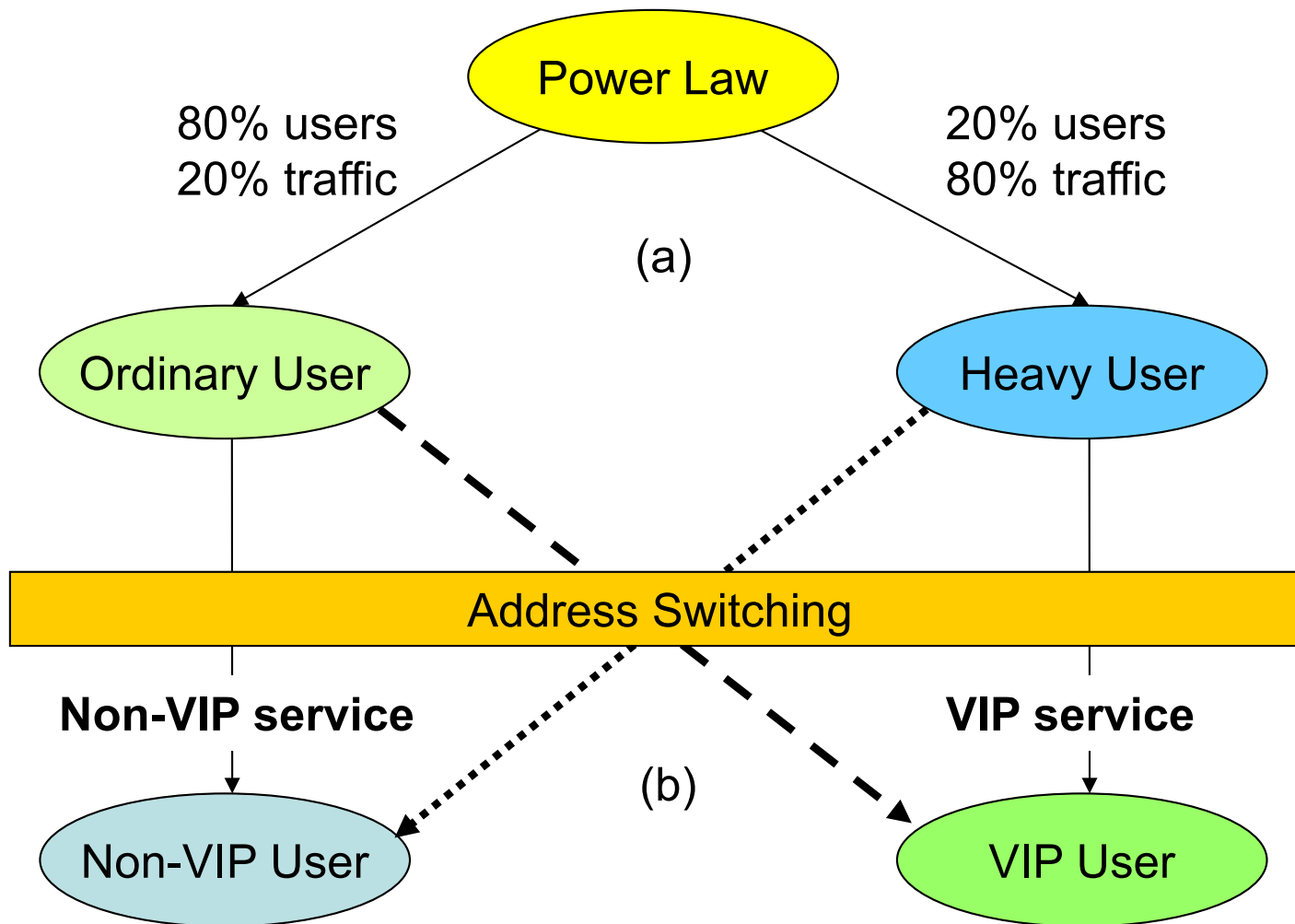


# Switching technology



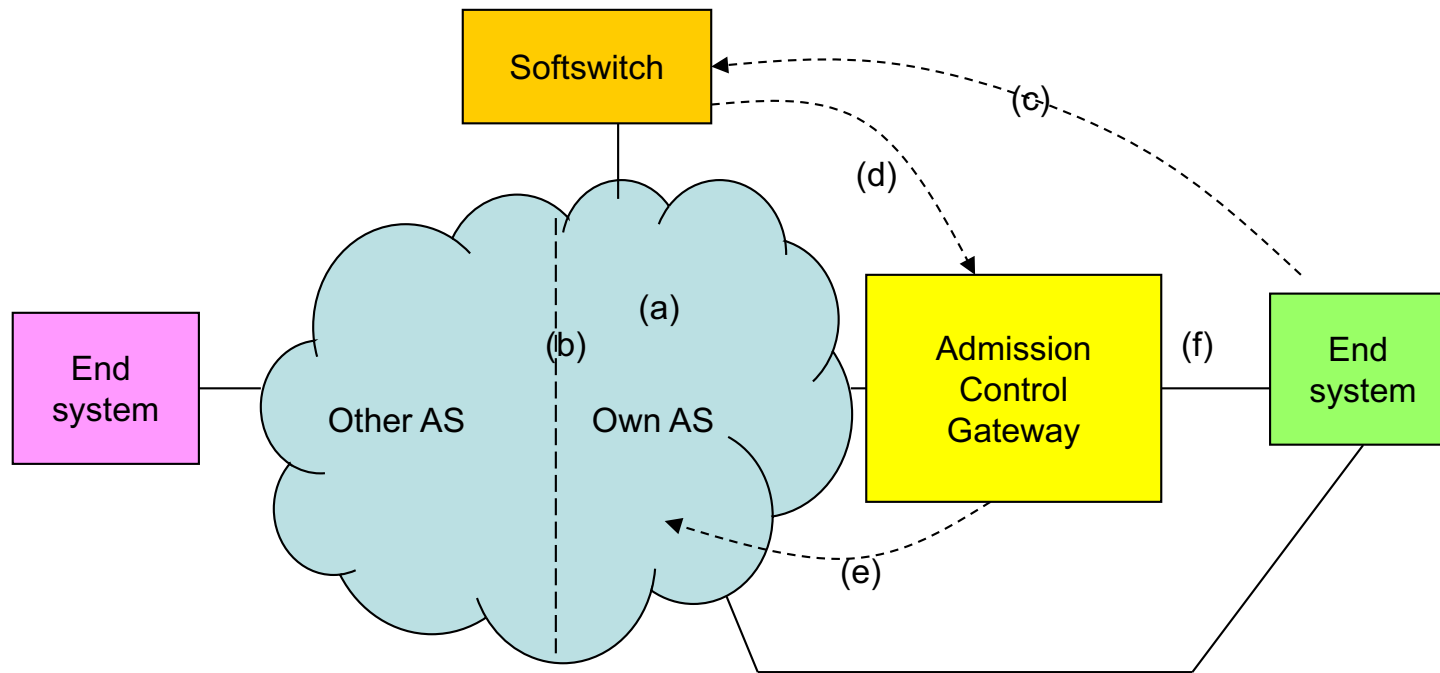
AaaS (Address as Service)

# Address-switching concept



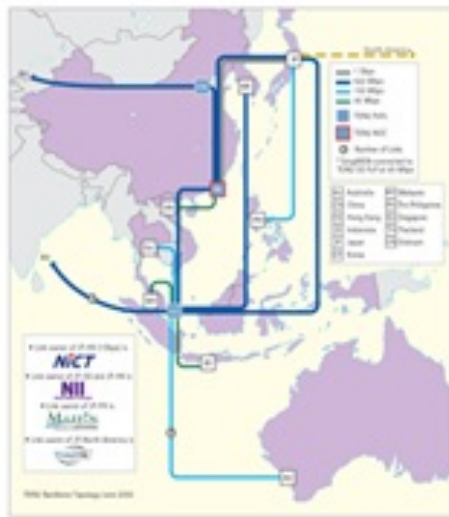


# Address switching via BGP RR



- Bandwidth is limited resource
- Reserve bandwidth per VIP address block
- Dynamic assign VIP address block with admission control

