



Transition from IPv4 to IPv6

Mario Baldi – Fulvio Rizzo

Politecnico di Torino





Nota di Copyright



This set of transparencies, hereinafter referred to as slides, is protected by copyright laws and provisions of International Treaties. The title and copyright regarding the slides (including, but not limited to, each and every image, photography, animation, video, audio, music and text) are property of the authors specified on page 1.

The slides may be reproduced and used freely by research institutes, schools and Universities for non-profit, institutional purposes. In such cases, no authorization is requested.

Any total or partial use or reproduction (including, but not limited to, reproduction on magnetic media, computer networks, and printed reproduction) is forbidden, unless explicitly authorized by the authors by means of written license.

Information included in these slides is deemed as accurate at the date of publication. Such information is supplied for merely educational purposes and may not be used in designing systems, products, networks, etc. In any case, these slides are subject to changes without any previous notice. The authors do not assume any responsibility for the contents of these slides (including, but not limited to, accuracy, completeness, enforceability, updated-ness of information hereinafter provided).

In any case, accordance with information hereinafter included must not be declared.

In any case, this copyright notice must never be removed and must be reported even in partial uses.





Introduction





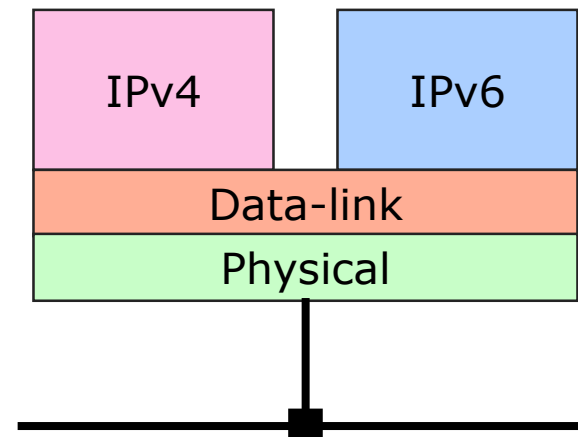
Assumption (and problem to solve)

IPv4 and IPv6 will coexist
(at least for a while)



Dual Stack Approach

- Both IPv4 and IPv6 capabilities
 - In all hosts and routers supporting IPv6
 - IPv4 support can be (gradually) removed (and included in new hosts) once all hosts have IPv6
- Hosts communicate natively with both
- Complete duplication of all protocol stack components
 - Routing protocols
 - Routing table
 - Access lists





Dual Stack Limitations

- It does not reduce the need for IPv4 addresses
 - Each host still needs an IPv4 address to use IPv4
- Applications have the responsibility whether to use IPv4 or IPv6





