Ongoing Innovation

IIJ
Internet Initiative Japan

IPv6 deployment at IIJ

Matsuzaki ‘maz’ Yoshinobu
<maz@iij.ad.jp>

IIJ network

- pure IP network
  - now dual stacked
  - no MPLS
- IIJ/AS2497 maintains its backbone in Japan and U.S..

History of IIJ's IPv6 service

- 1999
  - IPv6 over IPv4 tunnel service (experimental)
- 2000
  - IPv6 native service (leased line, experimental)
- 2001
  - IPv4/IPv6 dual stack service (leased line)
  - IPv6 data center service
  - IPv6 over IPv4 tunnel service for consumers

business model

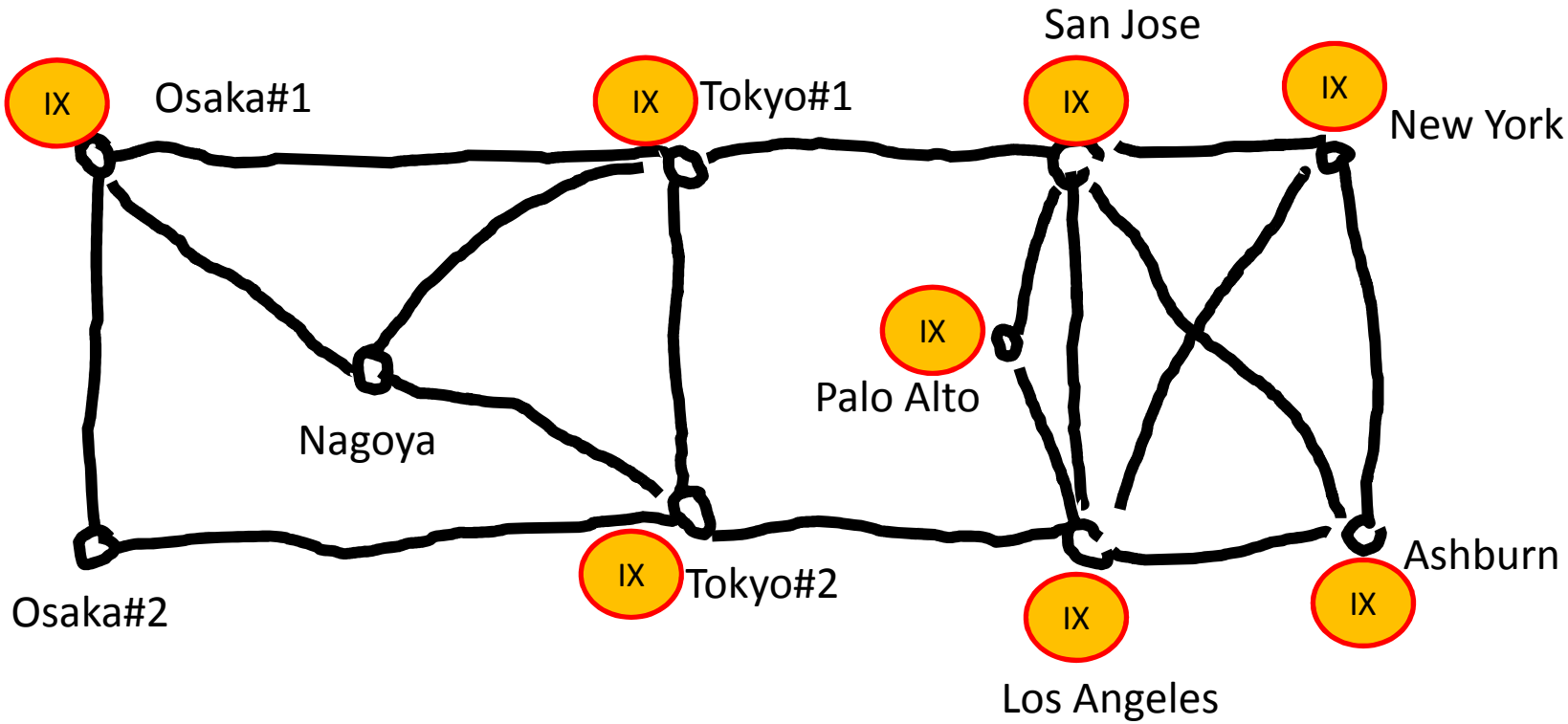
- IIJ sells bandwidth
  - Customers can select protocol(s) which should be forwarded on the link
    - IPv4 only
    - IPv4/IPv6 dual stack
    - IPv6 only
  - Or, customer can ask IPv6 over IPv4 tunnel for free.