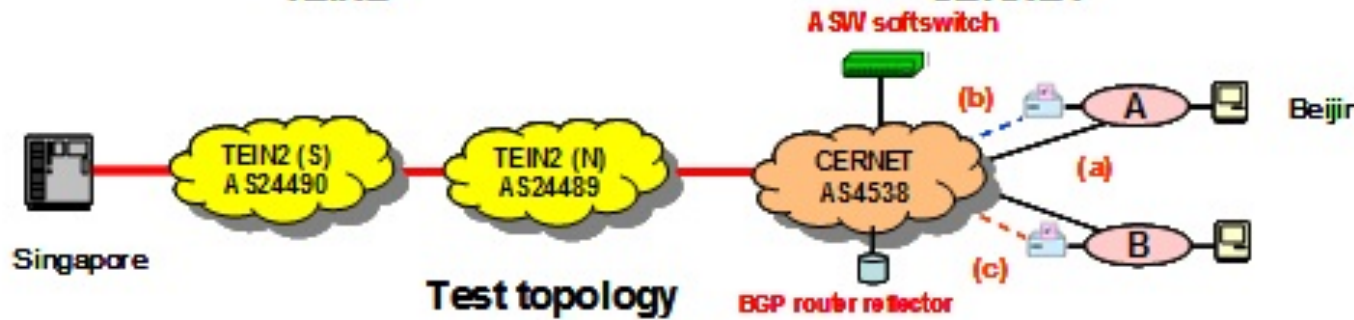
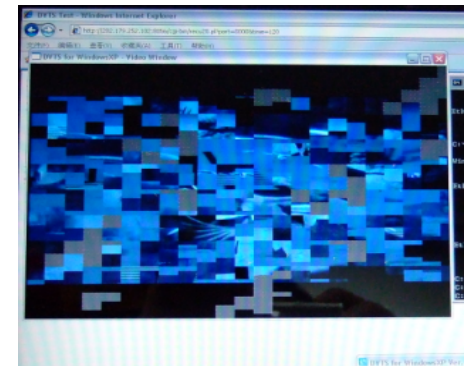
# Example



