IPv6 Functionality

Jeff Doyle
IPv6 Solutions Manager
jeff@juniper.net

Copyright © 2003 Juniper Networks, Inc.

Agenda

- ICMPv6
- Neighbor discovery
- Autoconfiguration

Agenda

- **ICMPv6**
- **Neighbor discovery**
- **Autoconfiguration**

Juniper NETWORKS

ICMPv6

- Many of the same functions as ICMPv4
  - ICMPv4 Protocol Number = 1
  - ICMPv6 Next Header Number = 58
- Adds new messages and functions
  - Neighbor discovery
  - Stateless autoconfiguration
  - Mobile IPv6

Copyright © 2003 Juniper Networks, Inc.

ICMPv6 Message Types

- Defined in RFC 2463

Type	Message
1	Destination Unreachable
2	Packet Too Big
3	Time Exceeded
4	Parameter Problem
128	Echo Request
129	Echo Reply

Copyright © 2003 Juniper Networks, Inc.

ICMPv6 New Message Types

- Defined in RFC 2461
- Used for Neighbor Discovery protocol

Type	Message
133	Router Solicitation (RA)
134	Router Advertisement (RA)
135	Neighbor Solicitation (NS)
136	Neighbor Advertisement (NA)
137	Redirect

Copyright © 2003 Juniper Networks, Inc.

Agenda

- ICMPv6
- **Neighbor discovery**
- Autoconfiguration

Juniper NETWORKS

IPv6 Neighbor Discovery

- RFC 2461
- Neighbor can be router or host
- Performs several functions
  - Link-layer address resolution
  - Router discovery
  - Local prefix discovery
  - Address autoconfiguration
  - Parameter discovery
  - Next-hop determination
  - Tracks neighbor and router reachability
  - Duplicate address detection
  - Redirects

Copyright © 2003 Juniper Networks, Inc.

Comparison to IPv4 Functions

- **Similar IPv4 functions**
  - **ARP**
  - **ICMP Router Discovery**
  - **ICMP Redirect**
- **IPv4 has no agreed-upon mechanism for neighbor unreachability detection**
  - **Detects failing routers and links**
  - **Detects nodes that change their link-layer address**
  - **Unlike ARP, detects half-link failures**

Copyright © 2003 Juniper Networks, Inc.

Improvements Over IPv4

- Router discovery part of base protocol
  - Hosts do not need to "snoop" routing protocols
- RAs and redirects carry link-layer addresses
  - No additional packet exchange needed
- RAs carry link prefixes
  - No separate mechanism to configure "netmasks"
  - Enables address autoconfiguration
  - Multiple prefixes can be associated with same link
- RAs can advertise link MTUs
  - Ensures all nodes on link use same MTU value
- Immune to reception of off-link ND messages
  - Hop limit always set to 255
  - IPv4 ICMP Redirects and Router Discovery messages can be sent from off-link

Copyright © 2003 Juniper Networks, Inc.

Router Discovery

- Router Advertisements sent periodically
  - Interval randomized to prevent synchronization
  - Configurable range defined by:
    - MaxRtrAdvInterval (default 600 seconds)
    - MinRtrAdvInterval (default 200 seconds)
  - RAs sent to All-Nodes multicast address (ff01::1)
- RAs sent in response to Router Solicitations
  - RS sent to All-Routers multicast address (ff01::2)
  - RA unicast to soliciting node

