Prefix-specific and Stateless Address Mapping (IVI) for IPv4/IPv6 Coexistence and Transition

draft-xli-behave-ivi-00

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

IETF-72, Dublin, behave, 29 July 2008

Outline

- Introduction
- IVI scheme
- Design considerations
- Testing result
- Transition
- Address Policy
- Conclusions

Introduction

- The experiences for the IPv6 deployment in the past 10 years strongly indicate that the IPv6 hosts need to communicate with the global IPv4 networks.
- In this document, we follow the basic specification of SIIT, but we define the address assignment and routing scheme (IVI).
  - It is stateless (or almost stateless) in both the IPv4-to-IPv6 mapping direction, as well as in the IPv6-to-IPv4 mapping direction
  - It supports address transparency.
  - It supports both IPv6 initiated communication and the IPv4 initiated communication without using NAT-traversal techniques.

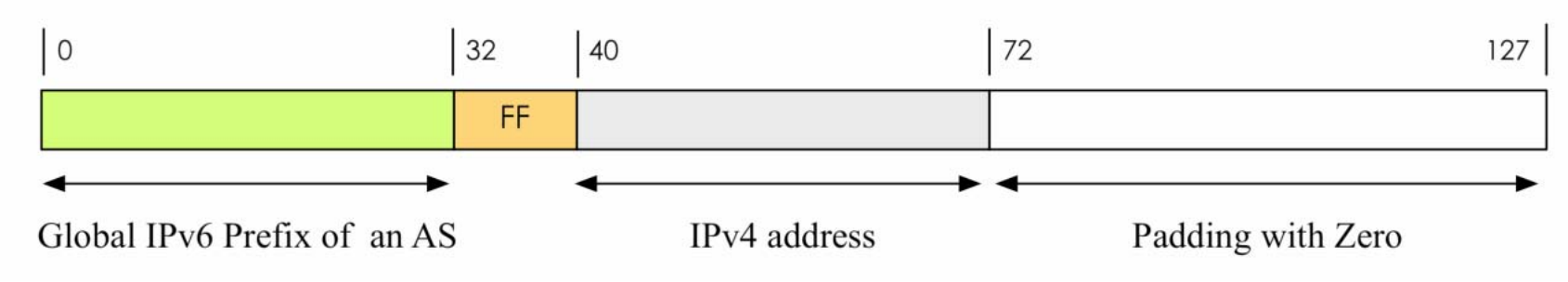
IVI Scheme

- The IVI is a prefix-specific and explicit bidirectional address mapping scheme.
  - Embed global IPv4 addresses into a subset of each ISP's IPv6 address block
  - Based on this mapping rule, each ISP can borrow a portion of its IPv4 addresses and use it in IPv6.
- The SIIT stateless translation is implemented in the IVI gateway.
- The IPv4 multiplexing techniques can be used.
- Ref:
  - http://www.ietf.org/internet-drafts/draft-xli-behave-ivi-00.txt

Terms and Abbreviations of IVI

- **General**
  - **IVI.**
  - **ISP(i)**
- **IPv4**
  - **IPG4:** An address set containing all IPv4 addresses, the addresses in this set are mainly used by IPv4 hosts at the current stage.
  - **IPS4(i):** A subset of IPG4 allocated to ISP(i).
  - **IVI4(i):** A subset of IPS4(i), the addresses in this set will be mapped to IPv6 via IVI rule and physically used by IPv6 hosts of ISP(i).
- **IPv6**
  - **IPG6:** An address set containing all IPv6 addresses.
  - **IPS6(i):** A subset of IPG6 allocated to ISP(i).
  - **IVIG46(i):** A subset of IPS6(i), an image of IPG4 in IPv6 address family via IVI mapping rule.
  - **IVI6(i):** A subset of IVIG46(i), an image of IVI4(i) in IPv6 address family via IVI mapping rule.
- **Components**
  - **IVI gateway**
  - **IVI DNS**

Address Mapping (1)

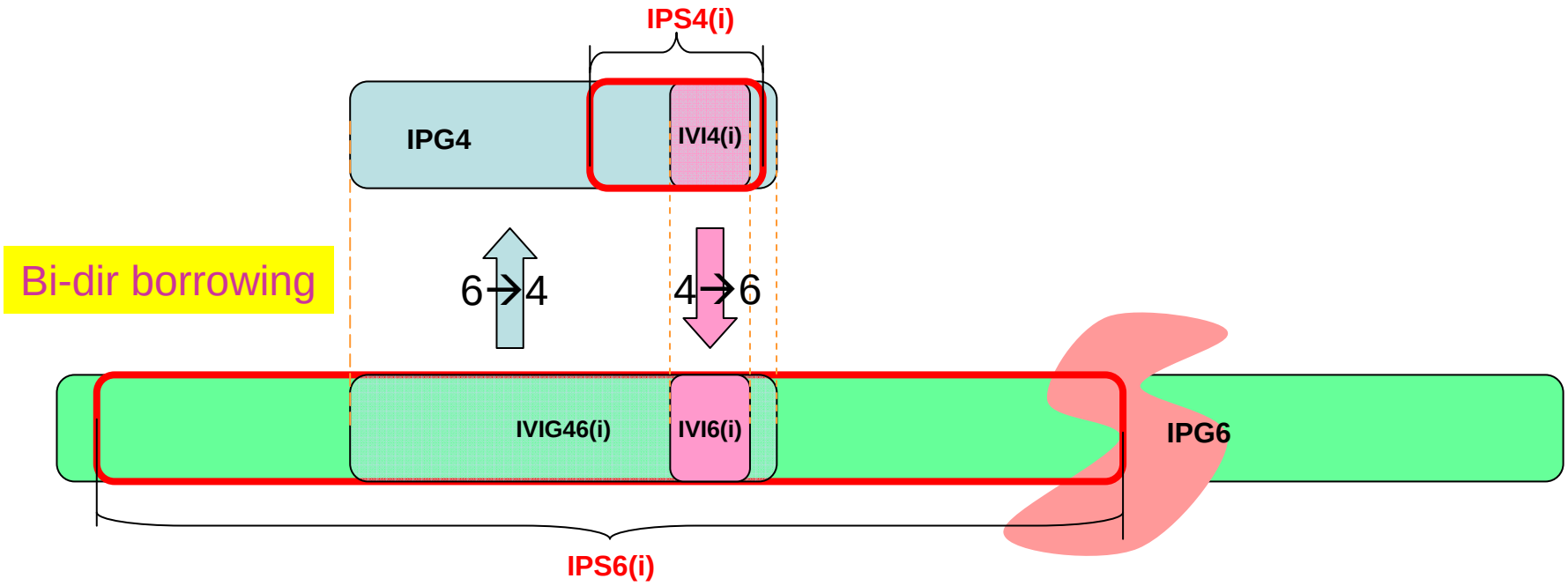


Mapping Rule: **IPv4 addresses are embedded from bit 40 to bit 72 of the IPv6 addresses of a specific /32.**

