Bi-directional Address Borrowing (IVI)

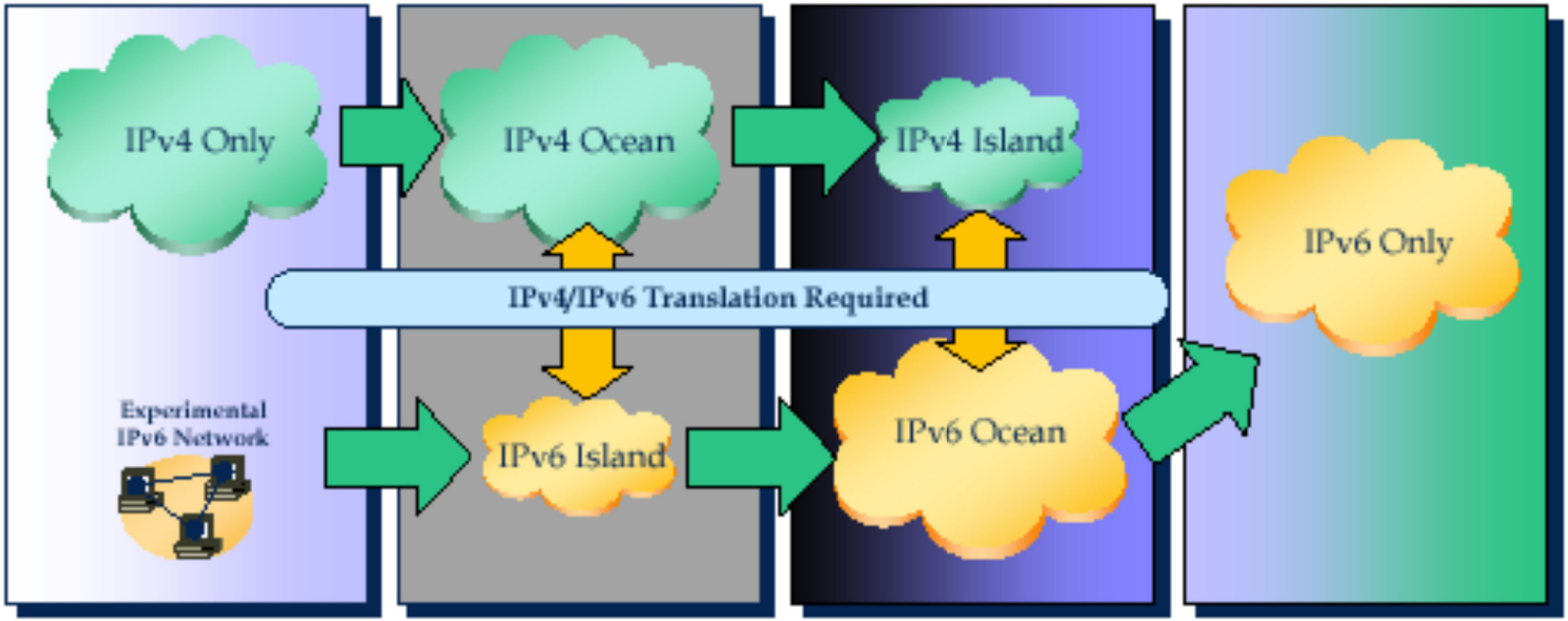
Xing Li, Maoke Chen, Congxiao Bao, Hong Zhang, Jianping Wu

2008-06-23

Problem statement

- During the transition time, the IPv6 hosts need to communicate with the global IPv4 network.
- The current transition methods do not fully support this requirement and it seems impossible to fulfill this requirement without restrictions (end-to-end, scalability, etc).
- However, it is possible to allocate a subset of the IPv6 addresses which can communicate with both the global IPv6 network and the global IPv4 network (This does not mean that the global IPv6 network can communicate with the global IPv4 network).
- This is the goal of this proposal.

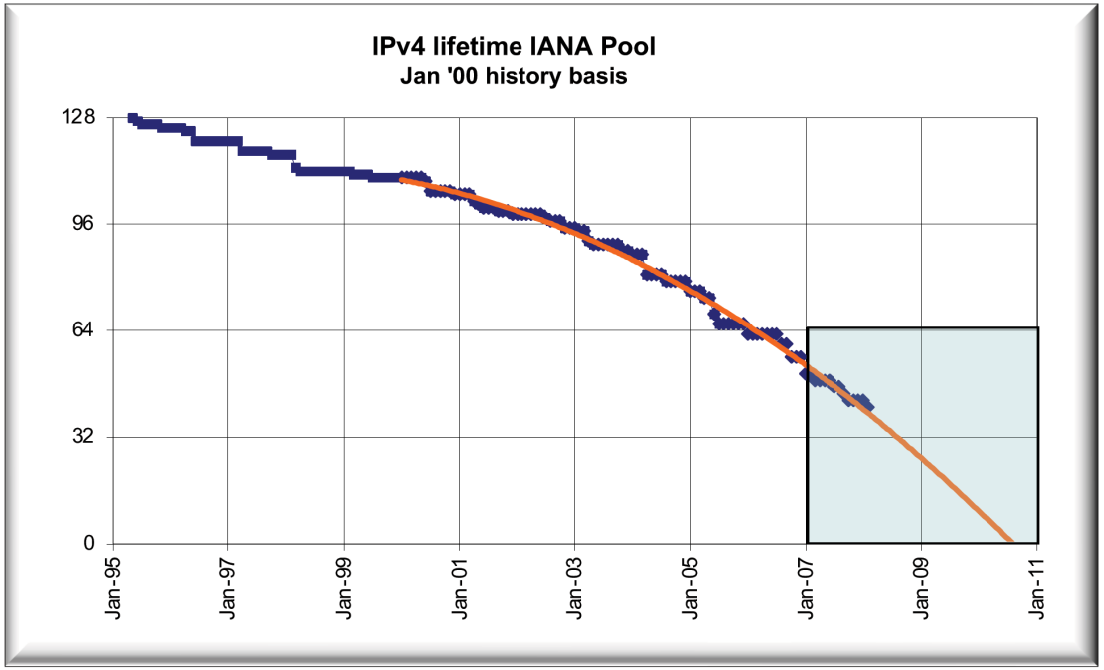
IPv6 transition roadmap



Transition methods

- Dual stack approach
  - IPv4 address depletion problem
- Tunneled architectures
  - No communication between two address families
- Translation architectures
  - Not scalable, lost end-to-end
- New transition method requirements
  - End-to-end address transparency, minimum state, globally deliverable, multi-homing support and effectively use of the global IPv4 addresses.

Remaining IANA IPv4 Pool



http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_8-3/ipj_8-3.pdf

Tony Hain

The existing mapping methods

2. Address Blocks

2.1. Node-scoped Unicast

::1/128 is the loopback address [RFC4291].

::/128 is the unspecified address [RFC4291].

Addresses within this block should not appear on the public Internet.

2.2. IPv4-Mapped Addresses

::FFFF:0:0/96 are the IPv4-mapped addresses [RFC4291].

Addresses within this block should not appear on the public Internet.

2.3. IPv4-compatible Addresses

::ipv4-address/96 are the IPv4-compatible addresses [RFC4291].

These addresses are deprecated and should not appear on the public Internet.

2.4. Link-scoped Unicast

fe80::/10 are the link-local unicast [RFC4291] addresses.

Addresses within this block should not appear on the public Internet.

2.5. Unique-Local

fc00::/7 are the unique-local addresses [RFC4193].

Addresses within this block should not appear by default on the public Internet.

Procedure for advertising these addresses are further described in [RFC4193].

- IPv4-mapped and IPv4-compatible addresses are not routable and cannot achieve the new transition goal

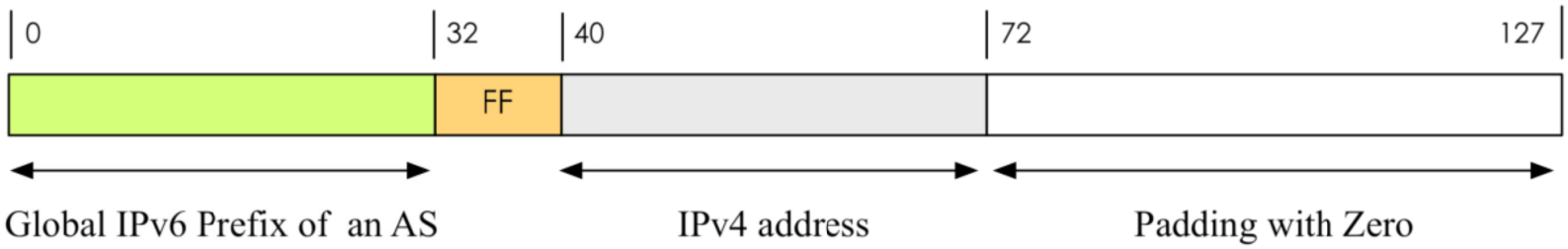
Key concepts of new transition method

- Goal
  - End-to-end address transparency, minimum state, globally deliverable, multi-homing support and effectively use of the global IPv4 addresses
- Key points
  - The global IPv4 space can be mapped into every /32 of the IPv6 space, this is a special IPv6 block in each /32 which has one-to-one mapping relationship between IPv4 and IPv6 in the corresponding /32 and globally routable in both address families.
  - A subset of global IPv4 address space can be borrowed and mapped into a subset of the above special IPv6 block. This IPv6 address subset can communicate with the global IPv6 directly and global IPv4 via a gateway.
  - Temporal and address/port multiplexing methods can be applied for higher utilization of the above subset of the IPv4 address block.

