



---

# IVI translation concepts and experience

Xing Li

2009-12-03

# Outline



- History
  - CNGI-CERNET2
- Lessons learned
  - Design philosophy
  - Killer application
- Transition plan
  - IVI concept
  - CNGI-CERNET2 plan
  - Future

# CERNET IPv6 history

---

- May 3, 1998, IPv6 on FreeBSD at Tsinghua University
- Sept. 1998 first BGP4+ peer to Sprint (MRTd)
- Nov. 1998, First 6Bone backbone node in China
- Mar., 2000 IPv6 BBS--6th Galaxy opened
- April 26, 2000 obtained sTLA (2001:250::/35 ) from APNIC
- Aug., 2000 《IPv6原理与实践》 published
- Sep. 25, 2000 CERNET joined IPv6 Forum
- Jan. 2002 DRAGON TAP IPv6 IX is running
- Jun. 2002 NSFCNET double stack is running
- Sept. 2003 CNGI-CERNET2 test nodes are running

# CNGI-CERNET2



# Be unique, be different

---

- Protocol selection
  - Pure IPv6
- Equipment
  - Multiple vendors
- Complexity
  - Multiple ASs
- Transition
  - IPv4 over IPv6 (IETF softwire)
  - IVI stateless translation (IETF behave)
- Architecture
  - Source address authentication (IETF SAVI)

# The lessons learned



- Plan
  - Performance
    - CNGI-CERNET2 (IPv6) is light loaded
    - CERNET (IPv4) is sometime congested
  - Charging
    - CNGI-CERNET2 (IPv6) is free
    - CERNET (IPv4) is not free.
  - Requirements
    - The Users need to run their applications in IPv6
- Reality
  - The users need to communicate with the IPv4 users, even the network is somehow congested and not free.

# What is the killer application of IPv6?

---

- The video?
  - The IPv4 can support it well.
- The P2P?
  - There are NAT transversal solutions
- My answer:
  - The transparent communication with the IPv4 Internet is the “**killer Application**” of IPv6.

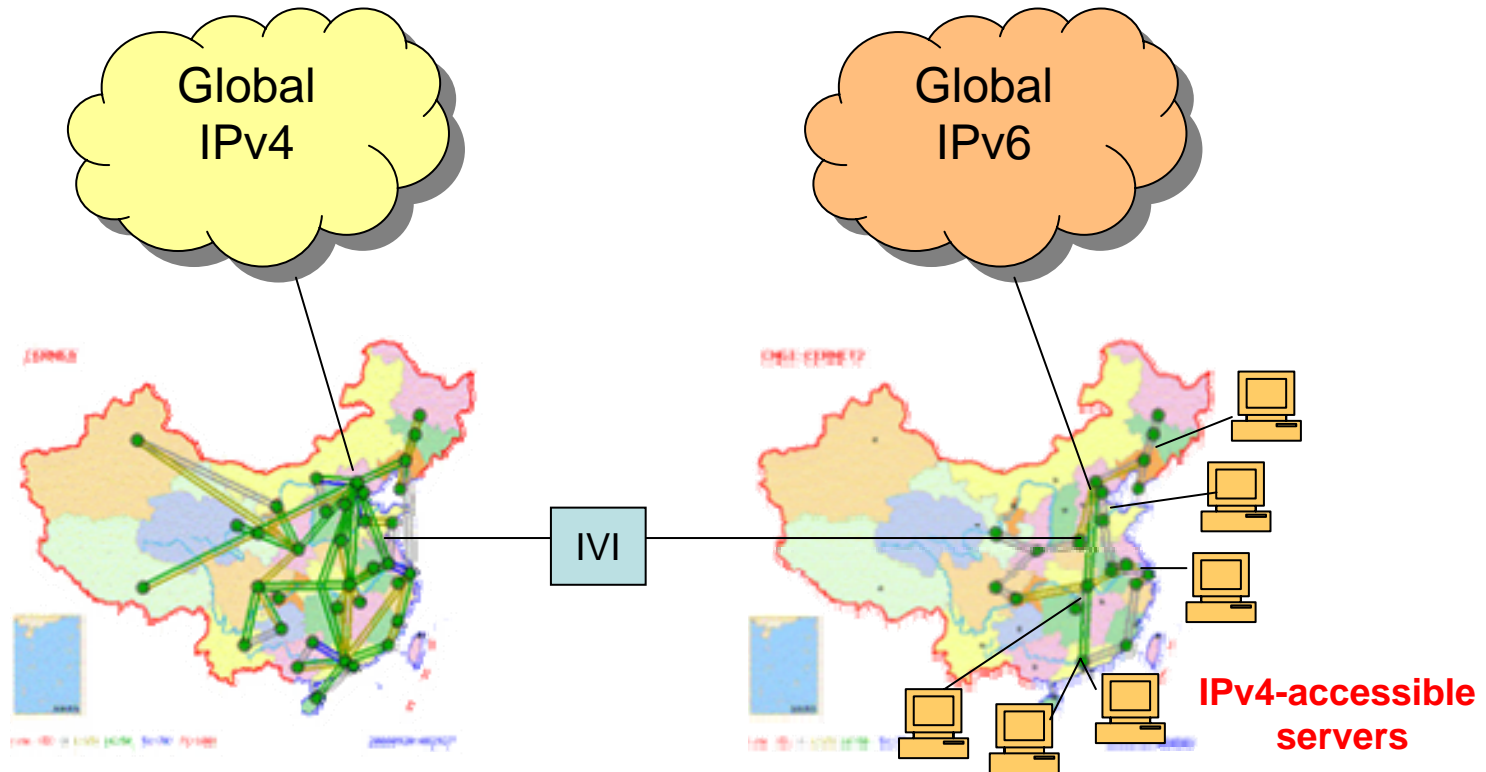
# How to achieve it?



- IPv4 address depletion
  - Dual stack is not practical
- The IPv4 over IPv6
  - Solve part of the transition problem
- The translation between IPv4 and IPv6
  - Difficult, but it is what users really want



# So we developed IVI



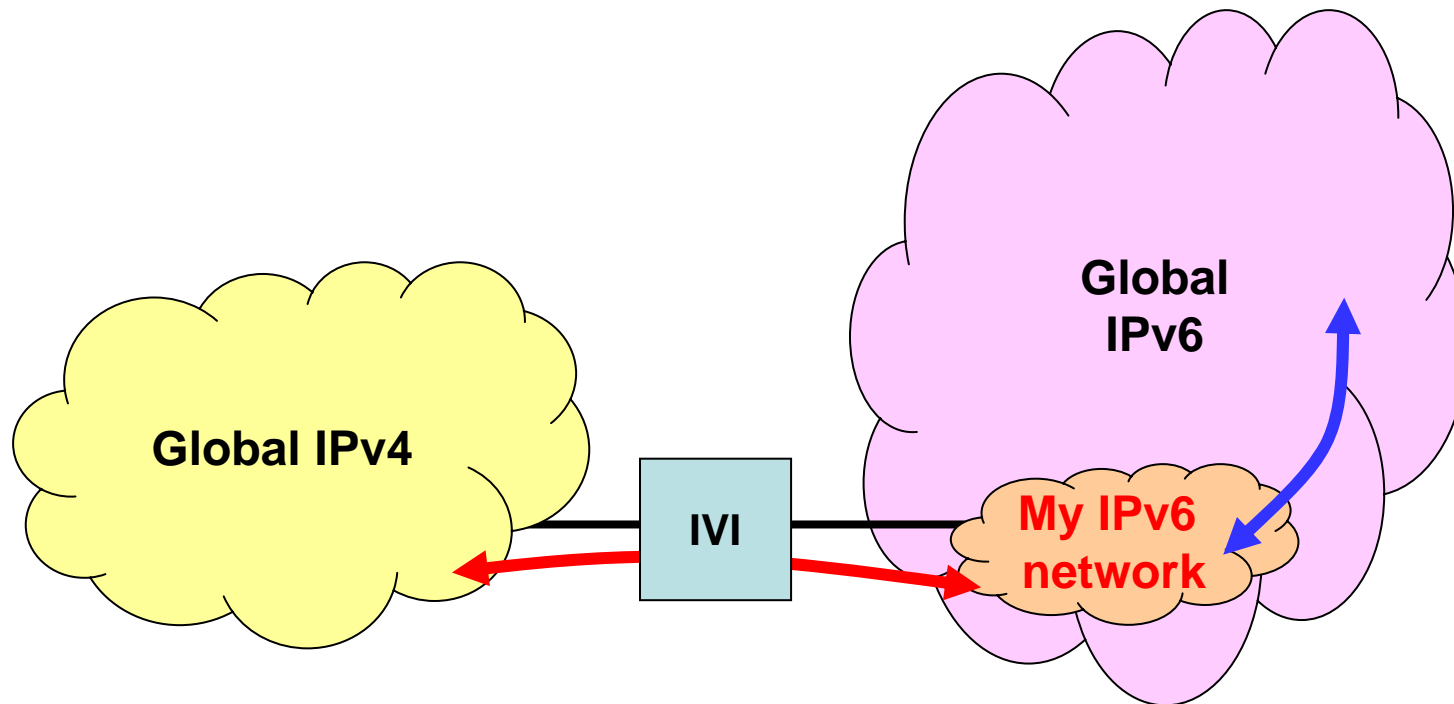
CERNET (IPv4)