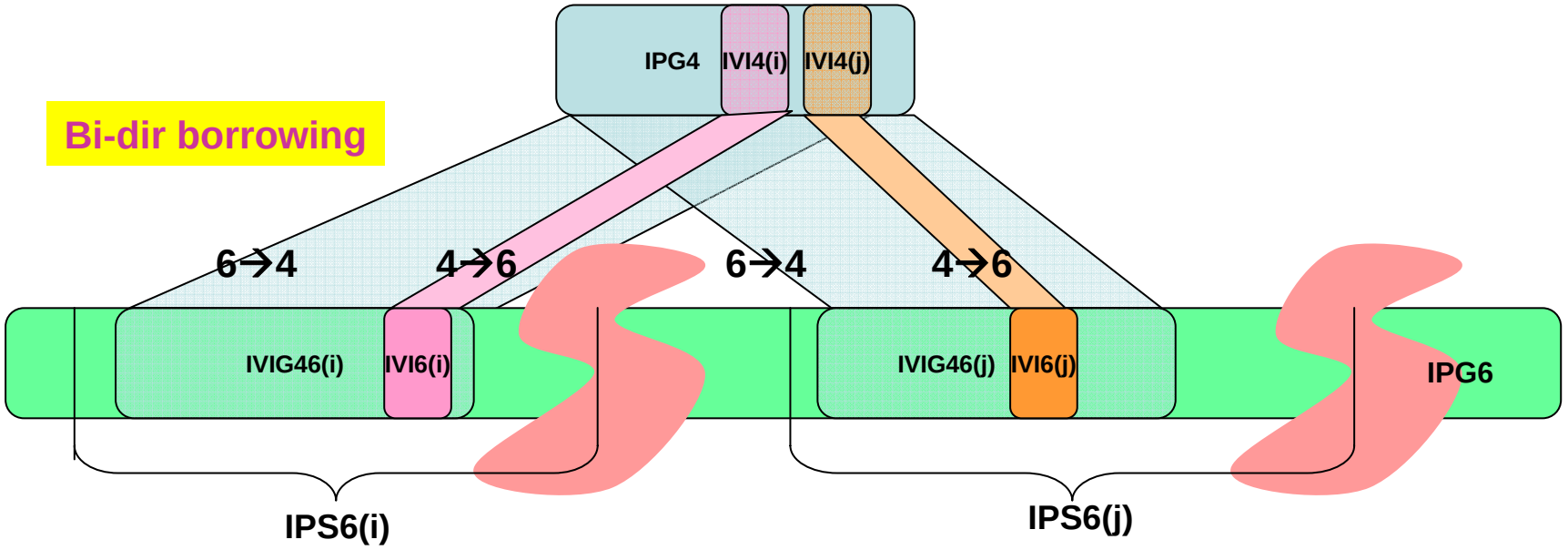
Example:

ISP's IPv6 /32 (ISP6)	**2001:250::/32**
image of global IPv4 (IVIG46):	**2001:250:ff00::/40**
borrowed IPv4 address (IVI4):	**202.38.108.0/24**
mapped IVI IPv6 address (IVI6):	**2001:250:ffca:266c::/64**

Address Mapping (2)

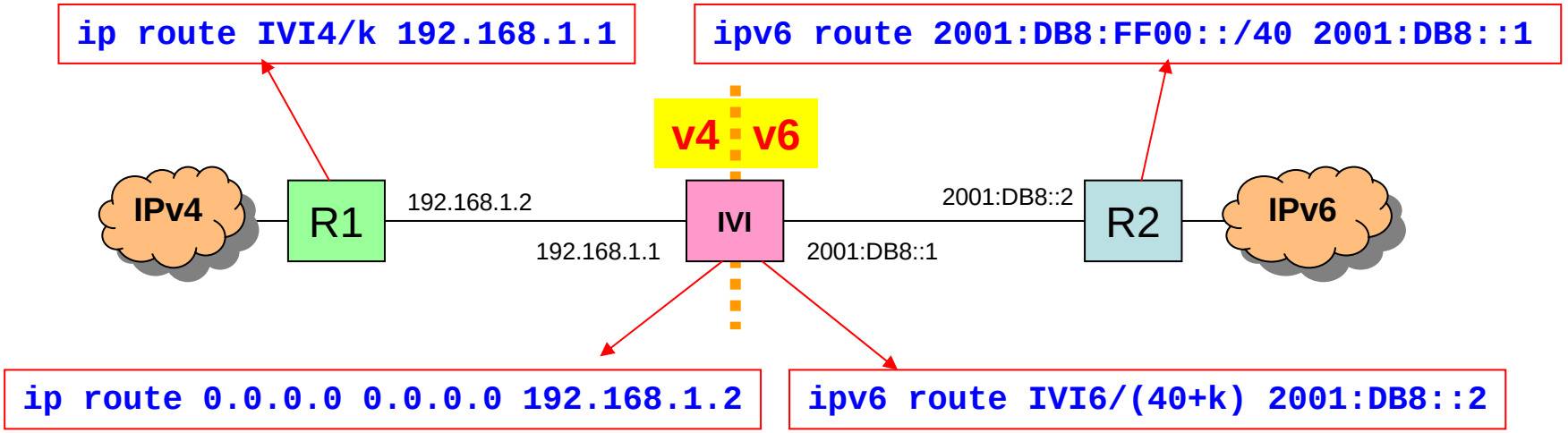


Address Mapping (3)



Routing and Forwarding

Routing and mapping configuration example

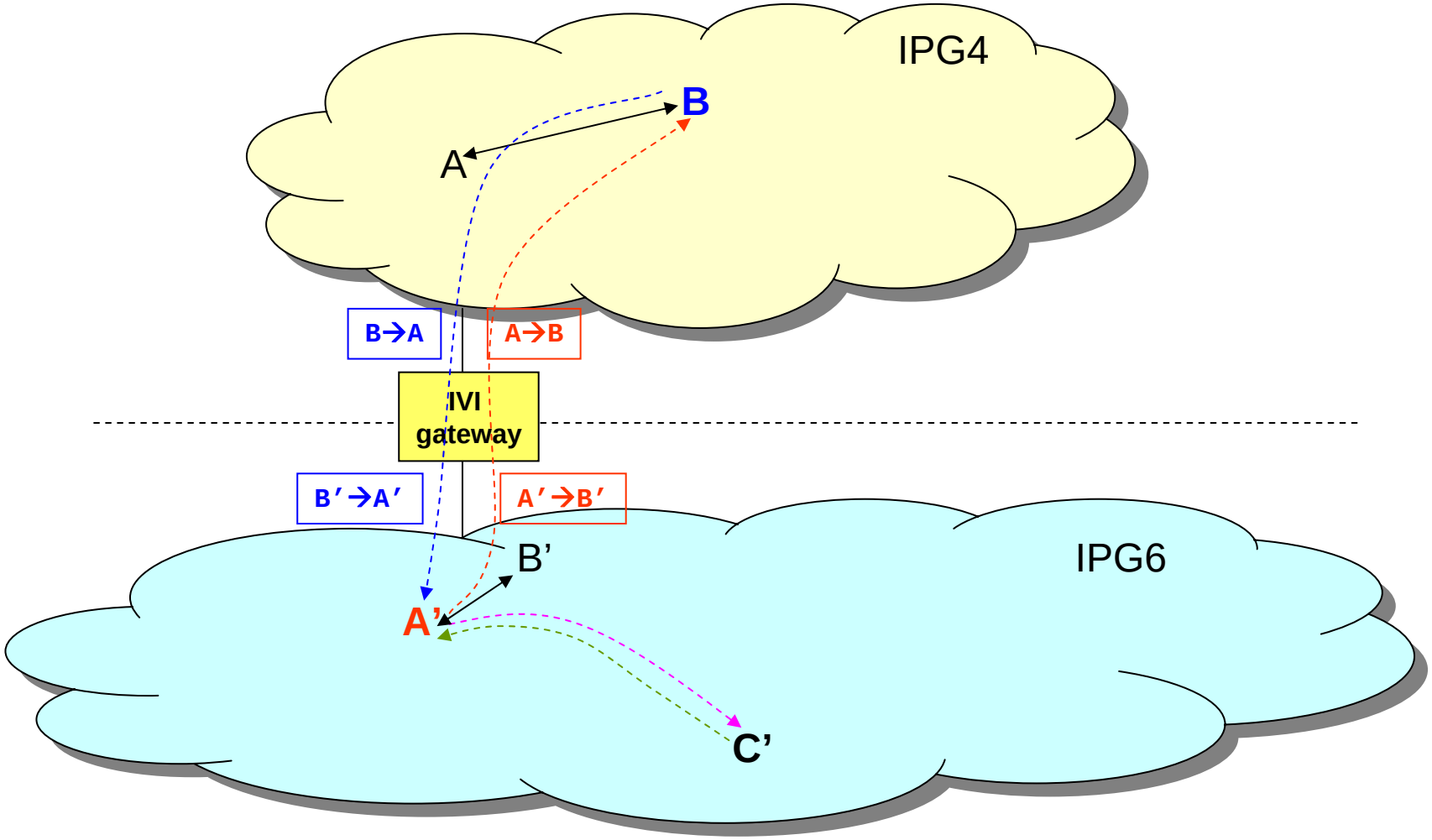



```
mroute IVI4-network IVI4-mask pseudo-address interface source-PF destination-PF
mroute6 destination-PF destination-PF-pref-len
```