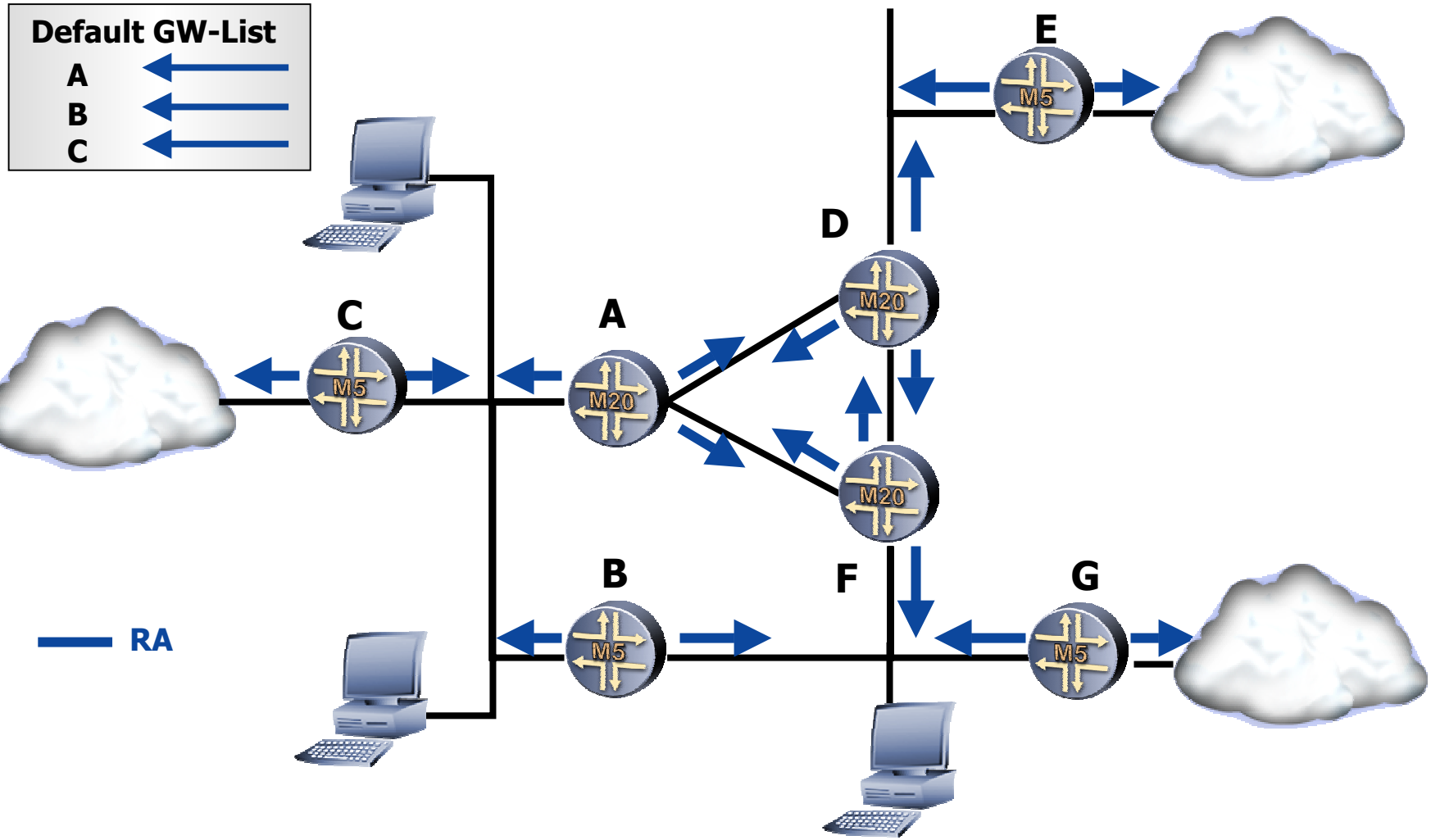
Copyright © 2003 Juniper Networks, Inc.

Router Advertisement Information

- Current hop limit
  - Value to be used by outgoing IP packets
- Address configuration flags
  - "M" and "O" bits
- Router lifetime
  - Lifetime for default router
- Reachable time/ Retrans timer
  - Used for router unreachability detection
- Source link-layer address (optional)
  - Can be omitted for in-bound load balancing
- MTU (optional)
  - If AdvLinkMTU is configured
- Prefix information (optional)
  - Used for address autoconfiguration