2,000 universities  
connected  
20M users

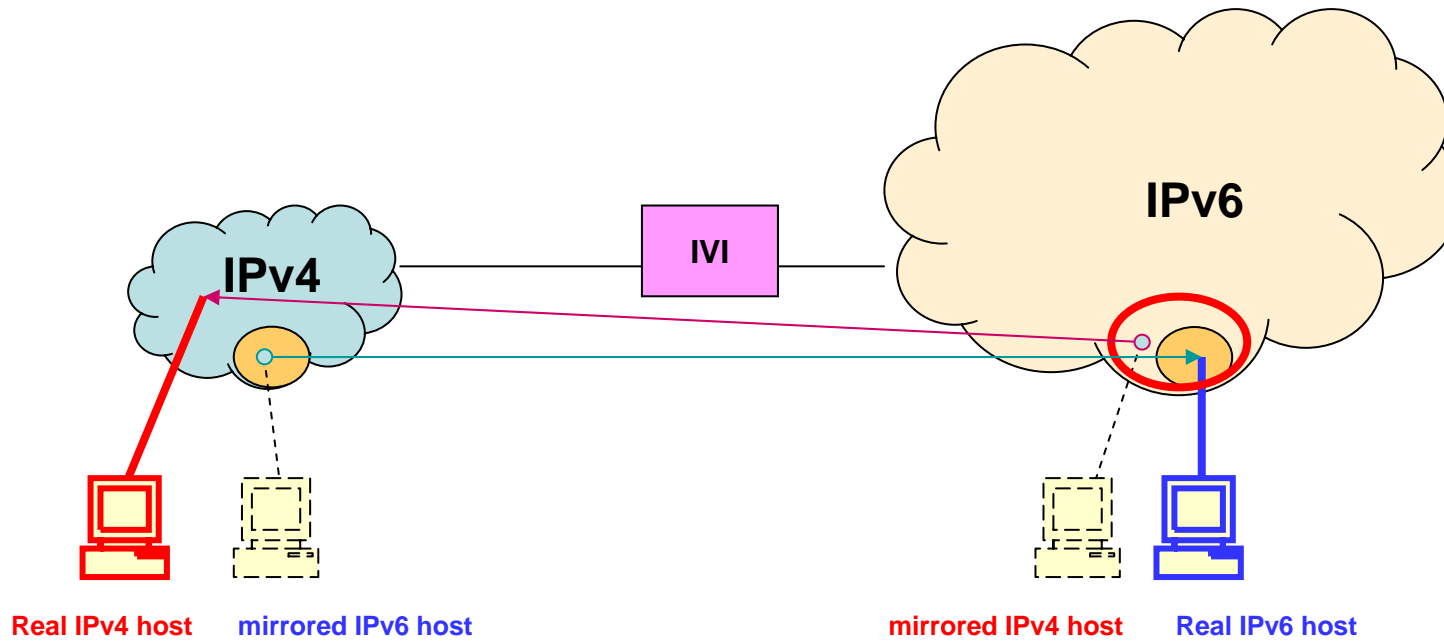
CNGI-CERNET2 (IPv6)

100 universities  
connected  
400K users

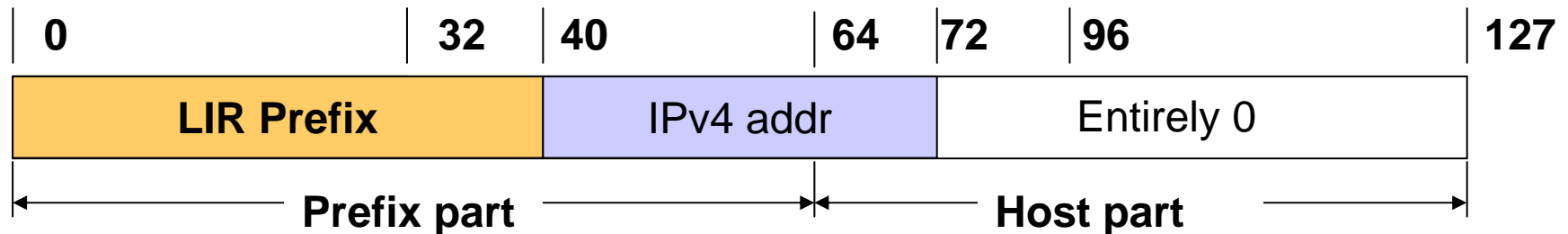
# The IVI model



# Conceptual example



# IVI address format

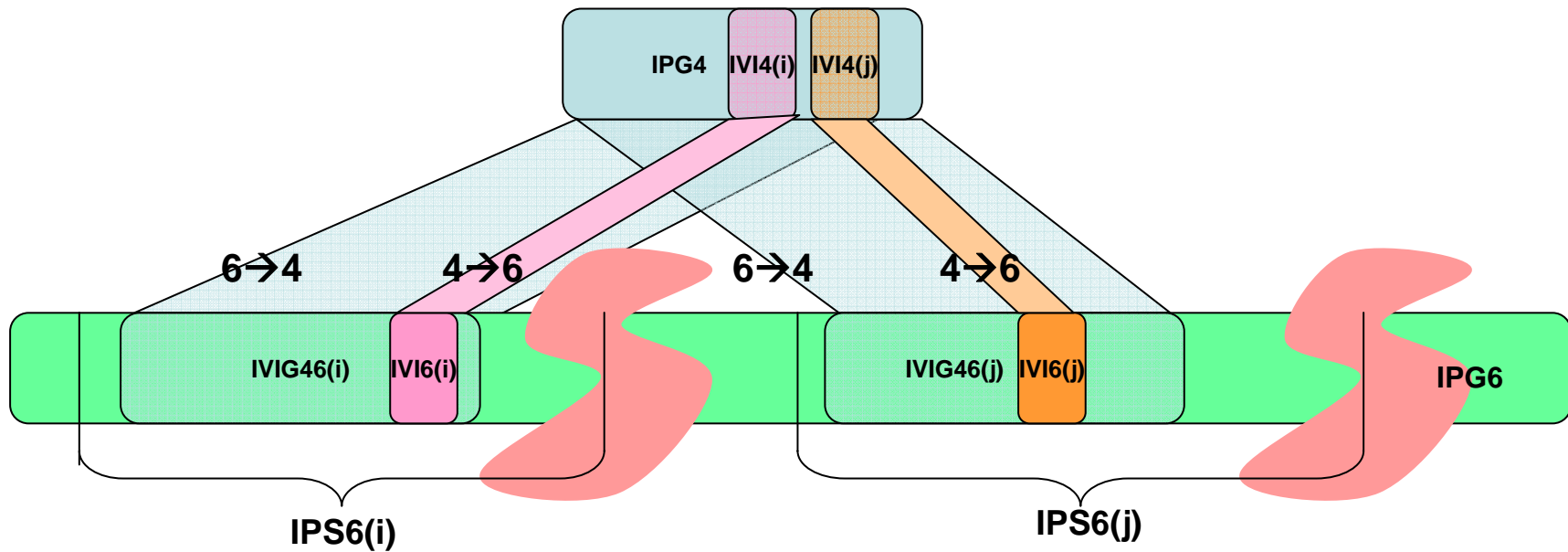


For example

LIR consists of ISP prefix (usually /32) and IVI flag  
CERNET/CNGI-CERNET2's selection

- LIR = 2001:da8:ff00::/40
- **ISP's IVI service IPv4 address mapping**
  - 202.38.108.0/24 → 2001:da8:ffca:266c:0000::/64
- ISP's non-IVI service IPv4 address mapping
  - 202.38.96.0/20 → 2001:da8:ffca:2660:0000::/60
- Other ISP's IPv4 address mapping
  - 0.0.0.0 → 2001:da8:ff00::/40
  - 18.181.0.31/32 → 2001:da8:ff12:b500:1f00::/72