TEIN2

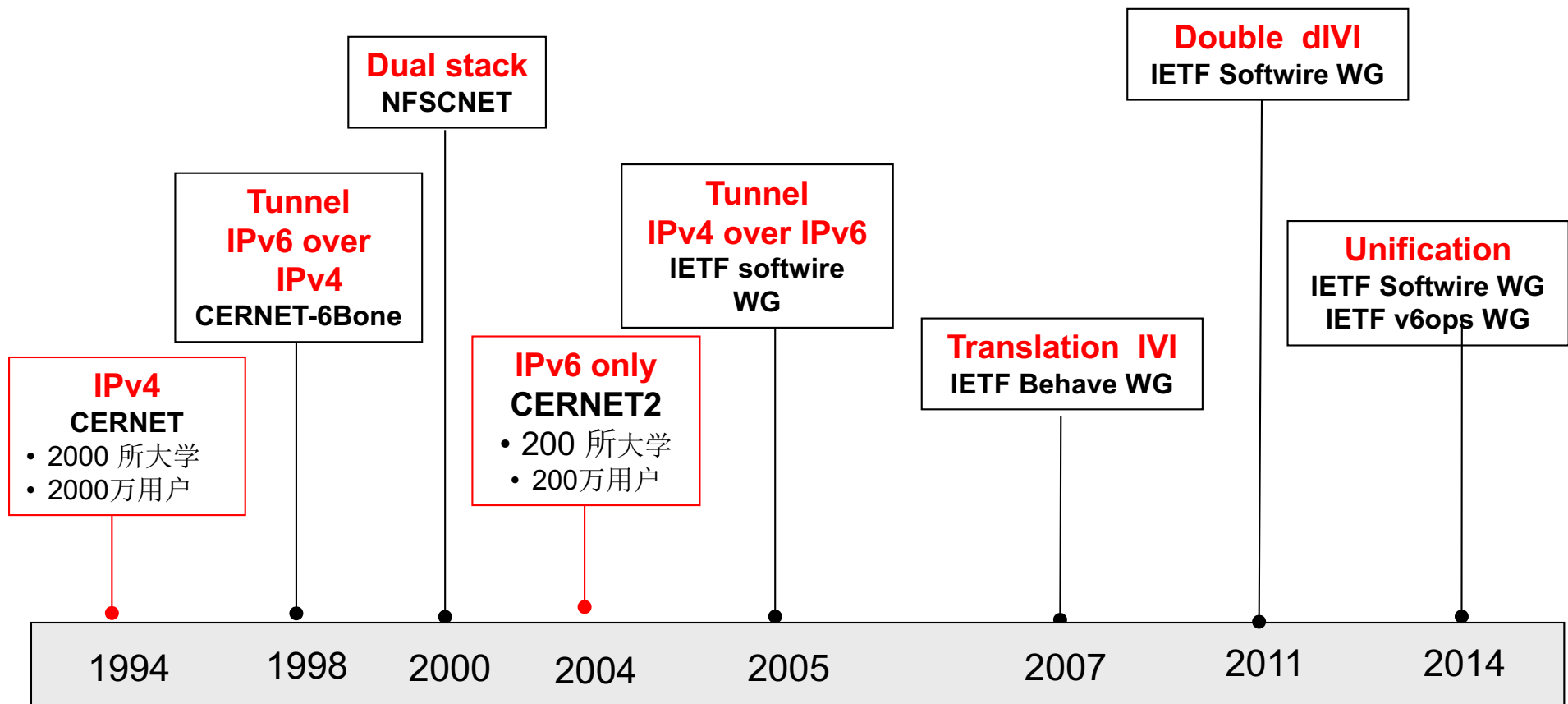


CERNET

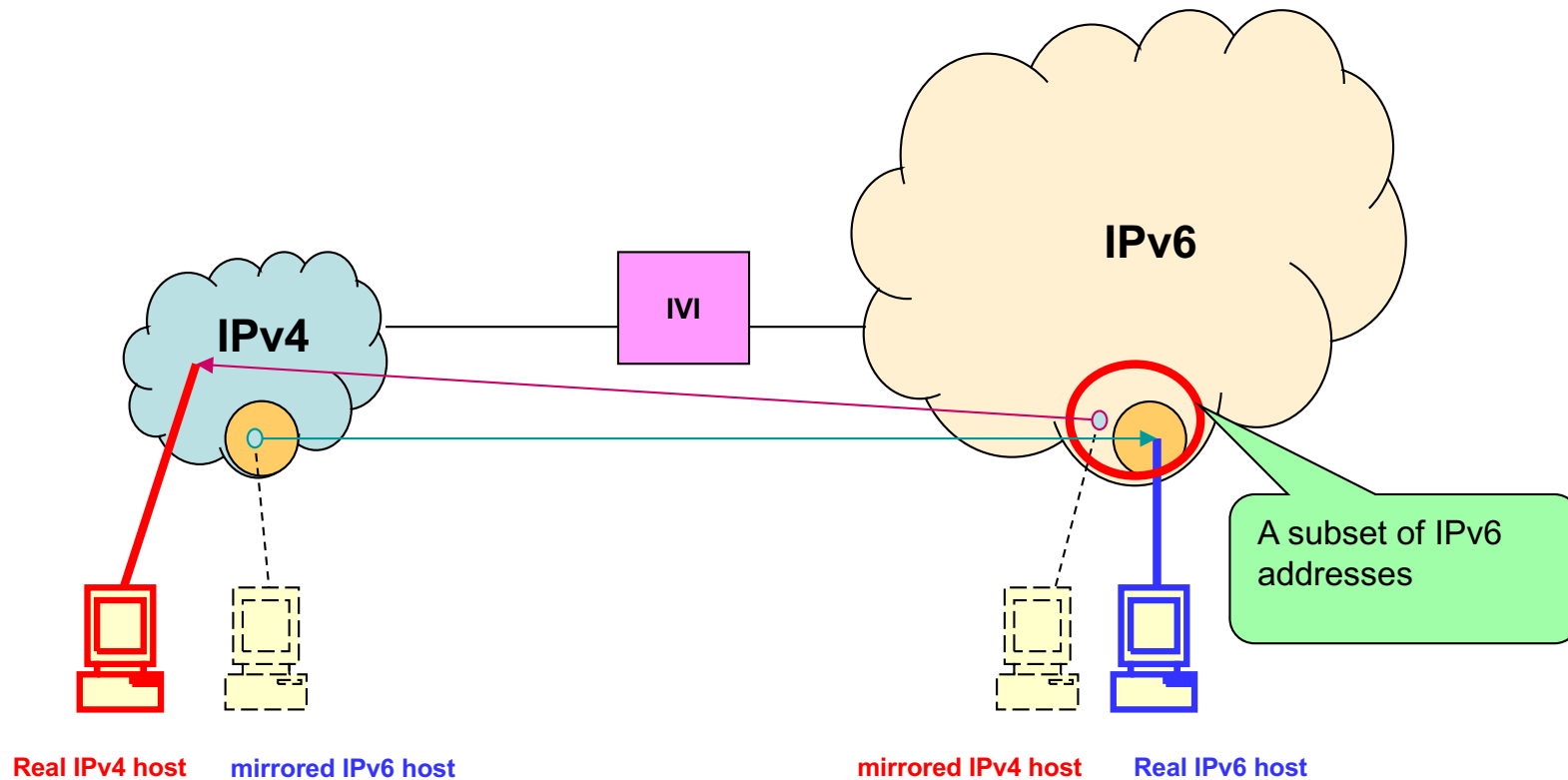


<http://www.address-switching.org/>

