Routing

Content

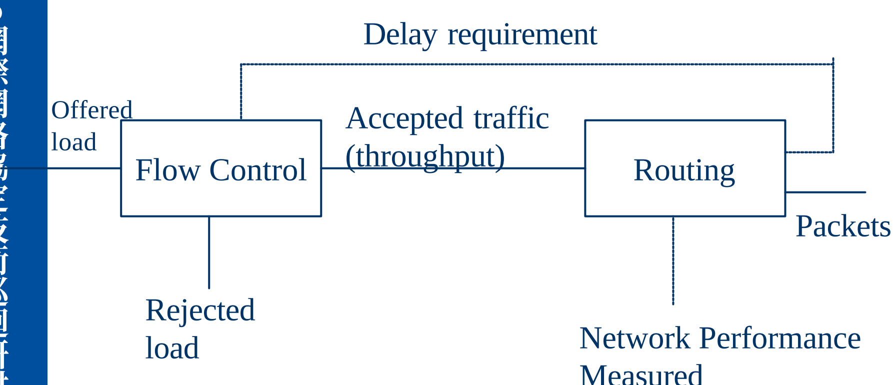
- Routing Basic
- Routing Protocols
  - Distance Vector Routing Protocols v.s. Link State Routing Protocols
  - Interior Gateway Protocol(IGP) v.s. Exterior Gateway Protocol(EGP)
- IPv6 Routing Concept
- RIPng for IPv6
- Multiprotocol Extensions for BGP-4

Routing

- Routing is the network layer function
  OSI Model 7 Layers
- Data grams v.s. virtual circuit routing
  ~ Packet Switching v.s. Circuit Switching
- Main issues in routing and flow control
  - throughput
  - average packet delay
  - throughput = offered load – rejected load
  - offered load may be rejected by the flow control mechanism

  ~ QoS, DiffServ , … a managed network

二〇〇二IPv6網際網路協定技術巡迴研討會



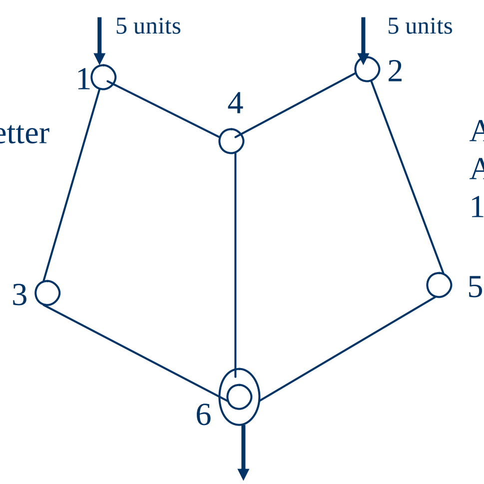
- The following examples illustrate the performance (delay) affected by the routing
- A good routing algorithm is to
  - Increase throughput → for the same average packet delay under high offered load;
  - Decrease average delay → under low and moderate offered load.
  - Now , the requirement of routing and classify computing by

  T-CAM → Network processor… To reach wire-speed

二〇〇二IPv6網際網路協定技術巡迴研討會

Example 1

Paths (1, 3, 6) (2, 5, 6) are better than (1, 4, 6) and (2, 4, 6)

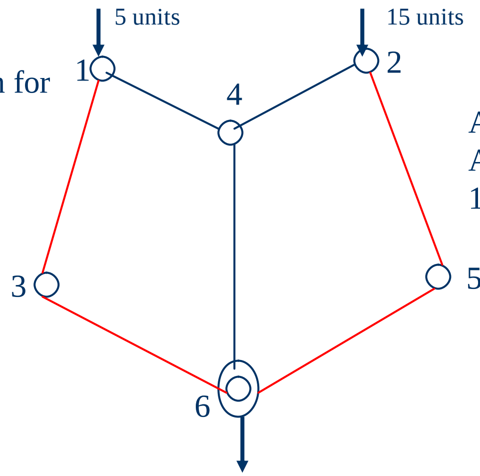


All links have A capacity of 10 units

Example 2

Best Path for
1 – 6

2 – 6



All links have
A capacity of
10 units

But how to make use of the rest banswidth …

二OO二IPv6網際網路協定技術巡迴研討會

Classifications of Routing Protocols

- Control – centralized v.s. distributed
- Route change – static routing v.s. adaptive routing
  - Static routing: the path used by the sessions of each origination-destination pair is fixed regardless of traffic condition.
  - Adaptive routing: Paths change occasionally in response to congestion.