Definition

- IVI
  - IV means 4 and VI means 6 in Roman representation, so IVI means mapping and transition between IPv4 and IPv6.
- IVI gateway
  - The mapping and forwarding gateway between two address families.
- IVI4(i)
  - Borrowed IPv4 address (IVI6 equivalent in IPv4)
- IPG4
  - Global IPv4 address (IVIG equivalent in IPv4)
- IVI6(i)
  - Mapped IPv6 address which is borrowed from IPv4 (IVI4 equivalent in IPv6)
- IVIG46(i)
  - Global IPv4 address (IPv4 equivalent in IPv6)
- IPG6
  - Natively assigned IPv6 address

Address mapping (1)

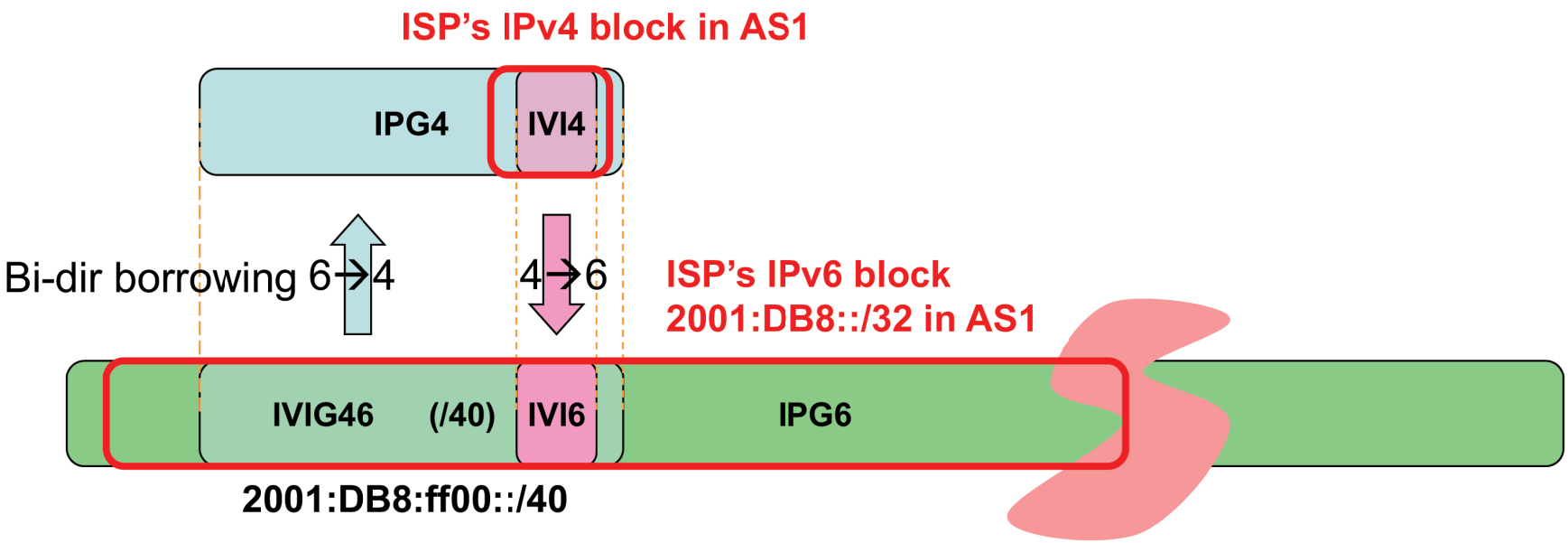


Mapping Rule: IPv4 addresses (both borrowed/non-borrowed) are embedded from bit 40 to bit 72 of the IPv6 addresses of a specific /32.

Example:

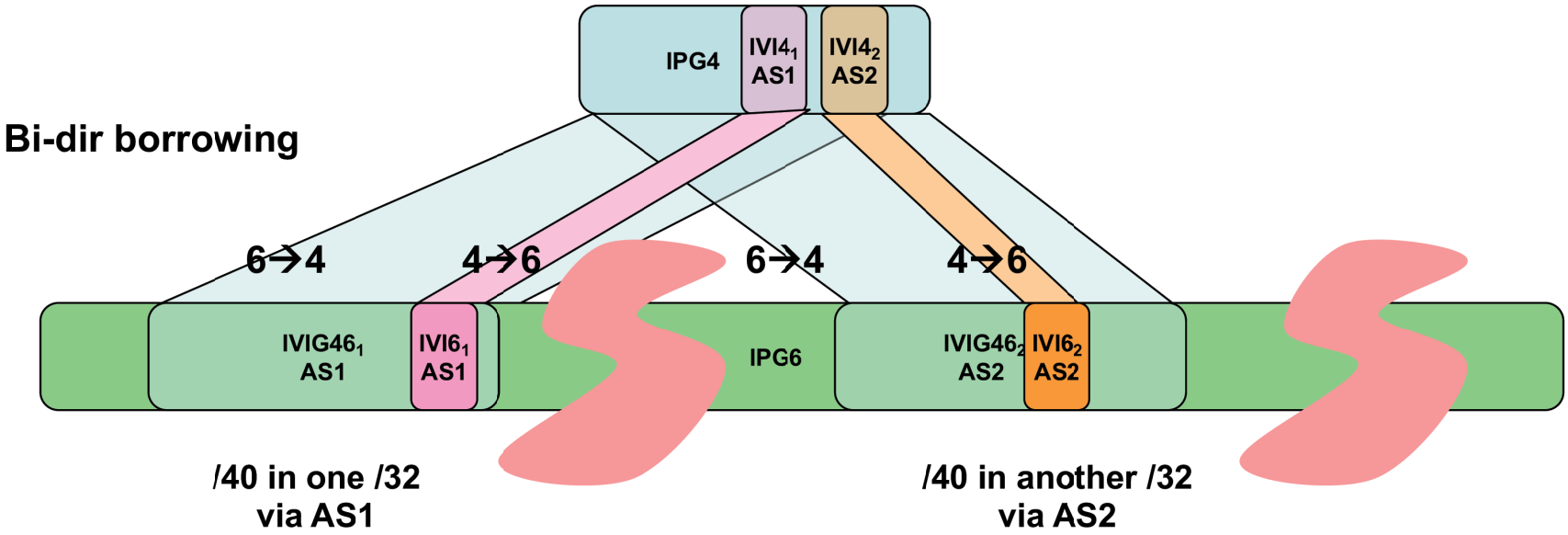
ISP's IPv6 /32	2001:250::/32
borrowed IPv4 address (IVI4):	202.38.108.0/24
mapped IVI IPv6 address (IVI6):	2001:250:ffca:266c::/64

Address mapping (2)



- IVI is deployed in one autonomous system (/32)

Address mapping (3)



- IVI is deployed in multiple autonomous systems (/32s)

Address mapping (4)

- There are address set (A6) (128 bits) and address set (A4) (32 bits).
- Each ISP_i gets $(A4)_i$, which is a subset of (A4), and $(A6)_i$, which is usually a /32 and it is a subset of (A6).
- Each ISP_i reserves a subset $(IVIG)_i$, which is a /40 from /32 $(A6)_i$.
- Each ISP_i borrows a subset $(IVI4)_i$ from $(A4)_i$. The corresponding IPv6 equivalent is $(IVI6)_i$, which is a subset of $(IVIG)_i$.
- The $(IVI6)_i$ can communicate with general IPv6 and the corresponding equivalent $(IVI4)_i$ can communicate with global IPv4.
- The relations of these sets are

$$IPv4 \quad space: \quad (A4) \supseteq (A4)_i \supseteq (IVI4)_i$$

$$IPv6 \quad space: \quad (A6) \supseteq (A6)_i \supseteq (IVIG)_i \supseteq (IVI6)_i$$

$$1-to-1 \quad mapping: \quad (A4) \Leftrightarrow (IVIG)_i$$

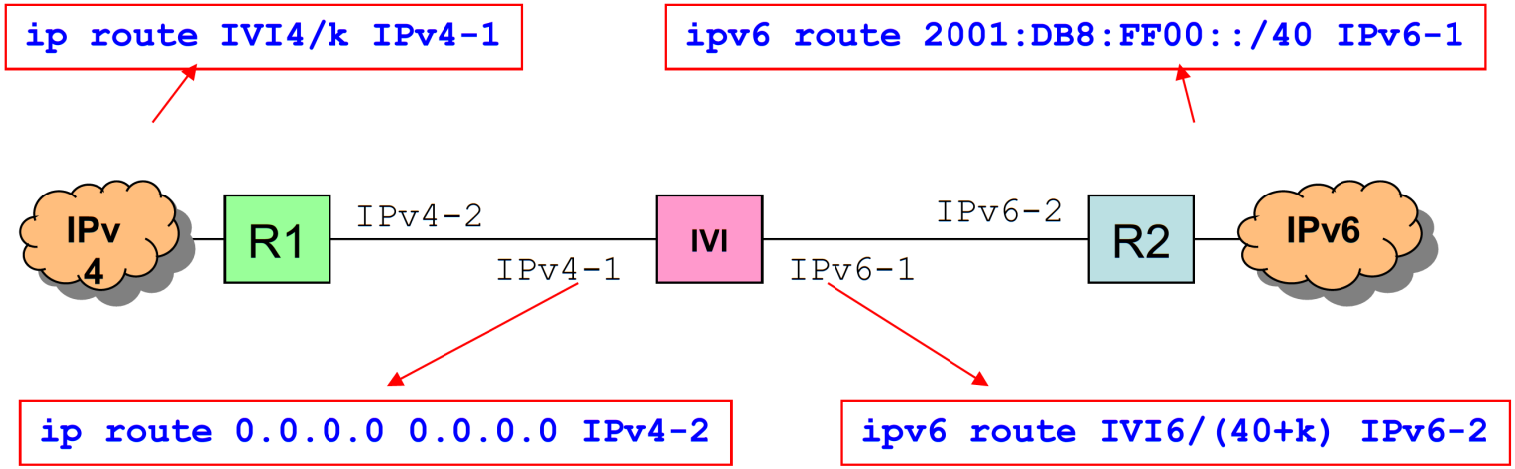
$$1-to-1 \quad mapping: \quad (IVI4)_i \Leftrightarrow (IVI6)_i$$

$$For \quad i \neq j, \quad (IVI6)_i \cap (IVI6)_j = 0$$

$$For \quad i \in I \quad j \in I, \quad \bigcup_i (IVI6)_i = \bigcap_j (IVIG)_j$$