History of IIJ's IPv6 backbone

- Initially we started with dedicated backbone
- 1998
  - PC based router(kame stack)
  - tunnel and ethernet
- 2000
  - cisco c72xx series
  - tunnel, ethernet and T1 line

brief backbone topology as of today

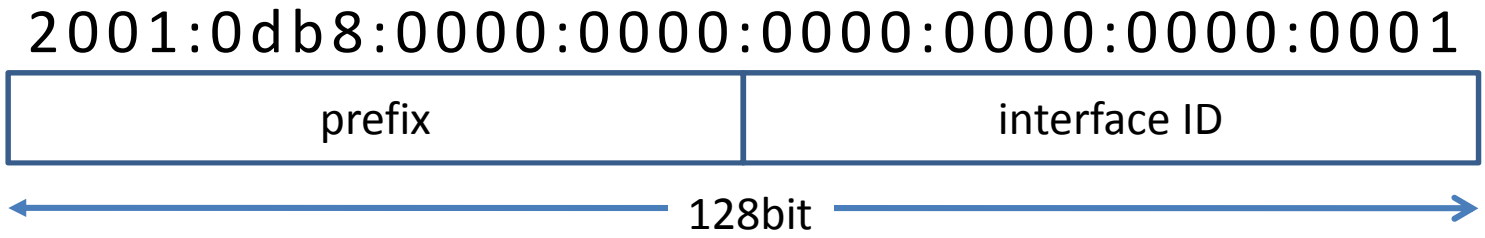
- dual stack as possible



addressing

- /128 for loopback interfaces
- /64 for links
  - /127 is used on several inter-router links
- /48 for customer sites
  - still considering the size
  - possible sizes are: /48, /52, /56, /60, /64

interface ID



- We assign unique 64bit ID to every router.
  - A router uses the ID as its interface id on every interface.
  - No correlate with IPv4 address
- exception
  - customer facing interface
  - /127

global unicast address

- Inter-router link does not require global address inside AS
  - OSPFv3 uses link-local address to exchange LSAs
  - only loopback interface needs to configure global address
- But we configure global address on every interface
  - as ping destination to check availability

link-local unicast address

- fe80::/64
- AS IS
  - We don’t touch
  - Most routers use Modified EUI-64 format address
- A virtual address for vrrp/hsrp is another story.
  - Customers might configure a static route at their equipments with this address, so the address should be assigned statically like fe80::1.

route aggregation inside AS

- not so aggressive at this moment
  - We allocate /56 to every POP for links inside POP
  - But currently, # of IPv6 prefix << # of IPv4 prefix

IPv6 Routing Table Summary - 1915 entries
5 local, 2 connected, 0 static, 0 RIP, 1329 BGP, 0 IS-IS, 579 OSPF, 0 EIGRP
Number of prefixes:
/8: 1, /16: 1, /19: 2, /20: 4, /21: 3, /22: 1, /24: 1, /25: 1
/26: 5, /27: 3, /28: 6, /30: 2, /31: 1, /32: 998, /35: 30, /40: 3
/41: 1, /42: 1, /43: 1, /44: 2, /45: 1, /46: 2, /47: 3, /48: 216
/49: 1, /53: 1, /56: 7, /64: 454, /126: 2, /127: 2, /128: 159

routing protocols

IPv4

- OSPFv2
  - mostly area 0
  - md5 authentication
- BGP4
  - peer through ipv4
  - route-reflector
  - md5 authentication

IPv6

- OSPFv3
  - area 0 only
  - ipsec authentication
- BGP4+
  - peer through ipv6 global
  - route-reflector (same as IPv4)
  - md5 authentication

IPv4/IPv6 separation in a router

- IPv4 and IPv6 are different protocols
  - should not affect each other
  - only sharing CPU, memory and bandwidth
- Quality of IPv4 network is still main concern in our business.

