Building Large IPv6 Access Networks

Ivan Pepelnjak (@ioshints, ip@ioshints.info)
NIL Data Communications

Revision history

2012-01-17	First draft
2012-01-19	Multi-access section: DHCPv6 and DHCPv6 Relay configurations Overview of first-hop security mechanism IPoE section: more details on DHCPv6 IA_PD and relaying Bulk Lease
2012-01-20	Completed PPPoX section New 6rd section
2012-01-23	6rd configurations Final fixes

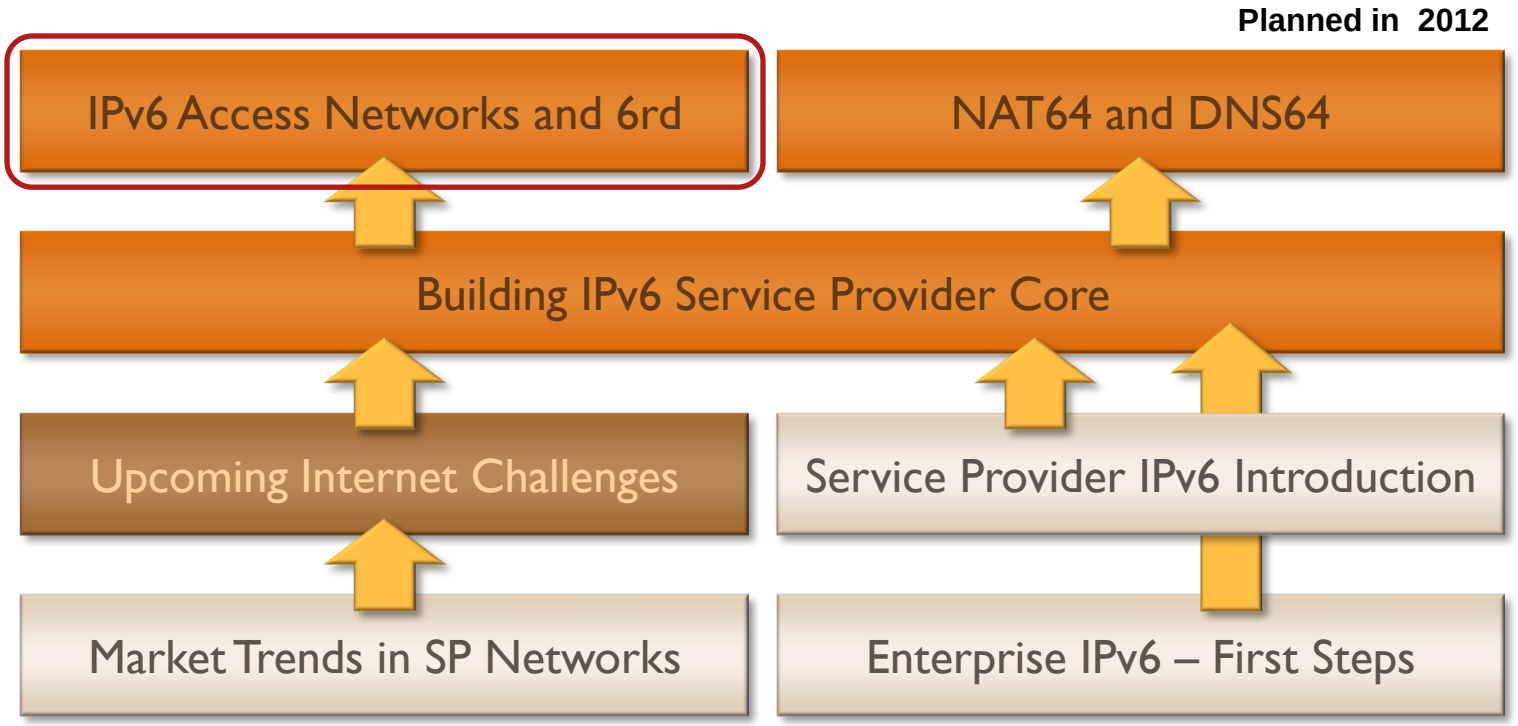
Who is Ivan Pepelnjak ... in 30 Seconds

- Networking engineer since 1985 (DECnet, Netware, X.25, OSI, IP ...)
- Technical director, later Chief Technology Advisor @ NIL Data Communications
- Started the first commercial ISP in Slovenia (1992)
- Developed BGP, OSPF, IS-IS, EIGRP, MPLS courses for Cisco Europe
- Architect of Cisco's Service Provider (later CCIP) curriculum
- Consultant, blogger (blog.ioshints.info), book author

Focus:

- Core routing/MPLS, IPv6, VPN, Data centers, Virtualization
- Rock climbing, mountain biking ;)

The Bigger Picture: IPv6 Webinars



Availability

- Live sessions
- Recordings of individual webinars
- Yearly subscription

Other options

- Customized webinars
- ExpertExpress
- On-site workshops

More information @ http://www.ipspace.net/Webinars

Agenda

- Design principles review
- Access network configurations

Business customers

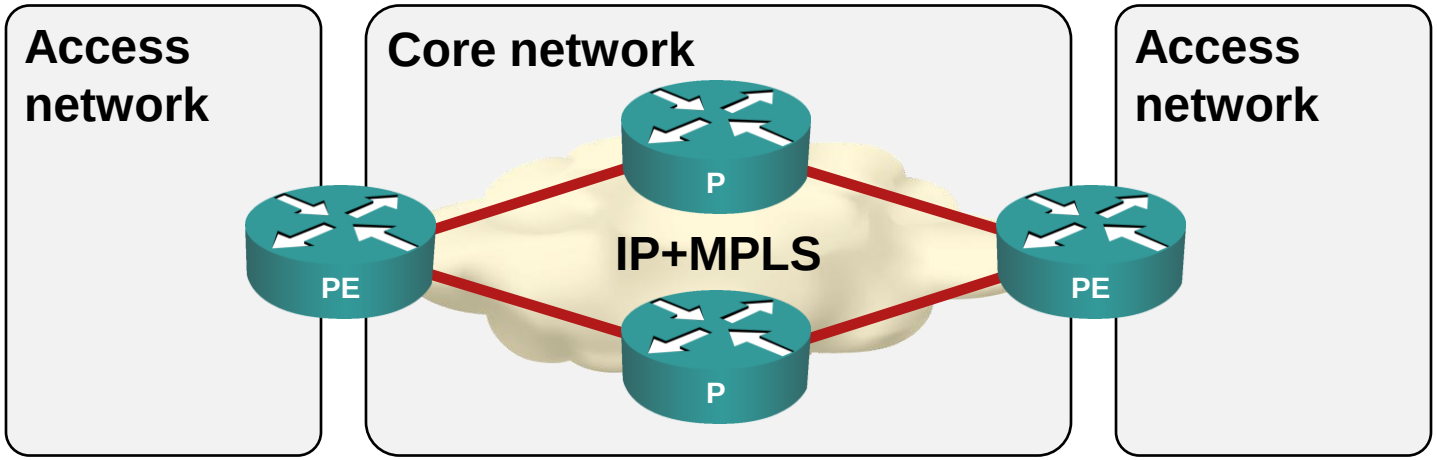
- Static routes
- EBGP

Residential customers

- Multi-access networks (WiFi)
- Carrier Ethernet (IPoE)
- DSL (PPPoX)
- 6rd

Design principles

Network Design Principles



Access network

- Customer connections
- Carrier Ethernet, DSL, Cable

Core network

- High-speed transport
- GE, 10GE, SDH, WDM

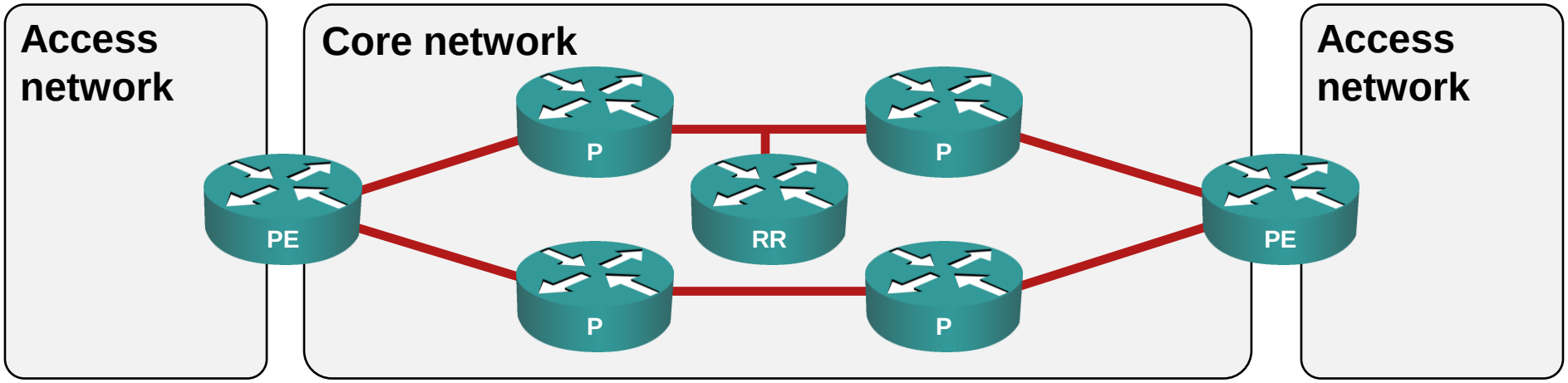
PE routers

- Connect access networks to the core
- Per-customer services
- QoS enforcement
- VPN functionality

P routers

- High-speed packet forwarding

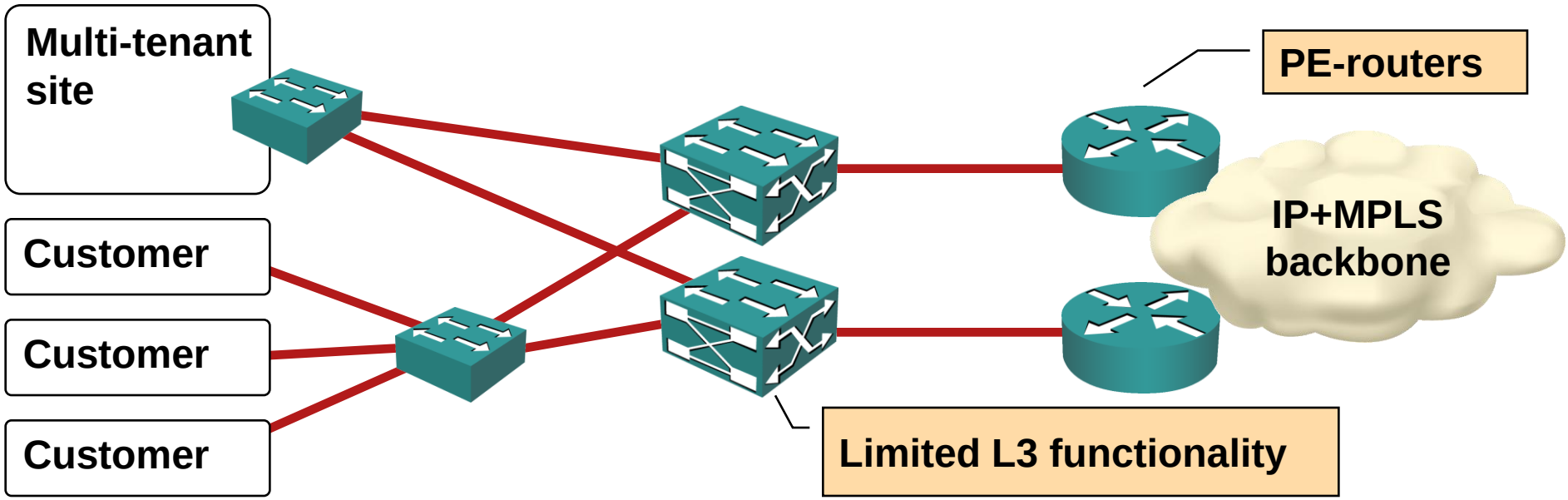
Core Network Design Principles



Goal: Highly stable core

- BGP on PE-routers
- Customer routes are carried in BGP
- Dedicated core IGP
- Only BGP next-hops and core links are in core IGP
- Route reflectors for BGP scalability

Access Network Design Principles



Internet access

- L2 switching or backhaul to concentration switches
- L3 switching on concentration switches
- IGP between L3 switches and PE routers

VPN access

- VLAN backhaul to the PE-routers
- Optional: Multiple sites in the same VLAN
- IGP between PE-routers and customer sites