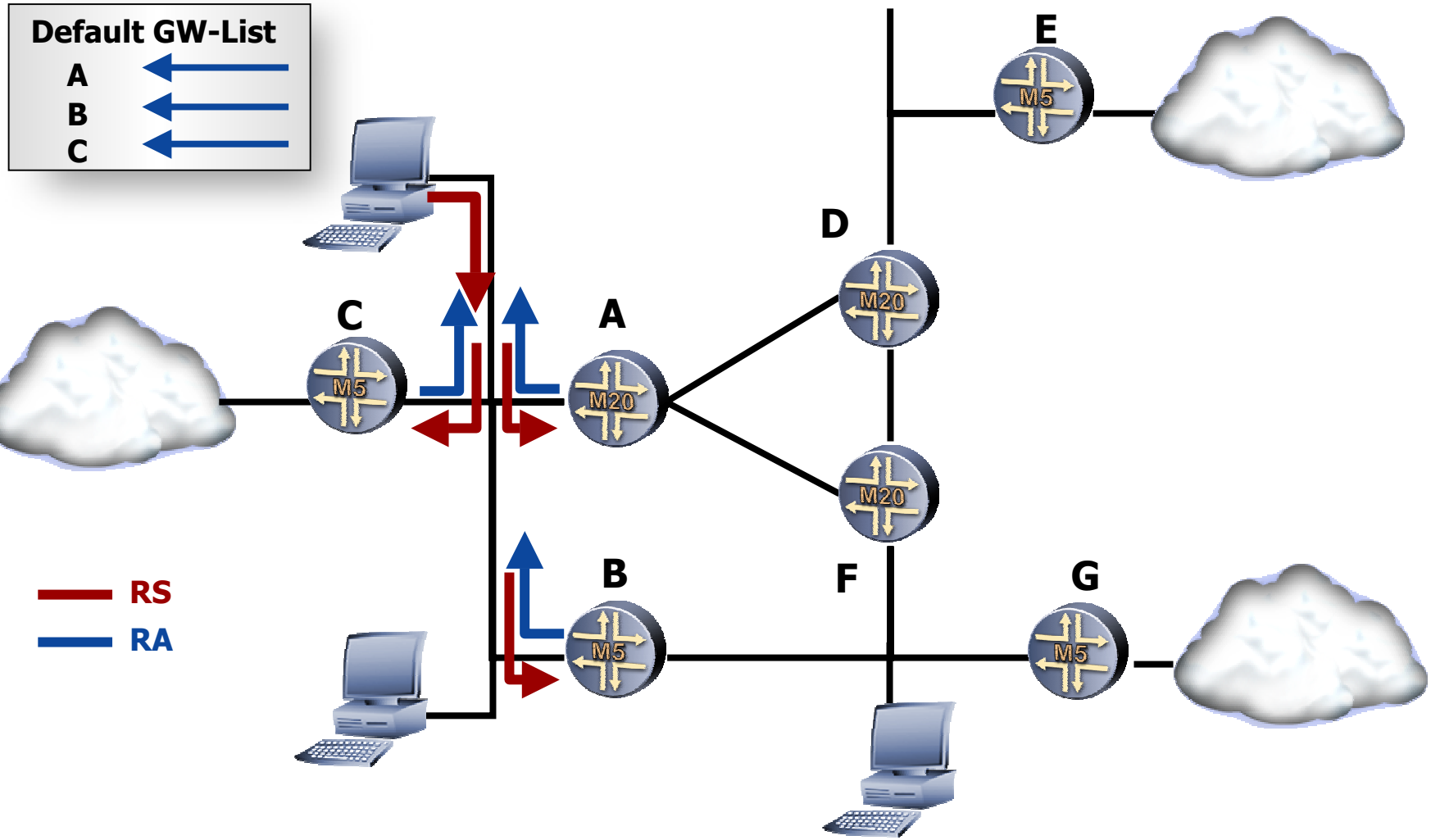
Copyright © 2003 Juniper Networks, Inc.

Unsolicited Router Advertisement



Copyright © 2003 Juniper Networks, Inc.