IVI Reachability Matrix

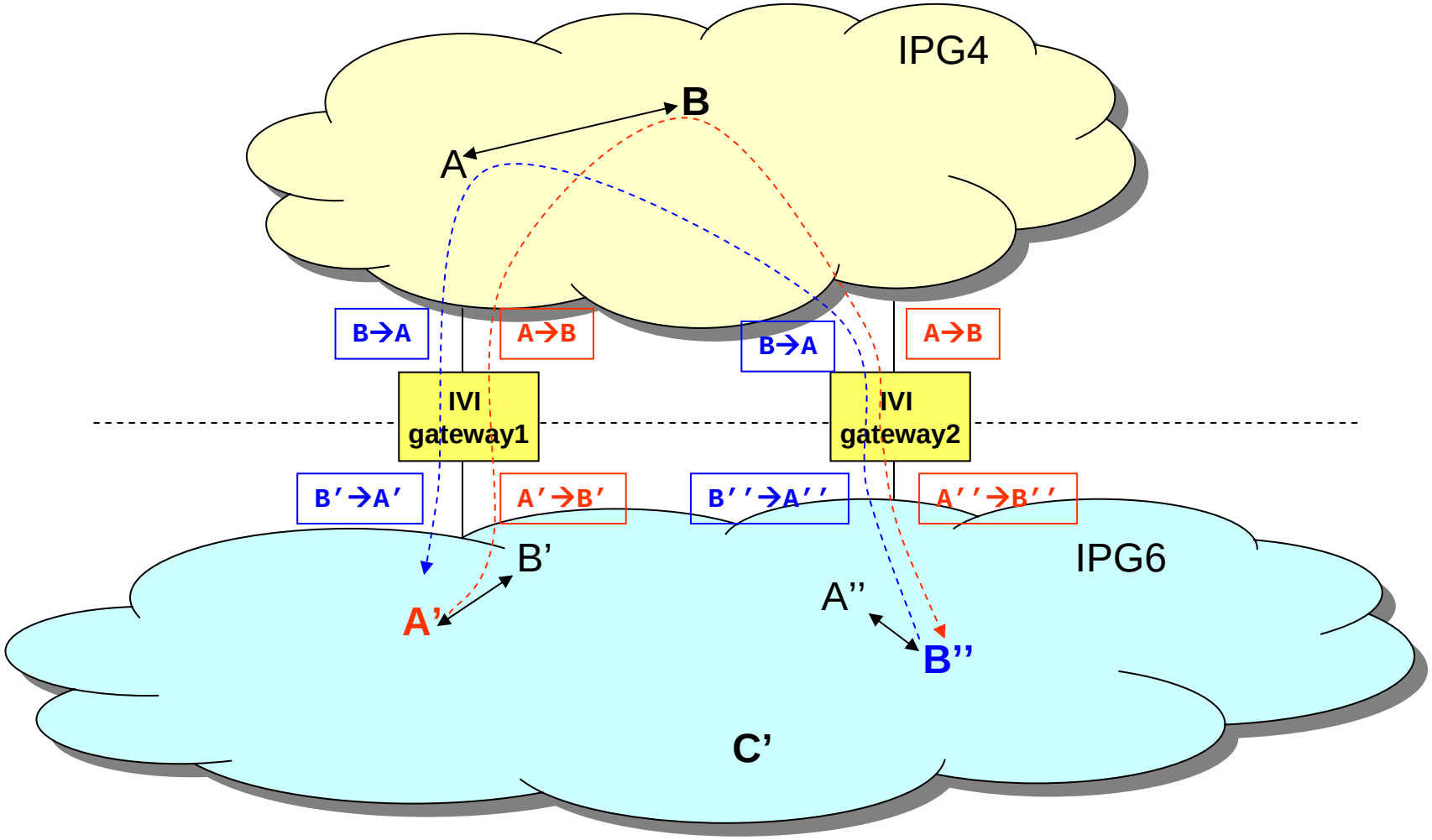
	IPG4	IVI	IPG6
IPG4	OK	OK	NO
IVI	OK	OK	OK
IPG6	NO	OK	OK

IVI Communication Scenarios (1)



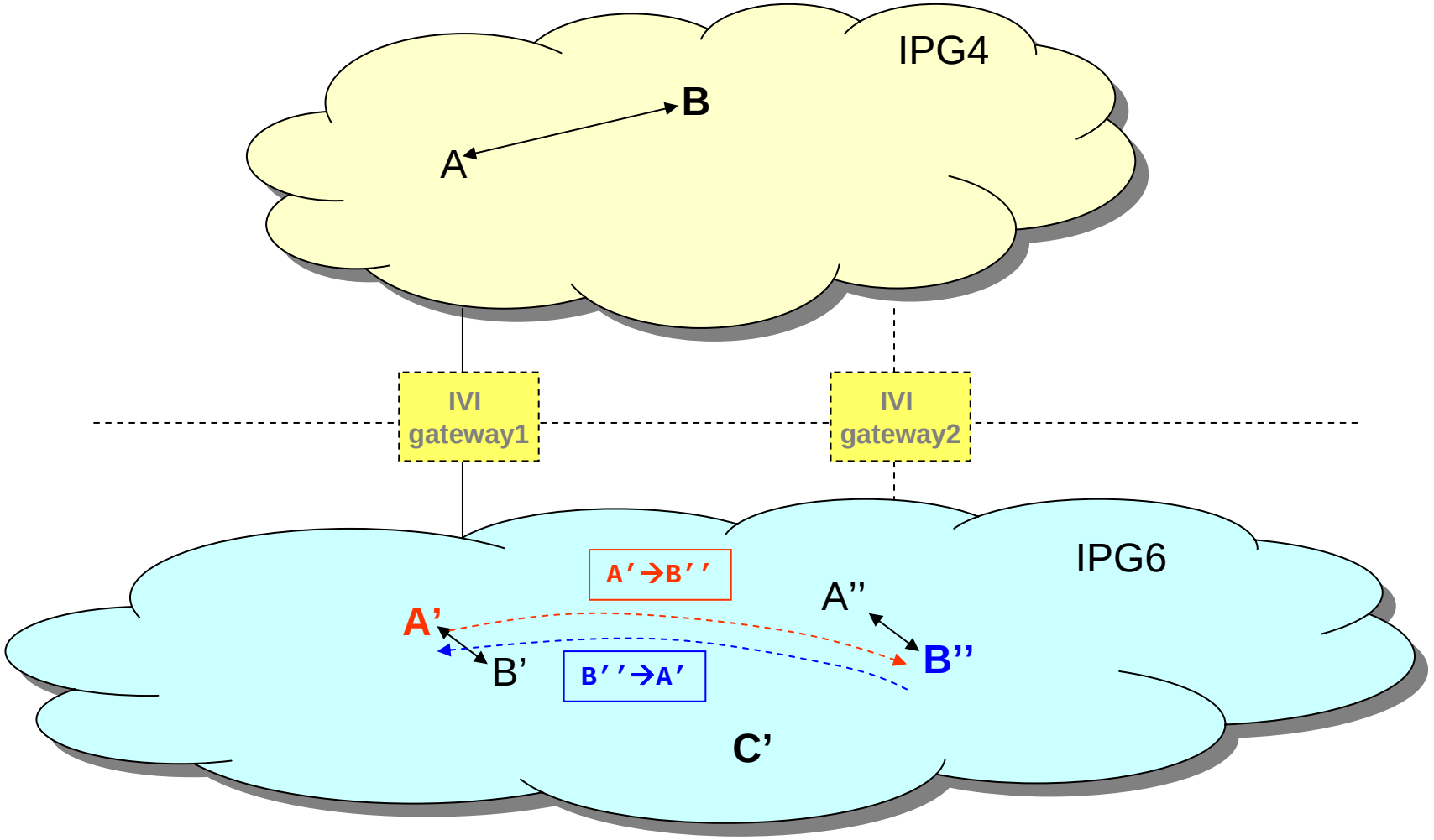
- A' ←→ B
- A' ←→ C'