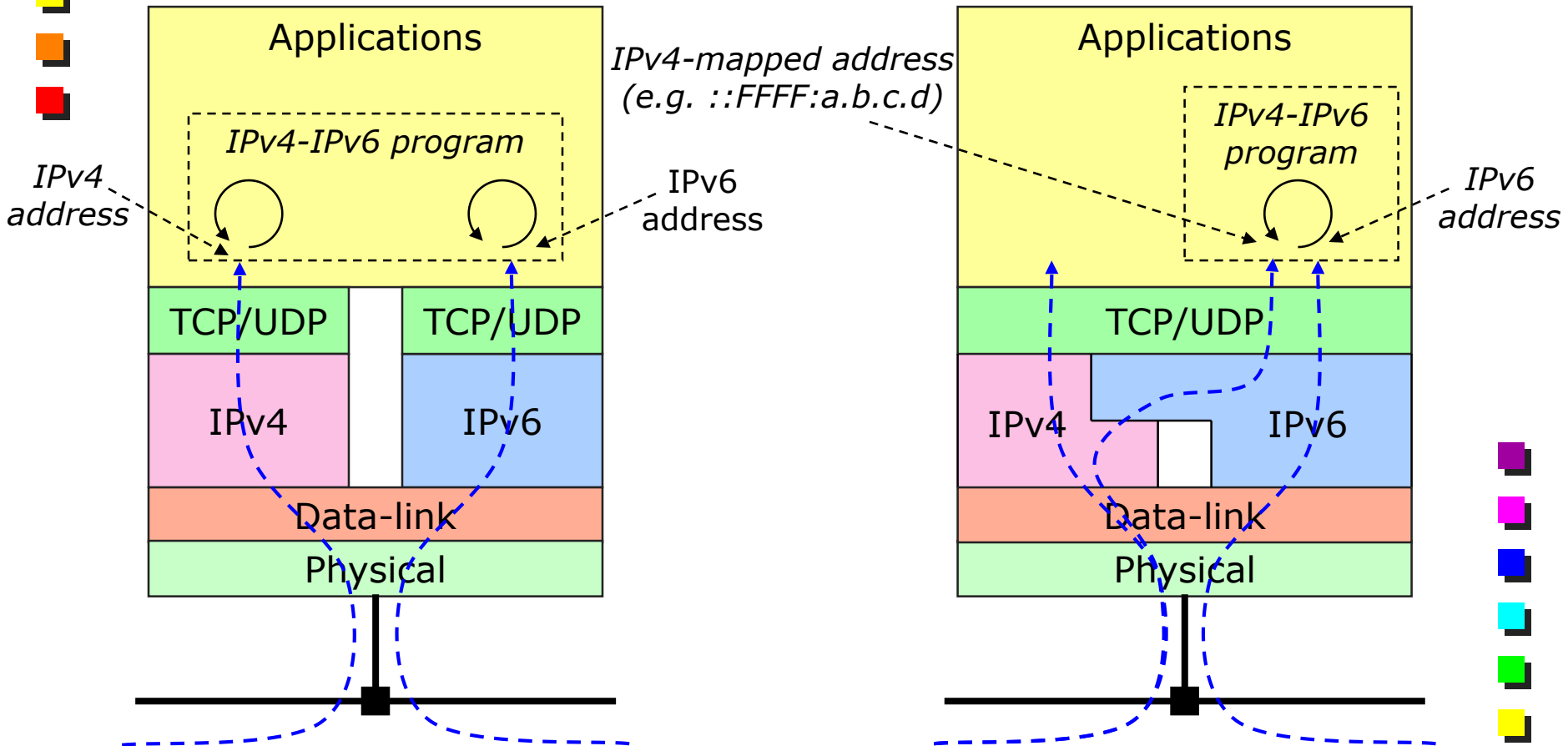
Dual Layer Approach

- An alternative
- Applications are not responsible for choosing the protocol to use
- Less modifications in the applications