# CERNET IPv6 transition experience



# Stateless translation concept



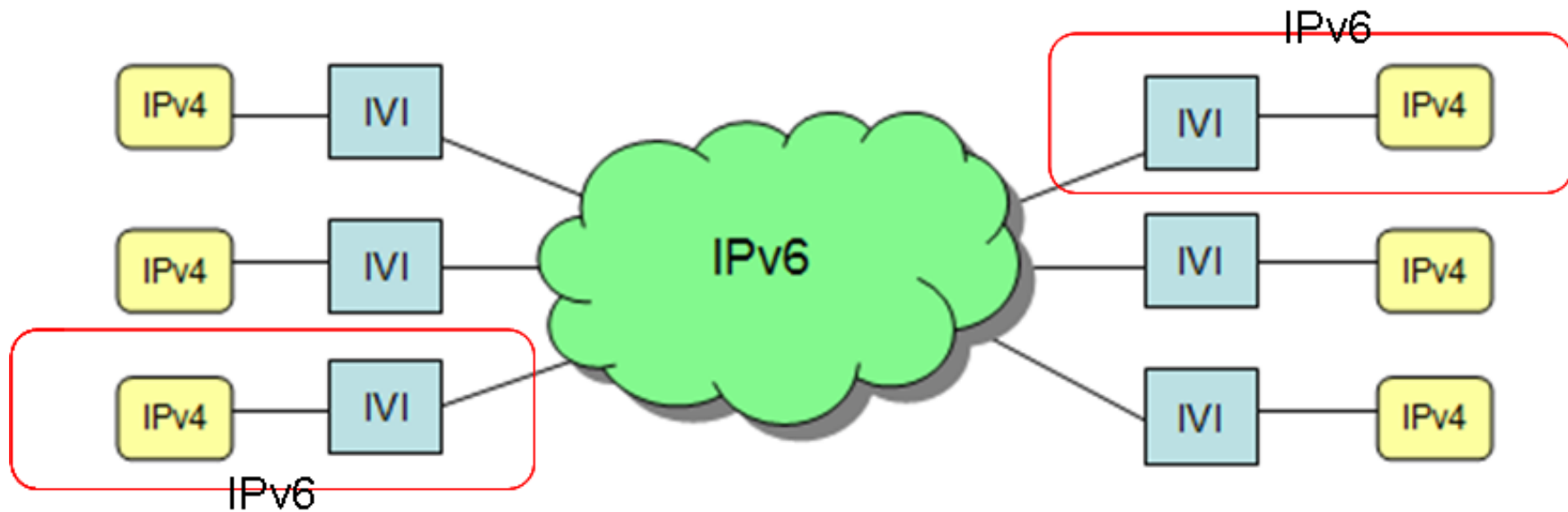
- IPv4/IPv6 translation makes IPv4 and IPv6 can communicate directly