Access Network Design and Configurations

Agenda

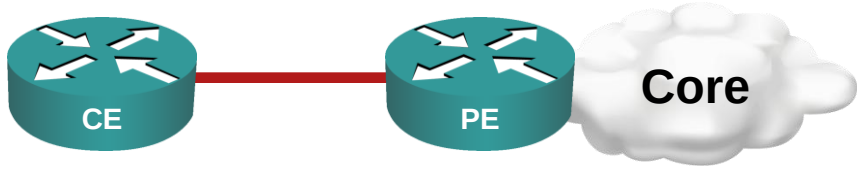
Business customers

- **Static routes**
- EBGP

Residential customers

- Multi-access networks (WiFi)
- Carrier Ethernet (IPoE)
- DSL (PPPoX)
- 6rd

Static Routing – Recommendations



Addressing:

- /48 per customer site
- Numbered PE-CE links (/64 per link)
- PE- or POP-aggregated prefixes

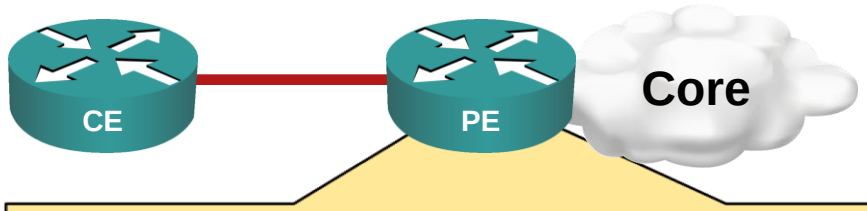
Routing:

- Redistribute static routes into BGP
- Set no-export on redistributed PA prefixes
- Aggregate in PE-router or at POP exit routers

Configuration recommendations:

- Use tags on static routes (simplifies redistribution route maps)
- RPF checks

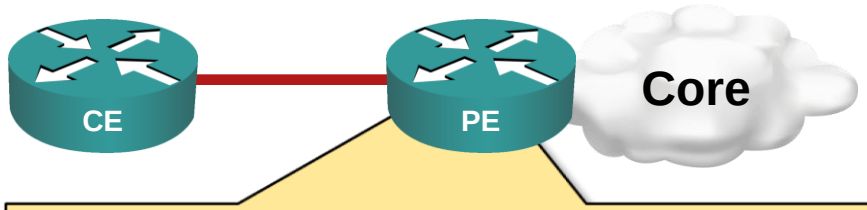
Static Routing – PE Interface Configuration




```
hostname PE-A
!
ipv6 unicast-routing
ipv6 cef
!
interface FastEthernet0/0.100
 description Link to Site-A (VLAN 100)
 encapsulation dot1Q 100
 ipv6 address FEC0:1:2300:1::1/64
 ipv6 nd cache interface-limit 4
 no ipv6 redirects
 no ipv6 unreachables
 ipv6 verify unicast source reachable-via rx
```

- Redirects are never needed
- Protect ND cache
- Unreachable needed only for PMTUD (only with MTU mismatch)

Static Routing – PE Routing Configuration




```
hostname PE-A
!
router bgp 65000
 address-family ipv6
  redistribute static route-map StaticV6ToBgp
  network FEC0:1:2300::/48
  aggregate-address FEC0:1:2200::/40 ipv6 unicast-routing
!
ipv6 route FEC0:1:2201::/48 FEC0:1:2300:1::2 tag 100
ipv6 route FEC0:1:2300::/48 Null0
!
ipv6 prefix-list PA seq 5 permit FEC0:1::/32 ge 33
route-map StaticV6ToBgp permit 10
 match tag 100
 match ipv6 address prefix-list PA
 set community no-export additive
!
route-map StaticV6ToBgp permit 20
 match tag 100
```

Connected prefixes

Customer prefixes

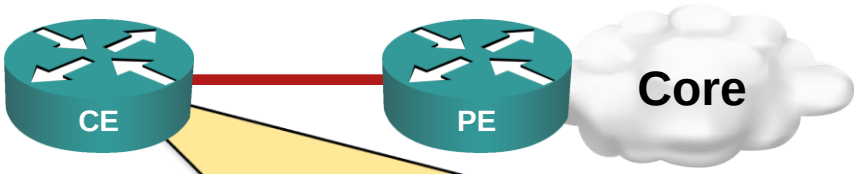
Tag customer static routes

Match PA customers

Never export PA routes

Match PI customers

Static Routing – CE Configuration




```
hostname Site-A
!
ipv6 general-prefix Site FEC0:1:2201::/48
ipv6 unicast-routing
ipv6 cef
!
interface Loopback0
 ipv6 address DHPrefix ::1/64
!
interface FastEthernet0/0
 ipv6 address DHPrefix ::1:0:0:0:1/64
!
interface FastEthernet0/1
 description Uplink (VLAN 100)
 ipv6 address FEC0:1:2300:1::2/64
!
ipv6 route ::/0 FEC0:1:2300:1::1
```

General prefixes simplify configuration and renumbering

Static default route to PE

CE router security (packet filters, firewall) not covered in this webinar

Agenda

Business customers

- Static routes
- **EBGP**

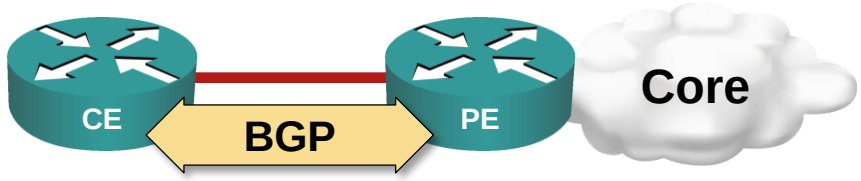
Residential customers

- Multi-access networks (WiFi)
- Carrier Ethernet (IPoE)
- DSL (PPPoX)

EBGP – Recommendations

Addressing:

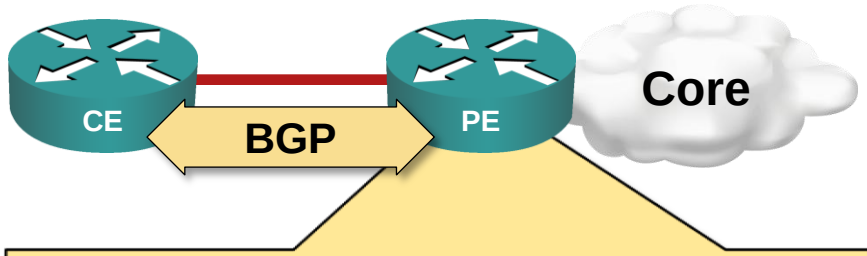
- PI or PA address space
- Numbered PE-CE links (/64 per link)



Routing:

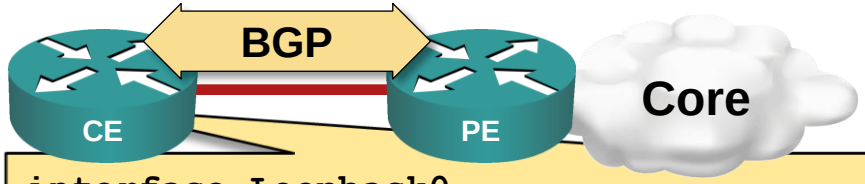
- Use two BGP sessions in dual-stack environment
- EBGP session over LLA is possible (but not recommended)
- BGP next-hop processing identical to IPv4
- All IPv4 BGP best practices still apply (inbound filters, RPF checks ...)

EBGP – PE configuration




```
hostname PE-B
!
ipv6 unicast-routing
ipv6 cef
!
interface Serial1/1
 description Link to Site-B
 encapsulation ppp
 ipv6 address FEC0:1:1001:1::1/64
!
router bgp 65000
 neighbor FEC0:1:1001:1::2 remote-as 65100
 !
 address-family ipv6
  neighbor FEC0:1:1001:1::2 activate
 exit-address-family
```

EBGP – CE configuration




```
interface Loopback0
 ipv6 address FEC0:1:3001:0::1/128
!
interface FastEthernet0/0
 ipv6 address FEC0:1:3001:1::1/64
!
interface Serial1/0
 description Link to the MPLS SP
 encapsulation ppp
 ipv6 address FEC0:1:1001:1::2/64
!
router bgp 65100
 bgp log-neighbor-changes
 neighbor FEC0:1:1001:1::1 remote-as 65000
 !
 address-family ipv6
  network FEC0:1:3001::/48
  neighbor FEC0:1:1001:1::1 activate
!
ipv6 route FEC0:1:3001::/48 Null0
```

EBGP – CE Router Printouts

```
Site-B#show bgp ipv6 unicast summary
BGP router identifier 10.0.1.4, local AS number 65100
BGP table version is 2, main routing table version 2
1 network entries using 144 bytes of memory
1 path entries using 76 bytes of memory
1/1 BGP path/bestpath attribute entries using 124 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 344 total bytes of memory
BGP activity 1/0 prefixes, 1/0 paths, scan interval 60 secs

Neighbor        V      AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
FEC0:1:1001:1::1
                4   65000       9      10        2    0    0 00:05:52        0

Site-B#show bgp ipv6 unicast
BGP table version is 2, local router ID is 10.0.1.4
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> FEC0:1:3001::/48 ::                       0         32768 i
```

EBGP – PE Router Printouts

```
PE-B#show ipv6 route

LC   FEC0::CCCC:2/128 [0/0]
      via Loopback0, receive
C    FEC0:1:1001:1::/64 [0/0]
      via Serial1/1, directly connected
L    FEC0:1:1001:1::1/128 [0/0]
      via Serial1/1, receive
B    FEC0:1:3001::/48 [20/0]
      via FE80::C803:8FF:FEC4:8, Serial1/1
L    FF00::/8 [0/0]
      via Null0, receive

PE-B#show bgp ipv6 unicast
BGP table version is 2, local router ID is 10.0.1.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> FEC0:1:3001::/48 FEC0:1:1001:1::2
                                             0             0 65100 i
```

Agenda

Business customers

- Static routes
- EBGP

Residential customers

- Multi-access networks (WiFi)
- Carrier Ethernet (IPoE)
- DSL (PPPoX)
- 6rd

Review: Configuring Host IPv6 Parameters

Minimum set of parameters:

- Host IPv6 address
- First-hop router's IPv6 address
- DNS server IPv6 address (could use IPv4 DNS server in dual-stack environments)

Configuration mechanisms:

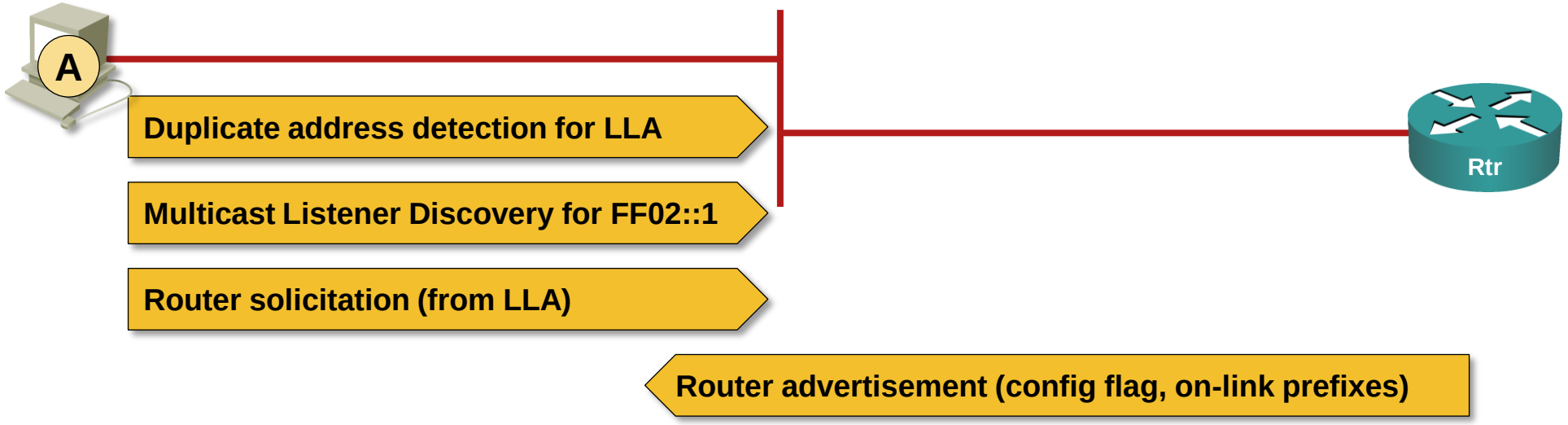
- Static configuration (servers, routers)
- Stateless Autoconfiguration (SLAAC) using Router Advertisements
- DHCPv6-based configuration