# IVI address mapping

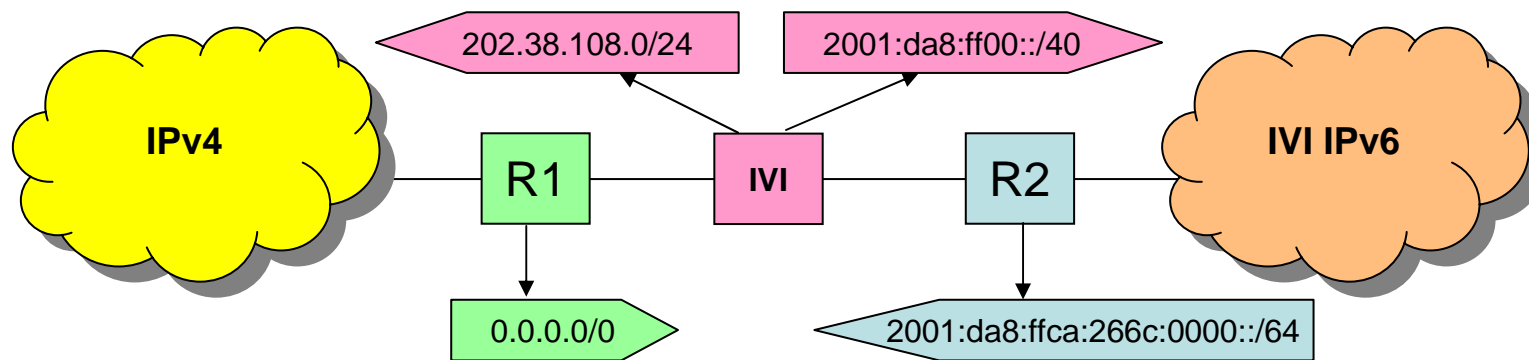


# IVI protocol translation



- The protocol translation is based on modified SIIT.
- Special issues are related to
  - MTU and fragmentation
  - Checksum recalculation
  - ICMP and ICMP error message handling
  - Etc.

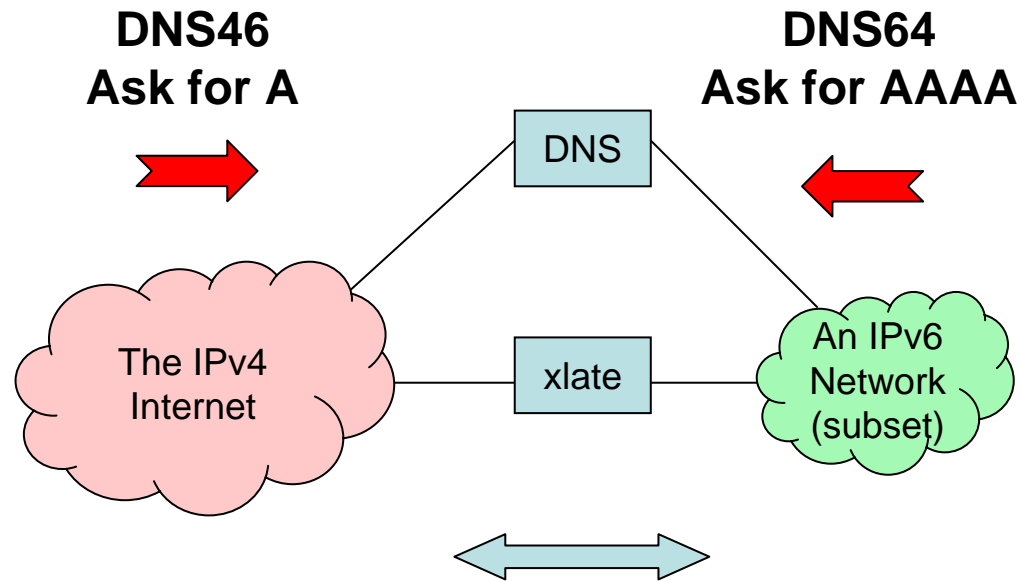
# IVI routing



**Longest prefix match**

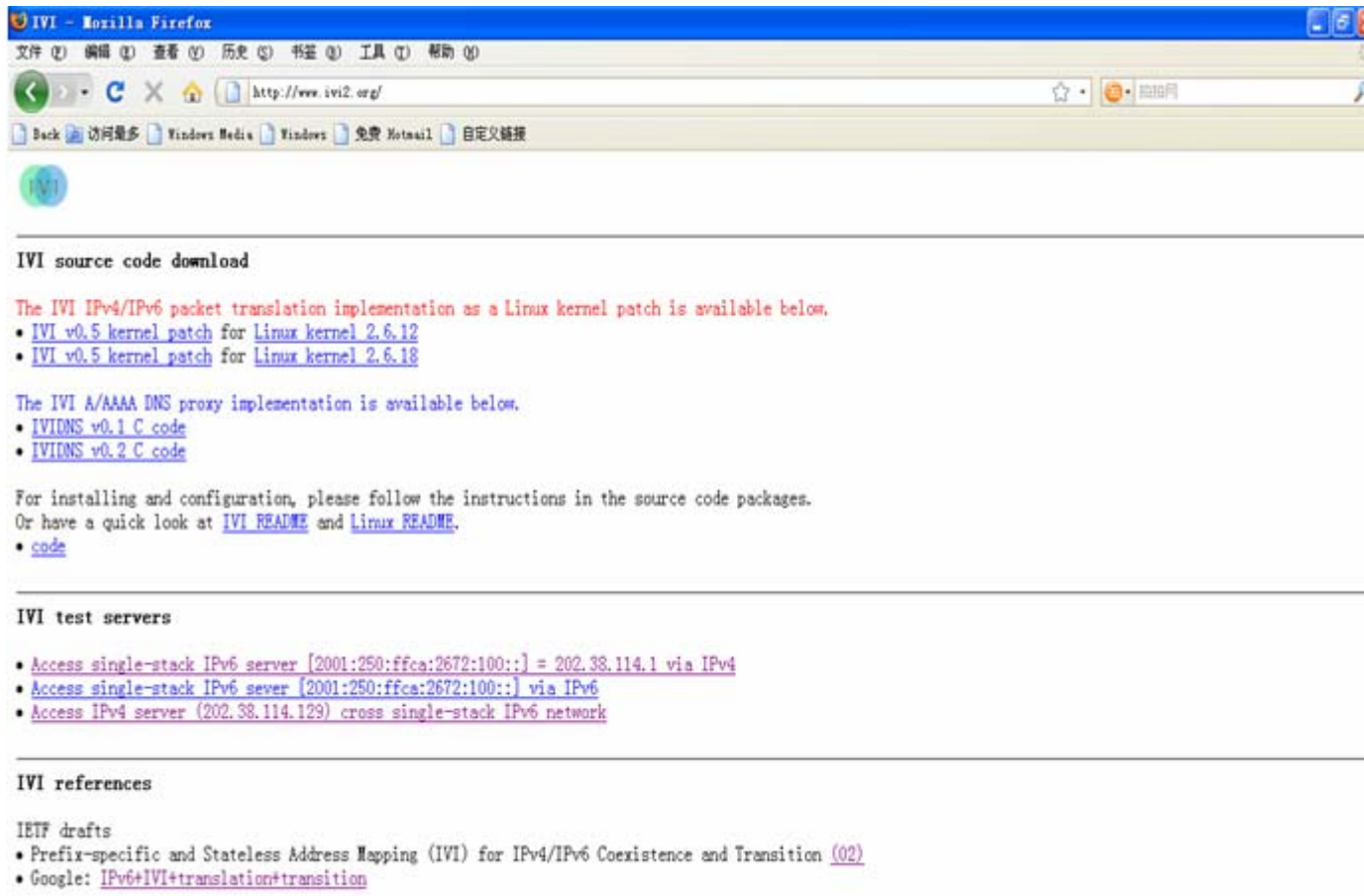
# DNS46/DNS64

---





# Linux patch source code



IVI - Mozilla Firefox

文件(F) 编辑(E) 查看(V) 历史(H) 书签(B) 工具(T) 帮助(H)

http://www.ivi2.org/

Back 访问最多 Windows Media Windows 免费 Hotmail 自定义链接

---

**IVI source code download**

The IVI IPv4/IPv6 packet translation implementation as a Linux kernel patch is available below.