router ID

- Routing protocols usually require 32bit ID
  - even a routing protocol for IPv6
  - We use IPv4 address of loopback interface as its router ID
- Every routers has IPv4 address in our network

OSPFv3 link cost

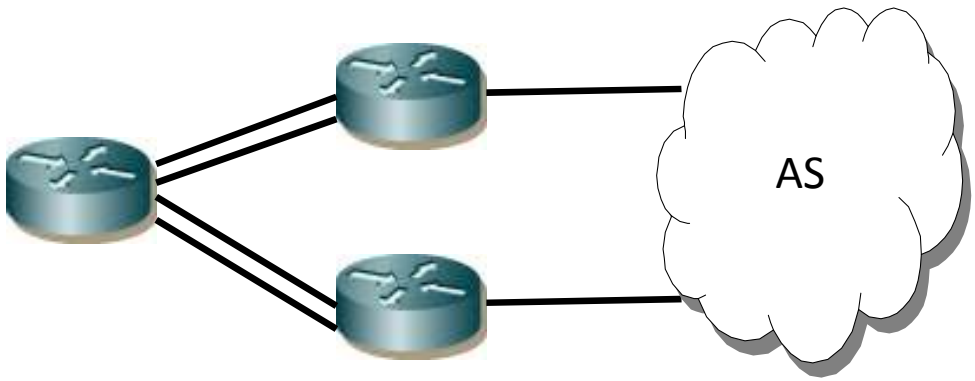
- We set the same link cost value as IPv4's.
  - The network topology is almost same.
  - working fine ☺
- When we were using RIPng as IGP (there was no choice at that time ☹), it was so much trouble.

BGP4+ nexthop attribute

- AS IS
  - We keep the nexthop attribute heard from eBGP.
    - no nexthop-self
    - same as IPv4 case at IIJ
- So we are importing prefixes of IXs and PNIs into OSPFv3.

BGP4+ multipath

- for testing purpose
  - The traffic level of IPv6 is low at this moment.
  - But we configure iBGP multipath at our backbone.
- still confusing multipath behavior
  - 2 BGP multipath and 2 IGP ECMP for each nexthop
    - 4path?
    - 2path?



management

- remote access to a router (ssh/telnet)
  - IPv4/IPv6 dual stack
  - out-of-band port has IPv4 address only
- other services
  - IPv4 only
  - AAA, snmp, syslog, ntp, flow export

availability check

- checking by ping
  - both IPv4 and IPv6
  - dual stack routers receive ping twice as much

monitoring

- We have to know about our network. ☺
  - traffic volume and so on

- We are waiting for
  - MIB for IP [RFC4293]
  - NetFlow v9 [RFC3954] for IPv6

ACLs

- Infrastructure ACL
  - not yet deployed, but we are still considering
  - We have allocated certain block for our infrastructure so that ACL rules become simple.
- Received ACL
  - nightmare with next header chain…