# Lessons learned

---

- The Internet could not have been so successful in the past years if IPv4 had contained any major flaw.
- The value of a network is proportional to the square of the number of connected users of the system ( $n^2$ )
- Incremental deployment is required.
- Any problem in CS can be solved by adding a layer of indirection.

# TE via IPv6



4aaS

<http://www.ivi2.org/>

# SASM

Search server by:[my IP/AS](#) | [IP](#) | [Domain](#) | [AS](#) | [Update](#) | [Performance test](#) | [Slow website report](#) | [Documentation](#)

**SASM**  Choose a protocol:

---

**166.111.1.1** belongs to **AS4538** from China. [3092 ipv4 web server\(s\)](#) found in this AS, **1248** of which are close to the input IP address. Web server 1  
- 10

AS information:  
AS Name: ERX-CERNET-BKB China Education and Research Network Center  
AS Path: 4538  
[Routing information from CIDR Report](#)

[learn.tsinghua.edu.cn](#) 166.111.4.17 Pagesize: 121140B Performance: 4.670MB/S [Download URL](#)

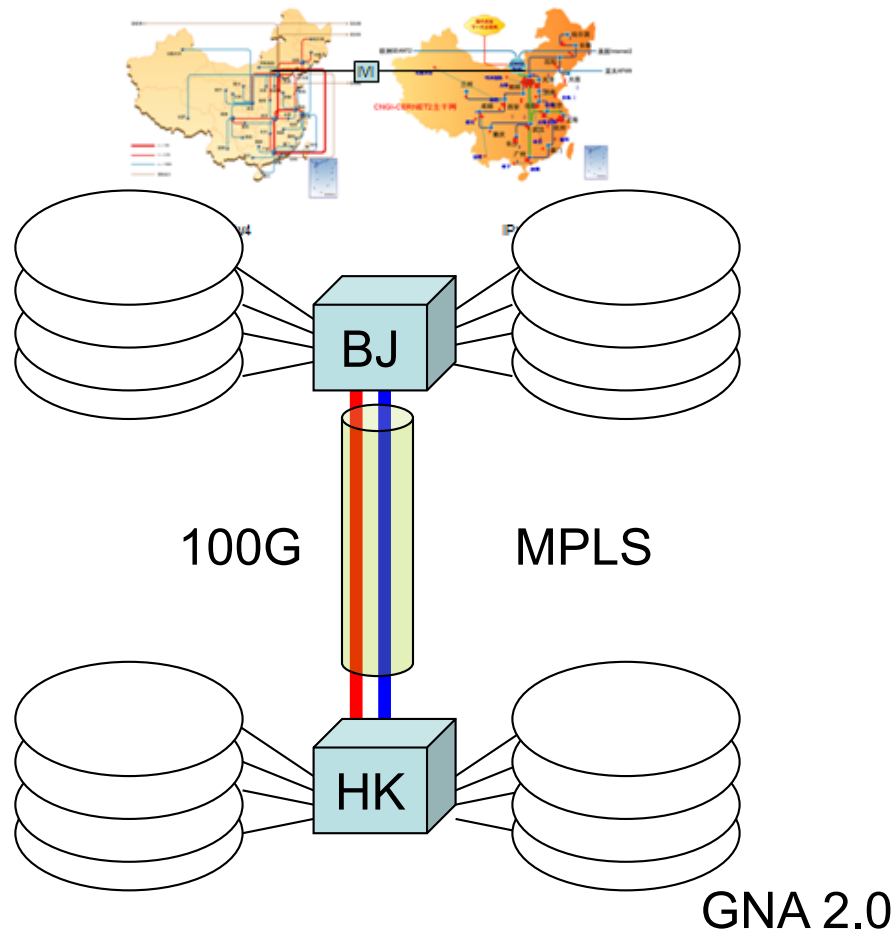
[清华大学新闻网](#)  
news.tsinghua.edu.cn 166.111.4.29 Pagesize: 402B Performance: 45.100MB/S [Download URL](#)

