IPv4/IPv6 Smooth Migration (IVI)

Xing Li etc.

2008-08-16

Abstract

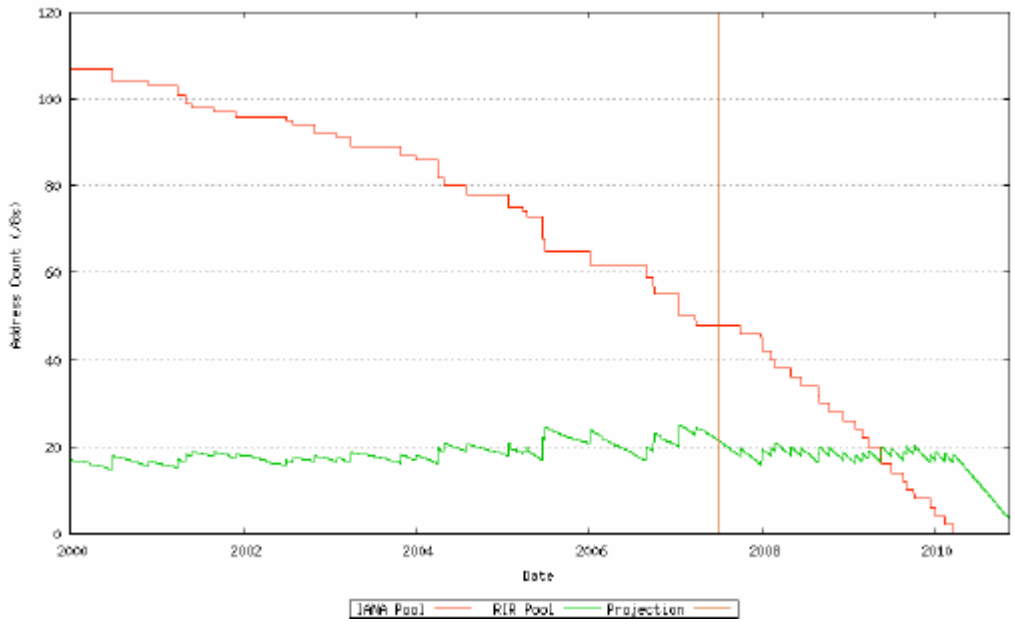
- This presentation will introduce the concept and practice of prefix-specific and bi-direction explicit address mapping (IVI) for IPv4/IPv6 migration.
- The comparison between IVI and other mechanisms will be compared.
- The impact to the IPv4/IPv6 address assignment and allocation policy will also be discussed.

Outline

- Introduction
- Requirements
- IVI concept and practice
- Comparisons
- Impact to the address allocation and assignment policy
- Remarks

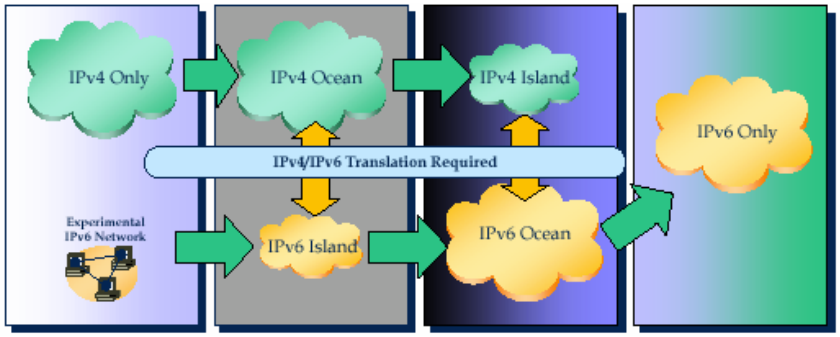
Major Issue

- IPv6 deployment is not fast enough to transition away from IPv4
- IPv4 run-rate predictions seem believable

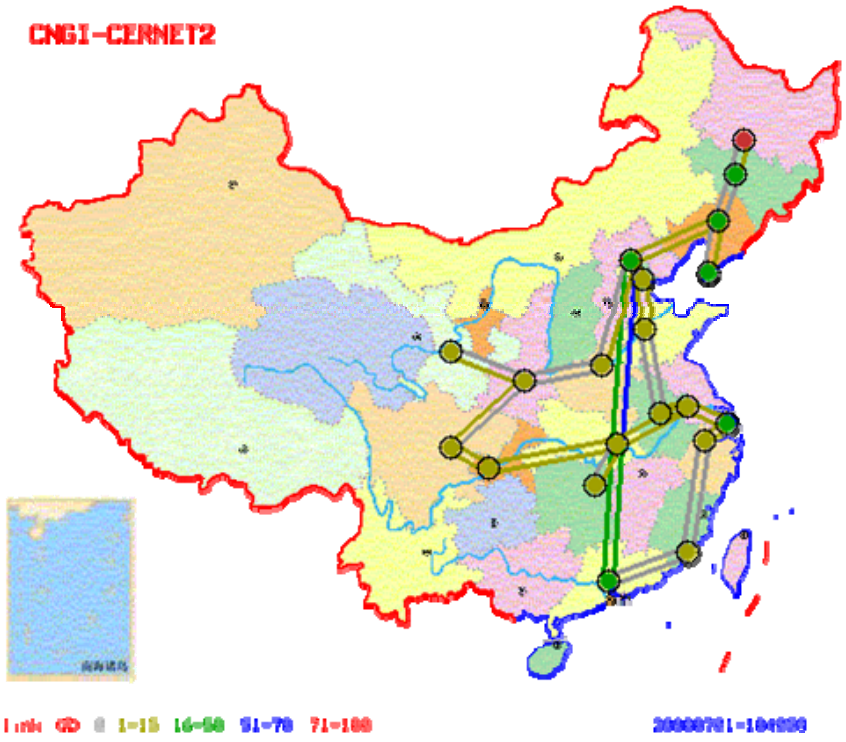


Fundamental Reason

- The false IPv6 selling arguments
  - Restoration of the “end to end” principle
  - Restoration of Address transparency
  - Multicast
  - Better QoS (flows)
  - Embedded IPSEC
  - Auto-configuration, Plug & Play, etc
- The fundamental reason should be
  - Extend the address space to keep the universal connectivity



CNGI-CERNET2



- CNGI-CERNET2 is an IPv6 single stack network.
- The original promotion concept
  - It is free and it is light loaded.
  - The users need to export their applications into IPv6.
- But this concept did not work well.
  - The connectivity is the most important issue.
- So we developed IVI
  - IV means 4
  - VI means 6
  - IVI means 4|6 coexistence and transition
  - IVI is symmetric and both v6 and v4 initiated communication are supported