Less common than dual stack




Dual stack vs. Dual Layer



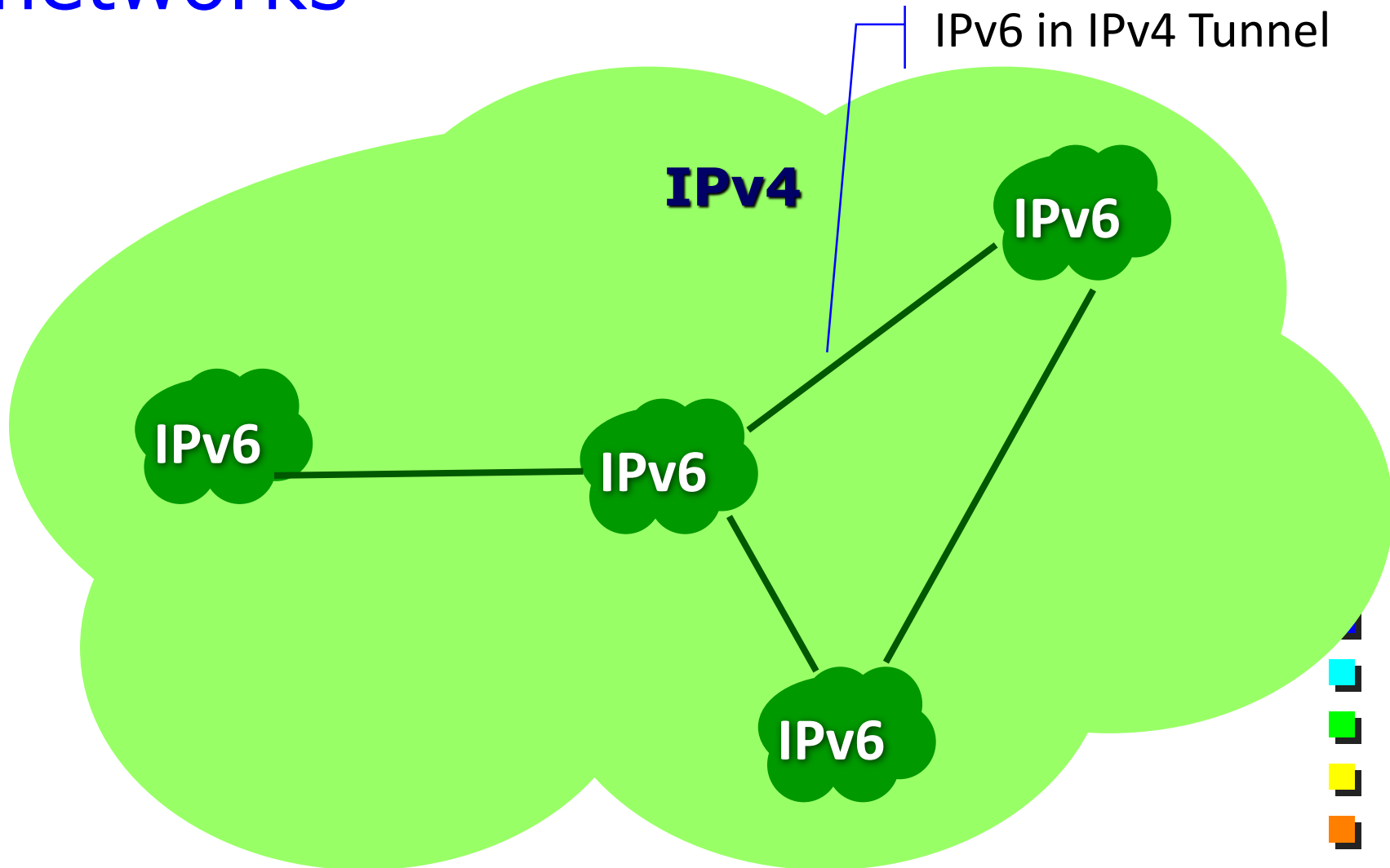


Not all hosts will be dual stack

- IPv6 hosts shall communicate with IPv6 hosts through an IPv4 network
 - Same for IPv4 hosts through an IPv6 network
 - IPv6 hosts shall communicate with IPv4 hosts
 - Translation mechanisms must be used
 - Not targeting IPv4 hosts contacting IPv6 ones
 - Difficult to map the large IPv6 address space on smaller IPv4 address space
- 



First stage: isolated IPv6 networks



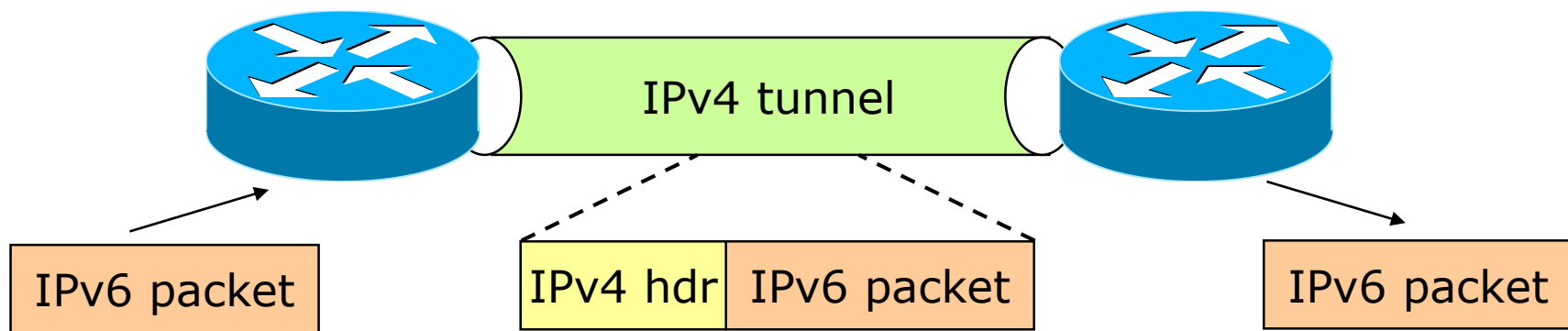


Traversing a IPv4-only Network

■ Tunnelling

- Encapsulation of IP (v6) packets into IP (v4) packets

■ Emulates a “direct” link among IPv6 devices

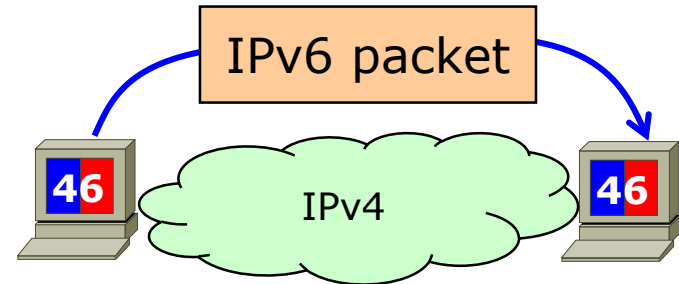




Tunnelling

- End points: hosts and routers
- Protocols
 - GRE (Generic Routing Encapsulation)
 - IPv6 in IPv4
 - Protocol type = 41
- Set up: manual and automatic
 - IPv4-compatible addresses, 6over4 (RFC 2529), 6to4, Tunnel Broker (RFC 3053), ISATAP, Teredo





Host-centered Solutions

Dual stack hosts exchange IPv6 packets through an IPv4 network






IPv4-compatible Addresses

- Sometimes improperly called “automatic tunneling”
 - Same name is used for other solutions
- IPv4-compatible addresses (:::/96) are used for IPv6 communication