- [IVI v0.5 kernel patch for Linux kernel 2.6.12](#)
- [IVI v0.5 kernel patch for Linux kernel 2.6.18](#)

The IVI A/AAAA DNS proxy implementation is available below.

- [IVIDNS v0.1 C code](#)
- [IVIDNS v0.2 C code](#)

For installing and configuration, please follow the instructions in the source code packages.  
Or have a quick look at [IVI README](#) and [Linux README](#).

- [code](#)

---

**IVI test servers**

- [Access single-stack IPv6 server \[2001:250:ffca:2672:100::\] = 202.38.114.1 via IPv4](#)
- [Access single-stack IPv6 sever \[2001:250:ffca:2672:100::\] via IPv6](#)
- [Access IPv4 server \(202.38.114.129\) cross single-stack IPv6 network](#)

---

**IVI references**

IETF drafts

- Prefix-specific and Stateless Address Mapping (IVI) for IPv4/IPv6 Coexistence and Transition ([02](#))
- Google: [IPv6+IVI+translation+transition](#)

# IVI ICMP extension

```
ivitraceroute6 www.mit.edu

src_ivi4=202.38.97.205 src_ivi6=2001:da8:ffca:2661:cd00::
dst_host=www.mit.edu
dst_ip4=18.7.22.83 dst_ivi6=2001:da8:ff12:716:5300::

traceroute to 2001:da8:ff12:716:5300:: (2001:da8:ff12:716:5300::)
30 hops max, 40 byte packets to not_ivi

 1  2001:da8:ff0a:0:100::      0.304 ms 0.262 ms 0.190 ms
    10.0.0.1
 2  2001:da8:ffca:7023:fe00::  0.589 ms * *
    202.112.35.254
 3  2001:da8:ffca:7035:4900::  1.660 ms 1.538 ms 1.905 ms
    202.112.53.73
 4  2001:da8:ffca:703d:9e00::  0.371 ms 0.530 ms 0.459 ms
    202.112.61.158
 5  2001:da8:ffca:7035:1200::  0.776 ms 0.704 ms 0.690 ms
    202.112.53.18
 6  2001:da8:ffcb:b5c2:7d00:: 89.382 ms 89.076 ms 89.240 ms
    203.181.194.125
 7  2001:da8:ffc0:cb74:9100:: 204.623 ms 204.685 ms 204.494 ms
    192.203.116.145
 8  2001:da8:ffcf:e7f0:8300:: 249.842 ms 249.945 ms 250.329 ms
    207.231.240.131
 9  2001:da8:ff40:391c:2d00:: 249.891 ms 249.936 ms 250.090 ms
    64.57.28.45
10  2001:da8:ff40:391c:2a00:: 259.030 ms 259.110 ms 259.086 ms
    64.57.28.42
11  2001:da8:ff40:391c:700::  264.247 ms 264.399 ms 264.364 ms
    64.57.28.7
12  2001:da8:ff40:391c:a00::  271.014 ms 269.572 ms 269.692 ms
    64.57.28.10
13  2001:da8:ffc0:559:dd00::  274.300 ms 274.483 ms 274.316 ms
    192.5.89.221
14  2001:da8:ffc0:559:ed00::  274.534 ms 274.367 ms 274.517 ms
    192.5.89.237
15  * * *
16  2001:da8:ff12:a800:1900:: 276.032 ms 275.876 ms 276.090 ms
    18.168.0.25
```

- Operation
  - IPv4 → IPv6
  - IPv6 → IPv4

```
ivitraceroute 202.38.108.2

 1  202.112.0.65 6 ms 2 ms 1 ms
 2  202.112.53.73 4 ms 6 ms 12 ms
 3  202.112.53.178 1 ms 1 ms 1 ms
 4  202.112.61.242 1 ms 1 ms 1 ms
 5  202.38.17.186 1 ms 1 ms 1 ms
    202.38 AS4538
 6  202.38.17.186 1 ms 1 ms 1 ms
    202.38 AS4538
 7  202.38.17.186 2 ms 2 ms 2 ms
    202.38 AS4538
 8  202.38.17.186 2 ms 2 ms 2 ms
    202.38 AS4538
 9  202.38.17.186 4 ms 4 ms 3 ms
    202.38 AS4538
10  202.38.108.2 2 ms 3 ms 3 ms
```

# Internet2 IVI trials

## Results for: IVI

grouped by location, scored using date,  
sorted by relevance

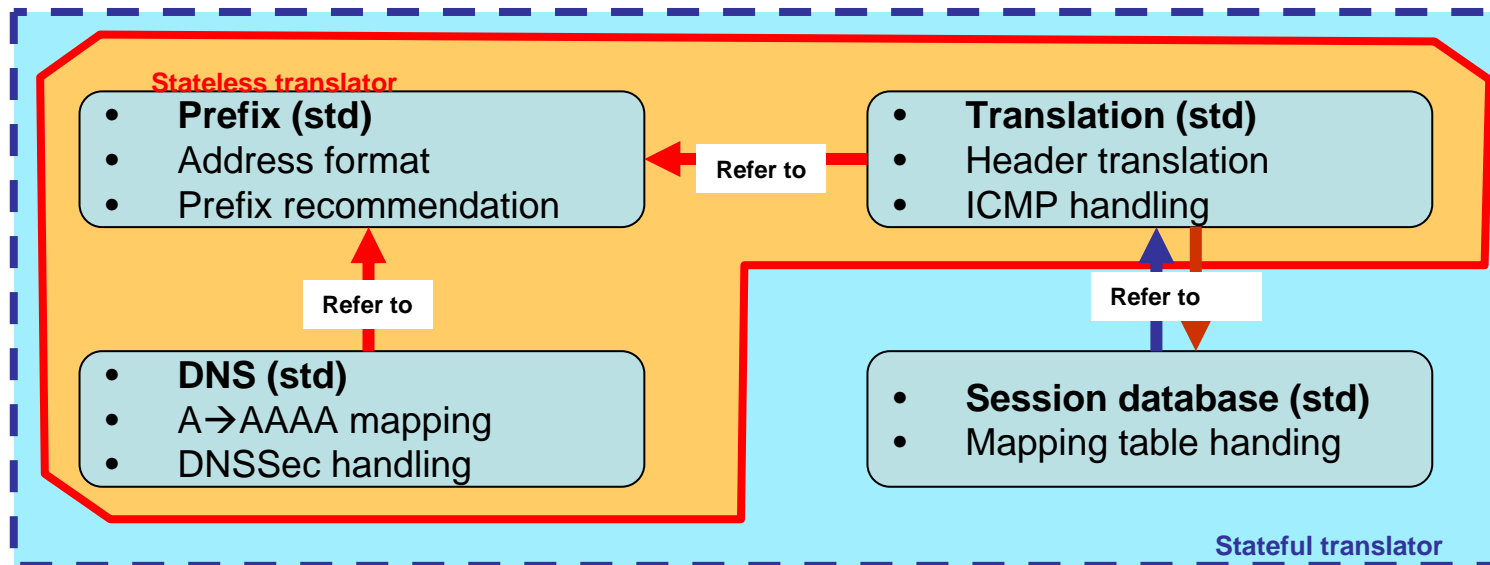
[score without date](#) [show summaries](#)  
[do not group](#)

1-6

- 54% [Summer 2009 ESCC/Internet2 Joint Techs Agenda](#)
- 54% [Winter 2009 ESCC/Internet2 Joint Techs](#)
- 45% [Summer 2009 ESCC/Internet2 Joint Techs Agenda](#)
- 45% [Summer 2009 ESCC/Internet2 Joint Techs Agenda](#)
- [View more hits from 'events.internet2.edu/2009'](#) (Group score: 44%)
- 54% [Internet2 - Digital Video Library](#)
- 50% [Internet2 - Digital Video Library](#)
- [View more hits from 'www.internet2.edu/video'](#) (Group score: 42%)
- 47% [Performance-related Sessions | Joint Techs](#)
- 46% [Internet2 Speakers Bureau: Presentations](#)
- 46% [Internet2 Speakers Bureau: Presentations](#)
- 46% [Internet2 Speakers Bureau: People: Mitsuru Kanda](#)
- 46% [Internet2 Speakers Bureau: People: Chris Small](#)
- [View more hits from 'events.internet2.edu/speakers'](#) (Group score: 46%)
- 35% [Program for Winter 2009 ESCC/Internet2 Joint Techs](#)
- 25% [Recovery Bill Div A](#)