IVI draft: http://www.ietf.org/internet-drafts/draft-xli-behave-ivi-00.txt

Requirements

- Exiting approaches
  - Dual stack approach
    - IPv4 address depletion problem
  - Tunneled architectures
    - No communication between two address families
  - Translation architectures (NAT-PT)
    - Not scalable, lost end-to-end
- Requirements
  - Technical
    - End-to-end address transparency, minimum state, globally deliverable and effectively use of the global IPv4 addresses.
    - Meet different requirements of server, client and P2P
    - Independent and incremental deployable
  - Non-technical
    - Encourage the migration

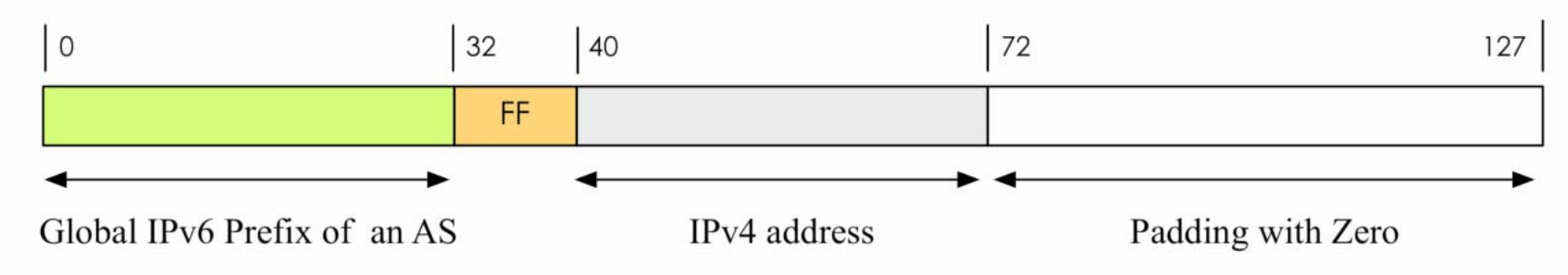
The Key Concepts of IVI

- Prefix Specific Addressing and Routing
  - Maintain a clean Internet addressing and routing architecture and globally deliverable
- Bi-directional and Explicit Mapping
  - Restore the end-to-end address transparency
  - Maintain the minimum state
- Extended Address Transparency
  - Support the both IPv6 initiated and IPv4 initiated communications for every IPv6 host (not every IPv6 address)
  - Effectively use the global IPv4 addresses
  - Meet different requirements of server, client and P2P
- Protocol translation
  - SIIT
  - ICMP extension
  - Multicast extension