Routing and Flooding

- Flooding and broadcasting
  - 不將收到之packet回送給來源端
  - 不重覆送出相同之packet
  - An important metric for the routing algorithms comparison → the shortest delay
- Shortest path routing
  - Each communication link is assigned a positive number (called length), the length may be any measurement criteria
  - Shortest path → Min-hop (if the length of each link=1)

二OO二IPv6網際網路協定技術巡迴研討會

Flooding

- 每一送出之Flooding packet均加上source ID並給予編號，每一node對一source相同編號的packet做傳送
- Source對其所傳送出去packet以遞增方式編號，而每一node對同一source只需記錄其所收到最大編號的packet即可

Routing Protocols

- Optimal routing
  - Traffic may be splitted at some strategic points to "smooth" delay and to increase network throughput
- Hot potato (deflection) routing
  - To minimize buffer overflow and to reduce the packet loss

二〇〇二IPv6網際網路協定技術巡迴研討會

Introduction to Graph

- A Graph G is defined as *(N, A)*, where *N* is a set of nodes, and a collection *A* of pairs of distinct nodes from *N*
- Undirected v.s. directed
- Spanning tree
- Minimum weight spanning tree
- Walk, path, loop

Importance of the Spanning tree

- A spanning tree is a subset of a network graph that includes all nodes
- There is only one path existed from the root node to each leaf node
- Broadcasting (flooding)
- Reverse path forwarding
- Minimum Spanning tree

二〇〇二IPv6網際網路協定技術巡迴研討會

STP(Spanning Tree Protocol)

-A standard techniquefor maintaining a network of multiple bridges or switches

-A part of IEEE 802.1D

-Avoid Loops and Establish redundant paths

-Default aging dynamic addresses is 5 min (300 secs)

-BPDU(bridge protocol data unit) :

Configuration BPDU, Topology Change Notification BPDU

通訊協定辨識碼
版本辨識碼
BPDU 型態
旗標
根橋接器辨識碼
根路徑費用
橋接器辨識碼
埠辨識碼
訊息年齡
訊息時限
問候時間
轉送時間

C-BPDU

通訊協定辨識碼
版本辨識碼
BPDU 型態

TCN-BPDU

Considerations of Network Routing

- The length (weight) of each link (arc) may be measured by *delay, cost, performance, … etc.*
- Each node collects the network status (path information) to calculate the shortest path to destinations
- Each node shall broadcast its link information → Flooding is necessary
- Each node decides the shortest path from its view point → short path algorithms

Routing Concept :

•Packet switching – forwarding function

•Path determination – routing protocol

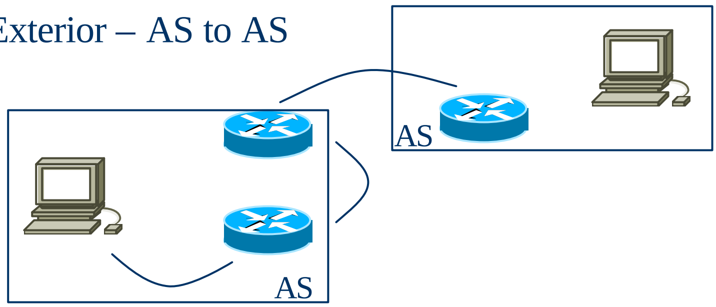
二〇〇二IPv6網際網路協定技術巡迴研討會

Content

- Routing Basic
- Routing Protocols Overview
  - Distance Vector Routing Protocols v.s. Link State Routing Protocols
  - Interior Routing (RIP, OSPF) v.s. Exterior Routing
- IPv6 Routing Concept
- RIPng for IPv6
- Multiprotocols Extension for BGP-4

Scope of IP Routing Protocols

- Host to router(GDP, IRDP)
- Interior – Router to Router
- Exterior – AS to AS



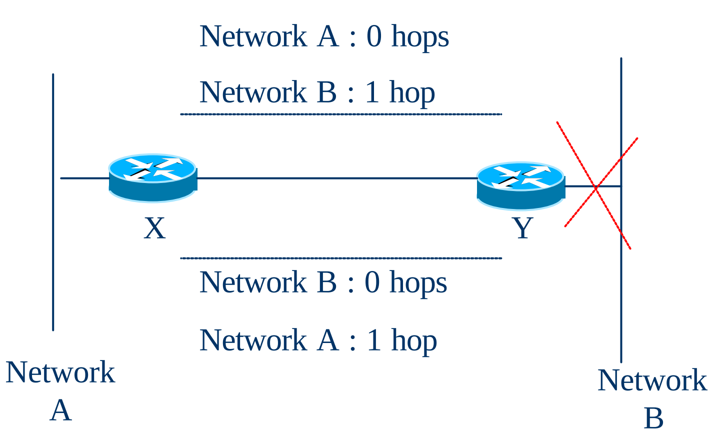
IP Networks and Routing

- Interior Routing Protocols
  - RIP (Routing Information Protocol)
  - OSPF (Open Shortest Path First)
  - Other Protocols
- Exterior Routing Protocols
  - EGP (Exterior Gateway Protocol)
  - BGP (Border Gateway Protocol)