IVI Communication Scenarios (2)



- A' ←(B&A)→ B''

IVI Communication Scenarios (3)



- A' ←→ B''

Design Considerations

- Address Mapping (general)
- Network-layer Header Translation (SIIT)
- Transport-layer Header Translation (SIIT)
- Fragmentation and MTU Handling (SIIT)
- ICMP Handling (SIIT + extension)
- Application Layer Gateway (SIIT)
- IPv6 Source Address Selection
- IPv4 over IPv6 Support
- IVI DNS
- Multiplexing of the Global IPv4 Addresses
- Multicast support

Address Mapping (general)

- IVI general address mapping
  - **2001:DB8:FF00::/40**
  - **2001:DB8:FFFF::/48,**
  - **2001:DB8:ABCD:FF00::/56**
  - **2001:DB8:ABCD:FFFF::/64**
  - **……**
  - **2001:DB8:XXXX:XXXX:XXXX:XXXX::/96**

ICMP + Extension

- The ICMP message may be generated by an intermediate router whose IPv6 address does not belong to IVIG46(i). Since ICMP translation is important to the path MTU discovery, the inverse mapping for unmapped addresses is defined in this document.
- In the current prototype, a pseudo IPv4 address is generated
  - First 16 bits are the IPv4 address of the IVI gateway
  - The last 16 bits are the AS number of the current domain. This prevents translated ICMP messages from being discarded due to unknown or private IP source.
- A small IPv4 address block should be reserved to identify the non-IVI mapped IPv6 addresses.
  - Similar to 4-byte AS AS23456

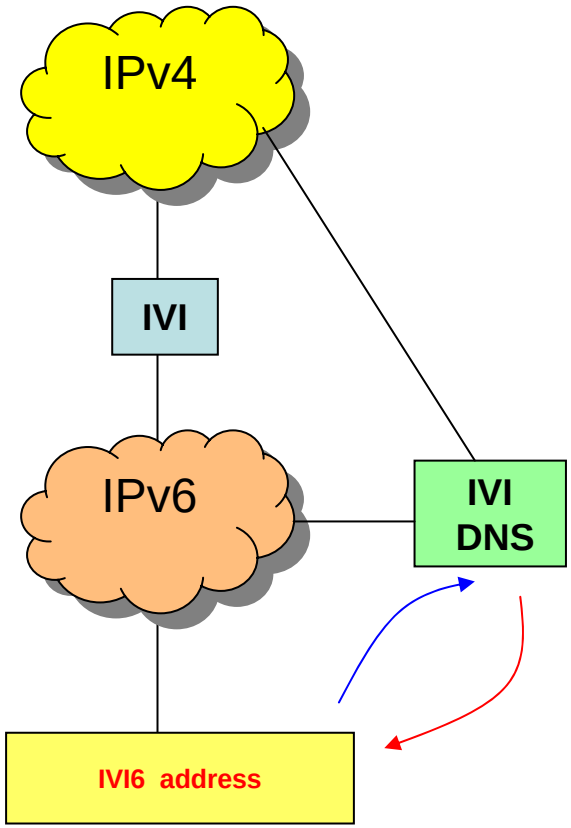
IPv6 Source Address Selection

- Since each IPv6 host may have multiple addresses, it is important for the host to use an IVI6(i) address to reach the global IPv4 networks.
  - The short-term work around is to use IVI6(i) as the default IPv6 address of the host.
  - The long-term solution requires that the application be able to select the source addresses for different services.
- IVI6 address configuration
  - DHCPv6 is required

IPv4 over IPv6 Support

- The IVI scheme can support the IPv4 over IPv6 service (NAT6-4-6), i.e. a stub IPv4 network can be connected to an IVI gateway to reach the IPv6 network and via another IVI gateway to reach the global IPv4 network
- A more interesting scenario is to integrate the functions of the first IVI gateway into the end-system. In this case, the application software are IPv4-based and there is no need to have ALG support in the IVI gateway when it is communicating with IPv4 hosts.

DNS Configuration and Mapping



- For providing primary DNS service for IVI4(i) and IVI6(i), 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 IVIG46(i) for IVI6(i), 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::
- Implementation scope
  - Host
  - DNS server provided via DHCPv6
  - ISP

Multiplexing of the Global IPv4 Addresses

- Temporal Multiplexing
  - Dynamic assignment of IVI6(i)
- Port Multiplexing
  - Combine address with the port number
- Spatial Multiplexing
  - Server 1:1 mapping
  - Home server 1:M mapping (via IPv4 initiated communication)
  - Client 1:N mapping (via IPv6 initiated communication)
- Multiplexing using IPv4 NAT-PT
  - Cascade IPv4 NAT-PT and IVI (1:1 mapping)

Port multiplexing – IPv6 initiated

- Example:
  - **202.38.108.5#100 ←→ 2001:250:ffca:266c:0500::81#100**
  - **202.38.108.5#101 ←→ 2001:250:ffca:266c:0500::82#100**
  - **202.38.108.5#102 ←→ 2001:250:ffca:266c:0500::83#100**
  - **202.38.108.5#103 ←→ 2001:250:ffca:266c:0500::84#100**
- In the case of port collision, map to an unused port.

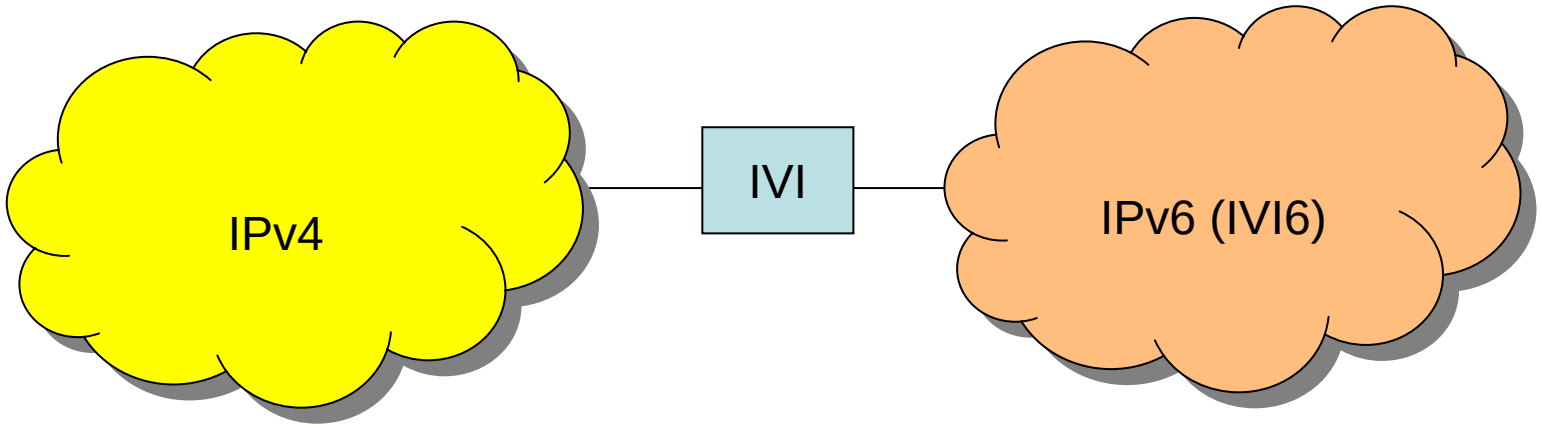
Port multiplexing – IPv4 initiated

- The remote IPv4 host can reach different IVI6s via different port number (pseudo-well-known port number)
  - **202.38.108.2#81 --> IVI61=2001:250:ffca:266c:0200::81#81**
  - **202.38.108.2#82 --> IVI61=2001:250:ffca:266c:0200::82#82**
  - **202.38.108.2#83 --> IVI61=2001:250:ffca:266c:0200::83#83**
  - **202.38.108.2#84 --> IVI61=2001:250:ffca:266c:0200::84#84**
- This can be provided via SRV DNS record.