Solicited Router Advertisement



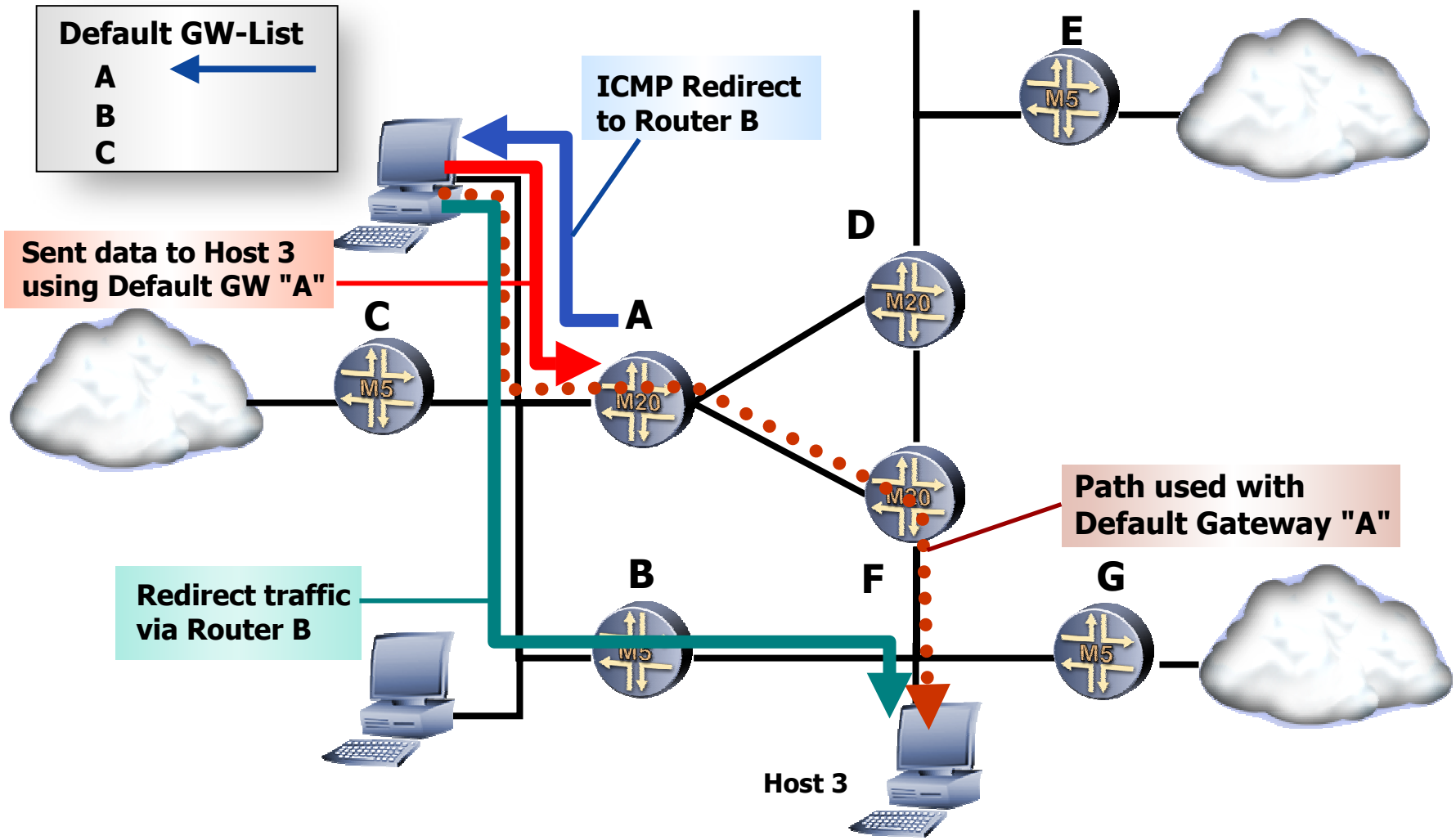
Copyright © 2003 Juniper Networks, Inc.

Choosing a Default Gateway

- Implementations may randomly select a default router
- Implementations may cycle through default list round-robin
- What happens when default router is the wrong router?

Copyright © 2003 Juniper Networks, Inc.

Redirect



Copyright © 2003 Juniper Networks, Inc.

Neighbor Cache

```
C:\Documents and Settings\Jeff Doyle>ipv6 nc
5: fe80::202:2dff:fe25:5e4c 00-02-2d-25-5e-4c permanent
4: fe80::260:83ff:fe7b:2df3 00-60-83-7b-2d-f3 stale (router)
4: fe80::210:a4ff:fea0:bc97 00-10-a4-a0-bc-97 permanent
4: 3ffe:3700:1100:1:210:a4ff:fea0:bc97 00-10-a4-a0-bc-97 permanent
4: 3ffe:3700:1100:1:d9e6:b9d:14c6:45ee 00-10-a4-a0-bc-97 permanent
4: 2001:468:1100:1:210:a4ff:fea0:bc97 00-10-a4-a0-bc-97 permanent
4: 2001:468:1100:1:d9e6:b9d:14c6:45ee 00-10-a4-a0-bc-97 permanent
3: 2002:c058:6301::c058:6301 192.88.99.1       permanent
3: 2002:836b:213c::836b:213c 131.107.33.60     permanent
3: 2002:4172:a85b::4172:a85b 127.0.0.1         permanent
3: 2002:836b:213c:1:e0:8f08:f020:6 131.107.33.60     permanent
3:  2001:708:0:1::624                  incomplete
2:   ::65.114.168.91 127.0.0.1         permanent
2: fe80::5efe:65.114.168.91 127.0.0.1         permanent
2: fe80::5efe:169.254.113.126 127.0.0.1       permanent
1:          fe80::1                    permanent
1:             ::1                  permanent
```


Copyright © 2003 Juniper Networks, Inc.