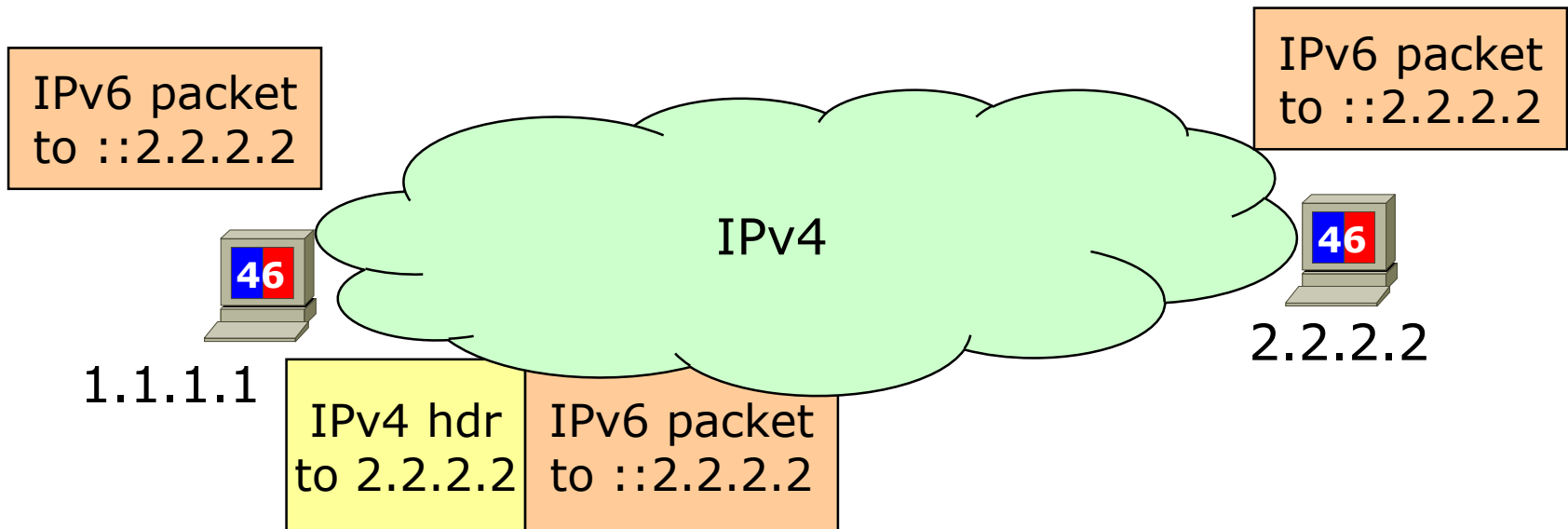
IPv4-compatible Addresses

- Application sends IPv6 packets to IPv6 addresses
 - E.g., ::2.2.2.2
 - Static route forwards packets to ::/96 through pseudo-interface
 - Automatic Tunneling Pseudo-Interface
 - Pseudo-interface encapsulates IPv6 packets in IPv4 packets and sends them out
- 



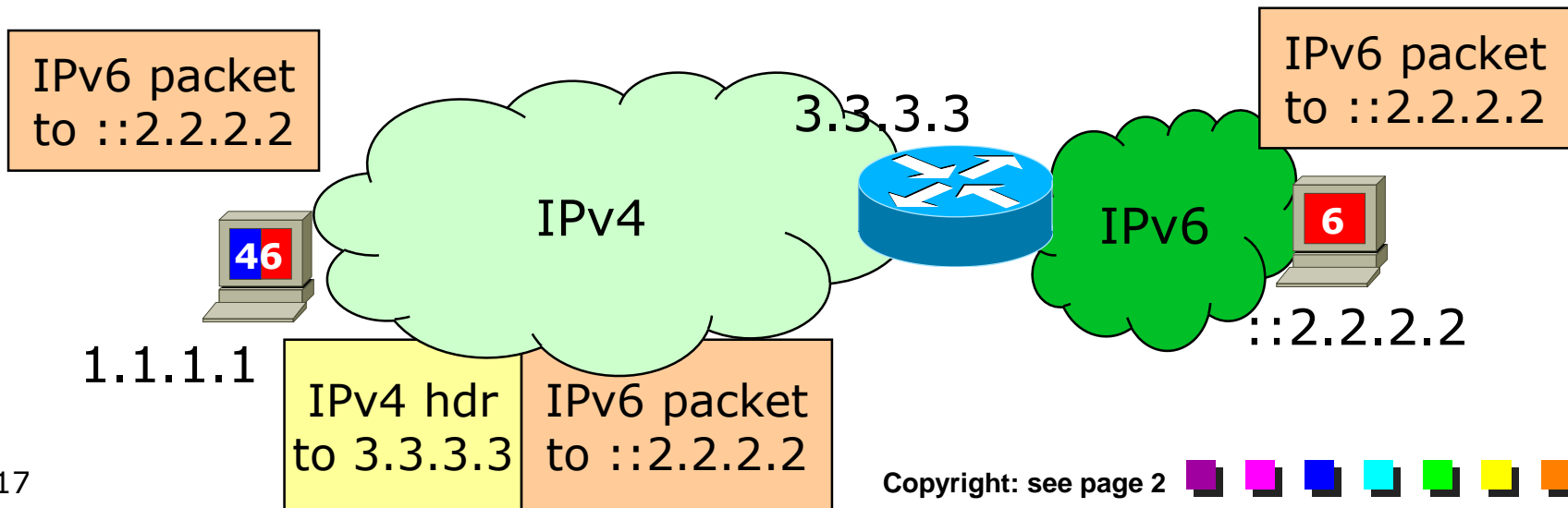
End-to-end Tunnel

Encapsulating IPv4 packets are sent to IPv4 address corresponding to IPv6 destination