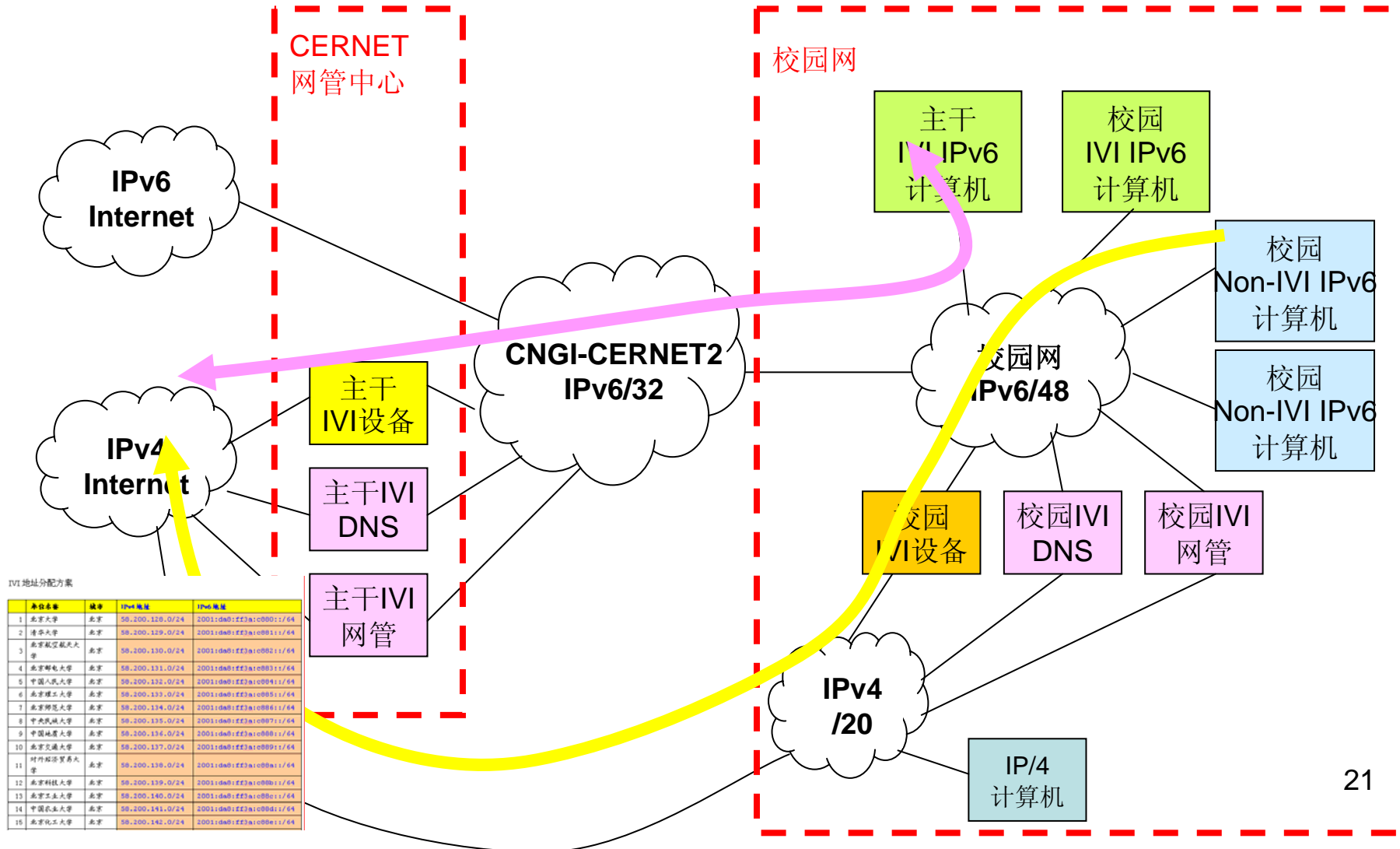
# IETF behave: document layout

- Framework (info)
- Scenarios
- Operation modes
- Building blocks



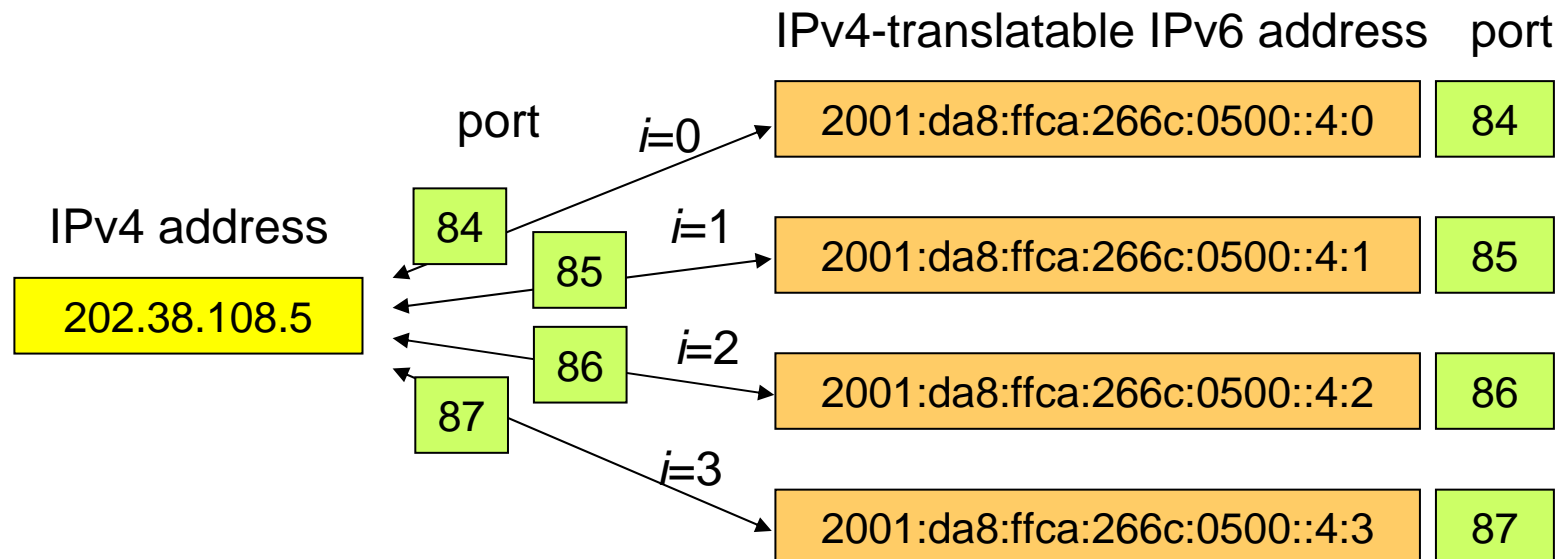
- Others (APL-ALGs, multicast, ...)

# CNGI-2 (CERNET)

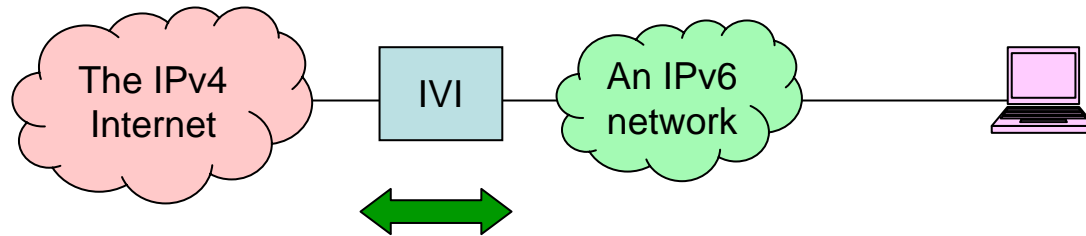


# Design concept of stateless 1:N IIVI

- Keep the basic features of the stateless translation (IVI)
- An IPv4 public address is shared by N IPv6 hosts (in this example N=4).
- The port number range for each IPv6 host is predefined.
- The port number range is encoded in the IPv6 address (IPv4-translatable address) and no signaling scheme is required.