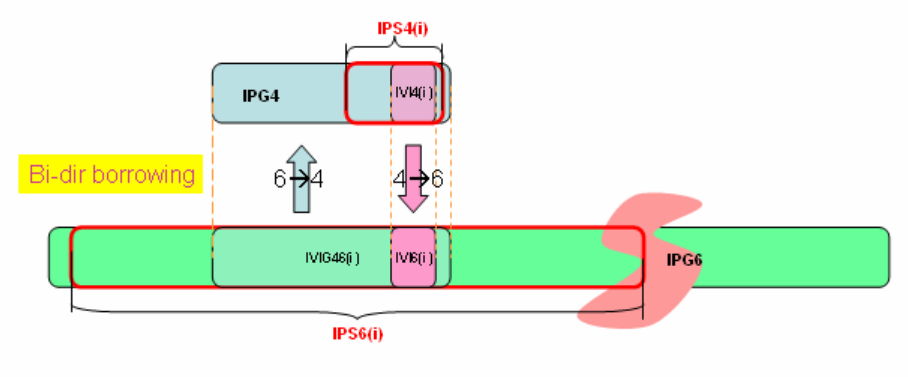
Terms and Abbreviations of IVI

- **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**

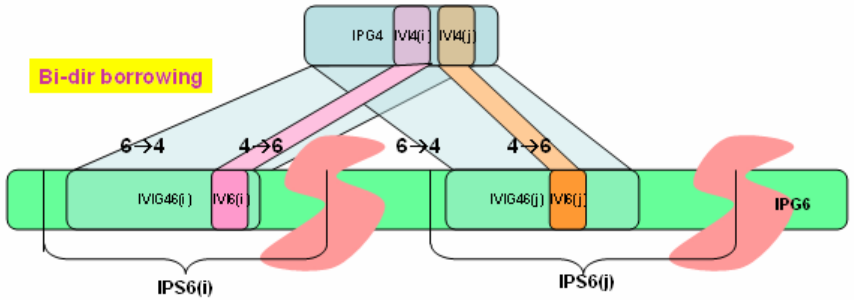
IVI Address Mapping



Mapping rule

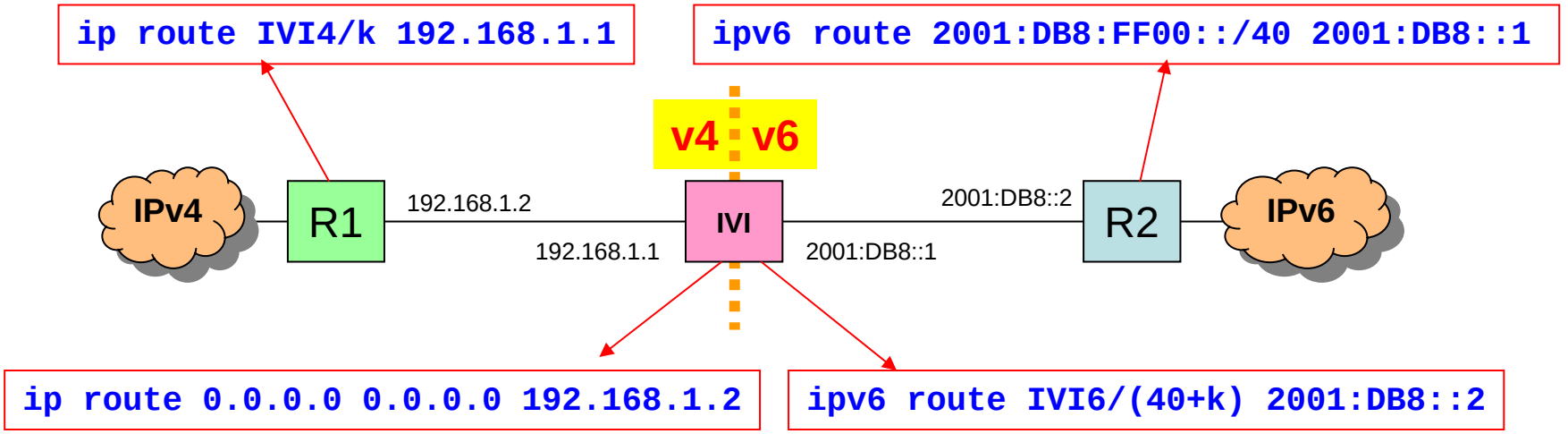


Single ISP deployed IVI



Multiple ISPs deployed IVI

Routing and Forwarding