IPv4-compatible Address with Dual Stack Router


- Pseudo-interface in sending host is configured to terminate tunnels on router
- Encapsulating IPv4 packets are sent to IPv4 address of dual stack router





6over4


仿真

- IPv4 network emulates a virtual LAN
 - Broadcast multiple access data link
 - IP Multicasting used for the purpose
 - Neighbor and router discovery enabled
 - IPv4 address is used for automatic IPv6 Interface ID generation of link local address
 - Not very used because IPv4 multicast support is not widespread
- 





6over4 Neighbor Discovery

- IPv6 multicast addresses are mapped on IPv4 multicast addresses
 - 239.192.x.y
 - x.y are the last 2 bytes of the IPv6 address
 - An example
 - IPv4 address: 1.1.1.1
 - IPv6 Link local address: fe80::101:101
 - Solicited node multicast address: ff02::1:ff01:101
 - 6over4 multicast address: 239.192.1.1
- 




ISATAP: Intra-site Automatic Tunnel Addressing Protocol

- IPv4 network as Non-Broadcast Multiple Access (NBMA) data link
 - No IP multicast support needed
- Interface ID derived from IPv4 address
 - Prefixed by 0000:5efe
 - E.g., fe80::5efe:0101:0101 for 1.1.1.1






(Lack of) Neighbor Discovery

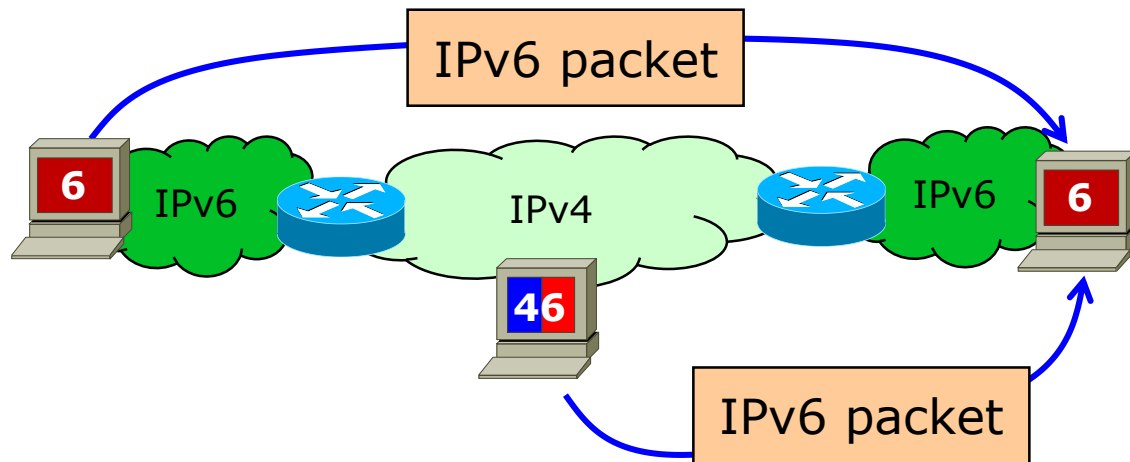
- Not needed for data-link address discovery as IPv4 address is embedded in IPv6 address
 - Last 4 bytes
 - PRL (Potential Router List) must be provided
 - Router discovery not possible
 - By configuration
 - Automatically acquired from DNS
 - Hostname not mandated
 - E.g., isatap.polito.it
- 



Automatic Configuration

- IPv4 address, DNS address and domain name obtained through DHCPv4
 - Generation IPv6 link-local address
 - Interface ID from IPv4 address
 - DNS query to obtain PRL
 - If not provided by DHCPv4 (proprietary)
 - Periodic Router Discovery to each router
 - On-link prefixes for autoconfiguration
- 