Review: Host Configuration Options

Parameter	ICMPv6 (ND/RA)	DHCPv6
Host IPv6 address	Yes (SLAAC)	Yes
First hop router's IPv6 address	Yes (RA)	No
DNS server's IPv6 address	Yes (RFC 6106)	Yes

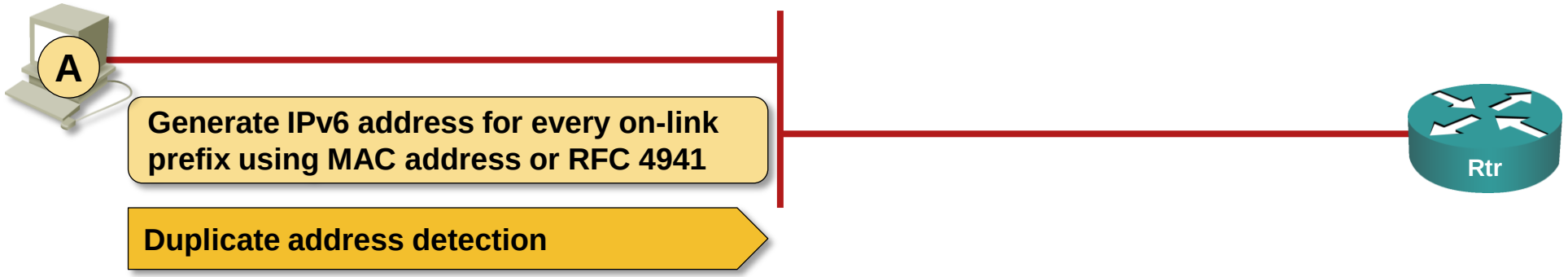
- RFC 6106 is not widely supported yet
- In most cases you need both RA and DHCPv6
- SLAAC with dynamic DNS registration is preferred to DHCPv6-based address allocation (user tracking)

Review: Host Configuration, Part 1



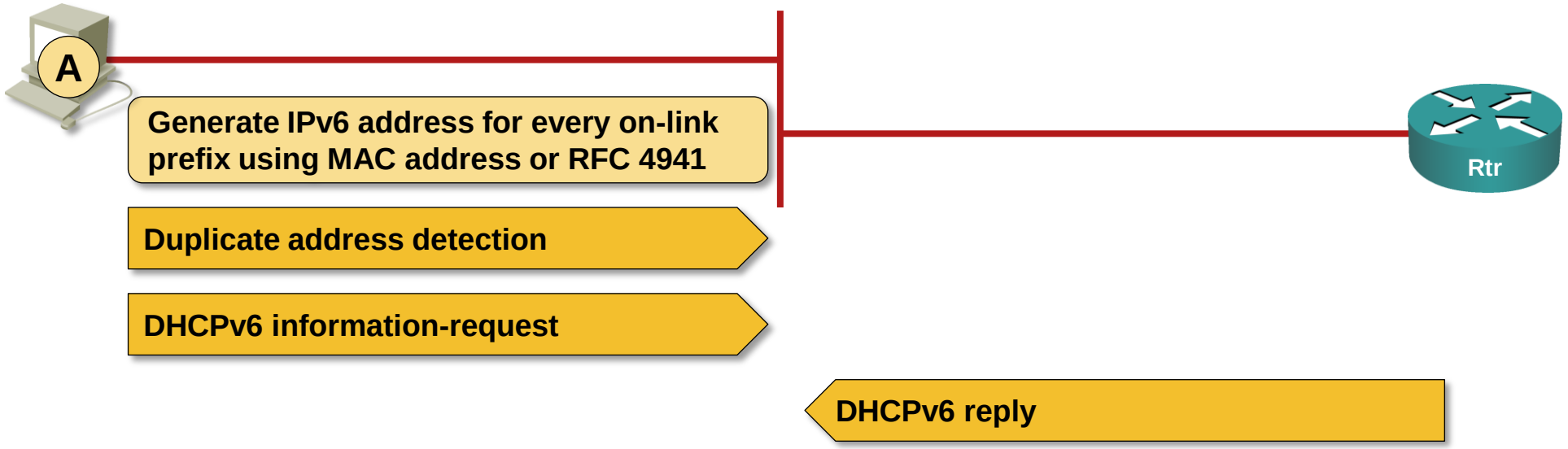
- Newly started host must first get a LLA (using its MAC address)
- Duplicate address detection is used to check LLA uniqueness
- Host joins the all-nodes multicast group (MLD needed for L2 switches)
- Host tries to find an adjacent router to get configuration mechanisms and on-link prefixes
- DHCPv6 *may* be used if no routers are present on the link

Review: Host Configuration – SLAAC



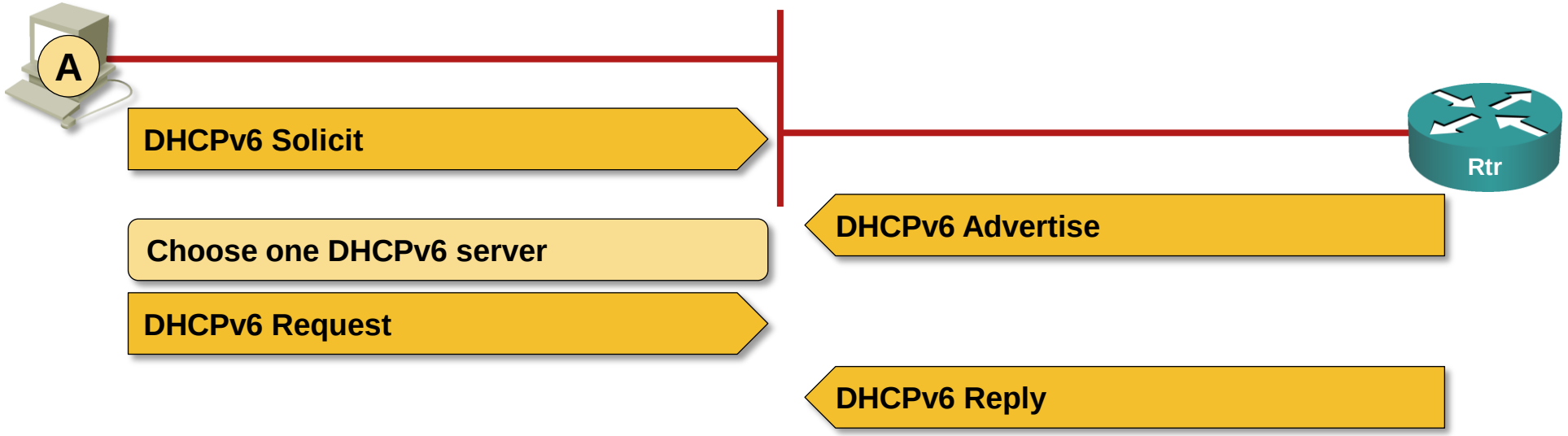
- Host generates an IPv6 address for every on-link prefix advertised by the routers
- Host uses duplicate address detection to check generated address uniqueness
- Non-RFC4941 SLAAC fails if the host encounters a duplicate IPv6 address (indicating duplicate Interface ID – MAC address)
- IPv4 DHCP can be used in dual-stack environments to specify DNS server IPv4 address

Review: Host Configuration – SLAAC + DHCPv6



- SLAAC is used to generate IPv6 addresses for on-link prefixes
- DHCPv6 request is sent to retrieve non-address parts of the configuration (DNS server IPv6 address)
- Router can reply to the DHCPv6 request or relay it to a central DHCPv6 server

Review: Host Configuration – DHCPv6 Only



- DHCPv6 is used to assign IPv6 addresses (and other parameters) to the hosts
- Two-step process like IPv4 DHCP
- Router can run a DHCPv6 server or relay DHCPv6 requests
- Rapid commit (one step process – *Solicit* message answered with *Reply* message) can be used if supported by the client and the DHCPv6 server

Residential Access – Individual Hosts

Common setup

- ND RA for default routing
- DHCPv6 for DNS server propagation
- SLAAC for address assignment

Addressing and routing

- Each subnet gets a /64 PA prefix
- One subnet per customer (VLAN or PPPoE link) or shared subnet
- Per-PE-router or per-POP aggregate is advertised in BGP

SLAAC or DHCPv6?

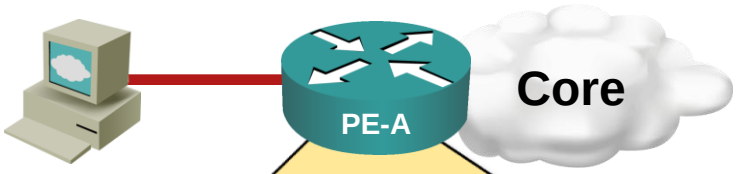
SLAAC benefits and drawbacks

- Simpler, more scalable (minimal PE-router involvement)
- No user authentication/authorization
- No control of end-user addresses (no accountability)

DHCPv6 benefits and drawbacks

- Port-based user authentication (Remote-ID/Subscriber-ID – Option 37/38)
- Static IPv6 addresses (uniquely identifies end-user)
- DHCPv6 snooping and IPv6 Source Guard recommended

PE-Router Configuration – SLAAC




```
hostname PE-A
!
ipv6 dhcp pool Clients
 dns-server FEC0::CCCC:4
 domain-name example.com
!
interface FastEthernet0/0.200
 description Carrier Ethernet access
 encapsulation dot1Q 200
 ipv6 address FEC0:1:2301::1/64
 ipv6 nd other-config-flag
 ipv6 nd router-preference High
 no ipv6 nd ra suppress
 ipv6 nd ra interval 5
 ipv6 dhcp server Clients
```

DNS server address is configured in DHCP pool

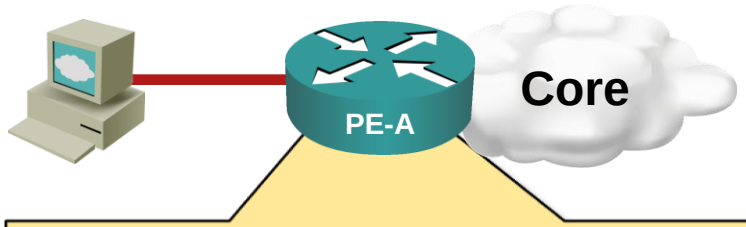
Other-Config-Flag is a hint to get information from DHCPv6

Prevents unintentional client-side mistakes

Just in case IOS picks wrong default value

Speeds up the SLAAC process

PE-Router Configuration – DHCPv6




```
hostname PE-A
!
ipv6 dhcp pool VLAN_100
 address prefix FEC0:1:2300:1:0:8000::/96 lifetime 30 10
 dns-server FEC0::CCCC:E
 domain-name example.com
!
interface FastEthernet0/0.100
 description Host Access LAN (VLAN 100)
 encapsulation dot1Q 100
 ipv6 address FEC0:1:2300:1::1/64
 ipv6 nd prefix FEC0:1:2300:1::/64 no-advertise
 ipv6 nd managed-config-flag
 ipv6 nd other-config-flag
 ipv6 nd router-preference High
 no ipv6 unreachables
 ipv6 dhcp server VLAN_100
 ipv6 verify unicast source reachable-via rx
```

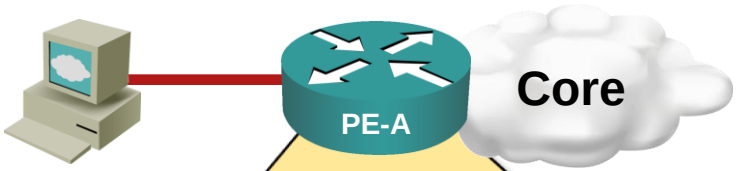
Address allocation range (can be less than /128)

IPv6 /64 prefix must be configured for routing purposes

Prefix hiding prevents SLAAC

Managed-Config-Flag should imply Other-Config-Flag, but ...