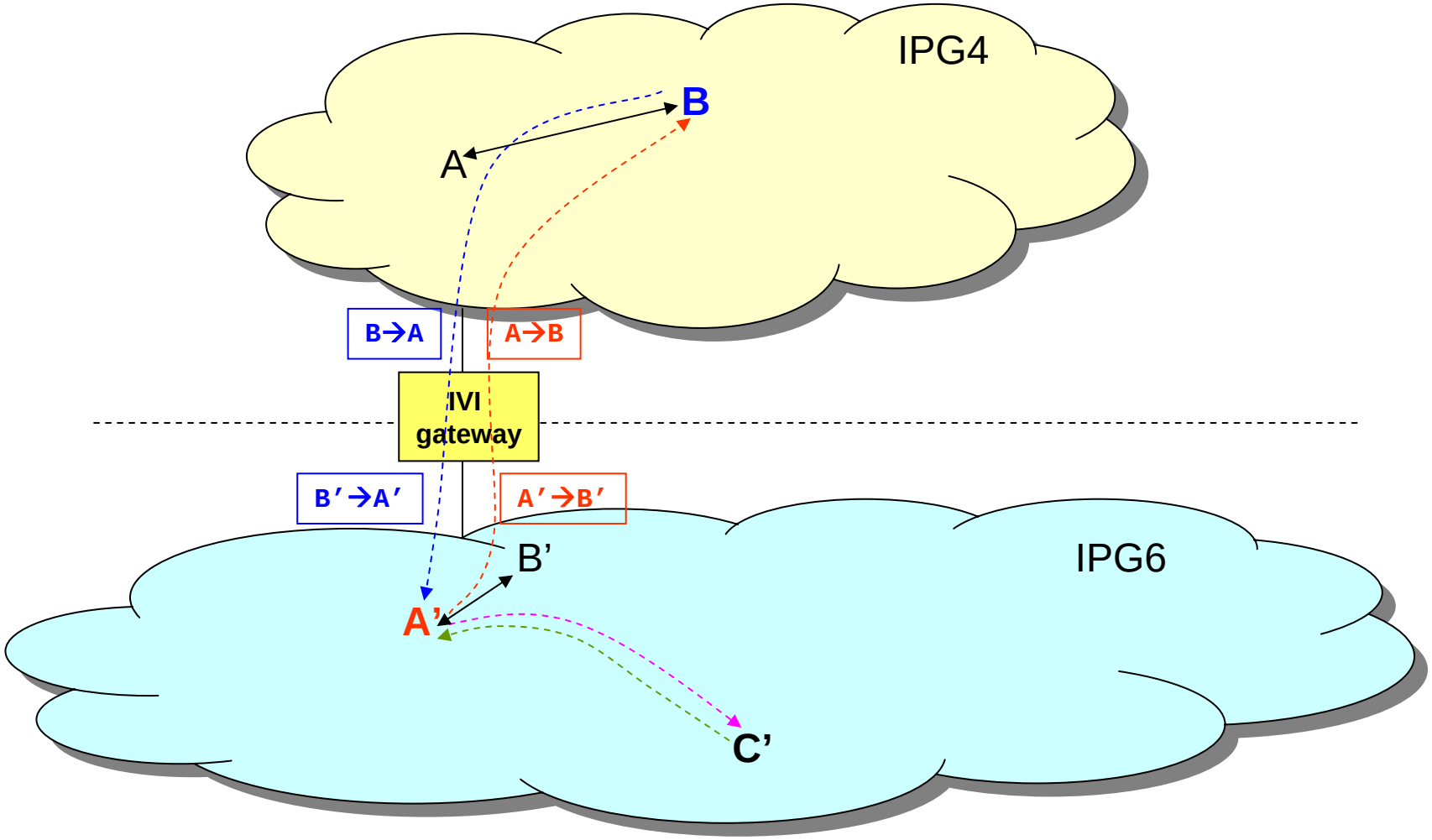
Longest prefix match

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

IVI Reachability Matrix

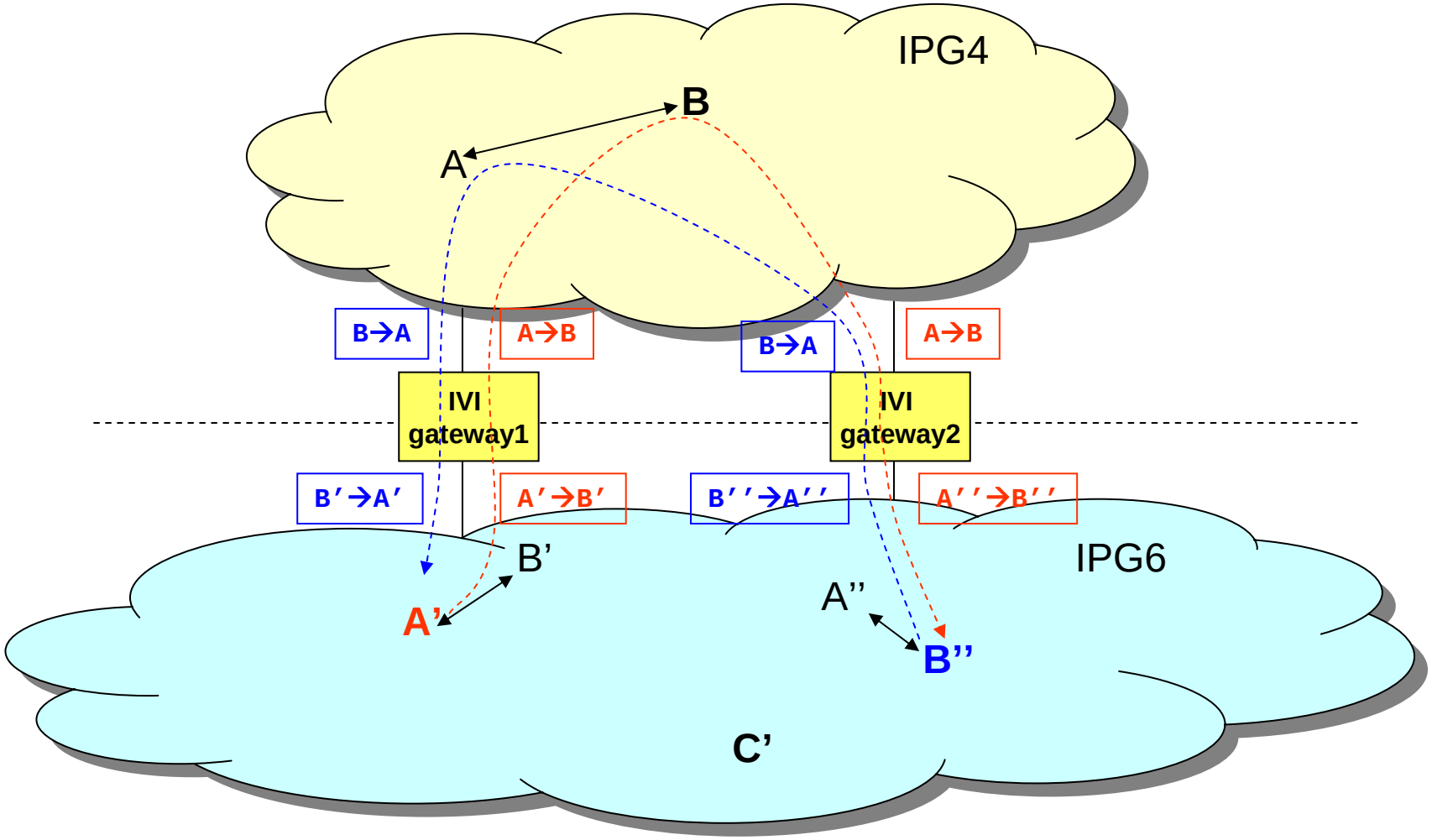
	IPG4	IVI	IPG6
IPG4	OK	OK	NO
IVI	OK	OK	OK
IPG6	NO	OK	OK

IVI Communication Scenarios (1)



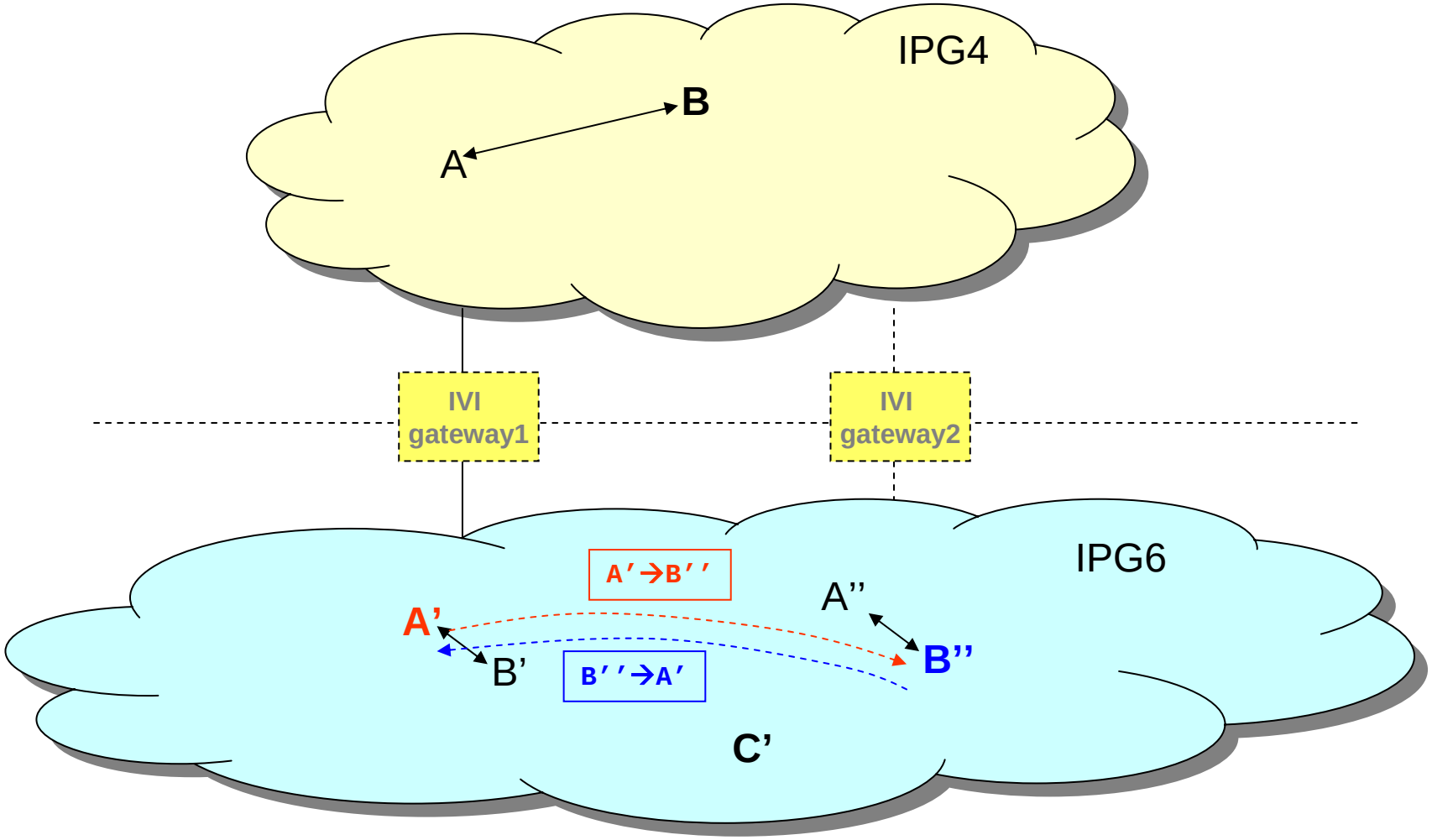
- A' ←→ B
- A' ←→ C'

IVI Communication Scenarios (2)