Network-centered Solutions

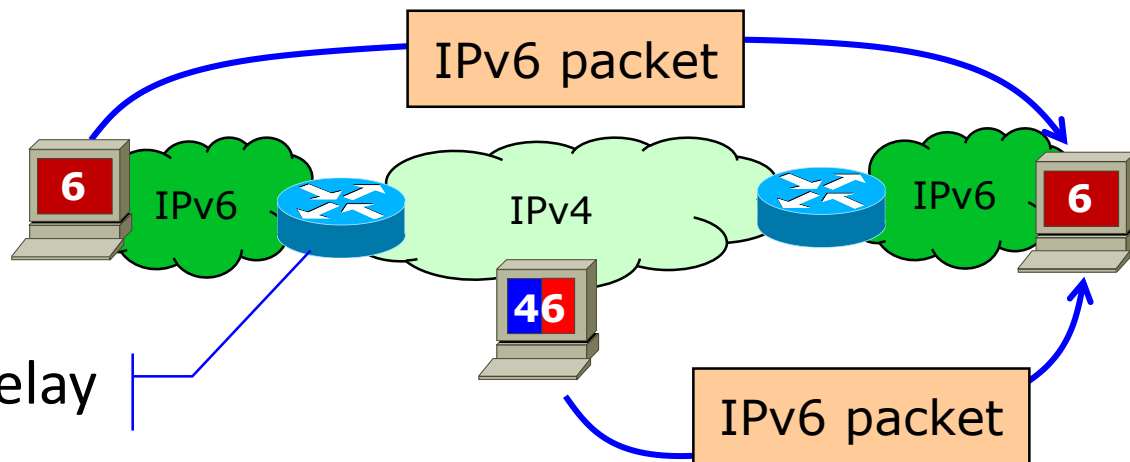
Dual stack and *native IPv6* hosts exchange IPv6 packets through an IPv4 network