PE-Router Configuration – DHCPv6 Relay




```
hostname PE-A
!
interface FastEthernet0/0.100
 description Host Access LAN (VLAN 100)
 encapsulation dot1Q 100
 ipv6 address FEC0:1:2300:1::1/64
 ipv6 enable
 ipv6 nd prefix FEC0:1:2300:1::/64 no-advertise
 ipv6 nd managed-config-flag
 ipv6 nd router-preference High
 ipv6 dhcp relay destination FEC0::CCCC:3
 ipv6 dhcp relay source-inteface Loopback0
 ipv6 verify unicast source reachable-via rx
```

IPv6 address of DHCPv6 server

Source IPv6 address of DHCPv6 relay message

DHCPv6 Relay Message

```
⊞ Internet Protocol Version 6, Src: fe80::c800:6ff:fe89:8 (fe80::c800:6ff:fe89:8), Dst: fec0::cccc:3 (fec0::cccc:3)
⊞ User Datagram Protocol, Src Port: dhcpv6-server (547), Dst Port: dhcpv6-server (547)
⊟ DHCPv6
    Message type: Relay-forw (12)
    Hopcount: 0
    Link address: fec0:1:2300:1::1 (fec0:1:2300:1::1)
    Peer address: fe80::c804:6ff:fe8a:6 (fe80::c804:6ff:fe8a:6)
  ⊟ Relay Message
      Option: Relay Message (9)
      Length: 90
      Value: 0514979800080002000000001000a00030001ca04068a0008...
    ⊟ DHCPv6
        Message type: Renew (5)
        Transaction ID: 0x149798
      ⊞ Elapsed time
      ⊞ Client Identifier: 00030001ca04068a0008
      ⊞ Option Request
      ⊞ Server Identifier: 00030001ca03068a0008
      ⊞ Identity Association for Non-temporary Address
  ⊞ Interface-Id
  ⊞ Remote Identifier
```

- Client DHCPv6 message is encapsulated in another DHCPv6 message
- Source address of IPv6 packet = outgoing interface or **ipv6 dhcp-relay source-interface**
- Link address = IPv6 address of incoming interface
- Peer address = IPv6 address of DHCPv6 client (from original DHCPv6 packet)

DHCPv6 Relay – Client Identification

```
DHCPv6
  Message type: Relay-forw (12)
  Hopcount: 0
  Link address: fec0:1:2300:1::1 (fec0:1:2300:1::1)
  Peer address: fe80::c804:6ff:fe8a:6 (fe80::c804:6ff:fe8a:6)
Relay Message
  Option: Relay Message (9)
  Length: 90
  Value: 0514979800080002000000010000a00030001ca04068a0008...
DHCPv6
    Message type: Renew (5)
    Transaction ID: 0x149798
  Elapsed time
  Client Identifier: 00030001ca04068a0008
  Option Request
  Server Identifier: 00030001ca03068a0008
  Identity Association for Non-temporary Address
Interface-Id
  Option: Interface-Id (18)
  Length: 9
  Value: 4661302f302e313030
  Interface-ID: Fa0/0.100
Remote Identifier
  Option: Remote Identifier (37)
  Length: 22
  Value: 000000090200000000064000a00030001ca0006890008
  Enterprise ID: ciscoSystems (9)
  Remote-ID: 020000000064000a00030001ca0006890008
```

Client ID in original DHCPv6 message

Interface ID (name) on the PE-router

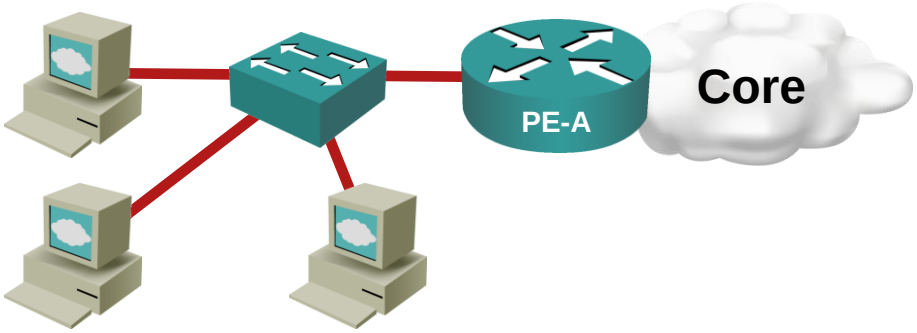
Unique client identifier (equivalent to Option 82 in IPv4)

First-Hop Security Mechanisms

Like IPv4, IPv6 is vulnerable to DoS and MITM attacks

- ND resource exhaustion attacks
- RA spoofing:
  - first hop router hijacks (MITM)
  - Managed-Config-Flag spoofing (MITM)
  - Insertion of non-routable prefixes (DoS)
- ND spoofing/hijacks (MITM or DoS)
- DHCPv6 server spoofing (MITM)

These attacks are possible only in shared-access L2 environments



Generic solutions:

- L3 access switches
- VLAN-per-customer
- Private VLAN (ND spoofing still possible)

IPv6-specific first-hop security features

- RA guard
- DHCPv6 Snooping
- IPv6 ND inspection (first in 12.2(50)SY)
- Secure Neighbor Discovery (SeND)

Agenda

Business customers

- Static routes
- EBGP

Residential customers

- Multi-access networks (WiFi)
- **Carrier Ethernet (IPoE)**
- DSL (PPPoX)
- 6rd

CPE Initialization: PE-CE Link Addressing

Duplicate address detection for LLA

Multicast Listener Discovery for FF02::1

Router solicitation (from LLA)

Router advertisement (other-config, on-link prefixes)

Generate IPv6 address for every on-link prefix using MAC address or RFC 4941

Duplicate address detection

DHCPv6 information-request

DHCPv6 reply (IPv6 DNS server address)

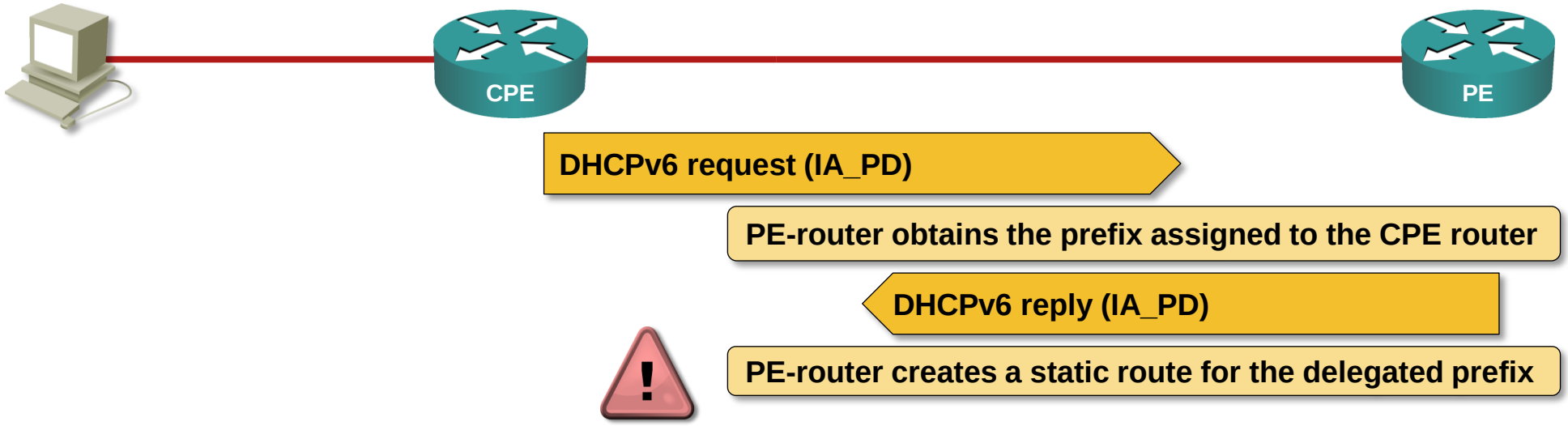
CPE router has IPv6 address on outside interface, still needs inside addresses (no NAT)

CPE Initialization: Prefix Delegation

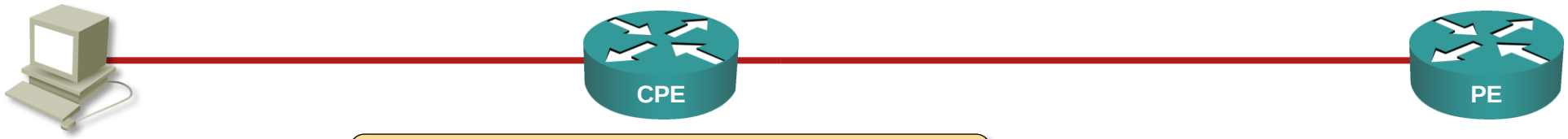
- CPE router needs one or more public /64 subnets to address inside interfaces
- IPv6 prefix assigned to CPE router should be stable
- Renumbering is possible (prefix lifetime in RA) but could take minutes or hours

Option#1: Out-of-band prefix allocation, manual CPE configuration, static routes

Option#2: Automatic prefix delegation with DHCPv6 IA_PD option (RFC 3633)



CPE Initialization: Host Addressing



CPE router addresses its inside interfaces using delegated prefix

RA (other-config, on-link prefixes)

Inside hosts use SLAAC to get public IPv6 addresses

Duplicate address detection

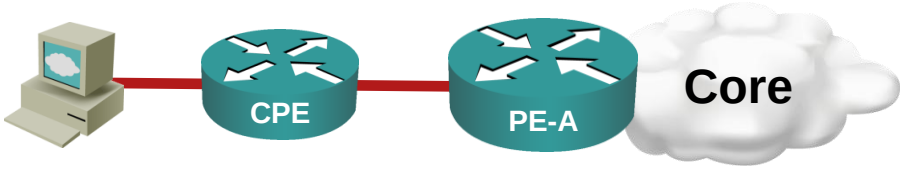
DHCPv6 information-request

DHCPv6 reply (IPv6 DNS)

IPv6 DNS server address is copied from the DHCPv6 reply sent by the PE-router

IPv6-over-Ethernet Requirements

Common setup



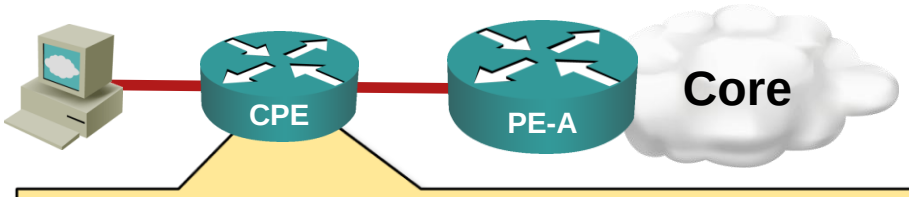
- ND RA for default routing
- DHCPv6 for DNS server propagation
- SLAAC or DHCPv6 for address assignment
- PA /64 prefix for every IPv6 subnet (VLAN)

Low-end IPv6-enabled CPE routers

- Prefix allocation with DHCPv6 IA_PD option (from /64 to /48)
  Recommended: /60 to /56
- Local pool of prefixes on PE-router or DHCP relay
- Individual prefixes or local pool prefix advertised in BGP

Prefixes delegated with DHCPv6 IA_PD must be stable

Low-End CPE Configuration – Addressing




```
hostname Site-A
!
ipv6 unicast-routing
ipv6 cef
!
interface Loopback0
 ipv6 address DHPrefix ::1/64
!
interface FastEthernet0/0
 ipv6 address DHPrefix ::1:0:0:0:1/64
!
interface FastEthernet0/1
 description Uplink to PE-A
 ipv6 enable
 ipv6 nd autoconfig default-route
 ipv6 dhcp client pd DHPrefix
 ipv6 nd ra suppress
```

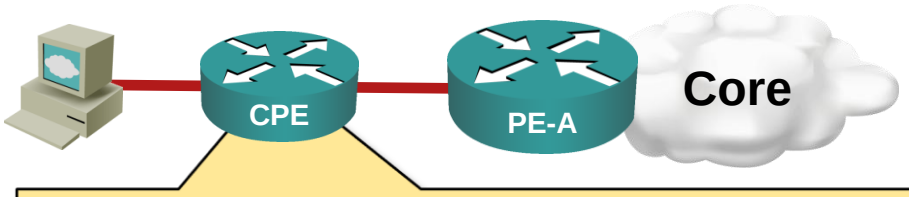
Fact: you will see lots of CPE routers sending RA messages