Routing and forwarding

Routing and mapping configuration example



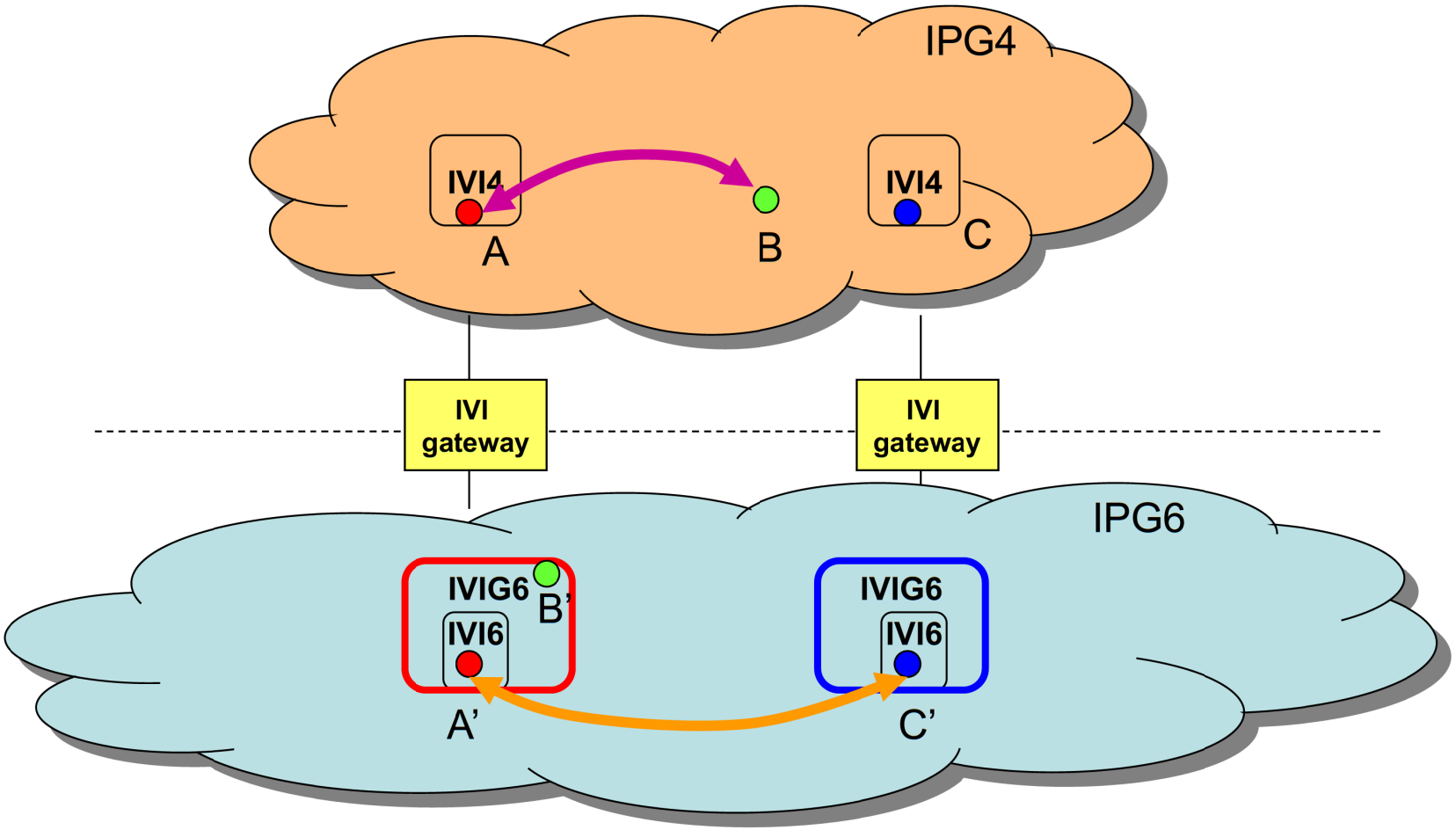

```
mroute IVI4-network IVI4-mask pseudo-address interface source-PF destination-PF
mroute6 destination-PF destination-PF-pref-len
```

Routing and forwarding

Summary of the reachability:

	IPv4	IVI	IPv6
IPv4	OK	OK	NO
IVI	OK	OK	OK
IPv6	NO	OK	OK

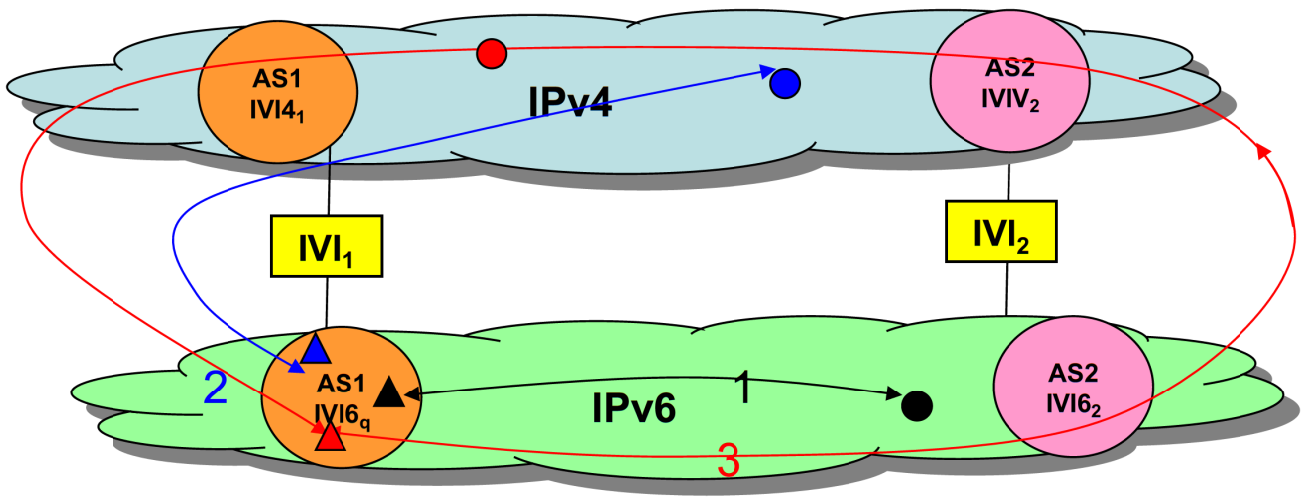
Address evolution



Routing and forwarding

- IPv4 global routing table
  - The borrowed IPv4 address is from a block of an aggregated global routable IPv4 address, therefore, the global IPv4 routing table will not be affected.
- IPv6 global routing table
  - Outside the specific autonomous system which supports IVI, only aggregated /32 appears, therefore, the global IPv6 routing table will not be affected.
- Multi-homing
  - In the case of one-to-one mapping between IVI4 and IVI6, the mapping is stateless and the multi-homing is supported.

Routing scenarios



1. For IPv6 address family communication
   - IVI6$_1$ to any IPv6 ➔ not via IVI box
2. For cross address family communication
   - IVI6$_1$ to any IPv4 using AS1's own /32 prefix → via the AS's IVI box
3. For cross address family communication
   - IVI6$_1$ to any IPv4 using other AS's /32 prefix
     - Deployed IVI → via other AS's IVI box, reachable, but asymmetric
     - Not deployed IVI → unreachable to a wrong destination

Header translation (from IPv4 to IPv6)

IPv4 Field	Translated to IPv6
Version (0x4)	Version (0x6)
IHL	(discarded)
Type of Service	(discarded)
Total Length	Payload Length = Total Length -IHL * 4
Identification	(discarded, cf. Subsection V-C)
Flags	(same as above)
Offset	(same as above)
Time to Live	Hop Limit
Protocol	Next Header
Header Checksum	(discarded)
Source Address	Apply IVI stateless address mapping
Destination Addr.	(same as above)
Options	(discarded)

Header translation (from IPv6 to IPv4)

IPv6 Field	Translated to IPv4 Header
Version (6)	Version (4)
Traffic Class	(discarded)
Flow Label	(discarded)
Payload Length	Total Length = Payload Length + 20
Next Header	Protocol
Hop Limit	TTL
Source Address	Apply IVI inverse address mapping
Destination Addr.	(same as above)
—	IHL = 5
—	Header Checksum recalculated