RIP

- RIP is a simple protocol based on “Bellman Ford” protocol → distance vector
- Overview of distance vector routing
  - Start up
  - Link breaks
  - Bouncing effect
  - Counting to infinity
- Improvement schemes → split horizon; triggered updates
  - Decrease the time to converge

二〇〇二Ipv6網際網路協定技術巡迴研討會



- Route to B fails
- Metric counts up to 15 hops

Improvement schemes

.Split Horizon :

- Routers do not advertise a route back to the original source
- Solves count to infinity
- Solves short routing loops

. Holddown :

- When routes to a network are in holddown, packets to that destination are thrown away
- Prevents general routing loops
- Sets a minimum convergence time

二〇〇二IPv6網際網路協定技術巡迴研討會

Comparison of RIPv1(rfc-1058, 1988) and RIPv2 (rfc-2453, 1998)

- Subnet routing – support subnet mask information
- Authentication – simple password protection defined in 2453, and MD5 is defined in rfc-2082
- Next hop indication
- Multicasting – define IP class D address for information advertisement.

In a word ….

Distance Vector Routing Protocol

- Routers pass routing information gathered from neighbors
- Routers tell neighbors about all known routes by Periodic broadcasting routing table to all interfaces

二〇〇二Ipv6網際網路協定技術巡迴研討會

OSPF

- Link state routing v.s. distance vector routing
  - Link state routing protocols are based on the “distributed map” concept
- Changing information of the network is achieved by flooding protocol
- Main issue – to maintain a synchronized copy of the link state database in all nodes of the network → secure map updates
- Shortest path first → Dijkstra algorithm

Why is a link state protocol better?

- Fast, loopless convergency
- Support of precise/multiple metrics
  - The largest throughput; the lowest delay; the lowest cost; the best reliability; …
  - Metric per system v.s. metric per packet (OSPFv2)
- Support multiple paths to a destination
  - Traffic splitting
- Separate representation of external routes

二〇〇二IPv6網際網路協定技術巡迴研討會

Protocols within RIP and OSPF

- RIP
  - RIP packets are carried over UDP/IP with port 520
  - Packets are sent every 30 seconds, or faster when triggered updates
  - If a route is not refreshed within 180 seconds (6x30), the distance is set to infinity
  - Each entry (one route) of RIP message is encoded over 20 bytes long (reservation part is used in RIPv2 for authentication, and etc.)

Protocols within RIP and OSPF

- OSPF
  - OSPF runs on top of of the IP layer with protocol type 89
  - Composes of 3 subprotocols
    - Hello: for checking the operation of the link and elect the designated/backup routers
    - Exchange: master/slave operation for exchange the routing information in DB
    - Flooding: to maintain the synchronization of the two databases

Other Routing Protocols

- Intermediate System to Intermediate System (IS-IS)
  - Defined by ISO in 1980s for DECnet (especially in the Backbone) and many concepts of IS-IS were adopted by OSPF
- IGRP (Internet Group Management Protocol)
  - Similar to ICMP and is a proprietary protocol defined by Cisco
  - Distance vector family protocol
  - Composite metrics: delay (D), bandwidth (B), reliability (R), load (L)

	Topology Update	Hello	Prefix
RIP	Routes	No	No
IGRP	Routes	No	No
EIGRP	Routes	Yes	Yes
OSPF	Links	Yes	Yes
IS-IS	Links	Yes	Yes

Choosing of Routing Protocols

- Most experts recommend the link state protocols are better than distance vector protocols
- For small networks – RIPv2 is enough
- Chance of IS-IS has now vanished but the IS-IS working group still active
- The only real contender to OSPF is probably EIGRP
- Insist on a “standard” protocol is always a good reason – however, marketing,…

Exterior Routing Protocols

- Splitting the internet into autonomous systems (AS)
- Concept of AS
  - The minimum AS is composed of exactly one router directly connecting one LAN to Internet
  - An AS can "self-routing" within its local network
- Hierarchical (two) level routing
- Exterior routing is to exchange routing information among ASs

二〇〇二Ipv6網際網路協定技術巡迴研討會

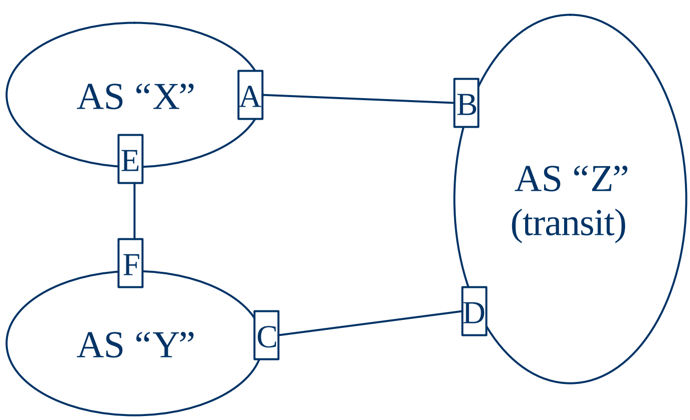
Exterior Gateways' Protocol (EGP)