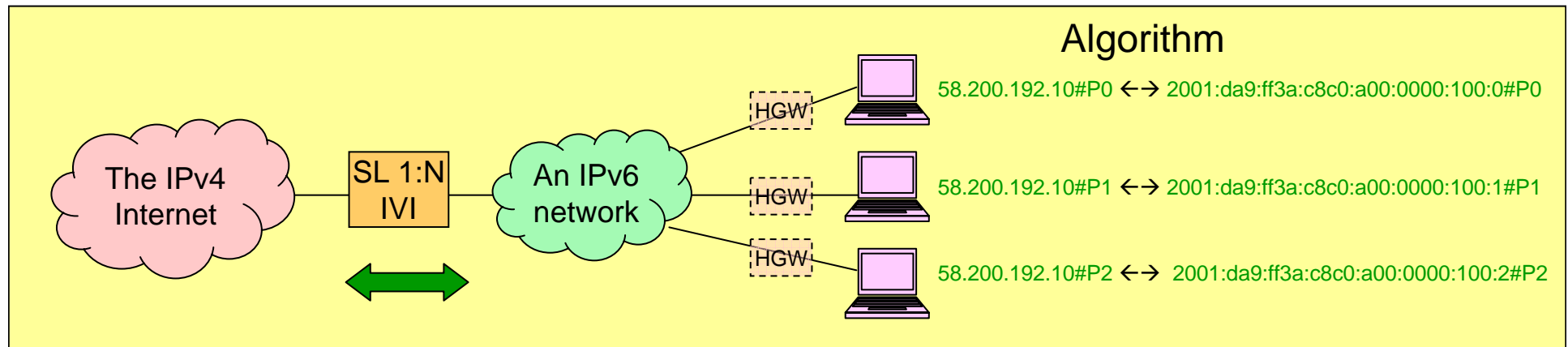


# stateless 1:N IIVI (dIVI) solution



Algorithm

202.38.114.1 ↔ 2001:250:ffca:2672:0100::0

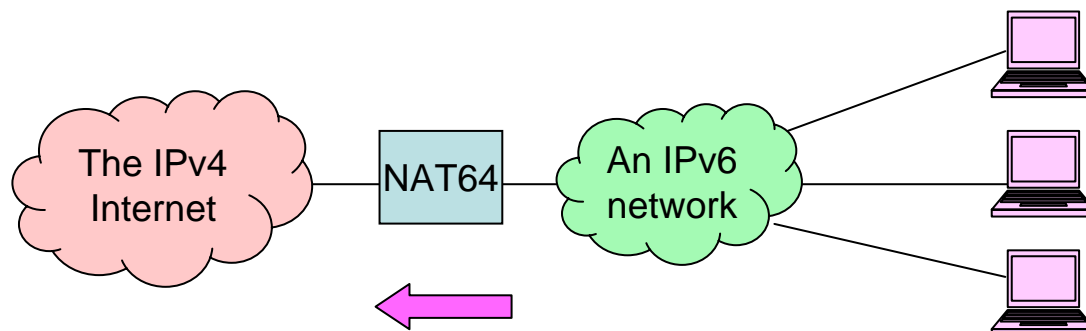


Algorithm

58.200.192.10#P0 ↔ 2001:da9:ff3a:c8c0:a00:0000:100:0#P0

58.200.192.10#P1 ↔ 2001:da9:ff3a:c8c0:a00:0000:100:1#P1

58.200.192.10#P2 ↔ 2001:da9:ff3a:c8c0:a00:0000:100:2#P2



State database

202.38.102.1#2000 ↔ 2001:da8::100#3000

202.38.102.1#2001 ↔ 2001:da8::101#3000

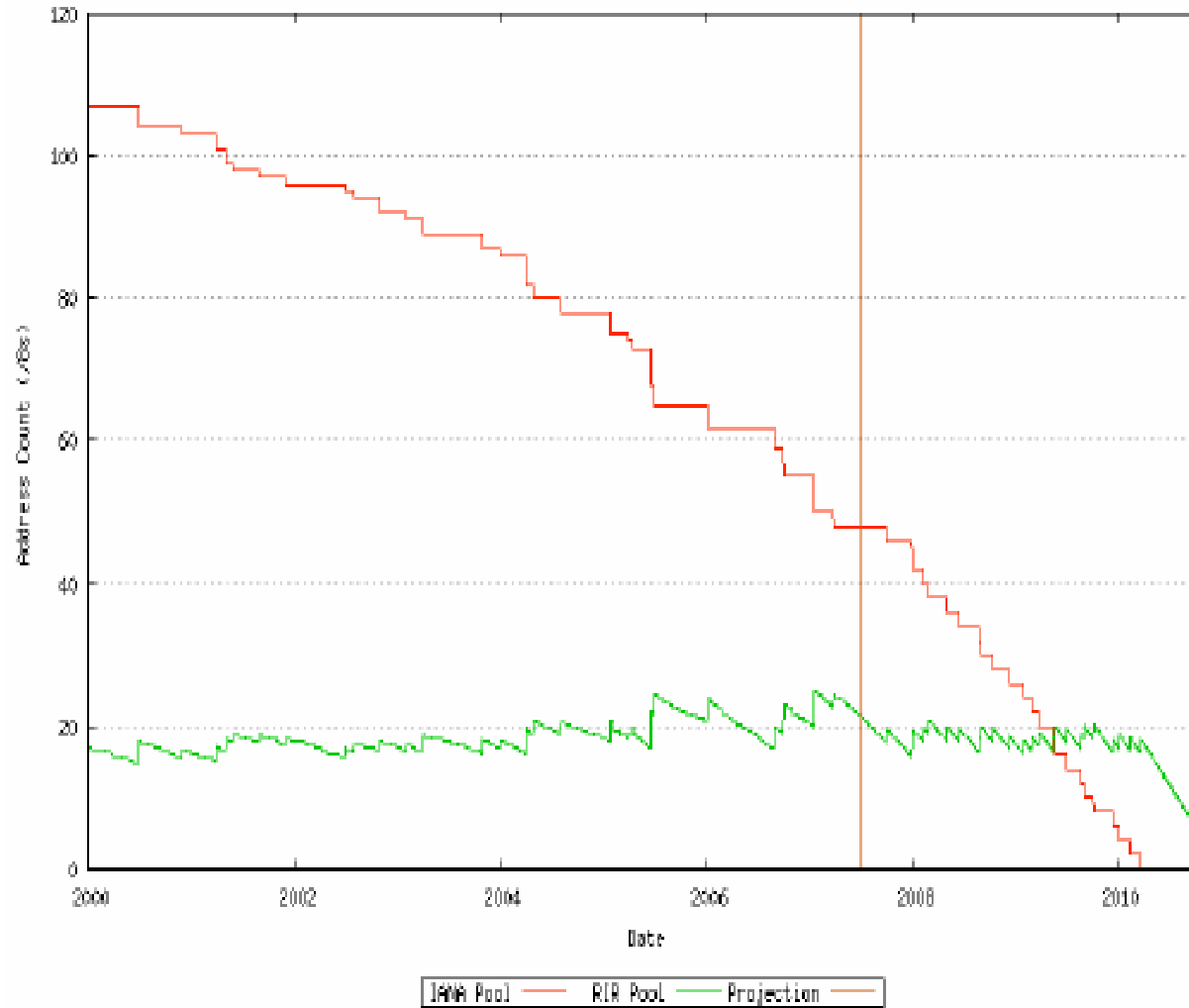
202.38.102.1#2002 ↔ 2001:da8::200#3000

# Comparisons

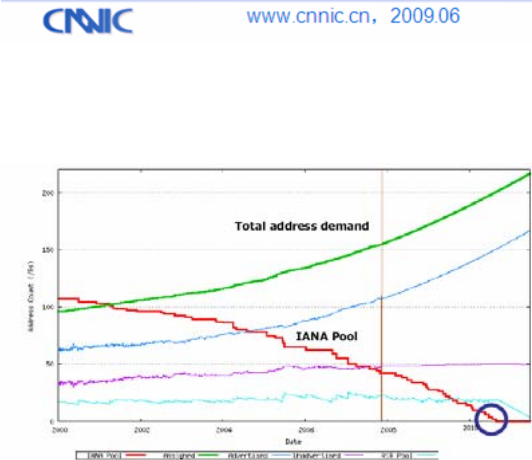
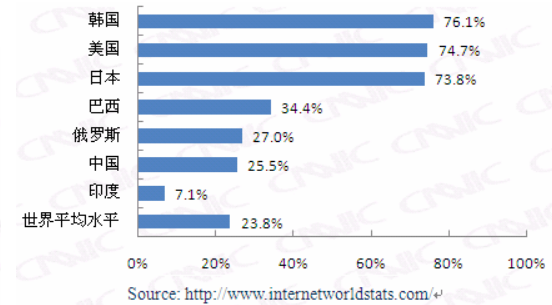
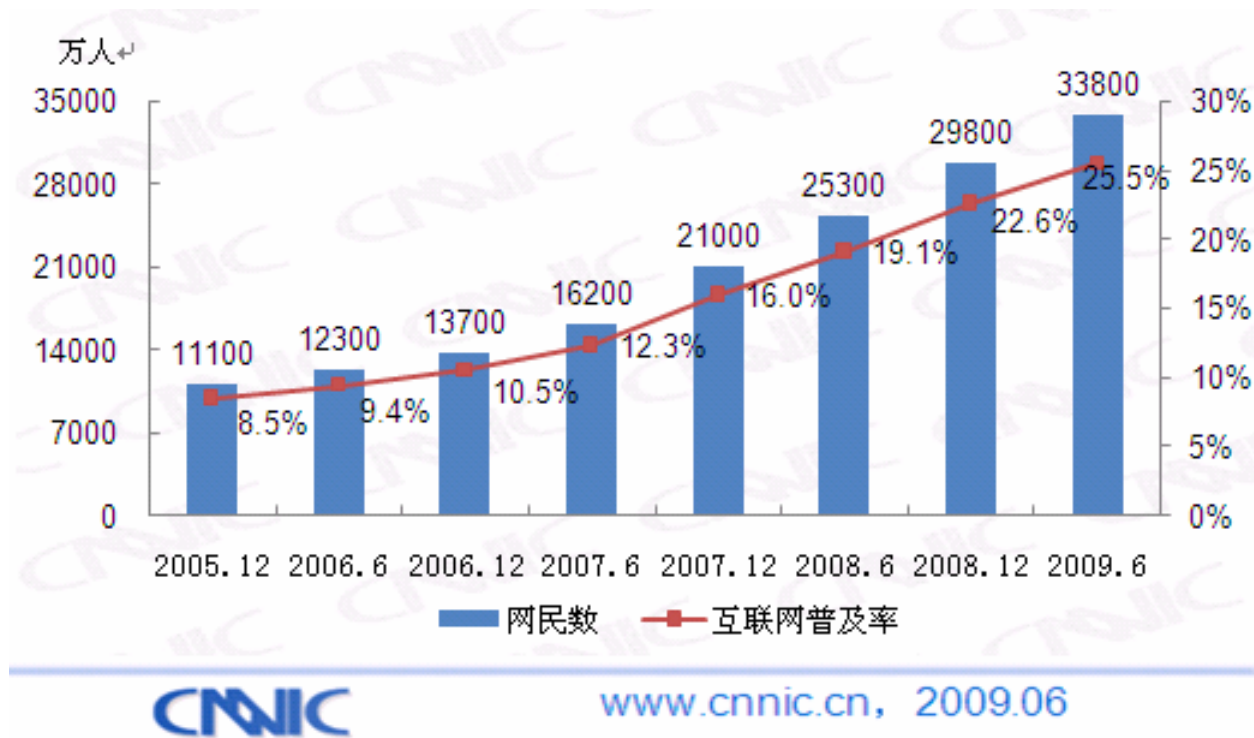
- Stateless 1:1 IVI
  - server/client full-function (bi-direction initiated)
  - Restricted IPv6 addresses (few, 1:1 mapping)
  - Stateless and scalable
  - Need DNS64 and ALG
- Stateless 1:N IVI (dIVI)
  - Server/client limited function (bi-direction initiated)
  - Restricted IPv6 addresses (many, 1:N mapping)
  - Stateless and scalable
  - Do not need DNS64 and ALG (dIVI)
- Stateful 1:N NAT64
  - Client limited function (IPv6 initiated only)
  - Unrestricted IPv6 addresses (many, many, M IPv6 hosts share n IPv4 addresses)
  - Stateful and less scalable
  - Need DNS64 and ALG