Stopgap: set RA preference to *high* on PE-routers

Solution: first-hop security

- Deploy switches with RA-guard
- VLAN per customer
- L3 access-layer switches

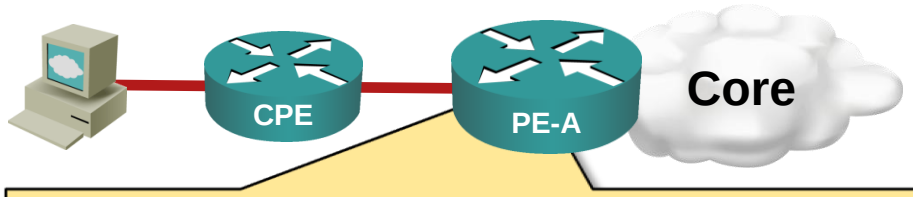
Low-End CPE Configuration – Inside DHCPv6




```
hostname Site-A
!
ipv6 dhcp pool Inside
 import dns-server
 import domain-name
 import information refresh
!
interface FastEthernet0/0
 ipv6 nd other-config-flag
 ipv6 nd ra interval 4 3
 ipv6 dhcp server Inside
```

- Import DHCP parameters received from the PE-router into a DHCP pool
- Configure *other-config-flag* and DHCP server on inside interfaces

PE-Router: Local Prefix Delegation With DHCPv6




```
hostname PE-A
!
ipv6 dhcp pool Clients
 prefix-delegation FEC0:1:2200:4700::/56 00030001CA04068A0008
 prefix-delegation pool DhcpPool
 dns-server FEC0::CCCC:4
 domain-name example.com
!
ipv6 dhcp database url
!
ipv6 route FEC0:1:2200::/40 Null0
ipv6 local pool DhcpPool FEC0:1:2200::/40 56
```

Delegate specific prefixes to clients with known DUID

Allocate other delegated prefixes from a pool

Save delegated prefixes in a DHCP database

Create an aggregate route for the delegated prefixes

Advertise only the aggregate route into BGP

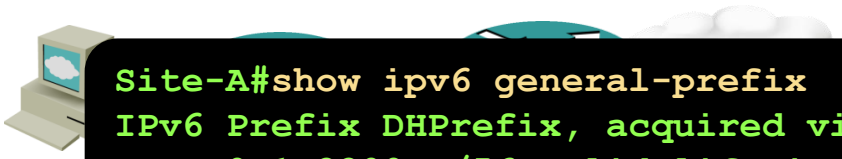
Delegation pool: /56 prefixes from a /40 IPv6 prefix

Initial IPv6 configuration on PE-router is identical to multi-access scenarios

DHCPv6 Client On CPE Router


```
Site-A#show ipv6 dhcp interface FastEthernet 0/1
FastEthernet0/1 is in client mode
  Prefix State is OPEN (0)
  Information refresh timer expires in 23:54:41
  Renew will be sent in 3d11h
  Address State is IDLE
  List of known servers:
    Reachable via address: FE80::C800:6FF:FE89:8
    DUID: 00030001CA0006890008
    Preference: 0
    Configuration parameters:
      IA PD: IA ID 0x00040001, T1 302400, T2 483840
        Prefix: FEC0:1:2200::/56
                preferred lifetime 604800, valid lifetime 2592000
                expires at Feb 18 2012 04:08 PM (2591681 seconds)
      DNS server: FEC0::CCCC:4
      Domain name: example.com
      Information refresh time: 0
  Prefix name: DHPrefix
  Prefix Rapid-Commit: disabled
  Address Rapid-Commit: disabled
```

IPv6 Routing Table On CPE Router



```
Site-A#show ipv6 general-prefix
IPv6 Prefix DHPrefix, acquired via DHCP PD
  FEC0:1:2200::/56 Valid lifetime 2591881, preferred lifetime 604681
   FastEthernet0/0 (Address command)
   Loopback0 (Address command)

Site-A#show ipv6 route | sect ^[SC]
S    ::/0 [2/0]
     via FE80::C800:6FF:FE89:8, FastEthernet0/1
S    FEC0:1:2200::/56 [1/0]
     via Null0, directly connected
C    FEC0:1:2200::/64 [0/0]
     via Loopback0, directly connected
C    FEC0:1:2200:1::/64 [0/0]
     via FastEthernet0/0, directly connected
C    FEC0:1:2300:1::/64 [0/0]
     via FastEthernet0/1, directly connected

Site-A#show ipv6 dhcp pool
DHCPv6 pool: Inside
  Imported DNS server: FEC0::CCCC:4
  Imported Domain name: example.com
```

IPv6 Routing Table On PE-router

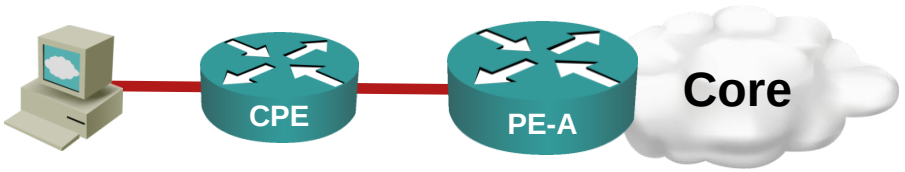
```
PE-A#show ipv6 local pool
Pool                    Prefix                                  Free  In use
DhcpPool                FEC0:1:2200::/40                        65535      1

PE-A#show ipv6 route static | begin ^S
S   FEC0:1:2200::/40 [1/0], tag 100
     via Null0, directly connected
S   FEC0:1:2200::/56 [1/0]
     via FE80::C804:6FF:FE8A:6, FastEthernet0/0.100
S   FEC0:1:2300::/48 [1/0]
     via Null0, directly connected
```

Remember:

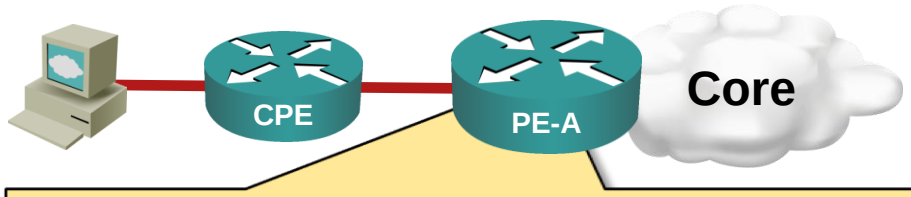
- Static route is automatically created for every delegated prefix
- You should configure a static summary route for the pool prefix
- Advertise only the summary route into IGP/BGP

DHCPv6 Bindings On PE-Router




```
PE-A#show ipv6 dhcp binding
Client: FE80::C804:6FF:FE8A:6
  DUID: 00030001CA04068A0008
  Username : unassigned
  VRF : default
  Interface : FastEthernet0/0.100
  IA PD: IA ID 0x00040001, T1 302400, T2 483840
    Prefix: FEC0:1:2200::/56
            preferred lifetime 604800, valid lifetime 2592000
            expires at Feb 18 2012 04:08 PM (2586792 seconds)
```

Adjusting Delegated Prefix Lifetime




```
hostname PE-A
!
ipv6 dhcp pool Clients
 prefix-delegation pool DhcpPool lifetime valid preferred
```

Problem

- Delegated IPv6 prefixes should be stable (expectation: unchanged)
- Delegation-based static routes might not survive interface flap or reload

Workaround

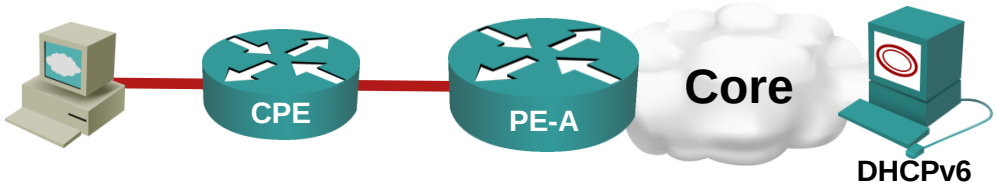
- Reduce delegated prefix validity to a few minutes (default: 30 days)

Caveats

- CPEs will renew the delegated prefix at half the *preferred* time (default: 3.5 days)

Prefix Delegation With DHCPv6 Relay

- Standard DHCPv6 Relay functionality is used
- DHCPv6 server can include IA_PD option in reply messages



- PE-router installs a static route to delegated prefix before forwarding the reply to the CPE-router

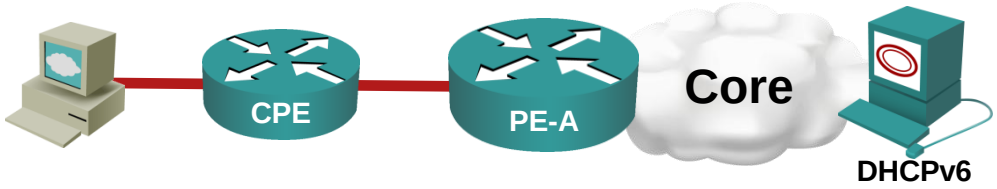
Surviving interface flaps or router reloads

- Static routes to delegated prefixes are lost after interface flap or reload
- Interface flap: PE-router restores static routes from DHCPv6 bindings
- Router reload: *Bulk Lease Query* (RFC 5460) downloads existing leases from DHCPv6 server (15.1(1)S, 12.2(58)SE, XE 3.2S)

Prefix Delegation With DHCPv6 Relay

Addressing

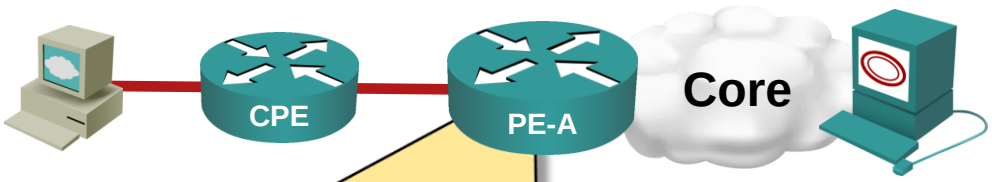
- Delegated prefixes must be stable
- DHCPv6 server can use DUID or RemoteID (Option 37) to identify the client



Routing

- Assign address blocks to POPs and PE-routers
- Redistribute static routes into BGP
- Aggregate on PE-routers and POP edge routers
- User migration supported with minimum routing overhead

DHCPv6 Relay PE-Router Configuration




```
hostname PE-A
!
ipv6 dhcp-relay bulk-lease retry 5
!
interface FastEthernet0/0.100
 description Carrier Ethernet access
 encapsulation dot1Q 100
 ipv6 address FEC0:1:2301::1/64
 ipv6 nd other-config-flag
 ipv6 dhcp relay destination FEC0::CCCC:4
 ipv6 dhcp relay source-interface Loopback0
!
router bgp 65000
 address-family ipv6
  redistribute static
  aggregate-address FEC0:1:2200::/40
```

Bulk lease is enabled by default if available (15.1S)

IPv6 address of DHCPv6 server

Use a stable interface with routable IPv6 address

Automatic static routes don't have tags

Aggregate routes from per-PE-router prefix pool

Alternative: Static summary routes and selective redistribution into BGP

DHCPv6 Relay Bindings and Static Routes

```
PE-A#show ipv6 dhcp relay binding
Prefix: FEC0:1:2310::/56 (FastEthernet0/0.200)
  DUID: 00030001CA0208A70008
  IAID: 524289
  lifetime: 2592000
  expiration: 08:44:19 UTC Jan 7 2011

PE-A#show ipv6 route interface FastEthernet0/0.200
IPv6 Routing Table - default - 16 entries
Codes: C - Connected, L - Local, S - Static, U - Per-user Static route
       B - BGP, R - RIP, I1 - ISIS L1, I2 - ISIS L2
       IA - ISIS interarea, IS - ISIS summary, D - EIGRP, EX - EIGRP external
       ND - Neighbor Discovery
       O - OSPF Intra, OI - OSPF Inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
       ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
C   FEC0:1:2301::/64 [0/0]
     via FastEthernet0/0.200, directly connected
L   FEC0:1:2301::1/128 [0/0]
     via FastEthernet0/0.200, receive
S   FEC0:1:2310::/56 [1/0]
     via FE80::C802:8FF:FEA7:6, FastEthernet0/0.200
```

Agenda

Business customers

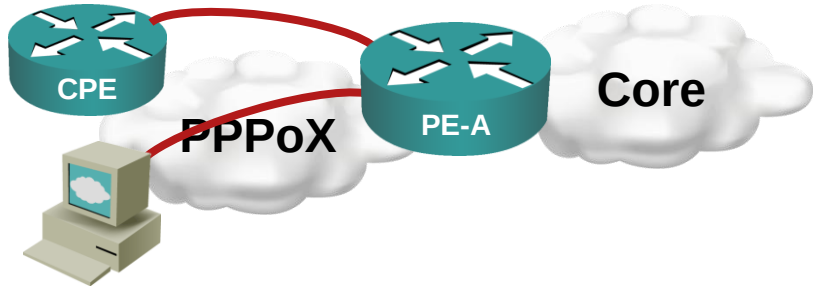
- Static routes
- EBGP

Residential customers

- Multi-access networks (WiFi)
- Carrier Ethernet (IPoE)
- **DSL (PPPoX)**
- 6rd

PPPoX Design Principles

- Every PPPoX session is a subnet
- Every PE-CE subnet needs a /64 prefix (end-host SLAAC)
- CPE routers need prefix delegation
- RADIUS server is commonly used for authentication