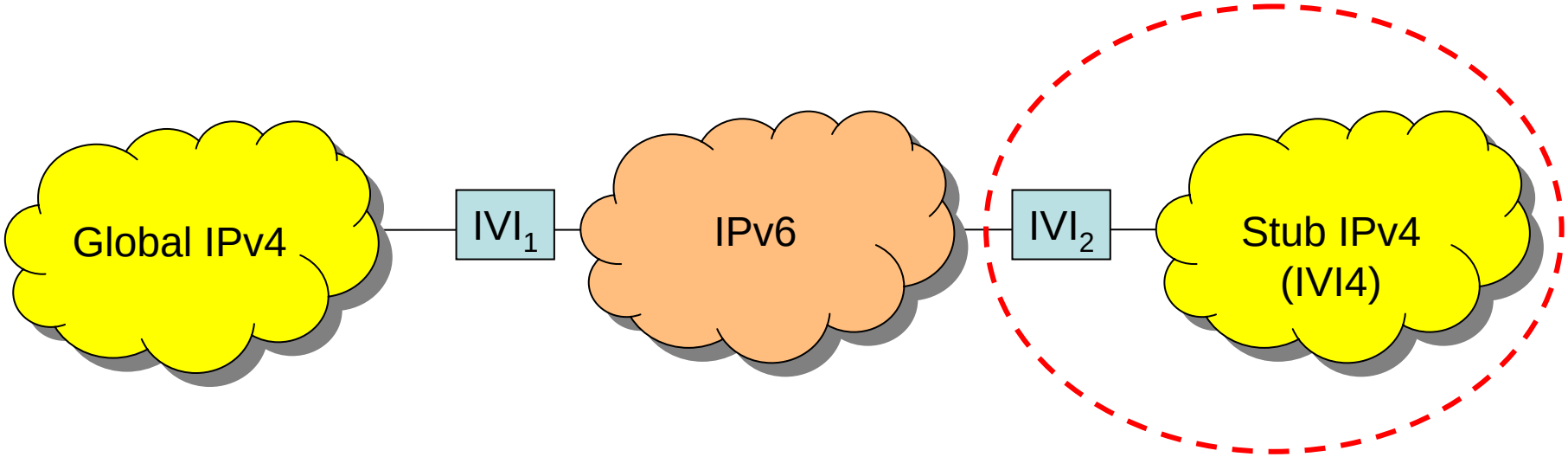
Multicast support

- SSM is supported for the IVI
  - no MSDP in IPv6
  - no embedded RP in IPv4
  - It is also possible to build a gateway for ASM
- 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

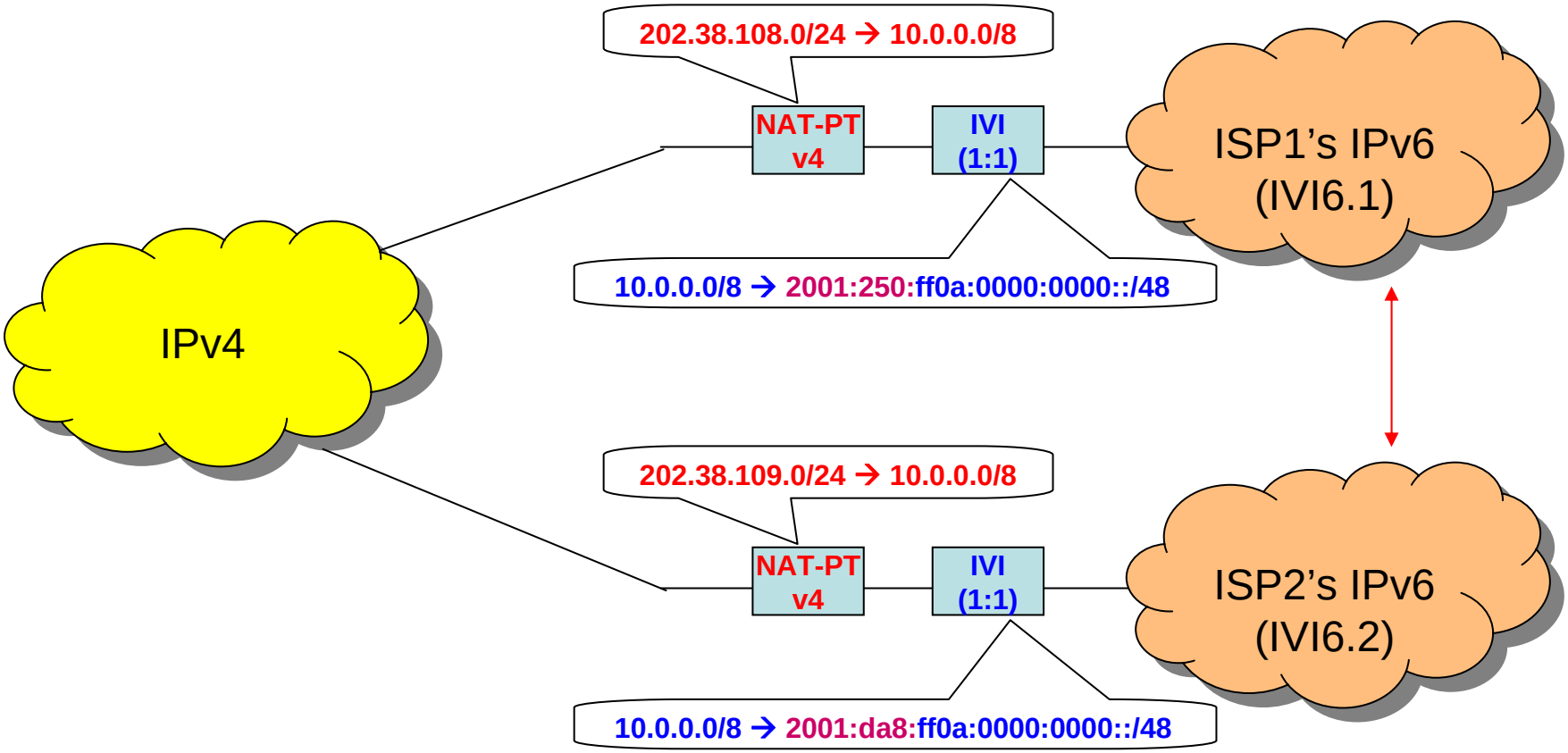
IVI Deployment Scenarios (1)