# IPv4 address depletion



# IPv4 address demand in China



# China telecom



网络传送：更广泛的互联互通

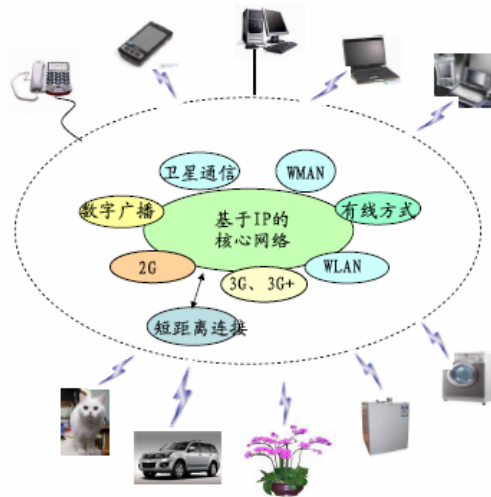


■ 物联网传送技术以互联网为核心

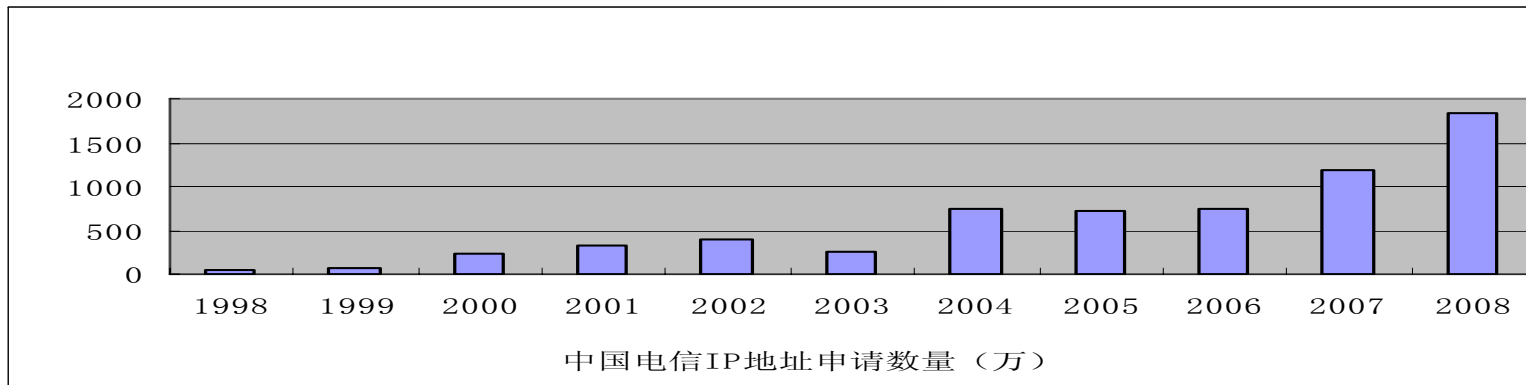
■ 引入IPv6是实现更广泛互联互通的基础

- 海量的物物接入，现有IP地址明显不足，IPv4地址将在2011年耗尽
- IPv6有助于创新应用

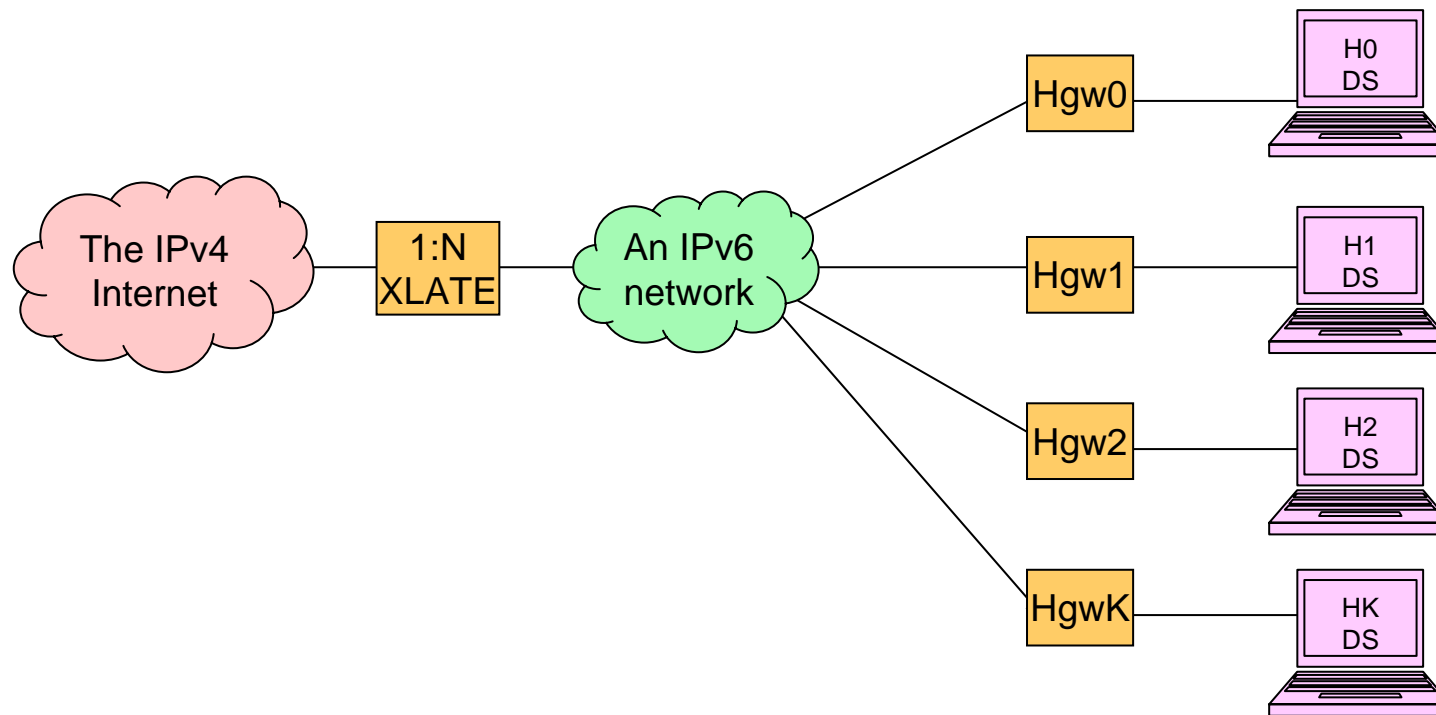
■ 中国电信全球第一个通过了面向运营商的国际权威IPv6就绪认证



◆ China Telecom lacks 16 M to 110M IPv4 addresses in the next 3 to 5 years

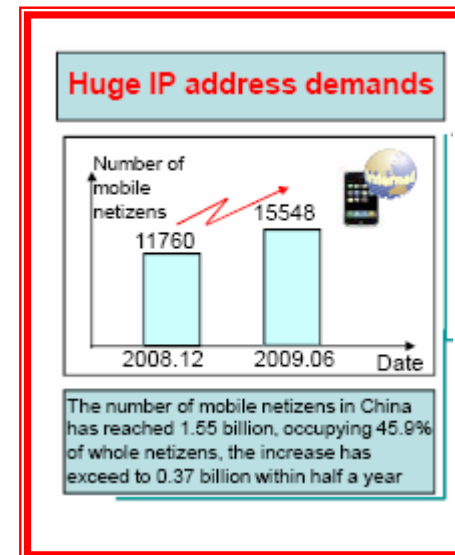
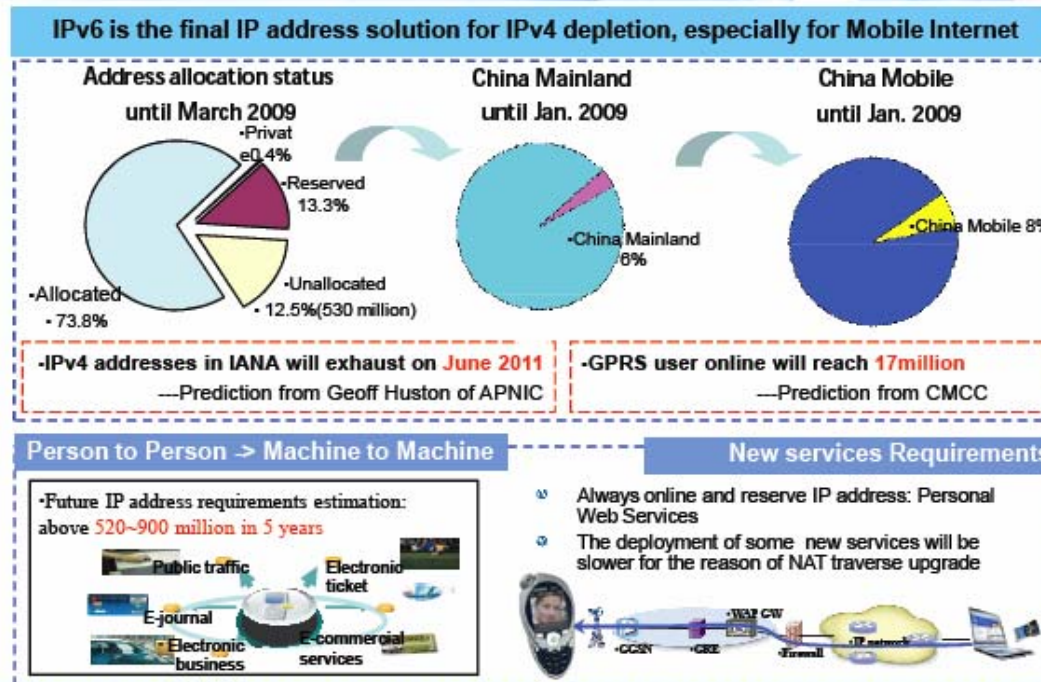