concerns

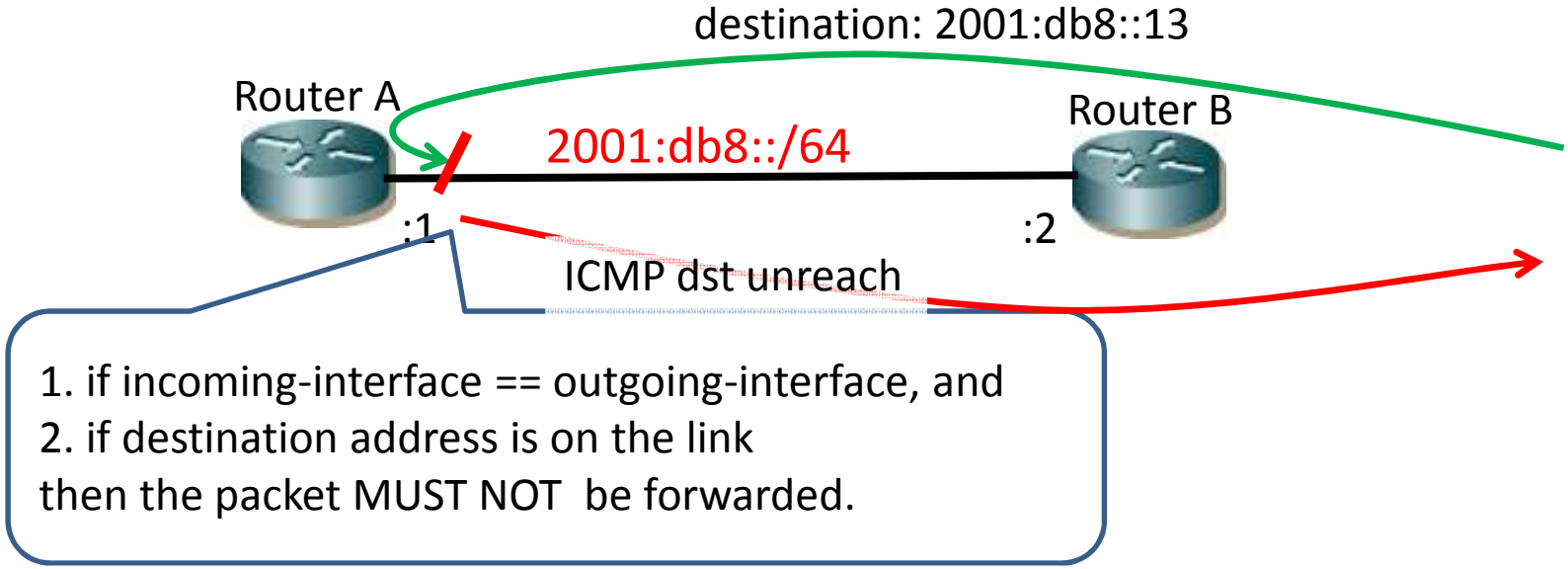
- license fee
- packet ping-pong on a point-to-point link
- loopback testing and DAD
- next header
- ndp flood
- subnet router anycast address
- “u” bit in the interface identifier

license fee

- more cost to use ‘IPv6 code’
  - an obstacle to deploy IPv6 in the world ☹

ping-pong on p2p link

- spec vs. implementation
  - ICMPv6[RFC4443] says



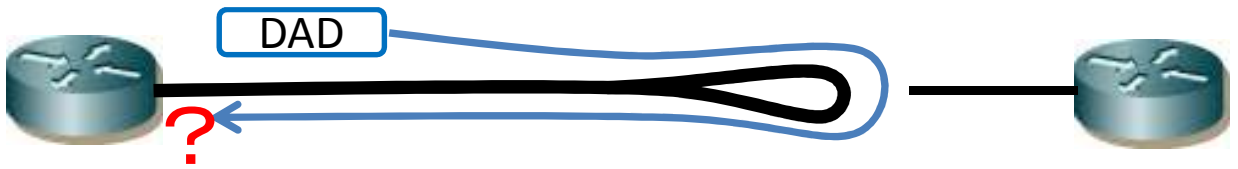
- I prefer /127 on inter-router links.

loopback testing

- a hard/soft loop for a SONET link checking
  - to check the link at installation
    - running for 24hours or so
    - error counters
  - for troubleshooting
    - alarms
- A router see itself through the looped link.

loopback testing and DAD

- An interface comes up when looped.
- IPv6 nodes have to perform DAD before use the address.
- Some routers get confused with this situation.
  - No IPv6 processing at the looped interface



next header

- We are annoyed by a packet filter

- for transit traffic
  - should we limit # of header chain?
- for traffic to a control plane
  - how to build appropriate rules…

ndp flood

- A remote user can send packets to every address in /64.
  - ndp flood at the segment
  - too many state on a router
  - continuous packets might cause DoS after expiring current ndp cache

subnet router anycast address

- one of the barriers to use /127
- Do we need this anycast address?
  - No.
  - It seems vendors disable this by default.
  - No one uses the network address in IPv4 anymore

“u” bit in the interface identifier

- universal/local bit
  - It is still kept for ‘future technology’

draft-ietf-6man-reserved-iids-01

> If the interface identifiers are generated from an unique token like an ethernet MAC address, they need to set bit 6 of the first octet to one. If they are not generated from an unique token they need to set bit 6 to zero.

- Who cares?
  - totally useless

END