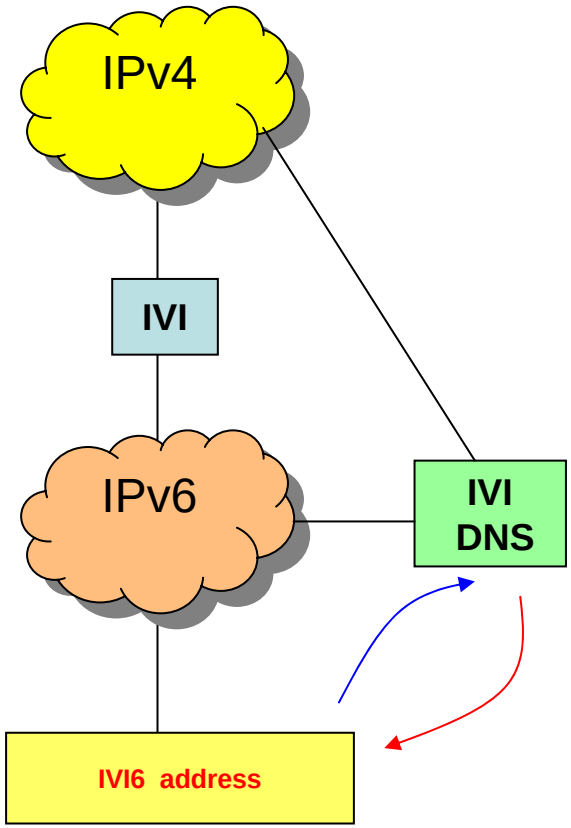
- A' ←(B&A)→ B''

IVI Communication Scenarios (3)



- A' ←→ B''

IVI DNS Configuration



- 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 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)

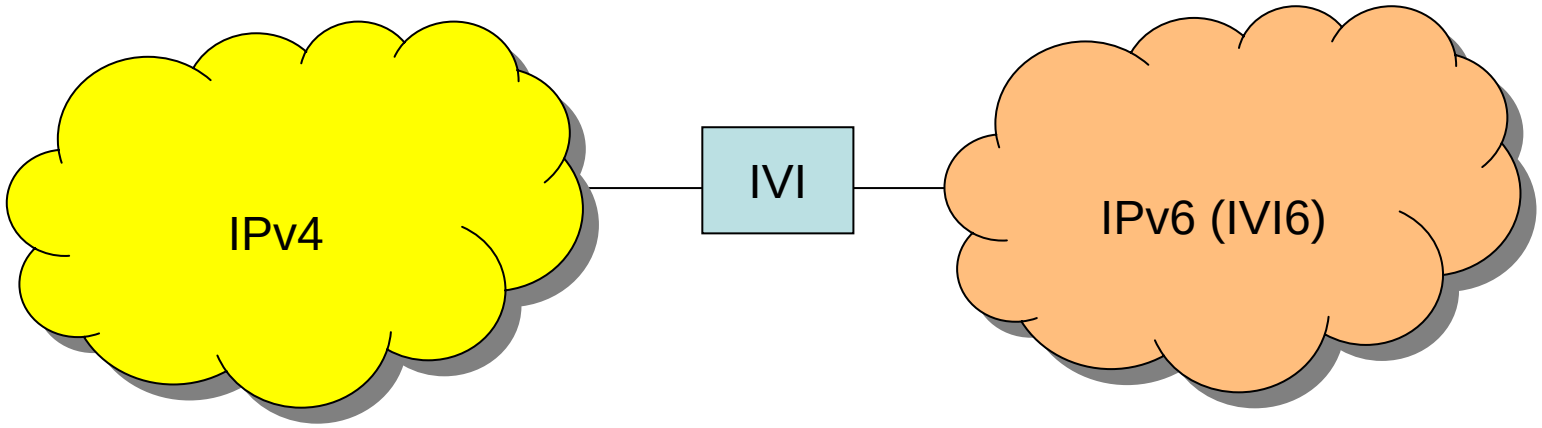
Extended Address Transparency

- End-to-end address transparency: the source and destination addresses of the packets could be used as unique labels for the end systems (RFC2755).
- Port multiplexing extends the address transparency
  - Basic NAT 2^{32}
  - Extended NAT 2^{48}

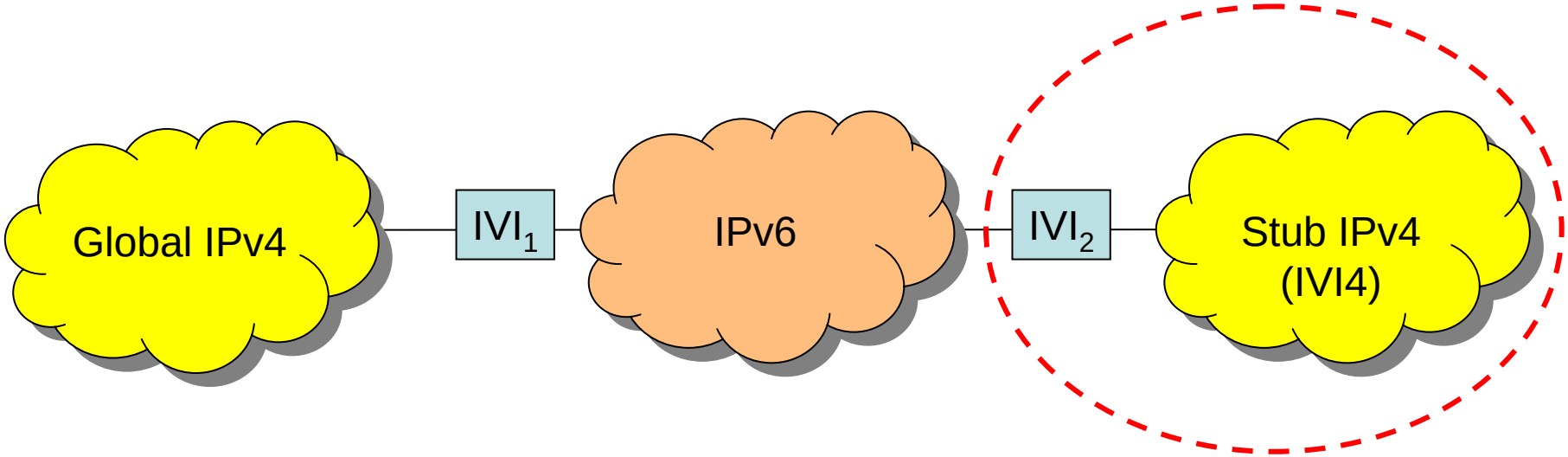
Port embedding

- IPv6 client initiates the communication to the IPv4 servers
  - Method 1: port collision avoidance
    - **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**
  - Method 2: embed port range into the IVI6 addresses
    - **2001:250:ffca:266c:0500:ratio:bias:pseudo-well-know-port**
- IPv4 client initiates the communication to the IPv6 servers
  - Method: provide pseudo-well-know-port via SRV DNS record (i.e. the remote IPv4 host can reach different IVI6s via different port number)
    - **202.38.108.2#81 ←→ 2001:250:ffca:266c:0200:3:0:81#81**
    - **202.38.108.2#82 ←→ 2001:250:ffca:266c:0200:3:1:82#82**
    - **202.38.108.2#83 ←→ 2001:250:ffca:266c:0200:3:2:83#83**
    - **202.38.108.2#84 ←→ 2001:250:ffca:266c:0200:3:3:84#84**