Local addressing

- Pool of prefixes used for PE-CE links
- A different pool of prefixes used for delegated prefixes

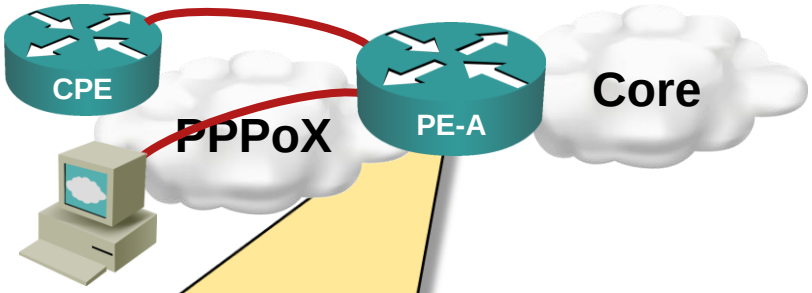
Doesn't work well with DHCPv6 prefix delegation and BRAS banks

Centralized addressing (RADIUS)

- *Framed-IPv6-Prefix* for PE-CE link
- *Framed-IPv6-Pool* selects local pool
- *Delegated-IPv6-Prefix* specifies DHCPv6 delegated prefix

More important for delegated prefixes than for PE-CE links

Local PE-CE Addressing




```
ipv6 dhcp pool Clients
 dns-server FEC0::CCCC:4
 domain-name example.com
!
interface Virtual-Template10
 mtu 1480
 peer default ipv6 pool PPP
 ipv6 enable
 ipv6 nd other-config-flag
 no ipv6 nd ra suppress
 ppp authentication chap pap
 ipv6 nd ra interval 5
 ipv6 dhcp server Clients
!
ipv6 local pool PPP FEC0:1:2300::/48 64
ipv6 route FEC0:1:2300::/48 Null0
```

DHCP pool needed for DNS server address

Local pool for PE-CE addresses

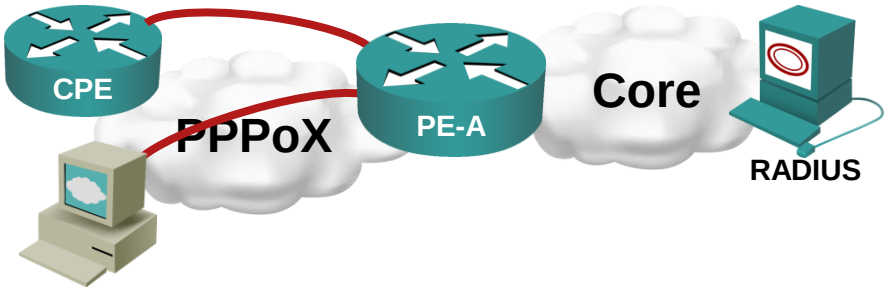
Unnumbered IPv6 interface

Allocate /64 prefixes to PE-CE links

Summary route for BGP

RADIUS-driven PE-CE addressing

- *Framed-IPv6-Pool* RADIUS attribute overrides **peer default ipv6 pool**
- *Framed-IPv6-Prefix* RADIUS attribute adds another PE-CE prefix
- You might get two PE-CE prefixes with *Framed-IPv6-Prefix* + **peer default ipv6 pool**
- Recommendation: Use local configuration or RADIUS attributes, not both
- No extra router configuration is needed



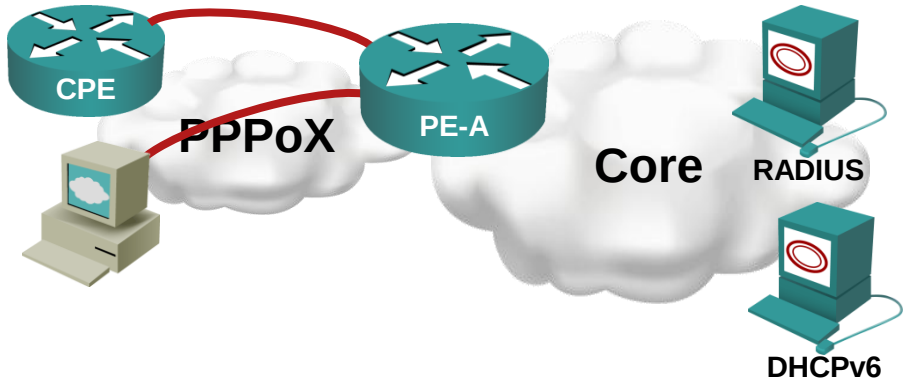

```
Site-A  Cleartext-Password := "Site-A"
        Service-Type = Framed-User,
        Framed-Protocol = PPP,
        Framed-IPv6-Prefix = "fec0:1:2400:1::/64"

User-B  Cleartext-Password := "User-B"
        Service-Type = Framed-User,
        Framed-Protocol = PPP,
        Framed-IPv6-Pool = "PPP"
```

DHCPv6 Prefix Delegation Options

Implementation options

- Local prefix delegation pools
- Central DHCPv6 server with DHCPv6 relay on PE-router
- RADIUS-driven prefix delegation



DHCPv6-RADIUS integration options

- RFC 4818 (15.1S and XE 3S)
- Two usernames per user (Cisco proprietary – 12.2S, 15.0M)
- Framed prefix = delegated prefix (Cisco proprietary – 15.0M)

DHCPv6-RADIUS Integration: Two Usernames

PPP LCP

CHAP u=xyz p=***

RADIUS RQ u=xyz p=***

OK, IPv6 prefix=xxx/64

CHAP ACK

SLAAC

DHCPv6 IA_PD

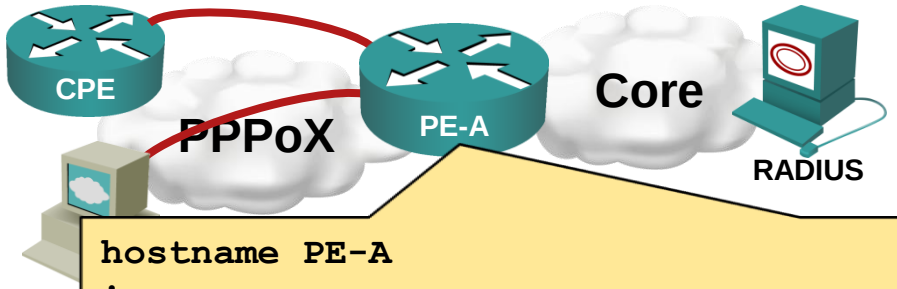
PE-A installs IPv6 prefix on PPPoE virtual access interface

RADIUS RQ u=xyz-dhcpv6

OK, IPv6 prefix=yyy/56

DHCPv6 IA_PD=yyy

DHCPv6-RADIUS Integration: PE Router Configuration




```
hostname PE-A
!
aaa authentication ppp default group radius
aaa authorization network default group radius
aaa authorization configuration IA_PD group radius
!
ipv6 dhcp pool PPP-Radius
 prefix-delegation aaa method-list IA_PD
 dns-server FEC0::CCCC:4
 domain-name example.com
!
radius-server host 10.17.0.2 auth-port 1812 acct-port 1813 key *****
```

Remember:

- Static route to CPE is created at prefix delegation time
- Routes to delegated prefixes are redistributed into BGP (use aggregates)

RADIUS Configuration (Two Usernames)

```
Site-A  Cleartext-Password := "Site-A"
        Service-Type = Framed-User,
        Framed-Protocol = PPP,
        Framed-IPv6-Prefix = "fec0:1:2400:1::/64",

Site-A-dhcpv6   Auth-Type := Accept, Service-Type !* any, NAS-Port-Id !* any
        cisco-avpair = "ipv6:prefix#1=fec0:1:2400:1100::/56",
```

Two users in the RADIUS user database

- **Framed-IPv6-Prefix** from the original user is used on the virtual access interface
- **Framed-IPv6-Prefix** from the second user is used for IA_PD delegation

DHCPv6 user has no password – security risk

- Must use **Auth-type := Accept** with FreeRADIUS
- Use additional check items to prevent PPP or login authentication

RADIUS-Generated DHCPv6 Prefix Delegations

```
PE-A#show ipv6 dhcp binding
Client: FE80::C802:8FF:FEA7:8
  DUID: 00030001CA0208A70008
  Interface : Virtual-Access2.1
  IA PD: IA ID 0x00090001, T1 302400, T2 483840
    Prefix: FEC0:1:2400:1100::/56
            preferred lifetime 604800, valid lifetime 2592000
            expires at Jan 07 2011 09:44 AM (2588153 seconds)
PE-A#show ipv6 route interface Virtual-Access2.1
IPv6 Routing Table - default - 16 entries
Codes: C - Connected, L - Local, S - Static, U - Per-user Static route
       B - BGP, R - RIP, I1 - ISIS L1, I2 - ISIS L2
       IA - ISIS interarea, IS - ISIS summary, D - EIGRP, EX - EIGRP external
       ND - Neighbor Discovery
       O - OSPF Intra, OI - OSPF Inter, OE1 - OSPF ext 1, OE2 - OSPF ext 2
       ON1 - OSPF NSSA ext 1, ON2 - OSPF NSSA ext 2
C   FEC0:1:2300::/64 [0/0]
     via Virtual-Access2.1, directly connected
C   FEC0:1:2400:1::/64 [0/0]
     via Virtual-Access2.1, directly connected
S   FEC0:1:2400:1100::/56 [1/0]
     via FE80::C802:8FF:FEA7:8, Virtual-Access2.1
```

RFC 4818: Delegated-IPv6-Prefix Attribute

PPP LCP

CHAP u=xyz p=***

RADIUS RQ u=xyz p=***

OK, IPv6 prefix=xxx/64
Delegated-IPv6=yyy/56

CHAP ACK

SLAAC

DHCPv6 IA_PD

DHCPv6 IA_PD=yyy

PE-A installs IPv6 prefix on PPPoE virtual access interface and remembers the delegated prefix for future use

RFC 4818: Delegated-IPv6-Prefix Attribute

```
Site-A  Cleartext-Password := "Site-A"
        Service-Type = Framed-User,
        Framed-Protocol = PPP,
        Framed-IPv6-Prefix = "fec0:1:2400:1::/64",
        Delegated-IPv6-Prefix = "fec0:1:2400:1100::/56"
```

- *Framed-IPv6-Prefix* or *Framed-IPv6-Pool* is used for PE-CE link addressing
- *Delegated-IPv6-Prefix* is used for DHCPv6 IA_PD prefix
- Available in IOS releases 15.1S and XE 3S
- Requires no changes in router configuration

Fallback procedures on RFC 4818-compliant PE-routers:

- Use *Delegated-IPv6-Prefix* if available
- Otherwise send RADIUS request for *user-dhcpv6* when CPE requests IA_PD

Framed-IPv6-Prefix = delegated prefix

Another Cisco IOS kludge

- Framed-IPv6-Prefix from RADIUS is used as delegated prefix
- Local IPv6 pool has to be used to address virtual access interfaces and enable SLAAC (mandatory for dial-in hosts)
- Works in 15.0+, 12.2SRE uses the framed-ipv6-prefix in ND RA and DHCPv6 reply (completely confuses the CPE router)
- DHCPv6 server still needed on the PE-router (including DNS server address)

```
ipv6 dhcp pool PPP-Radius
 dns-server FEC0::CCCC:4
!
interface Virtual-Template10
 peer default ipv6 pool PPP
 ipv6 enable
 no ipv6 nd prefix framed-ipv6-prefix
 ipv6 nd other-config-flag
 no ipv6 nd ra suppress
 ipv6 dhcp server PPP-Radius
```

Recommendations

Users connect to a single PE-router/BRAS

- Use local pool for PPPoE prefixes
- Use local DHCPv6 server with DHCP database
- Advertise the local pool prefixes into BGP

Options

- When using central DHCPv6 or RADIUS, allocate prefixes from a router-specific range
- Advertise local aggregates + out-of-local-range prefixes into BGP