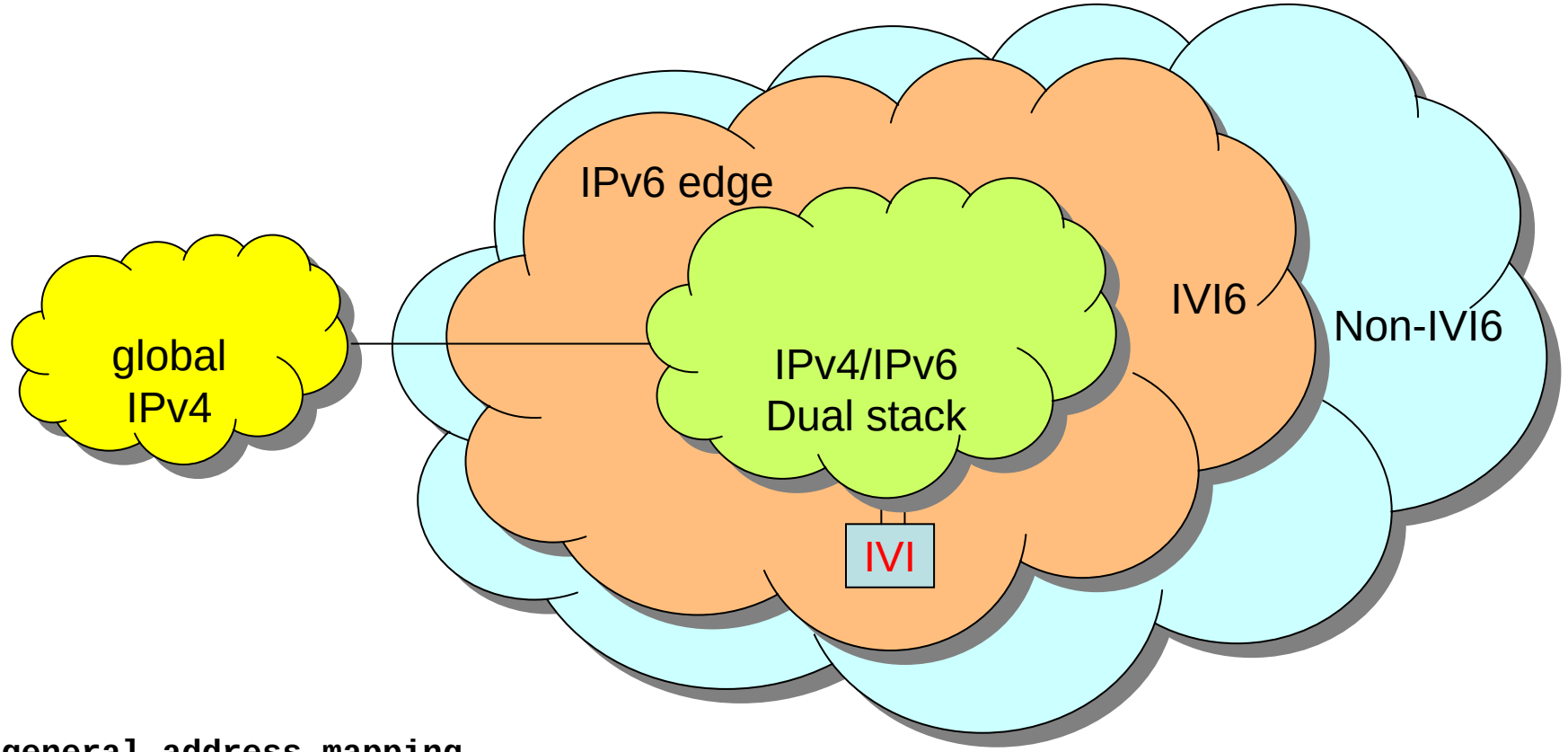
IVI Deployment Scenarios (2)



IVI Deployment Scenarios (3)



IVI Deployment Scenarios (4)



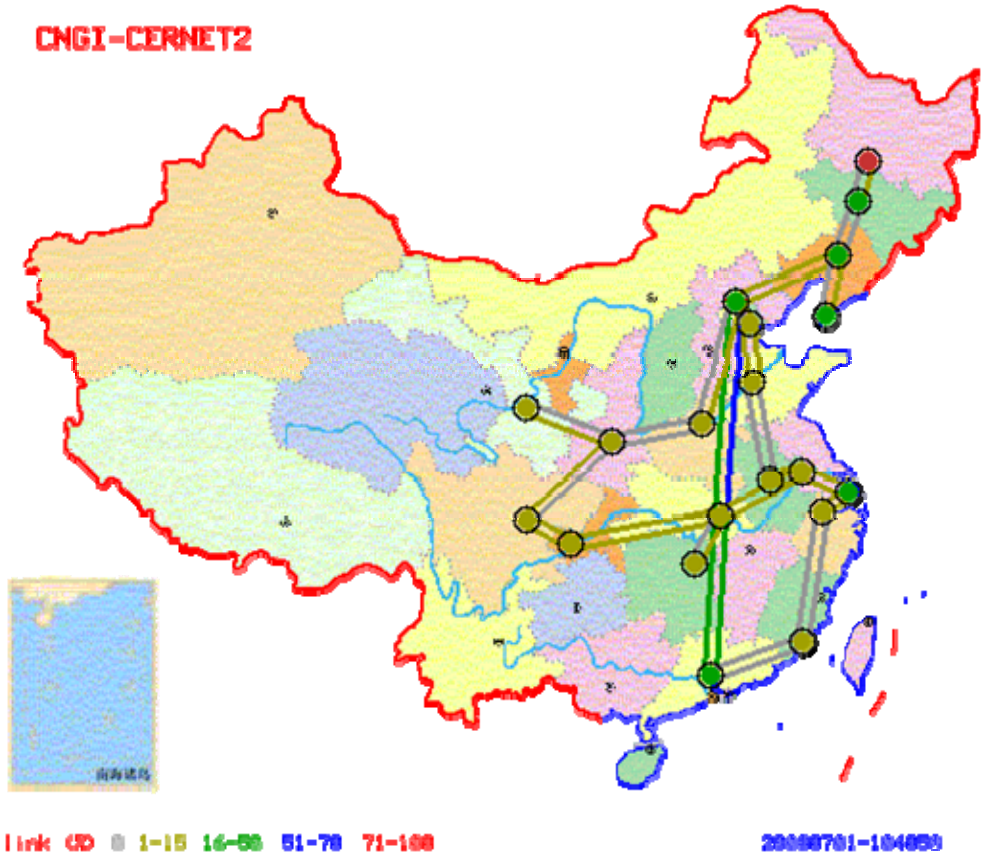
IVI general address mapping
2001:DB8:FF00::/40
2001:DB8:FFFF::/48,
2001:DB8:ABCD:FF00::/56
2001:DB8:ABCD:FFFF::/64
2001:DB8:XXXX:XXXX:XXXX:XXXX::/96

backbone scope (implemented)
site scope
sub-site scope
subnet scope
IPv4 mapped alike scope

Implementation and Preliminary Testing Results

- The IVI scheme presented in this document is implemented in the Linux OS
  - The source code can be downloaded [http://202.38.114.1/impl/].
- CERNET (IPv4 and partially dual-stack) and CNGI-CERNET2 (pure IPv6) since March 2006 (basic implementation).
  - IVI6 server for global IPv4
    - http://202.38.114.1/
  - IVI6 server for global IPv6
    - http://[2001:250:ffca:2672:0100::0]/
  - IVI server for stub IPv4
    - http://202.38.114.129/

IVI Hosts Installation in CNGI-CERNET2



From IVI6 host traceroute6 IVIG46

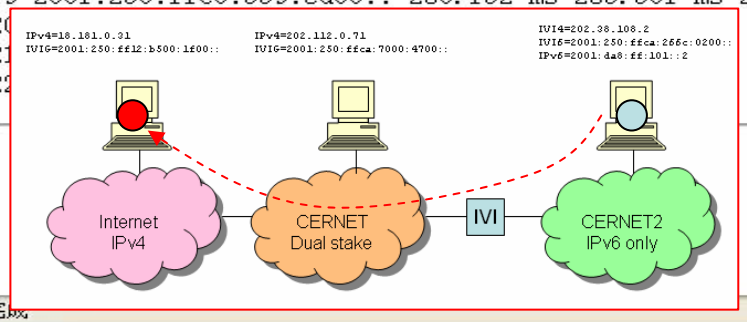
Mozilla Firefox

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 2001:250:ffca:7035:b100:: 11.638 ms 11.434 ms 11.356 ms 2001:250:ffca:7035:b100:: 202.112.53.177
9 2001:250:ffca:703d:9e00:: 5.074 ms 5.532 ms 5.399 ms 2001:250:ffca:703d:9e00:: 202.112.61.158
10 2001:250:ffca:7035:1200:: 5.325 ms 4.358 ms 5.162 ms 2001:250:ffca:7035:1200:: 202.112.53.18
11 2001:250:ffcb:b5c2:7d00:: 92.976 ms 91.484 ms 91.458 ms 2001:250:ffcb:b5c2:7d00:: 203.181.194.125
12 2001:250:ffc0:cb74:9100:: 209.784 ms 208.310 ms 224.348 ms 2001:250:ffc0:cb74:9100:: 192.203.116.145
13 2001:250:ffcf:e7f0:8300:: 206.548 ms 206.539 ms 206.649 ms 2001:250:ffcf:e7f0:8300:: 207.231.240.131
14 2001:250:ff40:391c:2d00:: 240.147 ms 239.321 ms 238.206 ms 2001:250:ff40:391c:2d00:: 64.57.28.45
15 2001:250:ff40:391c:2a00:: 263.962 ms 263.894 ms 261.707 ms 2001:250:ff40:391c:2a00:: 64.57.28.42
16 2001:250:ff40:391c:700:: 276.193 ms 276.179 ms 275.508 ms 2001:250:ff40:391c:700:: 64.57.28.7
17 2001:250:ff40:391c:a00:: 280.819 ms 280.744 ms 282.437 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                                                  288.081 ms 2001:250:ffcf:d28f:6e00:: 207.210.143.110
21                                                  286.591 ms 2001:250:ff12:a800:1900:: 18.168.0.25
22                                                  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 traceroute IVI4