# Home gateway implementation

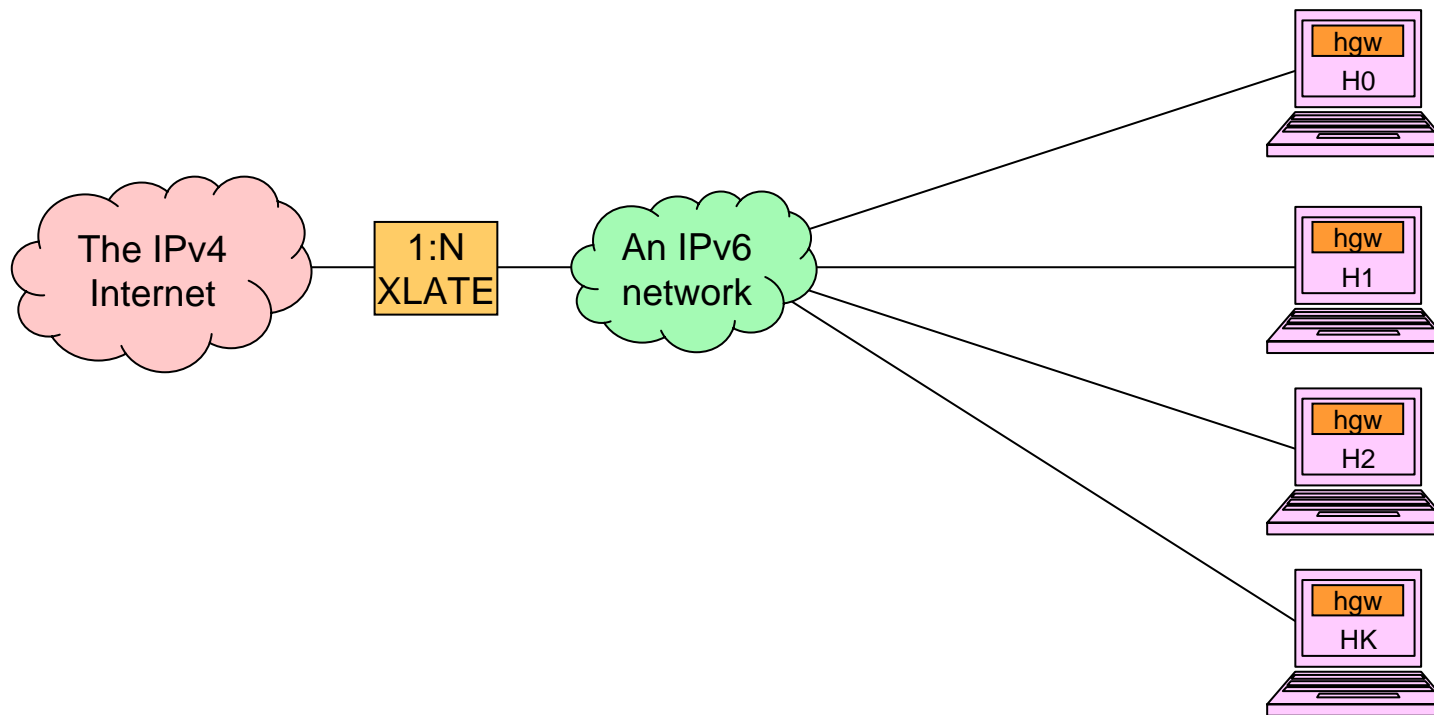


# China mobile

## Status of IPv4 address allocation and problems with Private IPv4



# End system implementation



# Transition mechanisms

---

- When IPv4 addresses are running out
  - IPv4 + NAT
    - Short term solution
  - Pure IPv6
    - Cannot reach the global IPv4, unacceptable
  - Dual stack
    - The cost increases, ISPs want others to deploy dual stack
  - IVI IPv6
    - The cost is the same as the single stack, but the IPv6 host can be reached by global IPv4

# Transition concept