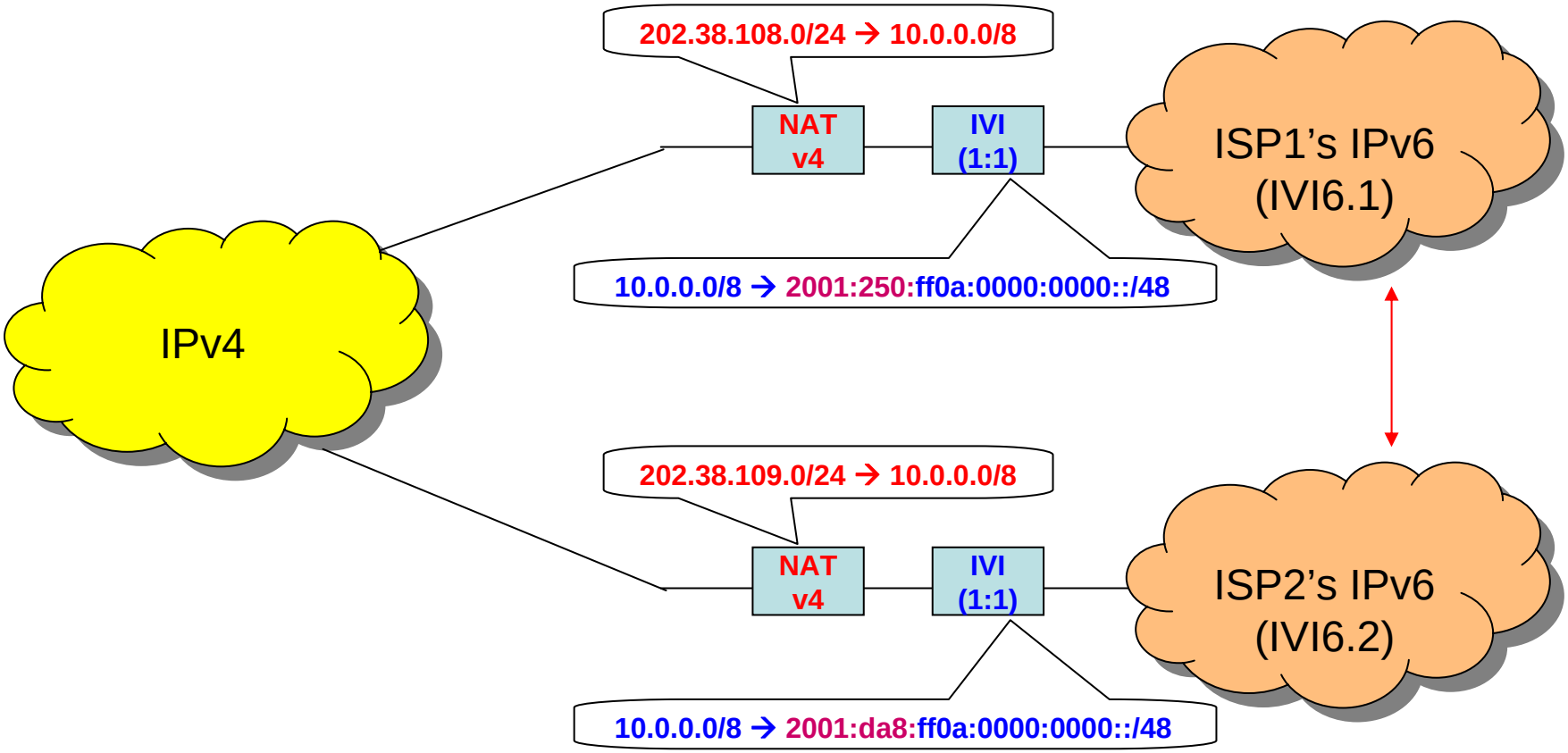
IVI Deployment Scenarios (1)