Users connect to a router pool

- Use local pools for virtual access interface prefixes
- Option: use RADIUS for users requiring fixed IPv6 subnet
- Use centralized DHCPv6 server or RADIUS integration
- Delegated prefixes should be from a POP-specific range
- Advertise connected and static routes into BGP or local IGP
- Advertise POP aggregates into network core (not individual prefixes)

Agenda

Business customers

- Static routes
- EBGP

Residential customers

- Multi-access networks (WiFi)
- Carrier Ethernet (IPoE)
- DSL (PPPoX)
- **6rd**

The History of 6rd

6in4 (RFC 4213)

- Encapsulation of IPv6 datagrams into IPv4 datagrams
- IP protocol 41
- Manually configured point-to-point tunnels

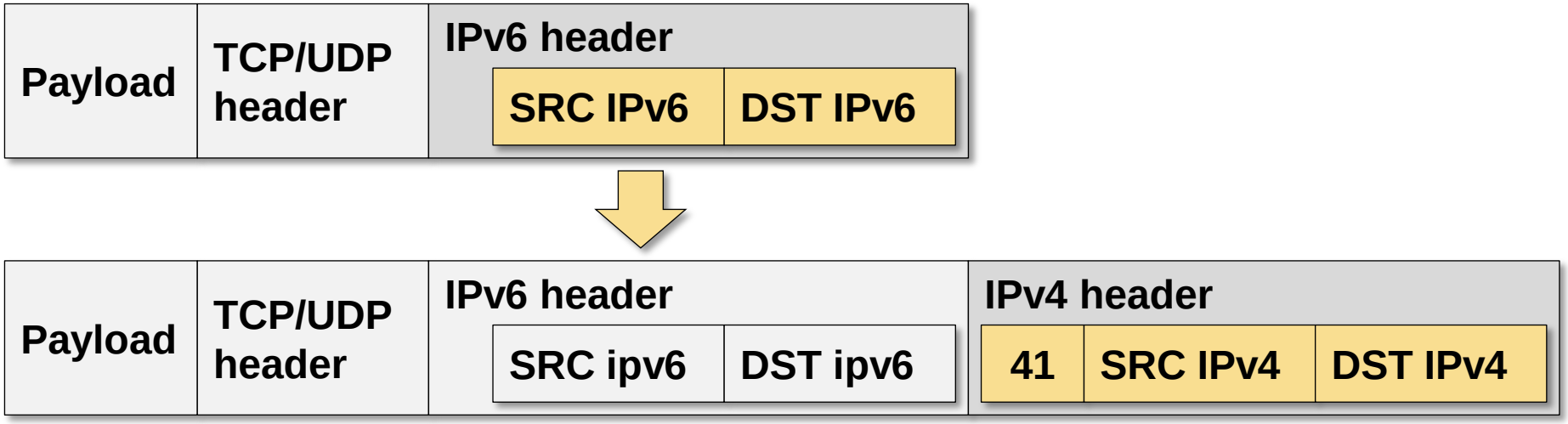
6to4 (RFC 3056 and RFC 3068)

- Uses 6in4 encapsulation
- Well-known IPv6 prefix
- Automatic allocation of /48 prefix to every public IPv4 address
- Automatic (asymmetric) tunnels

6rd (RFC 5969)

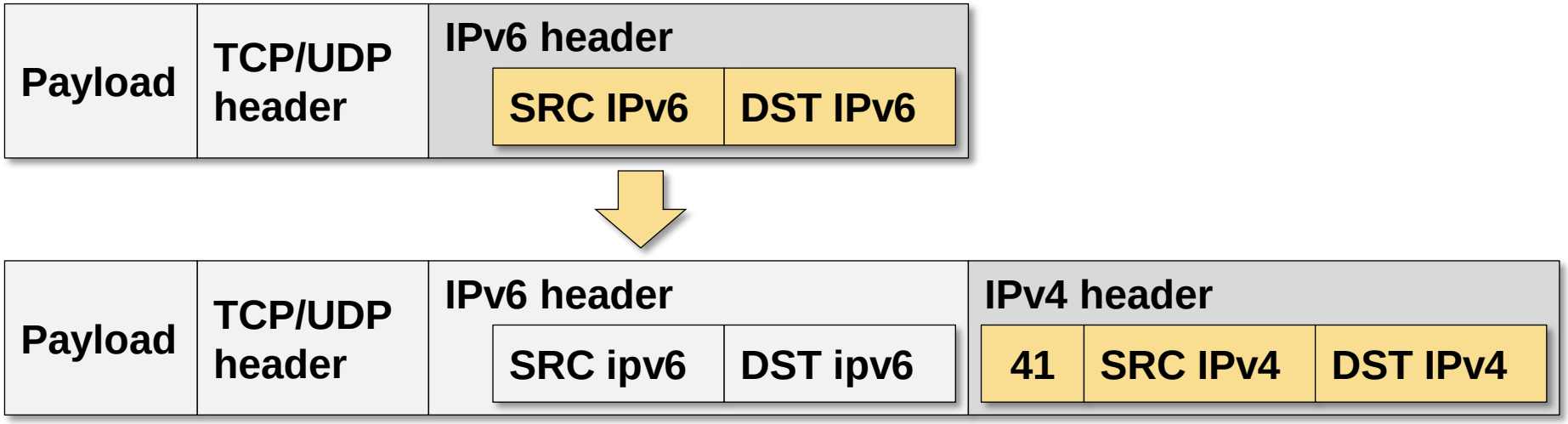
- Uses 6in4 encapsulation
- Configurable IPv4-address-to-IPv6-prefix algorithmic mapping
- 6rd on CPE configured manually or through DHCPv4 6rd option

6in4 Encapsulation



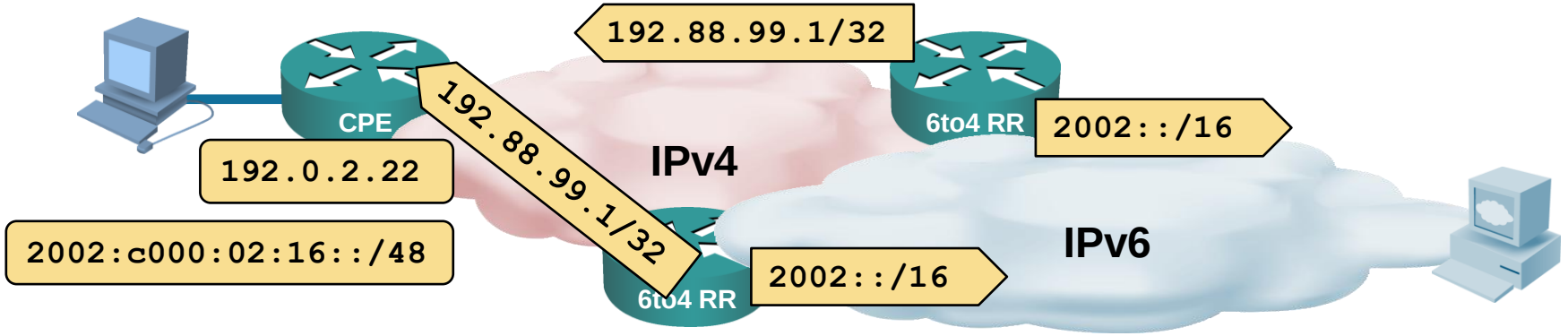
- Manually configured IPv6 addresses
- Manually configured tunnel endpoints (IPv4 addresses)
- Identical to unnumbered P2P GRE tunnels (different encapsulation mechanism)

6to4 Transition Mechanism



- Encapsulation identical to 6in4
- Well known prefix 2002::/16 used for 6-to-4 IPv6 addresses
- Every public IPv4 address corresponds to a /48 IPv6 prefix
- Destination IPv4 address for off-6to4 destinations = 192.88.99.1 (6to4 relay)
- Anycasted relay routers (192.88.99.1 in IPv4, 2002::/16 in IPv6)

6to4 in Action – Addressing and Routing



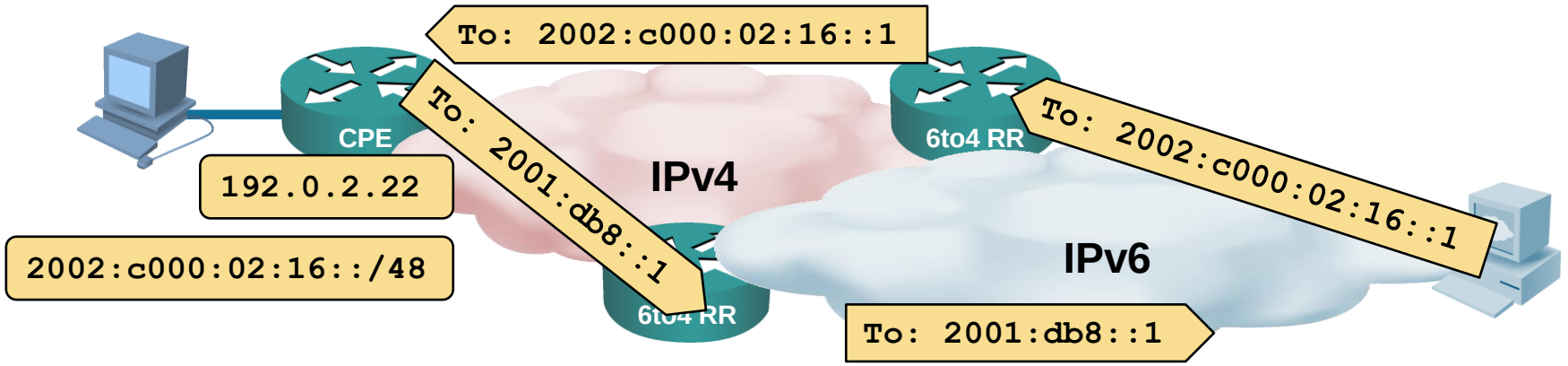
6to4 border router (CPE):

- Receives public IPv4 address
- Computes corresponding /48 prefix
- Allocates smaller IPv6 prefixes to IPv6 interfaces as needed

6to4 relay routers:

- Advertise 192.88.99.1/32 into IPv4 Internet
- Advertise 2002::/16 into IPv6 Internet

6to4 in Action – Packet Forwarding



6to4 border router (CPE):

- IPv6 packets to other 6to4 prefixes sent directly
- All other IPv6 packets sent to the closest 6to4 relay router

Other IPv6 hosts

- Send IPv6 packets for hosts in 2002::/16 to nearest 6to4 relay
- 6to4 relay forwards IPv4-encapsulated packets to 6to4 border router

Asymmetric routing results in performance issues (ISP cannot control the return path GW)

6to4 Drawbacks

- No control over 6to4 relays
- Asymmetric routing due to anycasts
- Very large address space consumption (/16)

6rd improvements:

- Uses PA address space (directly under SP control)
- Configurable delegated prefix length
- Configurable border router IPv4 address
- High-order IPv4 bits can be ignored, further reducing IPv6 address requirements

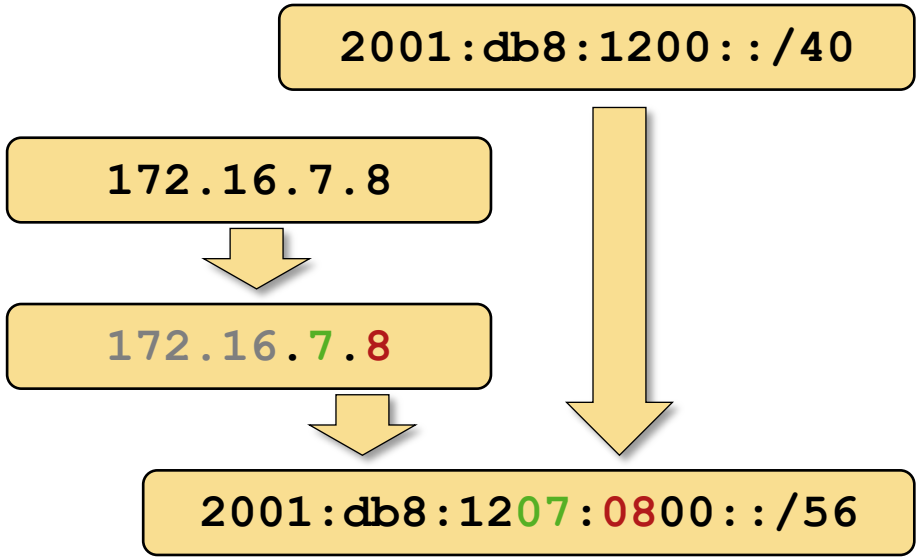
6rd Configuration Parameters

Parameter name	Description	6to4 value
IPv4MaskLen	Number of high-order IPv4 bits ignored during IPv6 prefix computation	0
6rdPrefix	IPv6 prefix used for 6rd delegated prefixes	2002::
6rdPrefixLength	Length of 6rdPrefix	16
6rdBRIPv4Address	IPv4 (anycast) address of the border router	192.88.99.1