- EGP is run over IP with protocol type number 8
- EGP Messages
  - Neighbor acquisition: to determine two adjacent gateways agree to become neighbors
  - Neighbor reachability: to monitor the links
  - Network reachability: exchange the reachability information

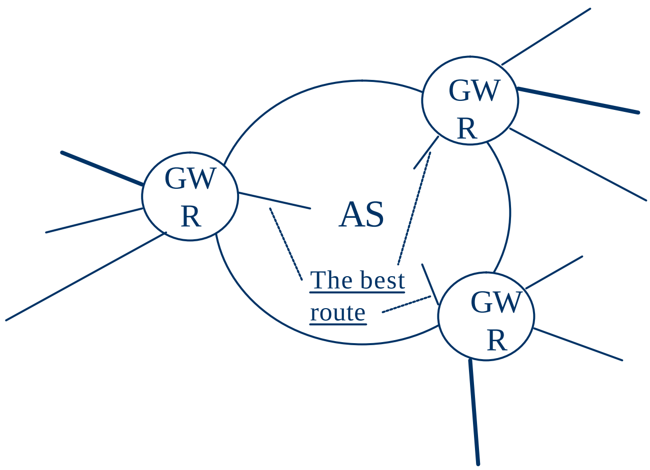
EGP Protocol

- NA → NR → NetR
- Neighbor acquisition is a two-way handshake procedure; while the partner may refuse
- Neighbor reachability uses “hello” and “I heard you” (I-H-U) packets to check the link
  - “Dual threshold” procedure is used to avoid oscillation
    - A reachable link is declared to be unreachable if fewer than i “I-H-U”s have been received for the last n “hello”s.
    - An unreachable link is declared to be reachable only if at least j “I-H-U”s have been received for the last m “hello”s.

EGP Protocol



EGP Protocol



二〇〇二Ipv6網際網路協定技術巡迴研討會

The Limits of EGP

- Handling of false information – rfc-1096, 1989, “EGP and Policy based Routing in the New NSFnet Backbone”
- Policy Routing
- Network Topology – from tree architecture to mesh architecture of the backbone network
- Message size and Fragmentation – EGP needs a robust fragmentation procedure to carry more routing information

→ BGP

BGP

- BGP is run over TCP/IP (with port number 179)
  - TCP provides a reliable data transmission link (with fair flow/congestion control), however,
    - Routing update packet to cure network congestion,…
    - Security issue (rfc-2385, 1998 – Protection of BGP Sessions via TCP MD5 Signature Option")
- Packet types of BGP
  - OPEN
  - UPDATE
  - NOTIFICATION
  - KEEPALIVE

BGP

- Initial exchange
  - Use OPEN packet to check the BGP version and the "hold time" (the number of seconds used by the "keep alive procedure")
  - Use UPDATE packet to exchange (list of) "withdrawn routes" and metrics information of each path
- Updates
  - Loop protection
  - Stable – the path shall not oscillate too rapidly between reachable and unreachable

BGP

- Keep alive
  - According to the “hold time” value, and the keep alive messages will not exchanged for zero hold time.
- Error Notifications
  - Message header error
  - OPEN message error
  - UPDATE message error
  - Hold time expired
  - Finite state machine error
  - Cease (terminate the association)

二〇〇二IPv6網際網路協定技術巡迴研討會

When to use BGP ?

-Dual or multi-homed

-Partial or full exterior routing is required downstream

-Anytime the AS path information is required to be propagated

If …

single homing, need not provide downstream routing …

Please use Default Route !

BGP :

-IBGP

-EBGP

Mesh problem :

-BGP Confederation

-BGP Route Reflectors

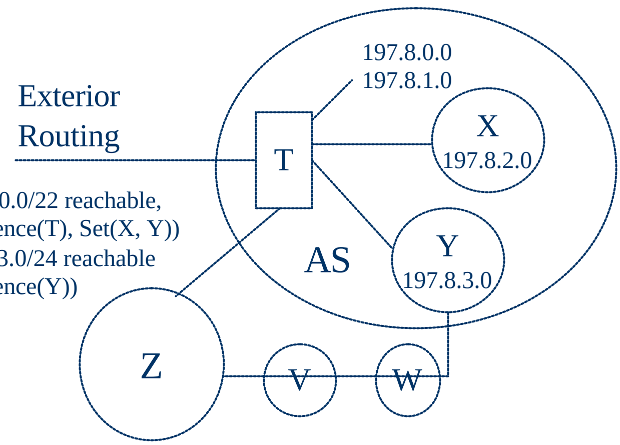
二〇〇二IPv6網際網路協定技術巡迴研討會

Policy in BGP

- The commercial networks will not necessarily agree to freely relay third parties’ traffic
- Route decision process
  - Evaluate the paths that have been learned from external routers
    - Content of AS path, local information (bandwidth to the next hop, local preference, …)