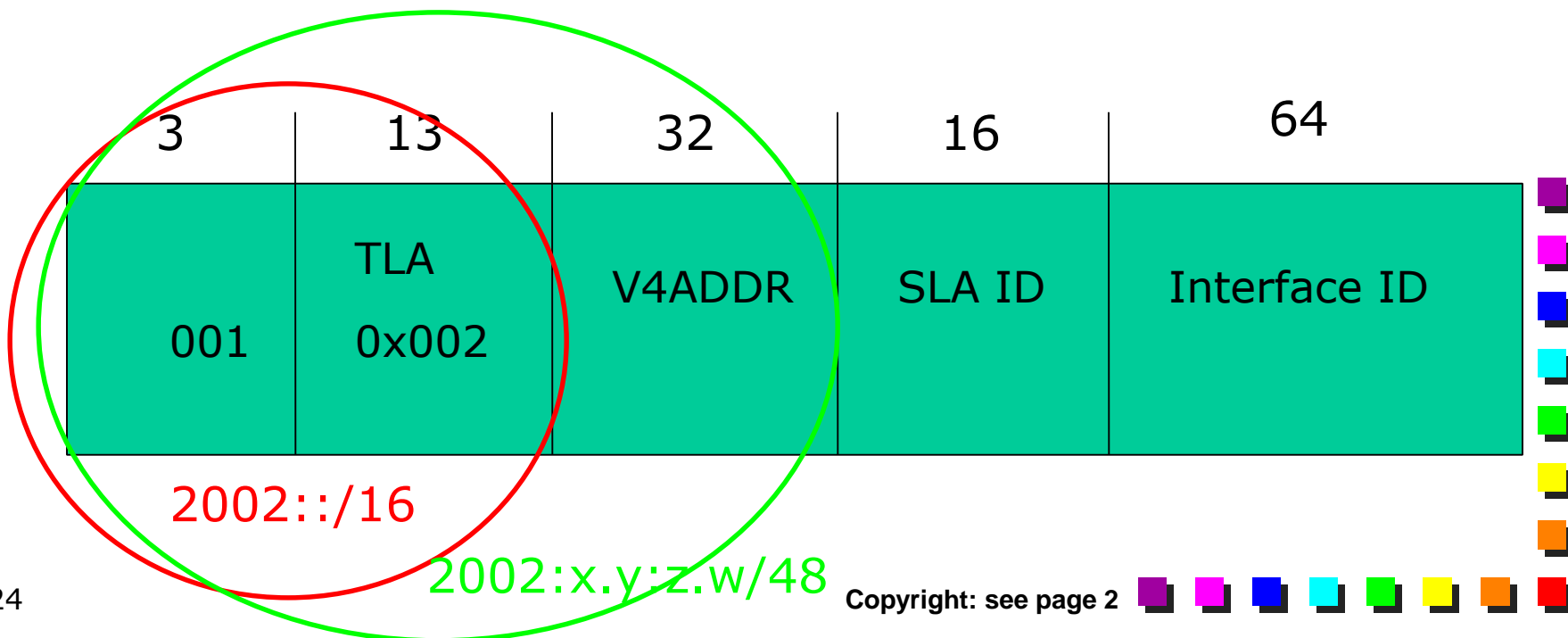


6to4

6to4 router or relay

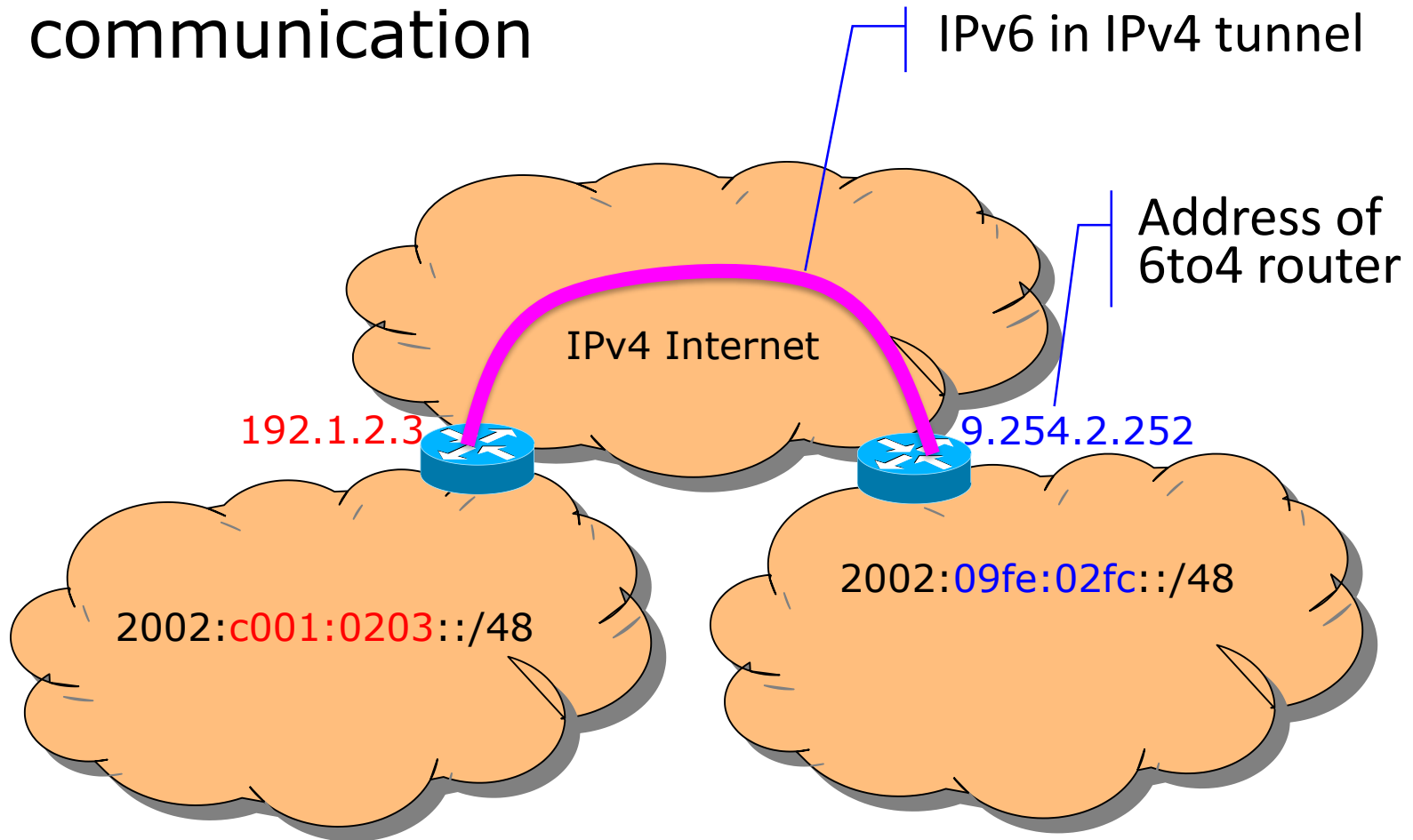


- Relay address embedded in IPv6 prefix



Basic 6to4 Scenario

Not meant for IPv4 host to IPv6 host communication



Mixed 6to4 Scenario

6to4 Relay must be default gateway of 6to4 routers

Address of **6to4 Relay**
(predefined *anycast* address)

192.88.99.1

Global Ipv6 Internet

IPv4 Internet

192.1.2.3

9.254.253.252

2002:c001:0203::/48

2002:09fe:fdfc::/48



Teredo

- Architecture similar to 6to4
- IPv6 packets are encapsulated in IP/UDP to be compatible with NAT
 - NAT cannot use ports with IPv6 in IPv4