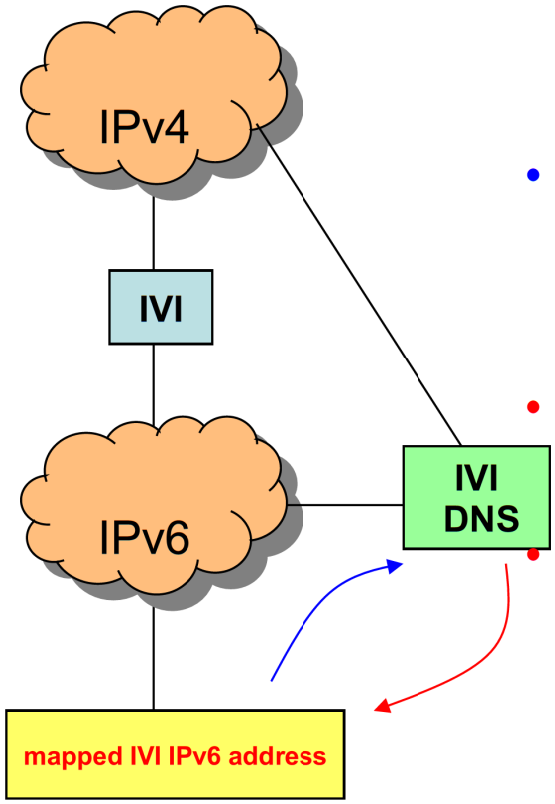
List of the IVI compatible applications

- Fully support
  - ssh
- partial support
  - http (addr in header does not work for virtual host, use DNS instead)
- Not support
  - ftp (Changing the behavior of the FTP server is a workaround, or via ALG)

Selection of the source address

- Since IPv6 host may have multiple addresses, it is important for the host use IVI6 address for reaching the global IPv4 networks.
- Short term work around: use IVI6 as the default IPv6 address of the host.
- Long term solution: application should be able to select the source address for different service.

DNS



- For providing primary DNS service of the borrowed IPv4 address (IVI4) and the corresponding mapped IVI IPv6 address (IVI6), each host will have both A and AAAA records
- Authoritative DNS server
  - Example
  - www.ivi2.org A 202.38.108.2
  - www.ivi2.org AAAA 2001:250:ffca:266c:200::
- For resolving of the IVI IPv6 mapped global IPv4 space (IVI6), use IVI DNS to do the dynamic mapping based on the IVI rule.
- Caching DNS server
  - Example
  - www.mit.edu A 18.7.22.83
  - www.mit.edu AAAA 2001:250:ff12:0716:5300::

Temporal multiplexing of the global IPv4 address

- IVI6 can be temporally multiplexed inside the ISP's /32, i.e. dynamic assignment of the IVI6 to the end systems when it requests the service.
- The corresponding fee can be charged, therefore, results in a more sustainable business model.

Address/port multiplexing of the global IPv4 address (1)

- IVI6 can be address/port multiplexed.
  - Since public-IPv4 addresses have become a precious resource, borrowing a public-IPv4 address with complete exclusivity in a user login session would be a waste. Therefore, a single borrowed IPv4 IVI address can be used for multiple IPv6 IVI addresses.
- This method is incremental
  - 1-to-1 stateless, gain=1
  - 1-to-2 almost stateless, gain=2
  - 1-to-N stateful, gain=N
- For example
  - IPv6 /32 2001:250::/32
  - borrowed IPv4 address (IVI4): 202.38.108.5/32
  - mapped IVI IPv6 address (IVI6): 2001:250:ffca:266c:0500::/72
  - Hence, if the gain=16, the IVI6s are from
    2001:250:ffca:266c:0500::0/72
    to 2001:250:ffca:266c:0500::15/72

Address/port multiplexing of the global IPv4 address (2)

- The tradeoffs
  - The number of the port number 2^{16}
  - The gain of the IPv4 address utilization
  - Half association or full association
  - The number of states
  - The average concurrent port used in a IPv6 host
  - The collision ratio of the port number
- For example (gain=4, case of collision, half association)
  - `202.38.108.5#100` ←→ `2001:250:ffca:266c:0500::0#100`
  - `202.38.108.5#101` ←→ `2001:250:ffca:266c:0500::1#100`
  - `202.38.108.5#102` ←→ `2001:250:ffca:266c:0500::2#100`
  - `202.38.108.5#103` ←→ `2001:250:ffca:266c:0500::3#100`

Spatial multiplexing of the global address

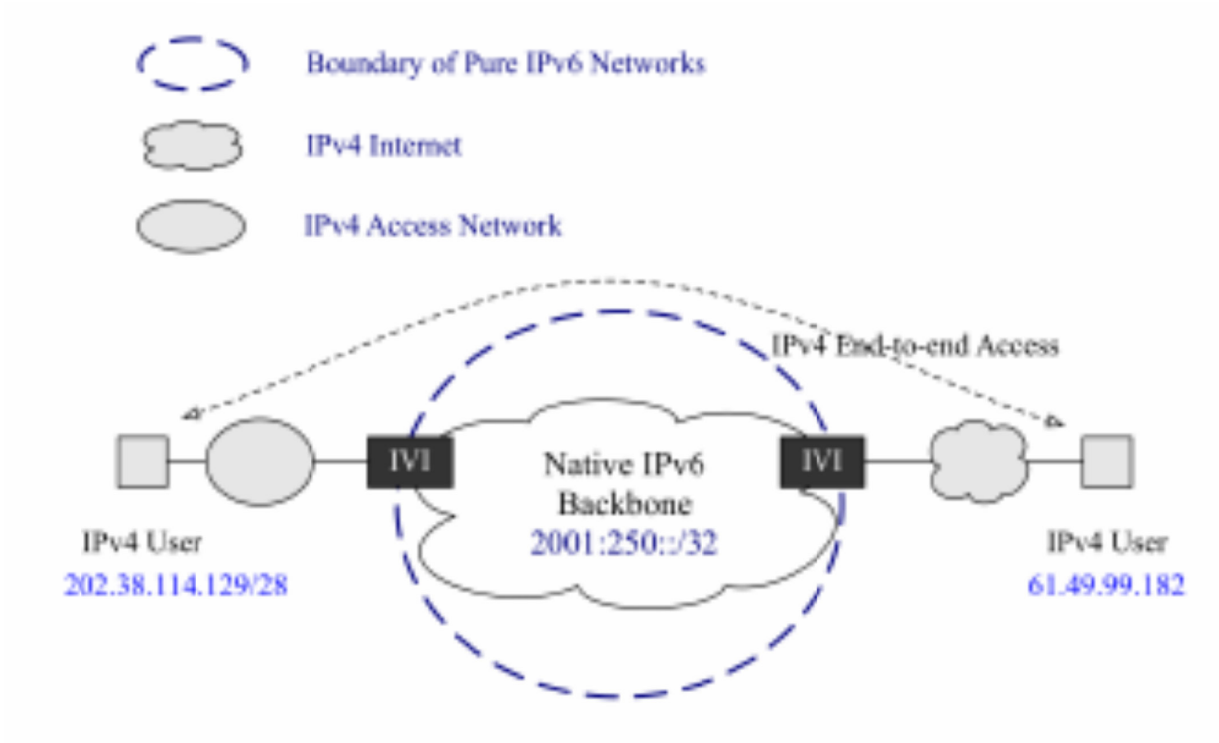
- For the server
  - 1-to-1 mapping between IVI4 and IVI6
    - Stateless
    - Support multi-homing
    - Provide service via well known port
- For the client
  - 1-to-N mapping between IVI4 and IVI6
  - Self initiated connection
    - The number of state can be controlled
    - Random port number, IVI gateway maintain the address/port mapping table to avoid collision
    - No modification is needed for the client OS and/or application
  - Peer initiated connection
    - The number of state can be controlled
    - Define pseudo-well-know port, which is unique for IVI4 and known to the peers
    - Modification is needed for the client OS and/or application