[www.meng.edu.cn](#) 166.111.4.39 Pagesize: 350456B Performance: 43.900MB/S [Download URL](#)

<http://search.sasm3.net/>

# Beijing-Hong Kong 100G

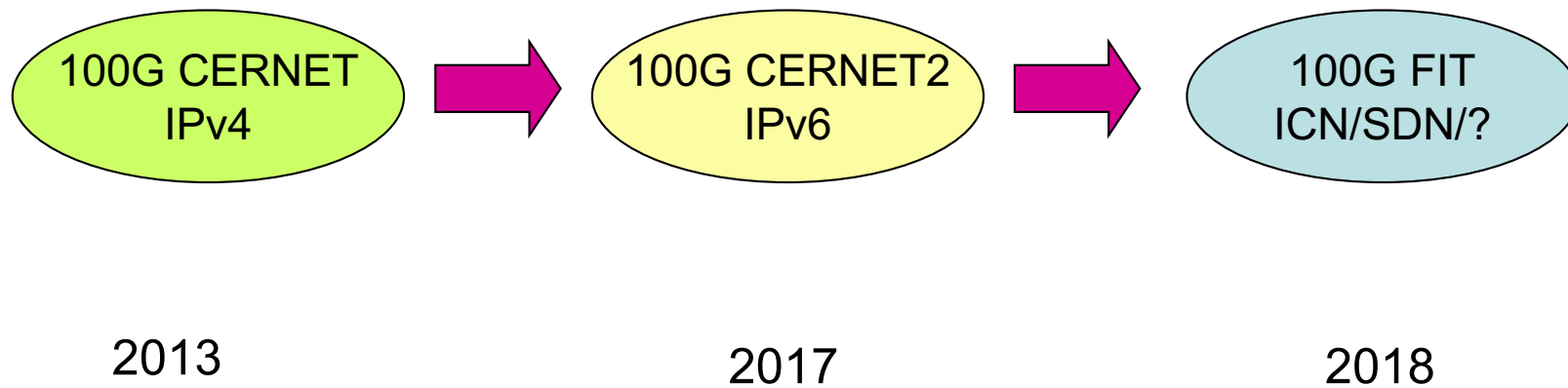
## APIs



- Routing
  - IPv4 BGP
  - IPv6 BGP
  - IPv6 multicast SSM
- Layer 2
  - point-to-point
- Layer 3 VPN
  - BGP VRF
- SDN
- Address switching
- IVI
- SASM



# Next steps



# Remarks



- 100G
- Keep the stateless nature of the Internet
  - Address switching
- Adding a layer of indirection
  - IPv6 and translation