BGP

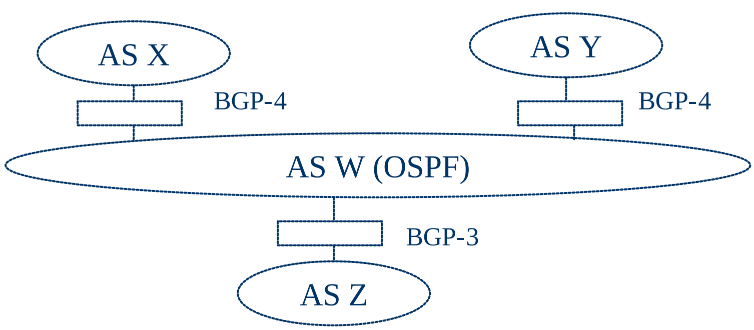
CIDR – “the longest match wins” rule?



Path (T): 197.8.0.0/22 reachable,
(Sequence(T), Set(X, Y))
Path(Y): 197.8.3.0/24 reachable
(Sequence(Y))

二〇〇二IPv6網際網路協定技術巡迴研討會

BGP



AS_Path: 197.8.0/22 reachable, Sequence (X)

AS_Path: 197.8.0.0 reachable, Sequence (X)
AS_Path: 197.8.1.0 reachable, Sequence (X)
AS_Path: 197.8.2.0 reachable, Sequence (X)
AS_Path: 197.8.3.0 reachable, Sequence (X)

Routing Table Explosion

```
203.X1.Y1.Z1    4  3462  663118  657135  2529751    0    0 1w1d        99694
203.X2.Y2.Z2    4  3462       0       0        0    0    0 never   Active
203.X3.Y3.Z3    4  3462 1106074   12856  2529754    0    0 1w1d       110841
==============================================================
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
* i6.1.0.0/16       203.X1.Y1.Z1            1     50      0        9680 6461 7170 1455 i
* i                 203.X2.Y2.Z2           60             0        9680 6461 7170 1455 i
* i                 203.X3.Y3.Z3           60             0        9680 6461 7170 1455 i
*>i                 203.X4.X5.X6           60             0        9680 6461 7170 1455 i
```


二〇〇二IPv6網際網路協定技術巡迴研討會

Content

- Routing Basic
- Routing Protocols Overview
  - Distance Vector Routing Protocols v.s. Link State Routing Protocols
  - Interior Routing v.s. Exterior Routing
- IPv6 Routing Concept
- RIPng for IPv6
- Multiprotocols Extension for BGP-4

IPv6 Routing Concept

- The loopback address “0:0:0:0:0:0:0:1” is a virtual interface shall not send outside of a single interface.
- Link-local address
  - Used on a single link for the purpose such as auto-address configuration, neighbor discovery, or when no routers are present.
  - Routers must not forward any packets with link-local source or destination address to other links.

IPv6 Routing Concept

Link-local

10	54	64
1111111010	0	interface ID

Site-local

10	38	16	64
1111111011	0	subbetID	interface ID

IPv6 Routing Concept

- Site-local address
  - Used for addressing inside of a site without the need for global prefix.
  - Routers must not forward ant packet with site-local source or destination address outside of the site.

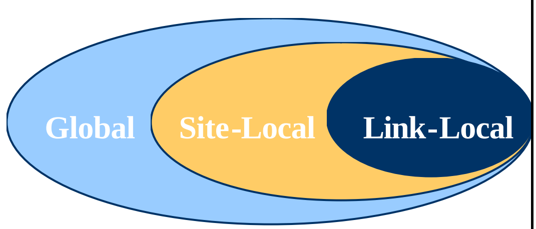
Addresses are assigned to interfaces

- **No change from IPv4 Model**

Interface 'expected' to have multiple addresses

Addresses have scope

- **Link Local**
- **Site Local**
- **Global**



Addresses have lifetime

- **Valid and Preferred lifetime**

- Unicast
  - Address of a single interface
  - Delivery to single interface
- Multicast
  - Address of a set of interfaces
  - Delivery to all interfaces in the set
- Anycast
  - Address of a set of interfaces
  - Delivery to a single interface in the set
- No more broadcast addresses

二OO二IPv6網際網路協定技術巡迴研討會

Address type	Binary prefix
IPv4-compatible	0000...0 (96 zero bits)
global unicast	001
link-local unicast	1111 1110 10
site-local unicast	1111 1110 11
multicast	1111 1111