Neighbor Address Resolution

- Equivalent function to IPv4 ARP
  - But multicast instead of broadcast
- Check Neighbor Cache for address
- If no address, create an Incomplete entry for target address
- Send Neighbor Solicitation to Solicited-Node Multicast address
- Target node sends Neighbor Advertisement with link-layer address
- Soliciting node changes Incomplete entry to Reachable

Copyright © 2003 Juniper Networks, Inc.

Solicited–Node Multicast Address

- All multicast-capable interfaces required to listen
- Formed by appending low-order 24 bits of target IPv6 address to prefix ff02:0:0:0:0:1:ff00::/104
- Addresses differing only in high-order bits will map to same solicited-node multicast
  - Useful when multiple addresses assigned to interface
  - Reduces number of multicast addresses a node must listen for

Example:

Interface Address #1 = 3ffe:3700:1100:1:200:bff:fec6:45ee

Interface Address #2 = 2001:468:1100:1:200:bff:fec6:45ee

Solicited-Node Multicast Address = ff02::1:ffc6:45ee

Copyright © 2003 Juniper Networks, Inc.

Next-Hop Discovery

- Check Neighbor Cache for existing next-hop entry for particular destination
- Check whether destination is on- or off- link
- On-link: Sent directly to destination
- Off-link: Sent to default router
- Identify link-layer address of next-hop

Copyright © 2003 Juniper Networks, Inc.

Neighbor Unreachability Detection

- Neighbor cache stores information about neighbors
  - IP address
  - Link-layer address
  - Reachability state
- Neighbor reachability states
  - INCOMPLETE
  - REACHABLE
  - STALE
  - DELAY
  - PROBE

Copyright © 2003 Juniper Networks, Inc.

Agenda

- ICMPv6
- Neighbor discovery
- **Autoconfiguration**

Address Autoconfiguration

- Stateless autoconfiguration
  - Requires only a router
  - Key advantage for applications such as Mobile IP
- Stateful autoconfiguration
  - When more control is desired
  - DHCPv6
- Stateless and stateful can be combined
  - "M" and "O" flags in RA
    - M flag: Stateless Address Autoconfiguration Y/N
    - O flag: Stateless Autoconfigure Other Parameters Y/N

Copyright © 2003 Juniper Networks, Inc.

Stateless Autoconfiguration

- Interface ID automatically derived
  - IEEE addresses use MAC-to-EUI-64 conversion
  - Other addresses use other means, such as random number generation
- Host creates a link-local address
- Host performs duplicate address check
- Host sends RS to the all-routers multicast address (ff01::2)
- Router unicasts RA with prefix information
- Host adds prefix to Interface ID to form global unicast address

Copyright © 2003 Juniper Networks, Inc.

MAC-to-EUI-64 Conversion

1. First three octets of MAC becomes Company-ID
2. Last three octets of MAC becomes Node-ID
3. 0xfffe inserted between Company-ID and Node-ID
4. Universal/Local-Bit (U/L-bit) is set to 1 for global scope

Copyright © 2003 Juniper Networks, Inc.

MAC-to-EUI-64 Conversion Example

- **MAC Address:** `0000:0b0a:2d51`
  - **In binary:**

    00000000 00000000 00001011 00001010 00101101 01010001

    U/L Bit

    Company-ID

    Individual Node-ID

- **Insert fffe between Company-ID and Node-ID**

  00000000 00000000 00001011 11111111 11111110 00001010 00101101 01010001

  = fffe

- **Set U/L bit to 1**

  00000010 00000000 00001011 11111111 11111110 00001010 00101101 01010001

  U/L Bit

- **Resulting EUI-64 Address:** `0200:0bff:fe0a:2d51`

Copyright © 2003 Juniper Networks, Inc.

Using the EUI-64 Interface ID

EUI-64 Address:

200:bff:fe0a:2d51

Link-Local Address:

fe80::200:bff:fe0a:2d51

Global Unicast Address:

3ffe:3700:1100:1:200:bff:fe0a:2d51

Copyright © 2003 Juniper Networks, Inc.

Solicited-Node Multicast Revisited

Interface Address #1 = 3ffe:3700:1100:1:200:bff:fec6:45ee

Interface Address #2 = 2001:468:1100:1:200:bff:fec6:45ee

Solicited-Node Multicast Address = ff02::1:ffc6:45ee

- **Last 24 bits are not changed by autoconfiguration or by solicited node multicast**

Copyright © 2003 Juniper Networks, Inc.