Mozilla Firefox

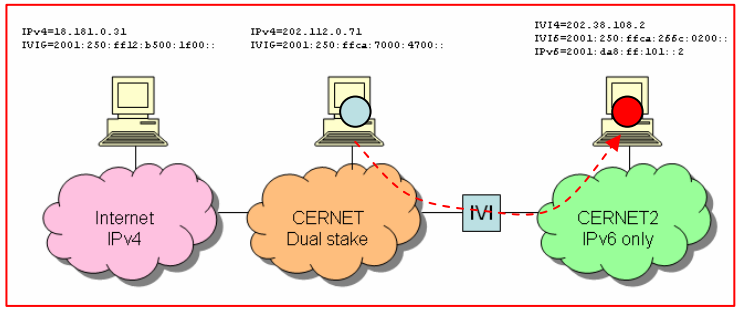
文件 (F) 编辑 (E) 查看 (V) 历史 (S) 书签 (B) 工具 (T) 帮助 (H)

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

Google

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

Features of IVI

1. No need to change the end system (IPv4 and IPv6).
2. Support v4-initiated and v6-initiated communications.
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 both IPv4 and IPv6 routing.
7. Support TCP, UDP, ICMP
8. Can handle fragmentation.
9. Support incremental deployment
10. Support multicast (SSM)

Address Policy and IVI Address Evolution

- IPv6 Address Assignment Policy
- IPv4 Address Allocation Policy
- Evolution of the IVI Addresses and Services

IPv6 Address Assignment Policy

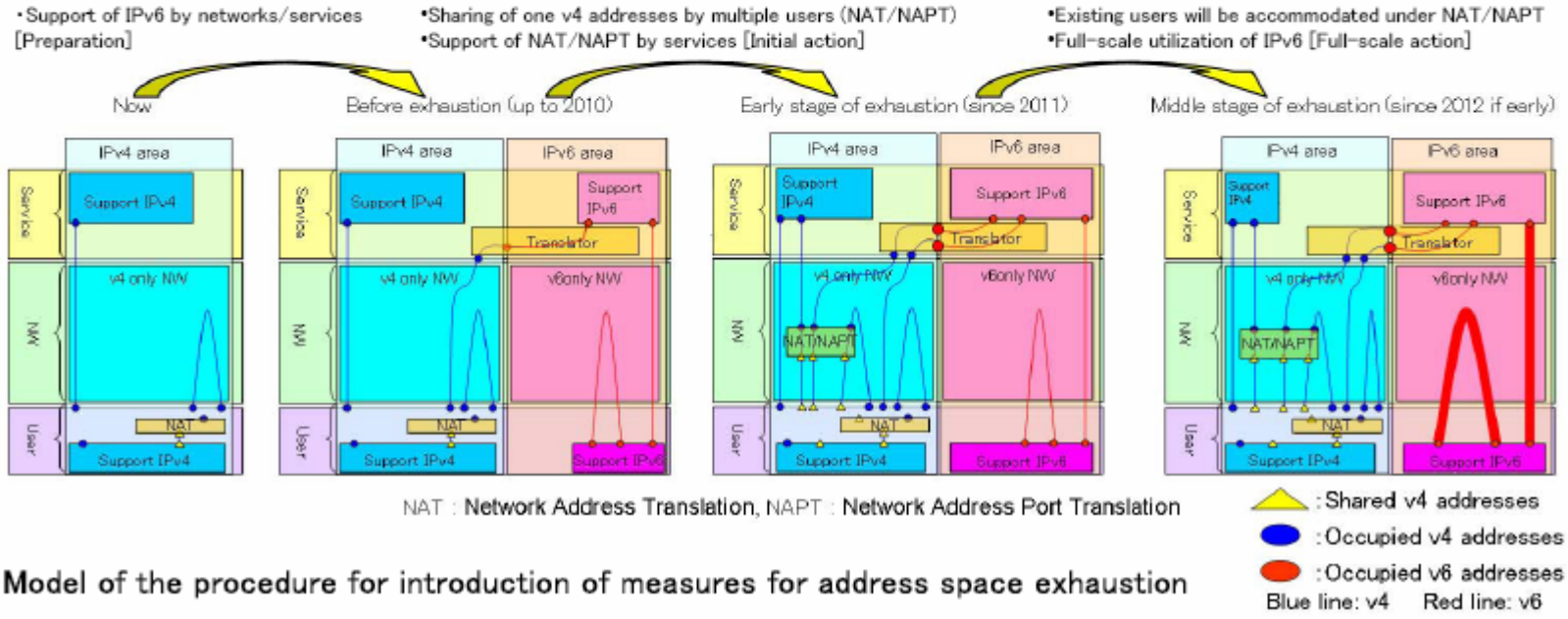
- Encourage ISPs to deploy their IPv6 networks and to install their IVI gateways.
  - Reserve 2001:DB8:ff00::/40 for each 2001:DB8::/32
  - Encourage ISPs to use a subset (i.e. IVI4(i)) of their own IPv4 address blocks and map it into IPv6 via the IVI scheme (i.e. IVI6(i)) for their initial deployment of IPv6.
    - For severs using the 1-to-1 mapping, and for clients using the 1-to-2^N mapping.
    - In this way, the scarce IPv4 addresses can be effectively used.
    - This IVI6 can communicate with the global IPv6 networks directly and communicate with the global IPv4 networks via IVI gateways.
- Encourage ISPs to increase the size of IVI4(i). When IVI4(i)=IPS4(i), the IPv4 to IPv6 transition for ISP(i) will be accomplished.

IPv4 Address Allocation Policy

- The remaining IPv4 address should be dedicated for the IVI transition use, i.e. using these blocks for the IVI6(i) deployment.
  - The users using IVI6(i) can access the IPv6 networks directly and the IPv4 networks via the IVI gateways.
- Based on multiplexing techniques, the global IPv4 addresses can be used effectively.
  - For example, with a reasonable port multiplexing ratio (say 16), one /8 can support 268M hosts. If 10 /8s can be allocated for the IVI use, it will be 2.6 billion addresses, possibly enough even for the unwired population in the world.
- The 43.0.0.0/8 could be a good candidate for the initial trial

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.