Multicast support

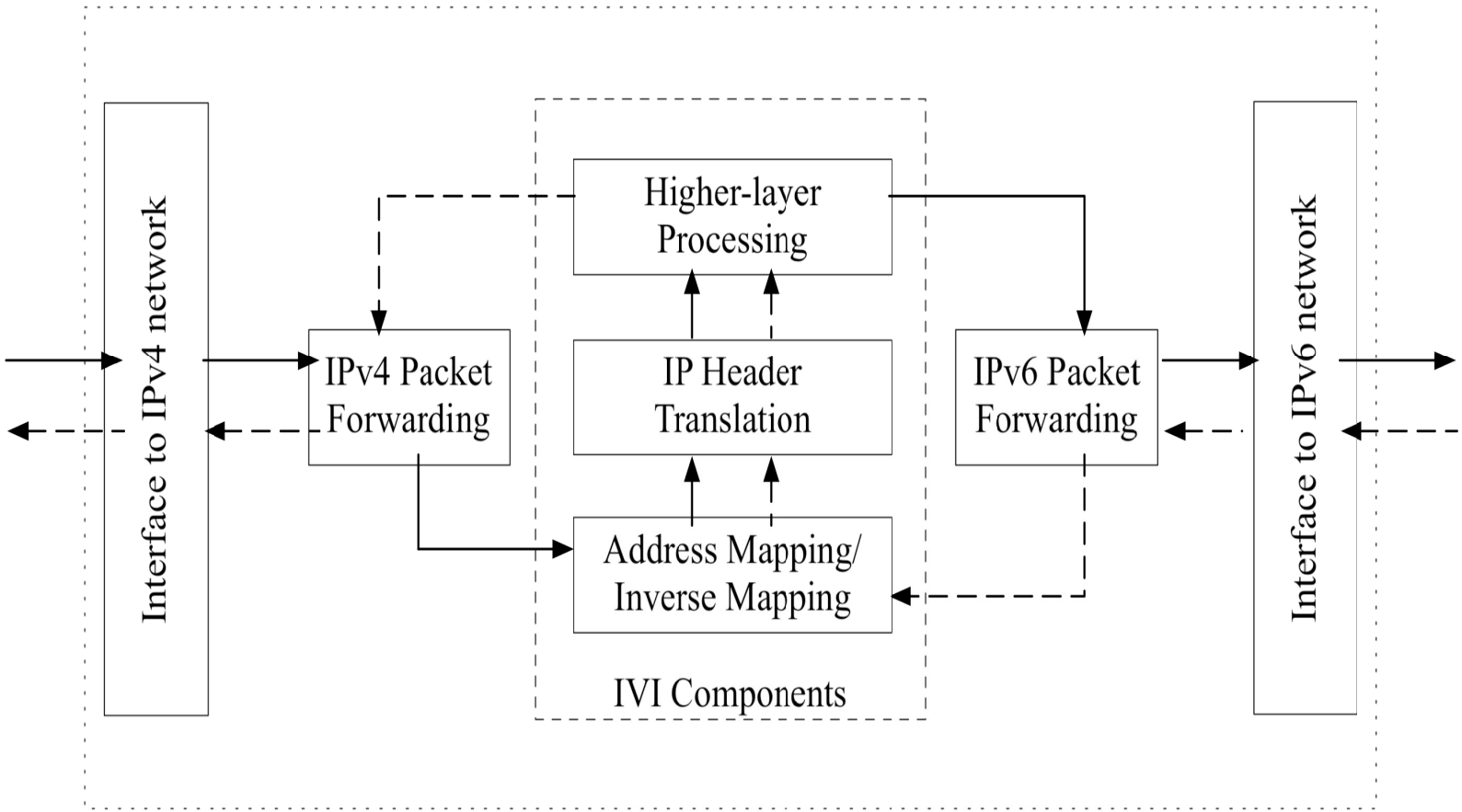
- Only SSM is supported for the IVI
  - no MSDP in IPv6
  - no embedded RP in IPv4
- Group address mapping rule (there will be 2^{24} group ID available)
  - `232.0.0.0/8` → `ff3e:0:0:0:0:0:f000:0000/96`
  - `232.255.255.255/8` → `ff3e:0:0:0:0:0:f0ff:ffff/96`
- For the cross address family SSM
  - the source address in IPv6 has to be IVI6 for the RPF scheme
- The inter operation of PIM-SM in IPv4 and IPv6
  - Application layer gateway
  - Static join using IGMPv3 and MLDv2