Address Autoconfiguration:
A Security Problem?

- Interface ID remains constant for a host
  - Even when prefix information changes
  - Unlike IPv4, where entire address changes
- Mobile users can be tracked
- Usage from always-on addresses can be tracked
- This is a concern for some, not for others
- Two solutions:
  - Always use stateful autoconfiguration (DHCPv6)
  - Use privacy addresses for outgoing connections

Copyright © 2003 Juniper Networks, Inc.

Privacy Addresses

- RFC 3041
- A new Interface ID is randomly generated
  - Whenever a new public address is autoconfigured
  - Periodically (period is configurable)
- Both autoconfigured public and private addresses are used
  - Public for incoming connections (DNS registered)
  - Private for outgoing connections

Copyright © 2003 Juniper Networks, Inc.

Stateful Autoconfiguration: DHCPv6

- Currently in Internet-draft
- Many changes from DHCPv4:
  - Configuration of dynamic updates to DNS
  - Address deprecation for dynamic renumbering
  - Authentication
  - Clients can ask for multiple IP addresses
  - Addresses can be reclaimed
  - Integration between stateful and stateless autoconfiguration
- Uses multicasting
  - All_DHCP_Agents: ff02::1:2
  - All_DHCP_Servers: ff05::1:3

Copyright © 2003 Juniper Networks, Inc.

Duplicate Address Detection

- Must be performed by all nodes
- Performed with both stateless and stateful autoconfiguration
- Performed before assigning a unicast address to an interface
- Performed on interface initialization
- Not performed for anycast addresses
- Link must be multicast capable
- New address is called "tentative" as long as duplicate address detection takes place

Copyright © 2003 Juniper Networks, Inc.

Duplicate Address Detection

1. Interface joins all-nodes multicast group
2. Interface joins solicited-node multicast group
3. Node sends one NS with
   - Target address = tentative IP address
   - Source address = unspecified (::)
   - Destination address = tentative solicited-node address

Copyright © 2003 Juniper Networks, Inc.

Duplicate Address Detection

- If address already exists, the particular node sends a NA with
  - Target address = tentative IP address
  - Destination address = tentative solicited-node address
- If soliciting node receives NA with target address set to the tentative IP address, the address must be duplicate

Copyright © 2003 Juniper Networks, Inc.

Configuration Example:
Router Discovery

```
[edit]
lab@Juniper5# show interfaces fe-2/1/0
unit 0 {
    family inet6 {
        address 2001:468:1100:1::1/64;
        address 3ffe:3700:1100:1::1/64;
    }
}

[edit]
lab@Juniper5# show protocols router-advertisement
interface fe-2/1/0.0 {
    other-stateful-configuration;
    prefix 3ffe:3700:1100:1::/64;
    prefix 2001:468:1100:1::/64;
}
```


Copyright © 2003 Juniper Networks, Inc.

Configuration Example: Windows XP Host

```
C:\Documents and Settings\Jeff Doyle>ipv6 if 4
Interface 4: Ethernet: Local Area Connection 2
  uses Neighbor Discovery
  uses Router Discovery
  link-layer address: 00-10-a4-a0-bc-97
    preferred global 2001:468:1100:1:d9e6:b9d:14c6:45ee, life 6d21h14m26s/21h12m4s (anonymous)
    preferred global 2001:468:1100:1:210:a4ff:fea0:bc97, life 29d23h59m25s/6d23h59m25s (public)
    preferred global 3ffe:3700:1100:1:d9e6:b9d:14c6:45ee, life 6d21h14m26s/21h12m4s (anonymous)
    preferred global 3ffe:3700:1100:1:210:a4ff:fea0:bc97, life 29d23h59m25s/6d23h59m25s (public)
    preferred link-local fe80::210:a4ff:fea0:bc97, life infinite
    multicast interface-local ff01::1, 1 refs, not reportable
    multicast link-local ff02::1, 1 refs, not reportable
    multicast link-local ff02::1:ffa0:bc97, 3 refs, last reporter
    multicast link-local ff02::1:ffc6:45ee, 2 refs, last reporter
  link MTU 1500 (true link MTU 1500)
  current hop limit 64
  reachable time 22000ms (base 30000ms)
  retransmission interval 1000ms
  DAD transmits 1
```


Copyright © 2003 Juniper Networks, Inc.