- all other prefixes reserved (approx. 7/8ths of total)
- anycast addresses allocated from unicast prefixes

```
Serial0/0:0.300 is up, line protocol is up
 IPv6 is enabled, link-local address is FE80::50:7310:D0C0:F
 Description: TP:DLCI-300  TY:33FR-323052 DlCI-200
 Global unicast address(es):
  2001:238:0:24::1, subnet is 2001:238:0:24::/64
 Joined group address(es):
  FF02::1
  FF02::2
  FF02::1:FFC0:F
  FF02::1:FF00:1
  FF02::9
 MTU is 1500 bytes
 ICMP error messages limited to one every 500 milliseconds
 ND reachable time is 30000 milliseconds
 Hosts use stateless autoconfig for addresses.
```


二〇〇二IPv6網際網路協定技術巡迴研討會

```
FastEthernet1/0 is up, line protocol is up
 IPv6 is enabled, link-local address is FE80::250:73FF:FE10:D0D0
 Description: 2001:238::250:73FF:FE10:D0D0 to Taipei(4500) E1-
2001:238:0:24:260:2FFF:FEA4:3E3D
 Global unicast address(es):
  2001:238::250:73FF:FE10:D0D0, subnet is 2001:238::/64
 Joined group address(es):
  FF02::1
  FF02::2
  FF02::1:FF10:D0D0
  FF02::9
 MTU is 1500 bytes
 ICMP error messages limited to one every 500 milliseconds
 ND reachable time is 30000 milliseconds
 Hosts use stateless autoconfig for addresses.
```

```
Tunnel0 is up, line protocol is up
 IPv6 is enabled, link-local address is FE80::D2F2:61
 Global unicast address(es):
  3FFE:3600::1:21, subnet is 3FFE:3600::1:21/127
 Joined group address(es):
  FF02::1
  FF02::2
  FF02::1:FFC0:8
  FF02::1:FF01:21
  FF02::1:FFF2:61
 MTU is 1480 bytes
 ICMP error messages limited to one every 500 milliseconds
 ND reachable time is 30000 milliseconds
 Hosts use stateless autoconfig for addresses.
```


二〇〇二IPv6網際網路協定技術巡迴研討會

```
interface FastEthernet1/0
 description 2001:238::250:73FF:FE10:D0D0 to Taipei(4500)
E1-2001:238:0:24:260:2FFF:FEA4:3E3D
 ip address 10.77.6.171 255.255.240.0
 speed 10
 half-duplex
 ipv6 enable
 ipv6 address 2001:238::/64 eui-64
 ipv6 nd suppress-ra
 ipv6 rip HiNet enable
 no cdp enable
!
```

Content

- Routing Basic
- Routing Protocols Overview
  - Distance Vector Routing Protocols v.s. Link State Routing Protocols
  - Interior Routing v.s. Exterior Routing
- IPv6 Routing Concept
- RIPng for IPv6
- Multiprotocols Extension for BGP-4

二〇〇二IPv6網際網路協定技術巡迴研討會

RIPng for IPv6

- **Protocol Suite :** TCP/IP
  Type : Application layer , interior gateway, distance vector
  Port : 521 (UDP)
- **The destination prefix is the usual 128-bit , IPv6 address prefix stored as 16 octets in network byte order**
- **Packets are sent every 30 seconds , or faster when triggered updates**

RIPng packet format

Command(1)	Version(1)	Must be zero(2)
Route Table Entry 1 (20)		
……		
Route Table Entry N (20)		

RIPng packet format

Command : 8 bits

Command

1 **Request** ：A request for the responding system to send all or part of its routing table

2 **Response**：A message containing all or part of the sender's routing table.This message may be sent in response to a request, or it may be an unsolicited routing update generated by the sender

Version : 1 ; 8 bits

Response Messages

A Response can be received for one of several different reasons:

- response to specific query
- regular update (unsolicited response)
- triggered update caused by a route change

二〇〇二IPv6網際網路協定技術巡迴研討會

Route Table Entry(RTE) format

IPv6 prefix (16)		
Route tag (2)	Prefix len (1)	Metric (1)

IPv6 prefix : 16 bytes

Route tag **:** 16 bits

Prefix len **:** 8 bits

Metric **:** 8 bits

Route Table Entry(RTE)

The determination of the number of RTEs which may be put into a given message , The formula is :

$$\#RTEs = INT\left\{\frac{MTU - sizeof(IPv6\_hdrs) - UDP\_hdrlen - RIPng\_hdrlen}{RTE\_size}\right\}$$