IPv4-IPv6-IPv4 mapping

- It is also possible to have mapped version of IPv4 over IPv6

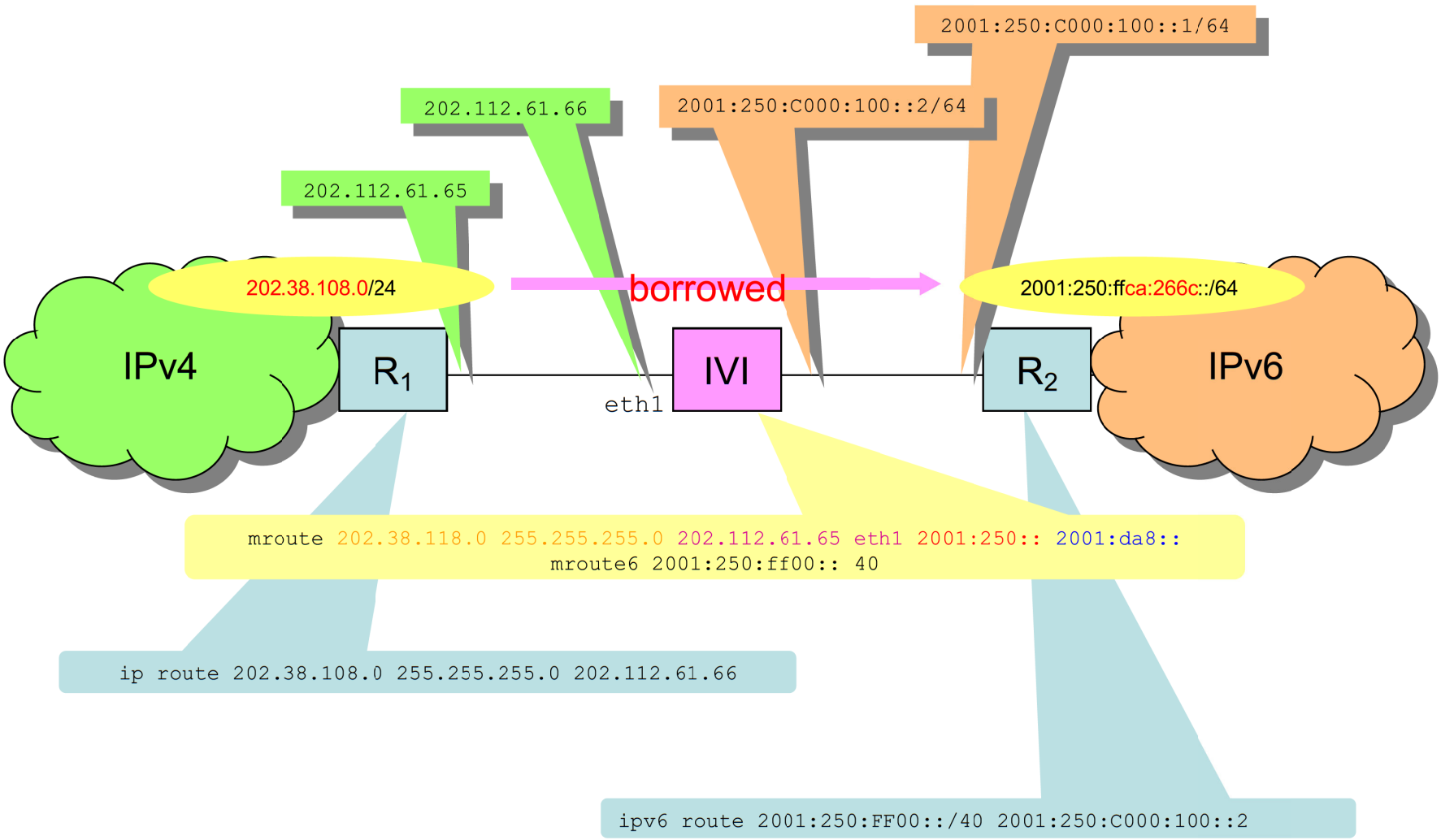


Implementation

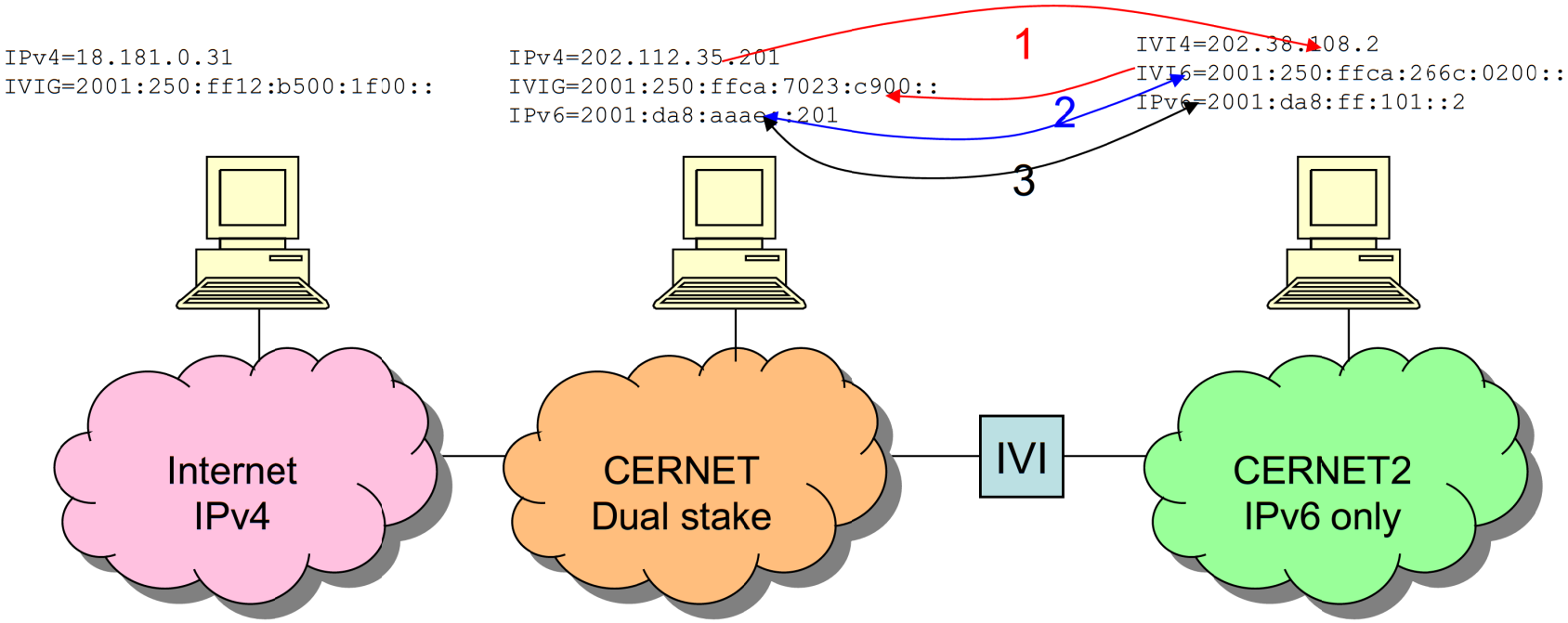


The code can be downloaded in http://linux.ivi2.org/impl/

Experiments

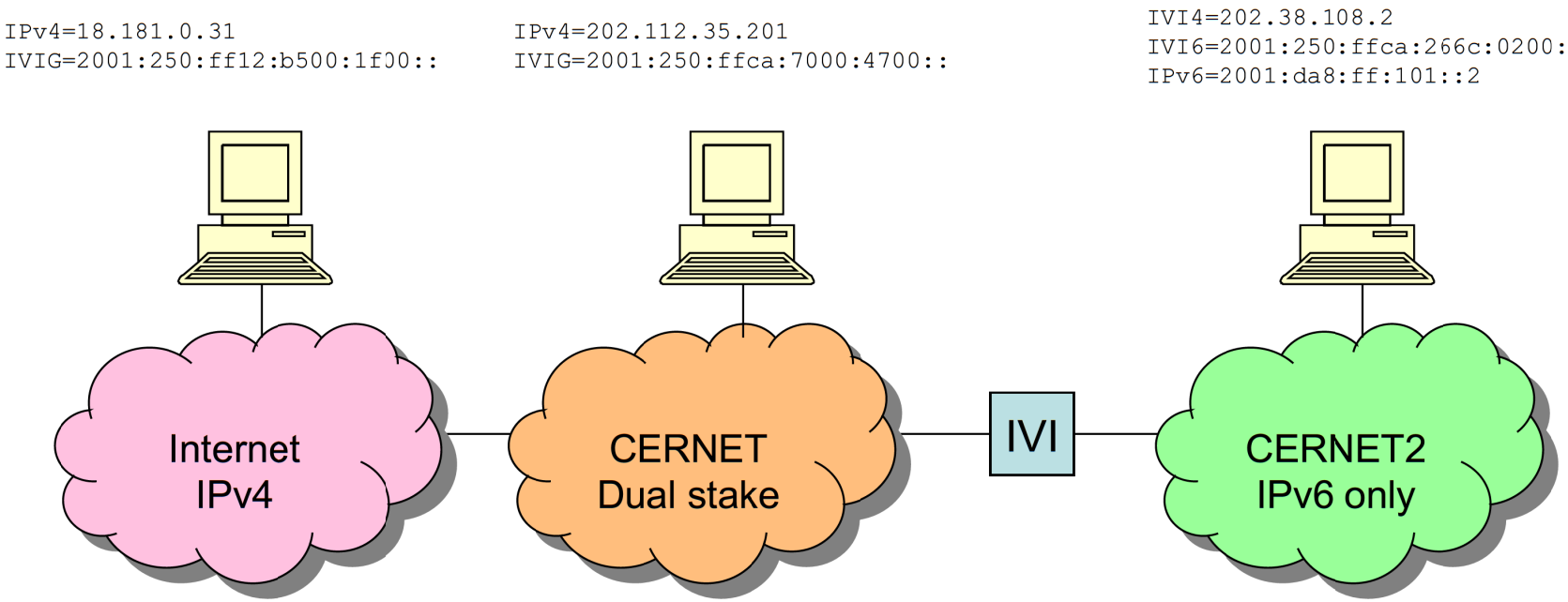


Connectivity test



(1)	server	2001:250:ffca:266c:200::	202.38.108.2
	client	2001:250:ffca:7023:c900::	202.112.35.201
(2)	server	2001:250:ffca:266c:200::	202.38.108.2
	client	2001:da8:aaae::201	not_ivi
(3)	server	2001:da8:ff:101::2	not_ivi
	client	2001:da8:aaae::201	not_ivi

traceroute test



From IVI6 host traceroute6 IVIG

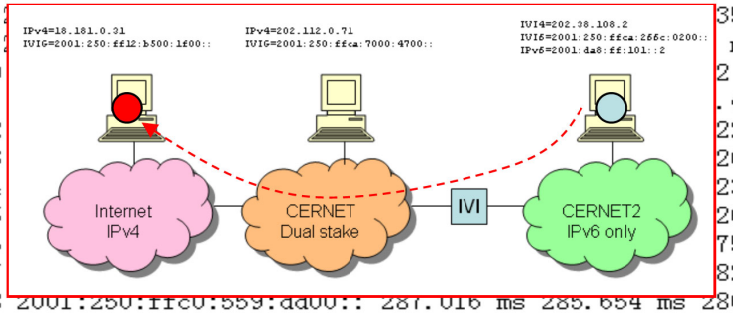
Mozilla Firefox

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

http://[2001:da8:ff:101::2]:8099/cgi-bin/nph-ivi6

Host N in in pure IPv6 environment

```
traceroute to 2001:250:ff12:b500:1f00:: (2001:250:ff12:b500:1f00::), 30 hops max, 40 byte packets to not_ivi
1 2001:250:ffca:266c:100:: 0.902 ms 0.884 ms 0.849 ms 2001:250:ffca:266c:100:: 202.38.108.1
2 2001:250:c000:63::1 1.210 ms 1.302 ms 1.378 ms 2001:250:c000:63::1 not_ivi
3 2001:250:c000:20::1 1.668 ms 1.766 ms 1.917 ms 2001:250:c000:20::1 not_ivi
4 2001:250:c000:2::2 2.915 ms 3.042 ms 3.095 ms 2001:250:c000:2::2 not_ivi
5 2001:250:ff0a:0:100:: 4.302 ms 4.283 ms 4.284 ms 2001:250:ff0a:0:100:: 10.0.0.1
6 2001:250:ffca:703d:4100:: 6.878 ms 7.676 ms 7.658 ms 2001:250:ffca:703d:4100:: 202.112.61.65
7 * 2001:250:ffca:703d:f100:: 5.879 ms * * 202.112.61.65
8                                                    356 ms 2001:250:ffca:7035:b100:: 202.112.53.177
9                                                     ms 2001:250:ffca:703d:9e00:: 202.112.61.158
10                                                    2 ms 2001:250:ffca:7035:1200:: 202.112.53.18
11                                                   .458 ms 2001:250:ffcb:b5c2:7d00:: 203.181.194.125
12                                                   224.348 ms 2001:250:ffc0:cb74:9100:: 192.203.116.145
13                                                   206.649 ms 2001:250:ffcf:e7f0:8300:: 207.231.240.131
14                                                   238.206 ms 2001:250:ff40:391c:2d00:: 64.57.28.45
15                                                   261.707 ms 2001:250:ff40:391c:2a00:: 64.57.28.42
16                                                   75.508 ms 2001:250:ff40:391c:700:: 64.57.28.7
17                                                   82.437 ms 2001:250:ff40:391c:a00:: 64.57.28.10
18 2001:250:ffc0:559:dd00:: 287.016 ms 285.654 ms 286.070 ms 2001:250:ffc0:559:dd00:: 192.5.89.221
19 2001:250:ffc0:559:ed00:: 286.132 ms 285.501 ms 289.742 ms 2001:250:ffc0:559:ed00:: 192.5.89.237
20 2001:250:ffcf:d28f:6e00:: 289.612 ms 287.616 ms 288.081 ms 2001:250:ffcf:d28f:6e00:: 207.210.143.110
21 2001:250:ff12:a800:1900:: 288.048 ms 285.286 ms 286.591 ms 2001:250:ff12:a800:1900:: 18.168.0.25
22 2001:250:ff12:b500:1f00:: 286.446 ms 286.453 ms 285.856 ms 2001:250:ff12:b500:1f00:: 18.181.0.31
```



完成

开始 Mozilla Firefox 7-ivi Microsoft PowerP... 84% 22:11

From IPv4 host trace IVI4

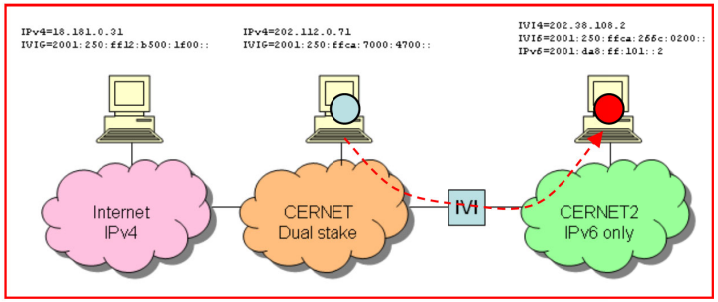
Mozilla Firefox

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

http://202.112.0.71:8033/cgi-bin/nph-ivi4

Host C in the IPv4 environment

```
1 * 202.112.0.65 1 ms *        0.0.0.0       0.0      AS0
2 * 202.112.53.73 1 ms 1 ms        0.0.0.0       0.0      AS0
3 202.112.53.178 0 ms 0 ms 0 ms        202.112.53.178       202.112      AS13746
4 202.112.61.242 9 ms 1 ms 1 ms        202.112.61.242       202.112      AS15858
5 202.38.17.186 1 ms 1 ms 1 ms        202.38.17.186       202.38      AS4538
6 202.38.17.186 1 ms 1 ms 1 ms        202.38.17.186       202.38      AS4538
7 202.38.17.186 2 ms 2 ms 2 ms        202.38.17.186       202.38      AS4538
8 202.38.17.186 2 ms 2 ms 2 ms        202.38.17.186       202.38      AS4538
9 202.38.17.186 2 ms 2 ms 2 ms        202.38.17.186       202.38      AS4538
10 202.38.108.2 2 ms 3 ms 4 ms        202.38.108.2       202.38      AS27650
```