- All four parameter must be configured on Customer Edge (CE) routers and 6rd Border Routers (BR)
- The CE and BR parameter values must match
- DHCPv4 6rd option can be used to pass the parameters to CE routers

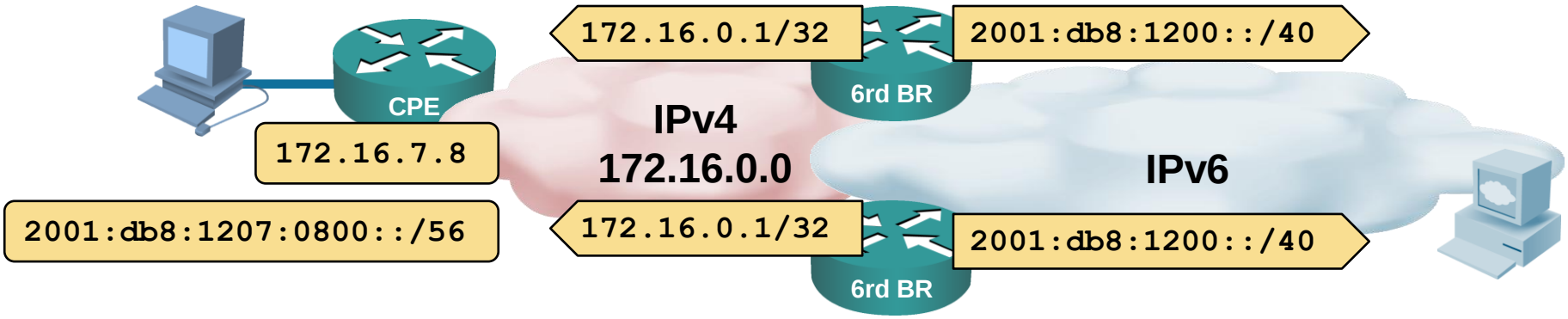
6rd Prefix Mapping Example

Parameter name	Value
IPv4MaskLen	16
6rdPrefix	2001:db8:1200::
6rdPrefixLength	40
6rdBRIPv4Address	172.16.0.1




```
interface Tunnel1
 tunnel source FastEthernet0/1
 tunnel mode ipv6ip 6rd
 tunnel 6rd prefix 2001:db8:1200::/40
 tunnel 6rd ipv4 prefix-len 16
```

Sample 6rd Network

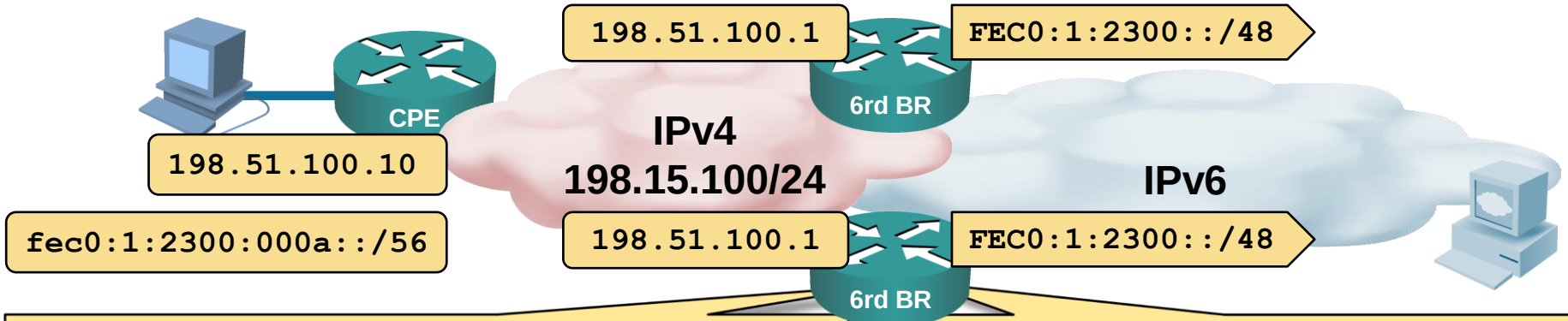


- 6rd address space belongs to the ISP
- Redundant 6rd border routers advertise /32 IPv4 prefix and 6rd Ipv6 prefix
- 6rd is stateless ➔ easy load balancing and failover
- Multiple tunnel interface configured on 6rd BR (one per IPv4 address range)

Worst case: CPE devices don't support IPv4MaskLen parameter

- Whole IPv4 Internet must be mapped into the 6rd prefix
- A /24 prefix is required to give each user a /56 prefix (/28 for /60 prefixes)

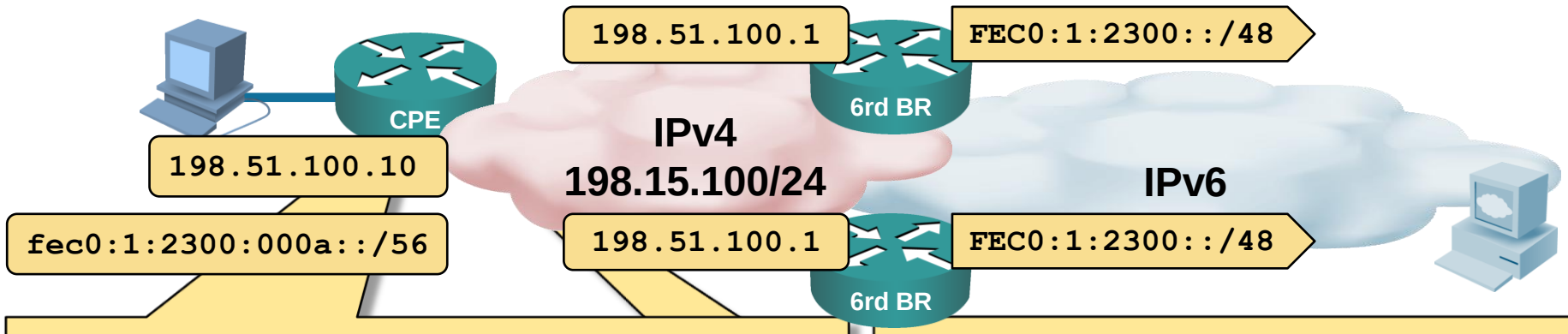
BR Configuration




```
ipv6 general-prefix 6rdPrefix 6rd Tunnel0
!
interface Tunnel0
 ipv6 address 6rdPrefix ::/128 anycast
 tunnel source 198.51.100.1
 tunnel mode ipv6ip 6rd
 tunnel 6rd ipv4 prefix-len 24
 tunnel 6rd prefix FEC0:1:2300::/48
!
interface FastEthernet0/0.100
 description Access network (VLAN 100)
 ip address 198.51.100.3 255.255.255.0
 glbp 1 ip 198.51.100.1
```

```
router bgp 65000
 address-family ipv6
  network FEC0:1:2300::/48
!
ipv6 route FEC0:1:2300:100::/56 Null0
ipv6 route FEC0:1:2300:300::/56 Null0
!
ipv6 route FEC0:1:2300::/48 Tunnel0
```

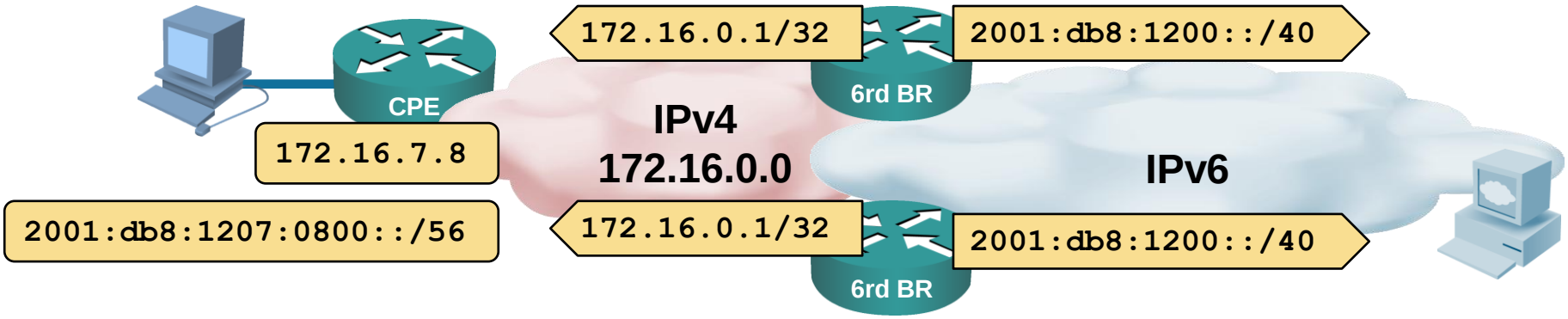
CPE Configuration




```
ipv6 general-prefix 6rdPrefix 6rd Tunnel0
!
interface Tunnel0
 ipv6 address 6rdPrefix ::/128 anycast
 tunnel source FastEthernet0/1
 tunnel mode ipv6ip 6rd
 tunnel 6rd ipv4 prefix-len 24
 tunnel 6rd prefix FEC0:1:2300::/48
 tunnel 6rd br 198.51.100.1
!
interface FastEthernet0/0
 description Inside interface
 ipv6 address DHPrefix ::1:0:0:0:1/64
```

```
!
interface FastEthernet0/1
 description Outside interface
 ip address dhcp client-id FastEthernet0/1
!
ipv6 route FEC0:1:2300::/48 Tunnel0
ipv6 route ::/0 Tunnel0 FEC0:1:2300:100::
```

6rd Use Cases



IPv4-only access network

- IPv6 over third-party access network (wholesale access solutions)
- IPv6 over legacy equipment (BRAS, transport or DSLAM not IPv6-capable)

IPv4 security mechanisms for IPv6 networks

- 6rd solves first-hop security issues with well-known IPv4 mechanisms

Conclusions

Conclusions

Addressing

- Each client subnet needs a /64
- Clients with CPEs need an *inside* /64 and an *outside* /64
- CPEs with multiple subnets need larger delegated prefixes
- *Outside* client interfaces are addressed with SLAAC or DHCPv6
- Inside CPE subnets are addressed with *delegated prefixes*

Security

- Use first-hop security, layer-3 access switches or VLAN-per-customer

Routing

- Aggregate as much as possible
- You don't need local pools to aggregate
- Use two-level aggregation with BRAS banks

Address/prefix allocation tracking

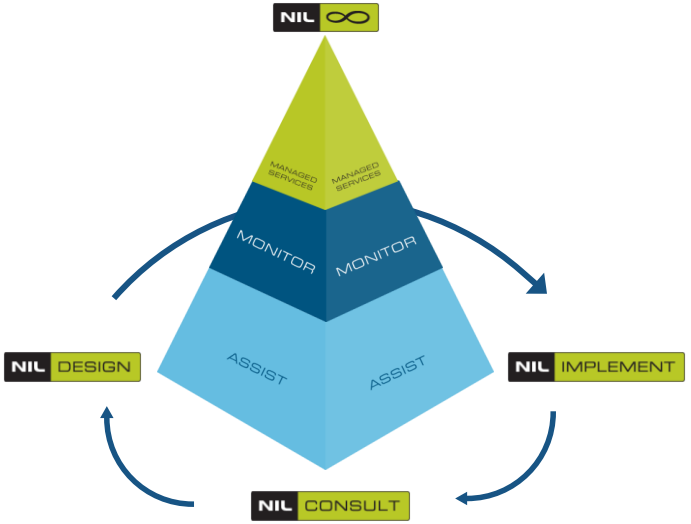
- Central DHCPv6 server or RADIUS
- Disable SLAAC
- Use tightly controlled CPE devices
- Worst case: use 6rd

Need help?

ExpertExpress for quick discussions, reviews or second opinions

NIL's Professional/Learning Services

- In-depth design/deployment projects
- IPv6-, Data Center-, virtualization- and cloud-related training
- Details: www.nil.com, flipit.nil.com



Questions?

Paperwork issues

- Follow-up email
- Please fill in the evaluation form (waiting in your browser)
- Recording available within 24 hours
- PDF materials always available for download
- Discount for future webinars – use wlp10 discount code
- Upgrade to yearly subscription
- Please spread the word!