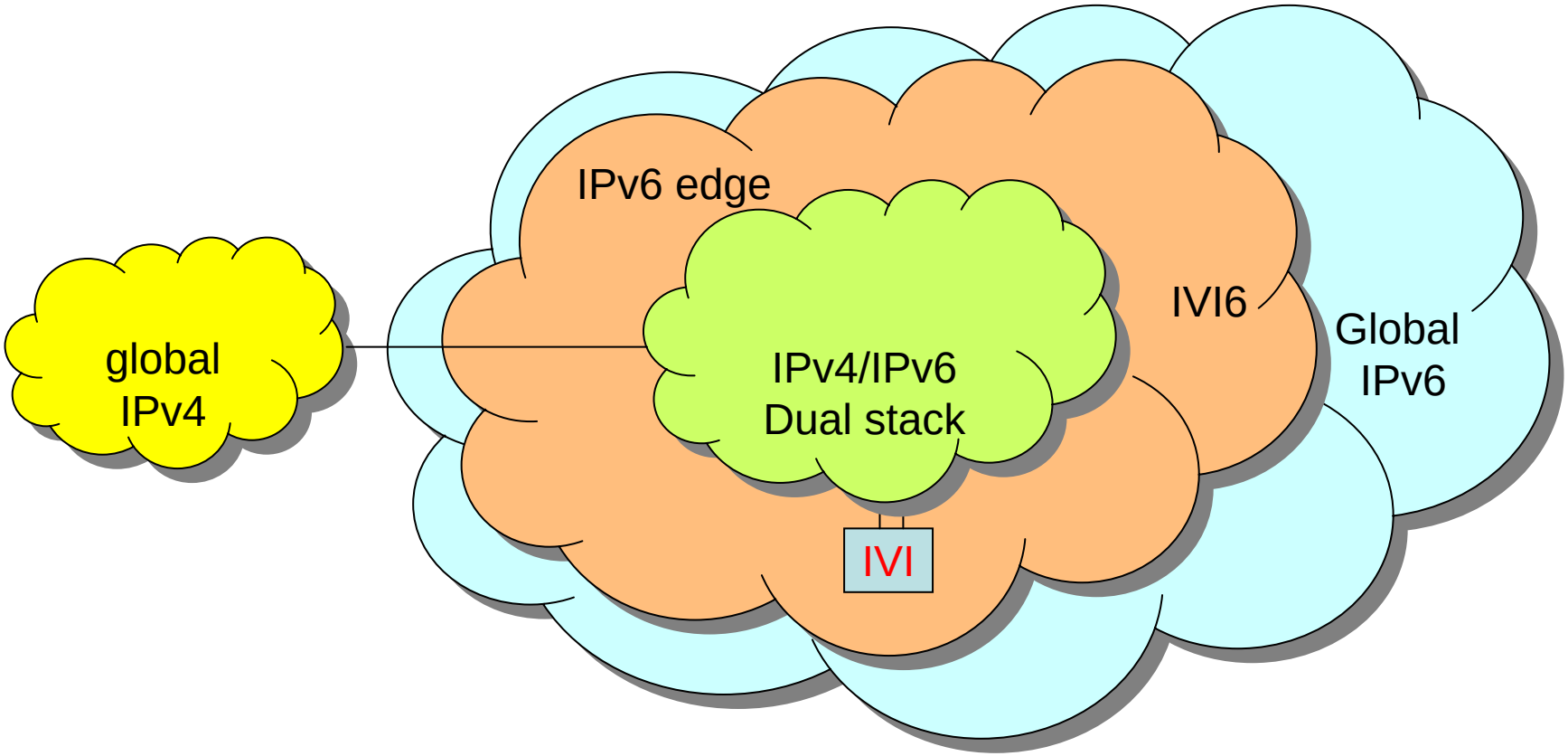
IVI Deployment Scenarios (2)



IVI Deployment Scenarios (3)



IVI Deployment Scenarios (4)

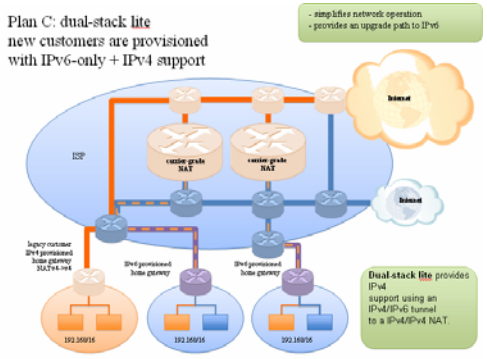


Implementation and Testing Results

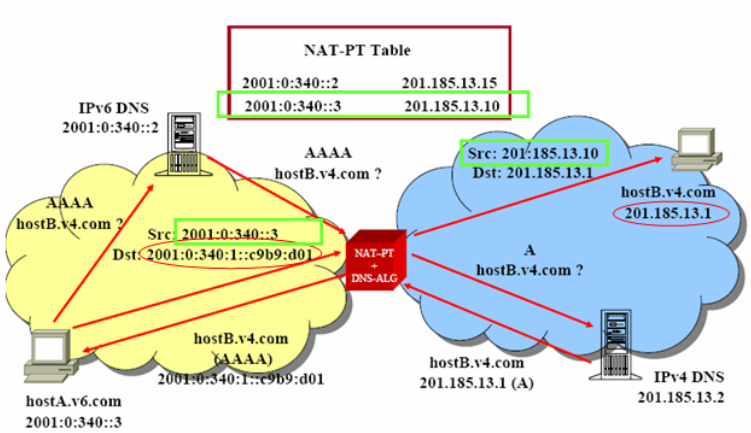
- The IVI scheme presented in this document is implemented in the Linux OS
  - The source code can be downloaded [http://202.38.114.1/impl/].
- CERNET (IPv4 and partially dual-stack) and CNGI-CERNET2 (pure IPv6) since March 2006 (basic implementation).
  - IVI6 server for global IPv4
    - http://202.38.114.1/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

Comparisons (1)