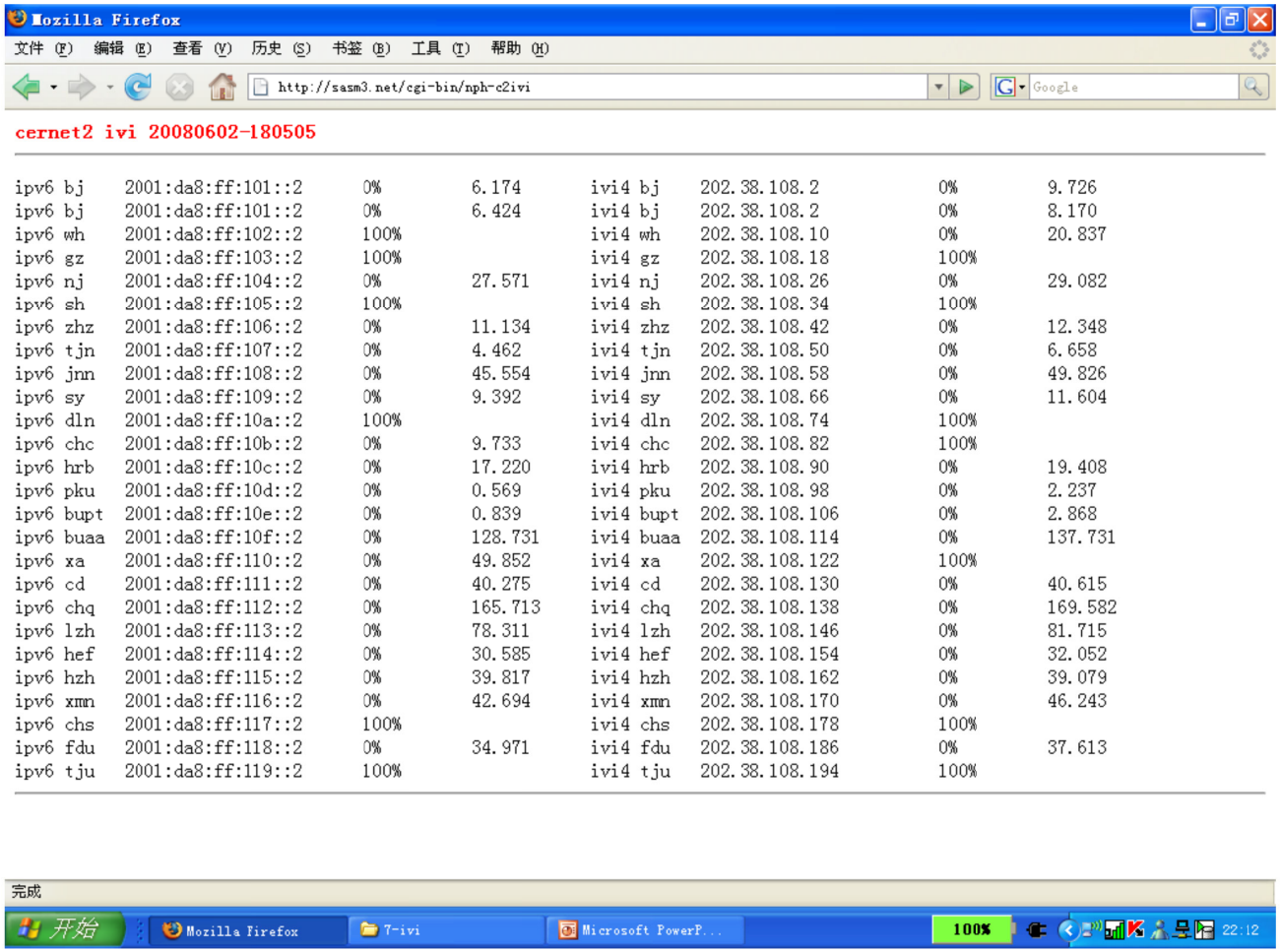


完成

开始 Mozilla Firefox 7-ivi Microsoft PowerP... PowerPoint 幻灯... 84% 22:14

Test Environment in CERNET and CNGI-CERNET2

- One IVI box is installed and 25 clients distributed in CNGI-CERNET2 Pops nation-wide.



Features of IVI (must)

1. No need to change the end system (IPv4 and IPv6).
2. Support v4-initiated and v6-initiated (?) short-lived local handle.
3. Support interaction with dual-stack hosts
4. The standard IPv4 NAT can easily be integrated into the system.
5. Do not violate standard DNS semantics.
6. No affect to IPv6 routing.
7. Support TCP, UDP, ICMP, TLS.
8. No strange behavior
9. Can handle fragmentation.
10. Do not introduce new network security vulnerabilities.

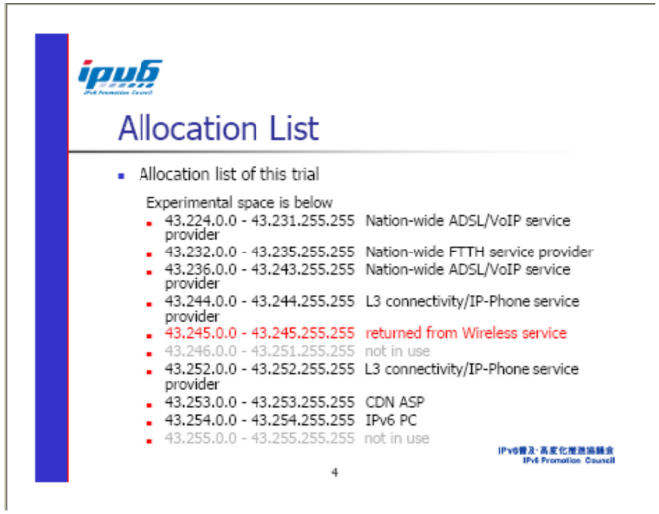
Features of IVI (should)

1. Support DNSSec (need to be tested)
2. support operational flexibility
3. Support central Management
4. Support fragmented packets bis (need to be tested)
5. Support richer application behavior
6. Support SCTP (need to be tested)
7. Support DCCP (need to be tested)
8. Support multicast (SSM)

Address policy discussion (1)

- We propose to allocate some portion of the remaining IPv4 address for the IVI use. Note that 43.0.0.0/8 could be a good candidate.

Large Space IPv4 Trial Usage Program
for Future IPv6 Deployment
ACTIVITIES UPDATE Vol.14



- A /8 with 16 gain results in 268M addresses (16777216*16=268435456).
- If we allocate 10 /8s for the IVI usage, it will be 2.6 billion for the unwired population (China, India, Africa, etc).

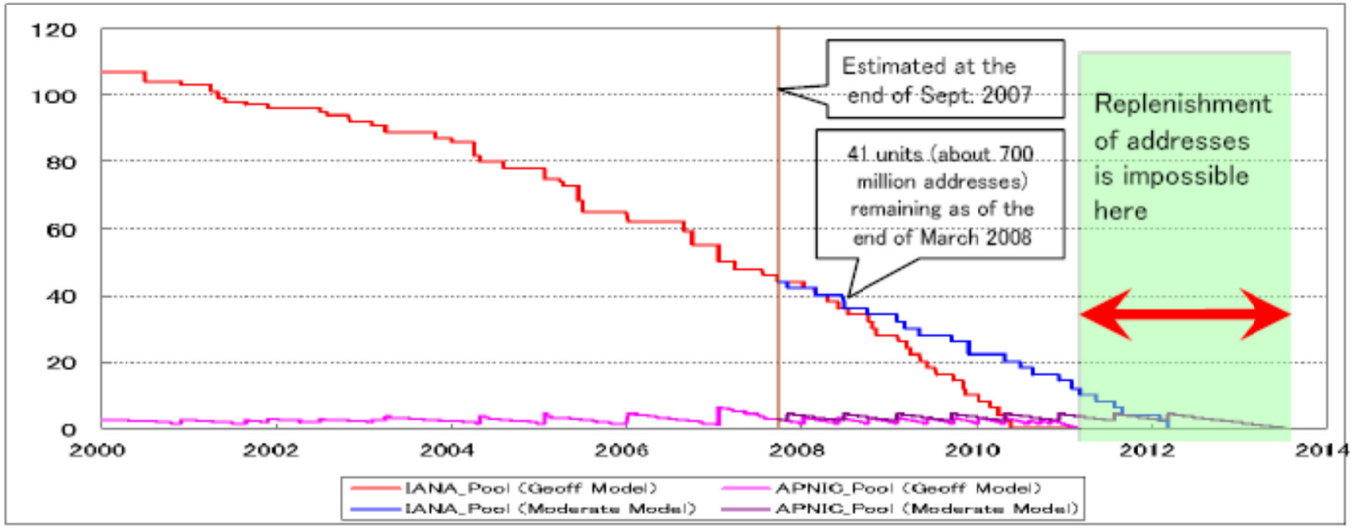
Address policy discussion (2)

- We propose a new IPv6 address allocation and assignment scheme
  - Reserve xxxx:yyyy:ff00::/40 for each xxxx:yyyy::/32
  - Borrow a subset of ISP's own IPv4 space and use it in IPv6, (i.e. IVI4 and IVI6, respectively) for the initial IPv6 deployment.
    - For severs use 1-to-1 mapping
    - For clients use 1-to-N mapping
  - More and more ISP will deploy IVI6s and these blocks can access both the global IPv4 and IPv6. When a threshold is reached (i.e. $\cup IVI4_i \approx IVIG$) , this implies that the transition from IPv4 to IPv6 is accomplished.
  - Then, the assignment of the non IVI6 can be started.