The next hop RTE

IPv6 next hop address (16)		
Must be)zero (2)	Must be zero(1)	0xFF

Timers and Split Horizon

- Two timers are associated with each route
  - Timeout timer
  - Garbage-collection time
- Split horizon
  - Split horizon is an algorithm for avoiding problems caused by including routes in updates sent to the nodes from which they were learned.


```
 sh ipv6 route rip :
IPv6 Routing Table - 328 entries
Codes: C - Connected, L - Local, S - Static, R - RIP, B - BGP
Timers: Uptime/Expires

R 2001:238:E80::4/127 [120/2]
  via FE80::260:5CFF:FE5E:8782, FastEthernet1/0, 08:51:20/00:02:55
R 2001:238:E80::6/127 [120/2]
  via FE80::260:5CFF:FE5E:8782, FastEthernet1/0, 08:51:20/00:02:55
R 2001:238:F00:13::/64 [120/2]
  via FE80::260:5CFF:FE5E:8782, FastEthernet1/0, 08:51:20/00:02:55
R 2001:238:F00:16::/64 [120/2]
  via FE80::260:5CFF:FE5E:8782, FastEthernet1/0, 08:51:20/00:02:55
R ::/0 [120/2]
```

```
Jul 19 04:26:08: RIPng: Sending multicast update on Serial0/0:0.300 for HiNet
Jul 19 04:26:08:        src=FE80::50:7310:D0C0:F
Jul 19 04:26:08:        dst=FF02::9 (Serial0/0:0.300)
Jul 19 04:26:08:        sport=521, dport=521, length=132
Jul 19 04:26:08:        command=2, version=1, mbz=0, #rte=6
Jul 19 04:26:08:        tag=0, metric=1, prefix=2001:238::/64
Jul 19 04:26:08:        tag=0, metric=2, prefix=2001:238:E80::4/127
Jul 19 04:26:08:        tag=0, metric=2, prefix=2001:238:E80::6/127
Jul 19 04:26:08:        tag=0, metric=2, prefix=2001:238:F00:13::/64
Jul 19 04:26:08:        tag=0, metric=2, prefix=2001:238:F00:16::/64
Jul 19 04:26:08:        tag=0, metric=2, prefix=::/0
Jul 19 04:26:08: RIPng: Sending multicast update on FastEthernet1/0 for HiNet
Jul 19 04:26:08:        src=FE80::250:73FF:FE10:D0D0
Jul 19 04:26:08:        dst=FF02::9 (FastEthernet1/0)
Jul 19 04:26:08:        sport=521, dport=521, length=32
Jul 19 04:26:08:        command=2, version=1, mbz=0, #rte=1
Jul 19 04:26:08:        tag=0, metric=1, prefix=2001:238:0:24::/64
Jul 19 04:26:15: RIPng: response received from FE80::260:70FF:FE0C:B008 on
FastEthernet1/0 for HiNet
Jul 19 04:26:15:        src=FE80::260:70FF:FE0C:B008 (FastEthernet1/0)
Jul 19 04:26:15:        dst=FF02::9
Jul 19 04:26:15:        sport=521, dport=521, length=52
Jul 19 04:26:15:        command=2, version=1, mbz=0, #rte=2
Jul 19 04:26:15:        tag=0, metric=1, prefix=2001:238::/64
Jul 19 04:26:15:        tag=0, metric=1, prefix=::/0
Jul 19 04:26:19: RIPng: response received from FE80::260:5CFF:FE5E:8782 on
FastEthernet1/0 for HiNet
```


二〇〇二IPv6網際網路協定技術巡迴研討會

Content

- Routing Basic
- Routing Protocols Overview
  - Distance Vector Routing Protocols v.s. Link State Routing Protocols
  - Interior Routing v.s. Exterior Routing
- RIPng for IPv6
- Multiprotocols Extension for BGP-4

Multiprotocol Extensions for BGP-4

- It can carry routing information for multiple Network Layer protocol

  (e.g., IPv6 , IPX , etc…)
- The extensions are backward compatible
  - a router that supports the extensions can interoperate with a router that doesn’t support the extensions

二〇〇二IPv6網際網路協定技術巡迴研討會

Multiprotocol Extensions for BGP-4

- To provide backward compatibility,as well as to simplify introduction of the multiprotocol capabilities into BGP-4; two new attributes are used:
  - Multiprotocol Reachable NLRI (MP_REACH_NLRI)
  - Multiprotocol Unreachable NLRI (MP_UNREACH_NLRI)

MP_REACH_NLRI (Type code 14)

Address Family Identifier		Sub-AFI	NEXTHOP len
NEXTHOP Address		...	
Num SNPAs	1st SNPA len	1st SNPA	...
Remaining SNPAs		...	
NLRI length		NLRI information	...
	...		