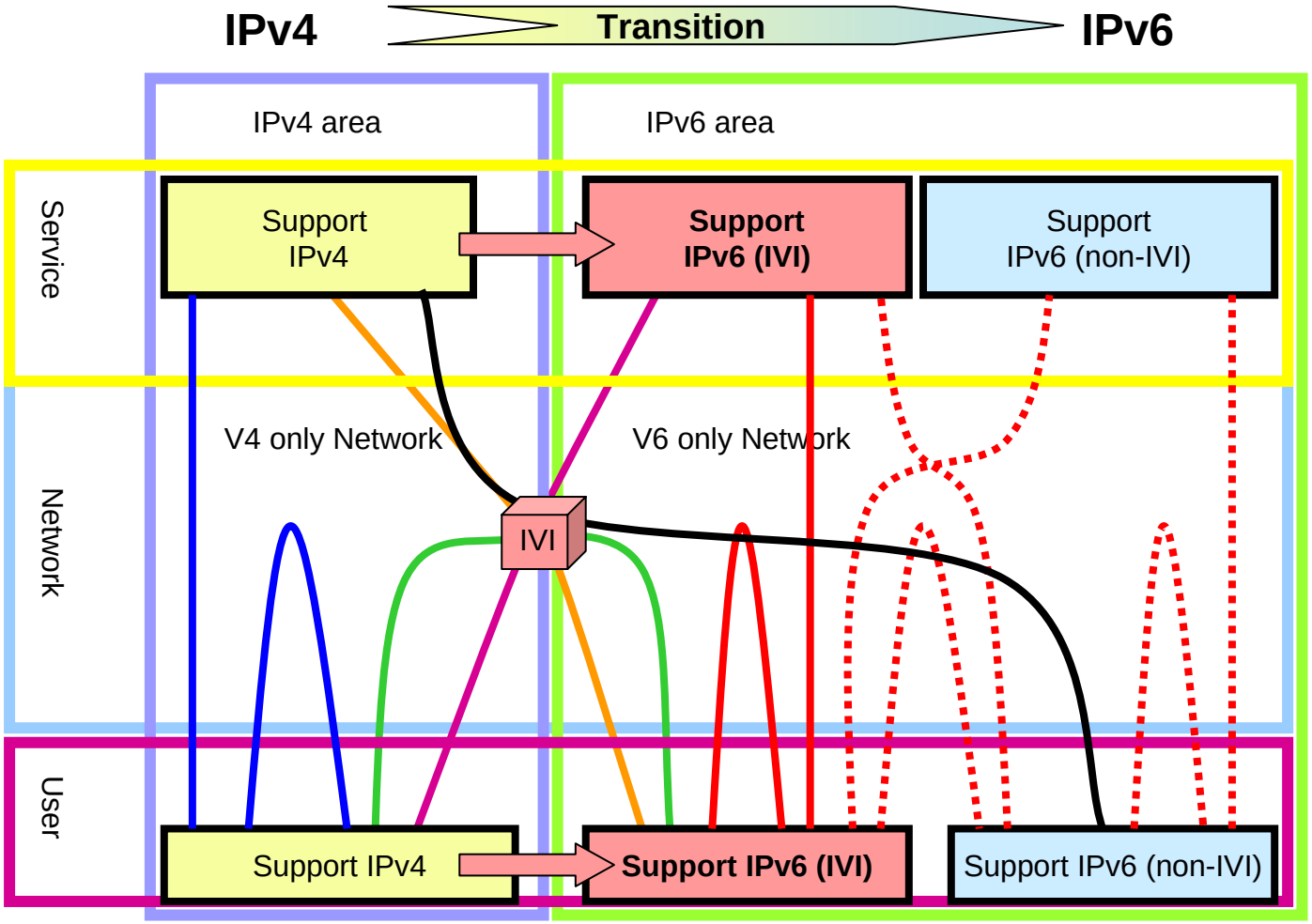


Model of the procedure for introduction of measures for address space exhaustion

6

From the June 2008 Report of the Japanese Study Group on Internet's Smooth Transition to IPv6

Evolution of the IVI Addresses and Services



Remarks for the transition (1)

- 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.

Remarks for the transition (2)

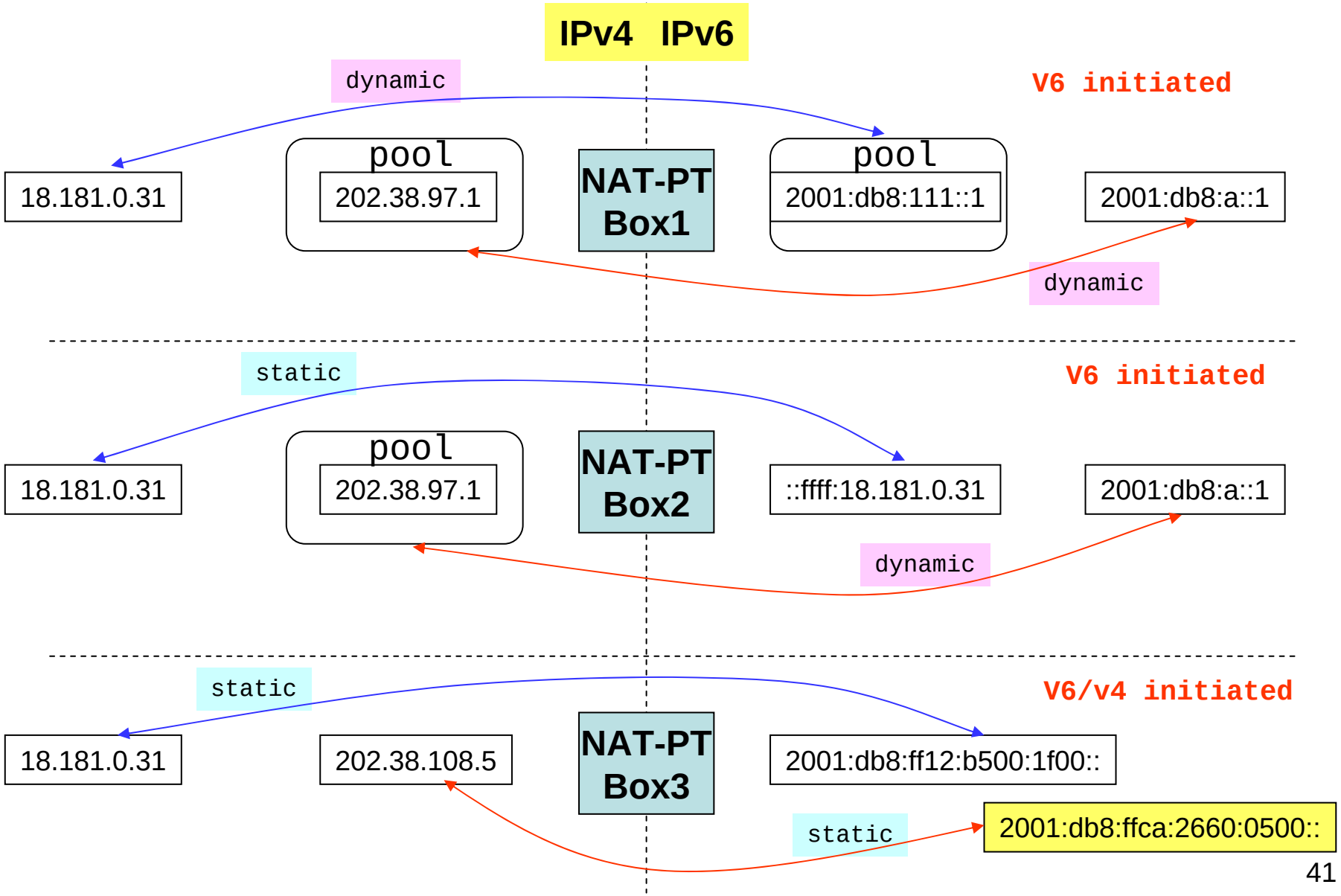
	Utilization of NAT/NAPT (Sharing of IPv4 addresses)	Reallocation of the assigned IPv4 addresses	Transition to IPv6	IVI
Feasibility within a time limit	✓	Doubtful	Extremely difficult	√
Service continuity	Limited	✓	✓	√
Permanent effect	Doubtful	NG	✓	√

Modified based on the June 2008 Report of the Japanese Study Group on Internet's Smooth Transition to IPv6

Discussion

- Why select a subset of the IPv6 addresses, rather than allow the whole IPv6 addresses to access the IPv4
- Mathematics of mapping
  - Because of the different size of the two address families, there must exist constrains.
- A subset is enough for the initial deployment
  - The IVI6 subnet is much, much larger than the global IPv4 whenIPv4 multiplexing techniques are used), even only a small portion of the public IPv4 addresses are borrowed by IVI.
  - Every IPv6 host can communicate with the global IPv4, not every IPv6 address (IPv4 class E address cannot communicate with class A, B, C).
- The standard NAT-PT methods also require the reservation of a similar size of the public IPv4 addresses in the pool.
  - These methods are maintaining a pool of public IPv4 addresses in NAT-PT box
- This subset supports the v6 and v4 initiated communications.
  - P2P
  - Pseudo-well-know-port, DNS SRV record

Comparisons (1:1 mapping example)



Conclusions

- The IVI is a prefix-specific and explicit bidirectional address mapping scheme.
- Both IPv6 initiated and IPv4 initiated communications can be supported.
- No affect to both IPv4 and IPv6 routing. It is scalable and reliable.
- The deployment can be done incrementally and independently.
- Depending on the mapping rule, the gateway can be in any part inside the ISP’s network.
- The IVI comes the closest to the end-to-end address transparency model.
- The IVI scheme encourages the transition.