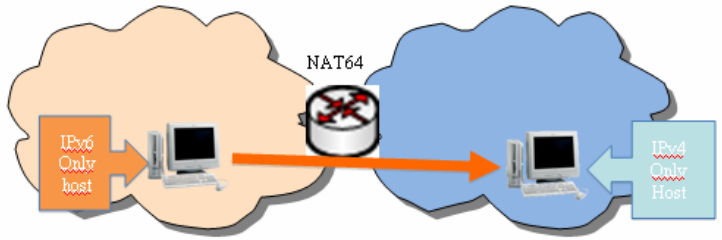
Dual-stack lite



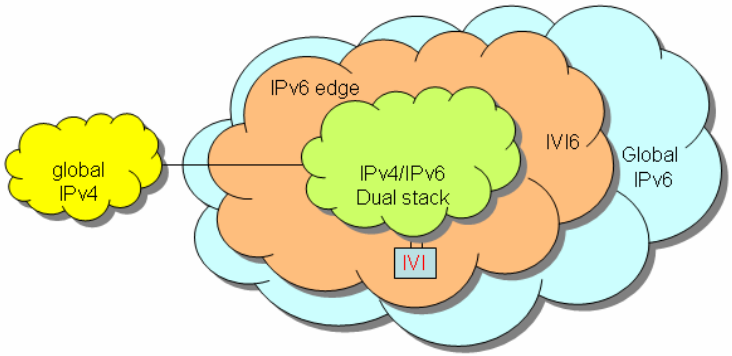
NAT-PT



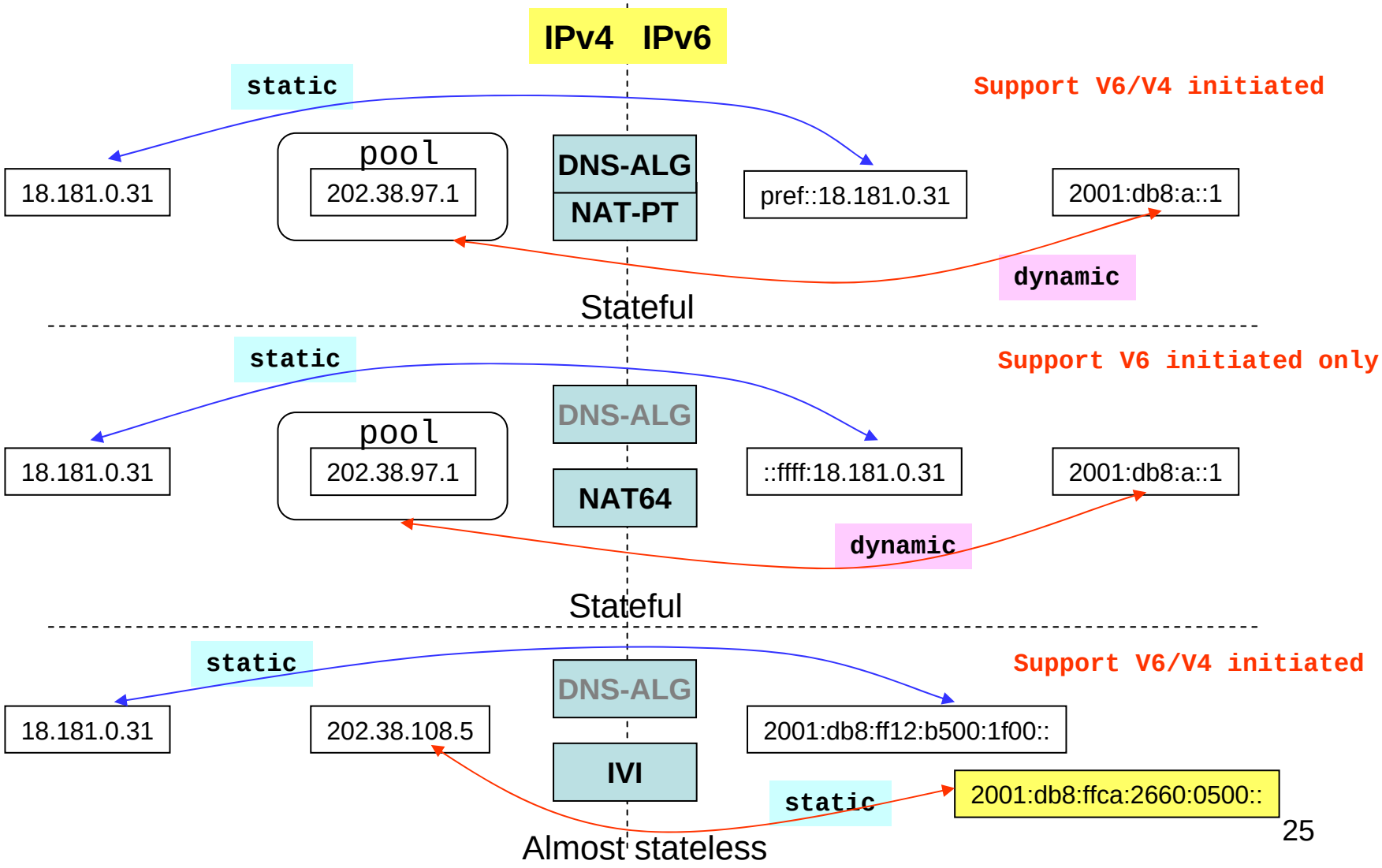
NAT64



IVI



Comparisons (2)



Comparisons (3)

- Dual-stack lite
  - ‘Carrier class’ is a euphemism for centralized. More semantics move to the core of the network. This is bad in and of itself. Net-heads call it ‘telco-think’ because it is the telco model of smarts in the core as opposed to the internet model of a simple, just forward packets, core and smart edges. - Randy
- NAT-PT
  - NATPT supports both v4 and v6 initiated, requiring a set of cumbersome techniques
- NAT64
  - NAT64 only supports v6 initiated communications
  - NAT64 and DNS64 are completelly decoupled
- IVI
  - End-to-end address transparency, minimum state, globally deliverable and effectively use of the global IPv4 addresses
  - DNS mapping is completelly decoupled
  - Meet different requirements of server, client and P2P
  - Independent and incremental deployable
  - Encourage the migration

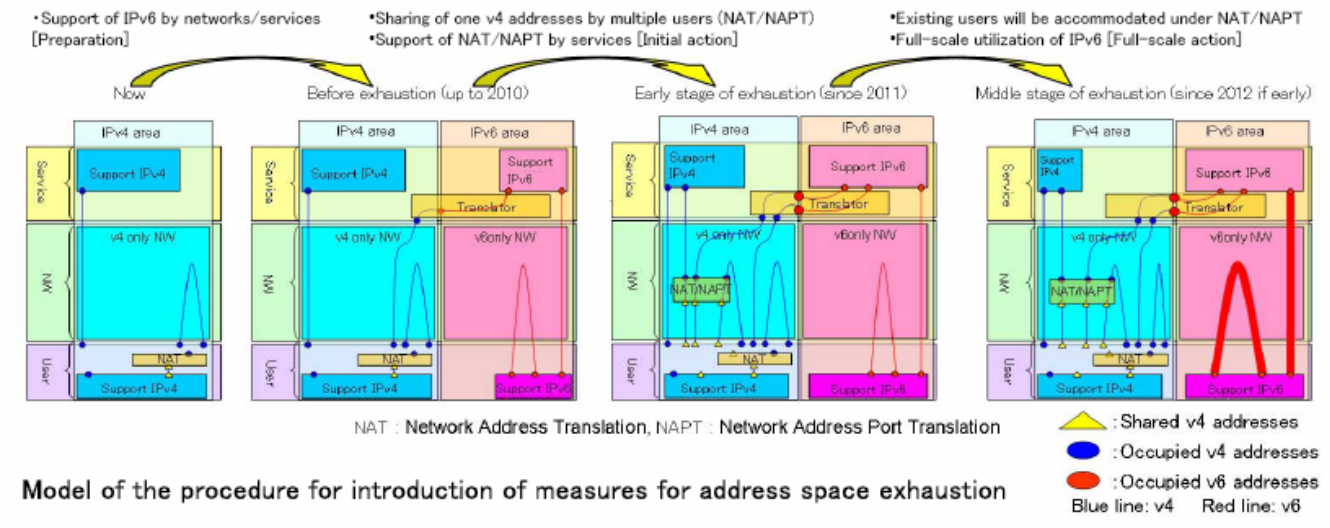
IPv6 Assignment Policy

- 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.
- 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 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

Remarks



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

The IVI migration path:

Every IPv6 host (not every IPv6 address) can communicate with the global IPv4 (both IPv6 initiated and IPv4 initiated).