二〇〇二IPv6網際網路協定技術巡迴研討會

MP_REACH_NLRI (Type code 14)

- This is an optional non-transitive attribute that can be used for the following purposes:
  - o to advertise a feasible route to a peer
  - o to allow a given router to report some or all of the Subnetwork Points of Attachment (SNPAs)

    that exit within the local system

MP_REACH_NLRI (Type code 14)

Address Family Identifier (2 octets)
Subsequent Address Family Identifier (1 octet)
Length of Next Hop Network Address (1 octet)
Network Address of Next Hop (variable)
Number of SNPAs (1 octet)
Length of first SNPA (1 octet)
First SNPA (variable)
………
Length of Last SNPA (1 octet)
Last SNPA (variable)
Network Layer Reachability Information (variable)

MP_UNREACH_NLRI (Type Code 15)

Address Family Identifier (2 octets)
Subsequent Address Family Identifier (1 octet)
Withdrawn Routes (variable)

Network Layer Reachability information(NLRI) encoding

- The Network Layer Reachability information is encoded as one or more 2-tuples of the form <length, prefix>,whose fields are described below:

Length (1 octet)
Prefix (variable)

二OO二Ipv6網際網路協定技術巡迴研討會

Network Layer Reachability information(NLRI) encoding

- The use and the meaning of these fields are as follows:
  - Length:
    The Length in bits of the address prefix. A length of zero indicates a prefix that matches all (as specified by the address family) address (with prefix , itself , of zero octets)
  - Prefix: The Prefix field contains address prefixes followed by enough trailing bits to make the end of the fall on an octet boundary

Subsequent Address Family Identifier

1 – Network Layer Reachability Information used for unicast forwarding

2 – Network Layer Reachability Information used for multicast forwarding

3 – Network Layer Reachability Information used for both unicast and multicast forwarding

BGP4+ Overview

- Added IPv6 address-family
- Added IPv6 transport
- Runs within the same process - only one AS supported
- All generic BGP functionality works as for IPv4
- Added functionality to route-maps and prefix-lists

```
router bgp 17419
 no synchronization
 bgp router-id A.Z.Y.X
 no bgp default ipv4-unicast
 bgp cluster-id 101058054
 bgp log-neighbor-changes
 neighbor 2001:238::260:5CFF:FE5E:8782 remote-as 17419
 neighbor 2001:238::260:5CFF:FE5E:8782 description peer-with-
HiNet_IPv6_AR01
 no auto-summary
 !
 address-family ipv6
 neighbor 2001:238::260:5CFF:FE5E:8782 activate
 neighbor 2001:238::260:5CFF:FE5E:8782 next-hop-self
 neighbor 2001:238::260:5CFF:FE5E:8782 send-community
 neighbor 2001:238::260:5CFF:FE5E:8782 soft-reconfiguration inbound
 network 2001:238::/35
 exit-address-family
!
ipv6 route 2001:238::/35 Null0
ipv6 route 2001:238:700::/64 2001:238:0:27::2
ipv6 route 2001:238:0:27::/64 Tunnel1
```



```
2001:2B8:FFFE::28
             4 17832    1804     531     3141    0    0 08:47:36      260
2001:2B8:FFFE::48
             4 17832    1864     531     3141    0    0 08:47:42      260
3FFE:3600::1:20 4 17715    1807     531     3141    0    0 08:47:41      262

  Network          Next Hop            Metric LocPrf Weight Path
*>i2001:200::/35    2001:238::260:70FF:FE0C:B008
                                        100      0 3425 2500 i
*               2001:2B8:FFFE::28
                                         80      0 17832 17832 17832 17579 3425 2500 i
*               3FFE:3600::1:20              80      0 17715 17715 17715 4725 2500 i
*               2001:2B8:FFFE::48
                                         80      0 17832 17832 17832 17579 3425 2500 i
*>i2001:208::/35    2001:238::260:70FF:FE0C:B008
                                        100      0 3425 7610 i
*               3FFE:3600::1:20              80      0 17715 17715 17715 6939 3425 7610 i
*               2001:2B8:FFFE::48
                                         80      0 17832 17832 17832 7610 i
*  2001:218::/35    2001:2B8:FFFE::28
                                         80      0 17832 17832 17832 3748 6939 2914 i
*               3FFE:3600::1:20              80      0 17715 17715 17715 2914 i
*               2001:2B8:FFFE::48
                                         80      0 17832 17832 17832 3748 6939 2914 i
*>               2001:238:800:3::1
                                     0           0 2914 i
*>i2001:220::/35    2001:238::260:70FF:FE0C:B008
```

- BGP4+ References

  RFC2858 Multiprotocol extension to BGP
  RFC2545 BGP MP for IPv6
  RFC2842 Capability negotiation

- RIPng RFC2080

二OO二IPv6網際網路協定技術巡迴研討會

Question ?

Thank you !

Any information to HiNet IPv6 ： www.ipv6.hinet.net ！