2.2 Estimated date of the IPv4 address exhaustion

Unless there will be no significant changes in the circumstances (there will be no change in the rules concerning international address allocation or address maintenance, and that those to whom addresses are allocated will not try to retain surplus addresses arbitrarily), the following is estimated:

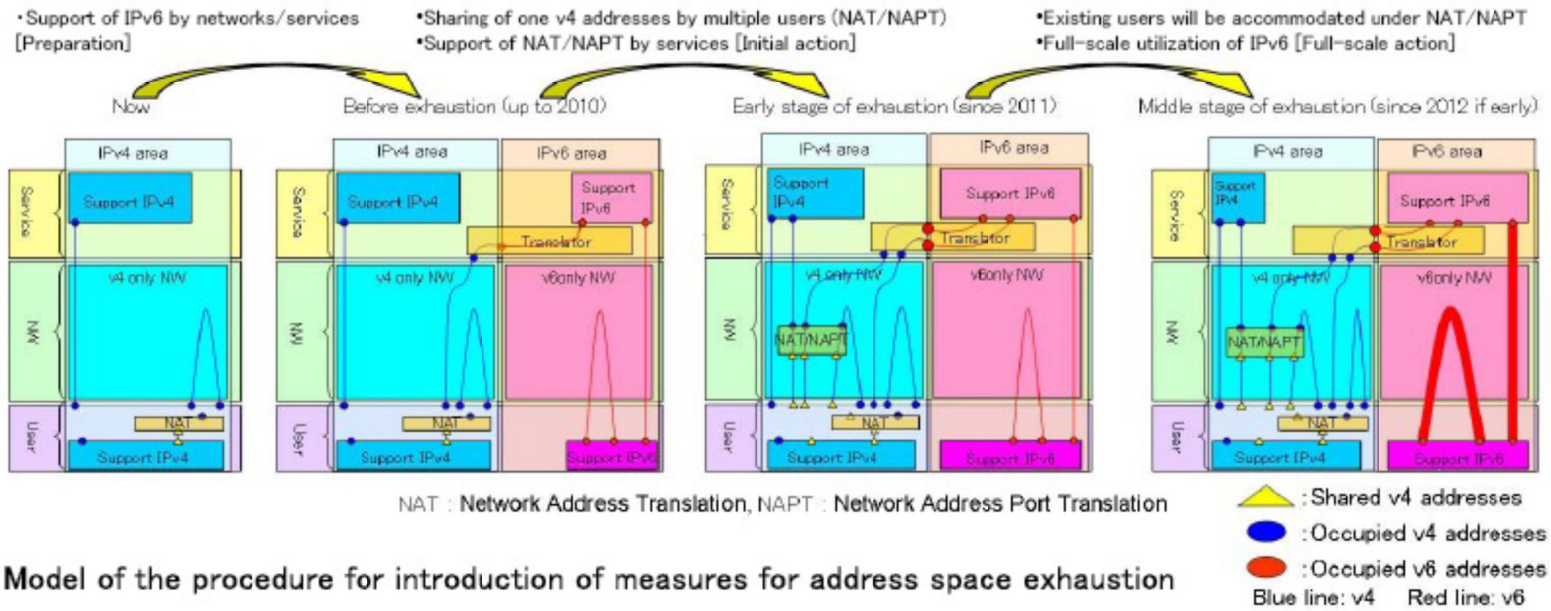
- IANA pool will be exhausted in the middle of 2010 to the beginning of 2012
- No more allocation of the addresses used in Japan at the beginning of 2011 to the middle of 2013



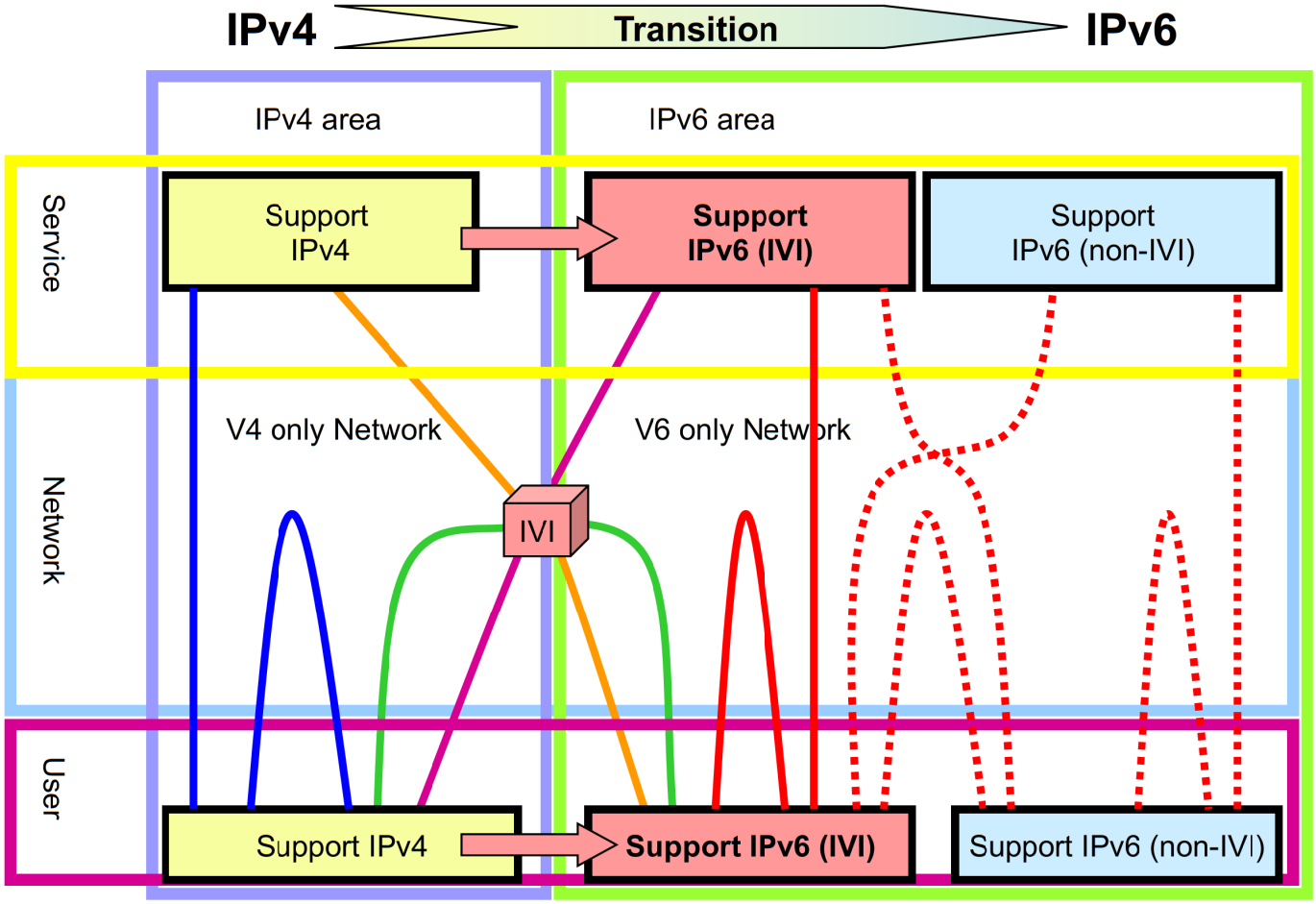
(The vertical axis is shown in units called "/8". "1" is equivalent to about 16.8 million addresses)

Measures for address space exhaustion

1. For continuous development of the Internet since 2011, the combination of **the transition to a new address system (IPv6)** and **sharing of one address by multiple users (using NAT/NAPT)** must be performed from three viewpoints of *feasibility within a time limit, continuity of service on the Internet, and continuance of effect*,
2. It is appropriate to **introduce the measures in three stages**: before exhaustion, early, and middle stages of exhaustion.



Model of the procedure for introduction of measures for address space exhaustion



- The existing IPv4 users may not have motivation to transit to IPv6.

- Provide IVI6(i) for new Internet users, so they can have IPv4 connectivity and new IPv6 services. Then the existing IPv4 users may want to use IVI6(i). Therefore, more and more IPv4 addresses are borrowed by IPv6 networks as IVI6(i)

- When the number of services and users which support IPv6 (via IVI) reaches a critical mass, non-IVI IPv6 addresses can be used

3. Actions for address space exhaustion

1. Because the initial action must be completed by the beginning of 2011, the three actions that are possible with the existing technology were examined
   - ① Share one address among multiple nodes (saving of IP addresses) ⇒ Use NAT/NAPT (Sharing of IPv4 addresses)
   - ② Use all addresses (maximum density of IP address usage) ⇒ Reallocate the distributed IPv4 addresses
   - ③ Utilize new address resources ⇒ Transition to IPv6
2. These are compared from the following viewpoints
   - ① Feasibility within a time limit: Problems that should be solved can be solved in about 3 years.
   - ② Continuity of service: The uses that can be provided in the current Internet will continue to be used.
   - ③ Continuance of effect: Reattempt actions are unnecessary because general users are influenced.

	Utilization of NAT/NAPT (Sharing of IPv4 addresses)	Reallocation of the assigned IPv4 addresses	Transition to IPv6
Feasibility within a time limit	✓	Doubtful	Extremely difficult
Service continuity	Limited	✓	✓
Permanent effect	Doubtful	NG	✓

- Essential action: Transition to IPv6
- Initial action: Utilization of NAT/NAPT

should be carried out in combination.

NAT : Network Address Translation, NAPT : Network Address Port Translation

	Utilization of NAT/NAPT (Sharing of IPv4 addresses)	Reallocation of the assigned IPv4 addresses	Transition to IPv6	**IVI**
Feasibility within a time limit	✓	Doubtful	Extremely difficult	√
Service continuity	Limited	✓	✓	√
Permanent effect	Doubtful	NG	✓	√

Actually, IVI combines the methods of NAT/NAPT, reallocation of the assigned IPv4 addresses and transition to IPv6 in a natural and harmonious way.