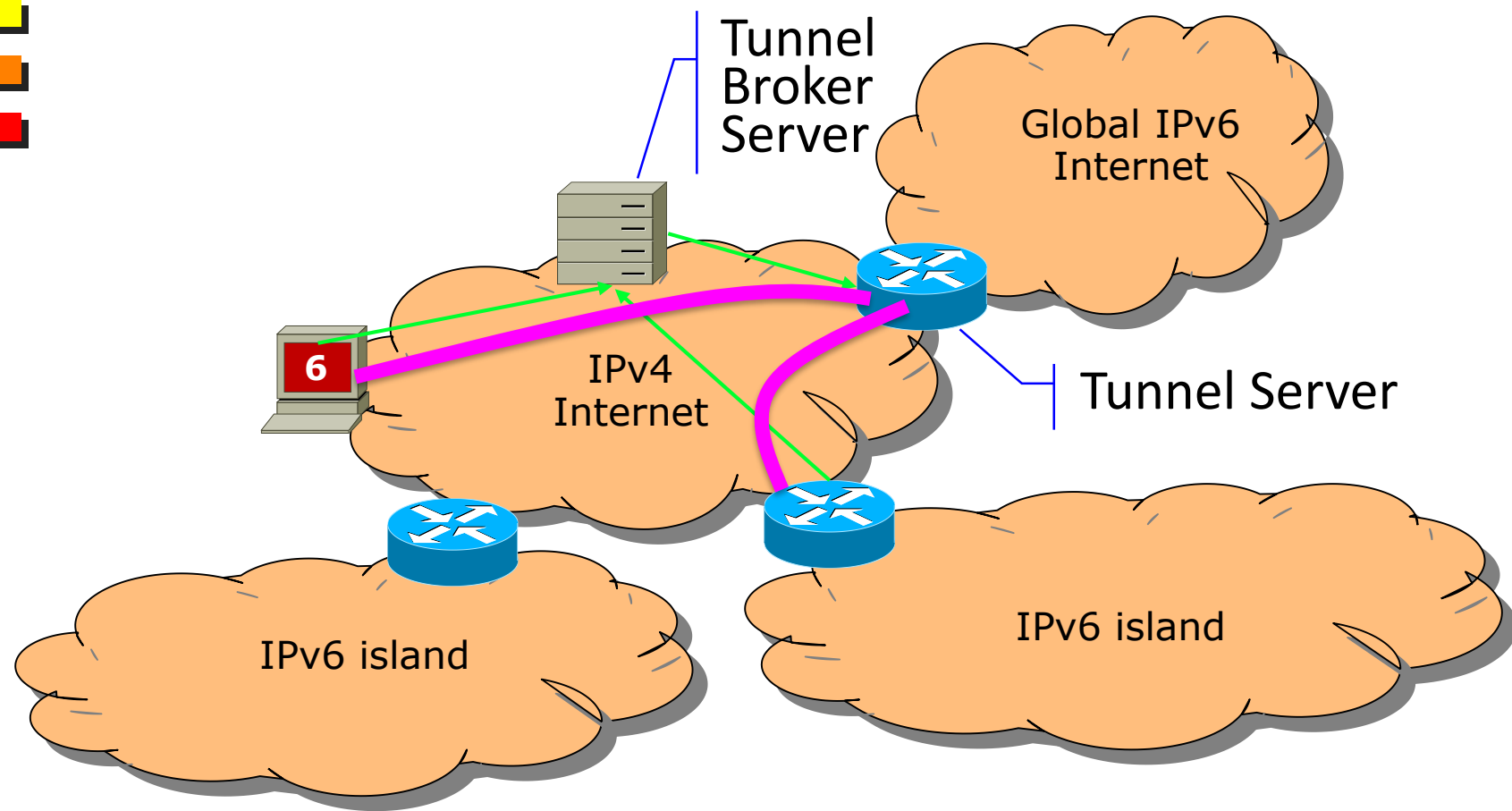
Tunnel Broker

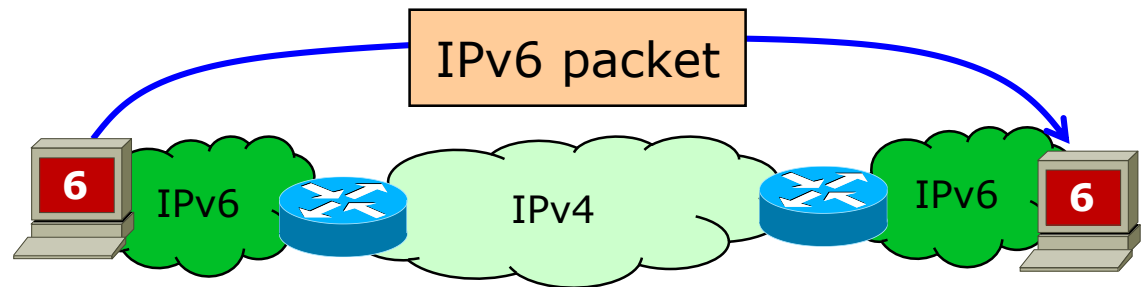
- Communication with a *tunnel broker server*
 - Identifies *tunnel server* and mediates tunnel setup
- IPv6 in IPv4 (a.k.a. proto-41) tunnels
- Tunnel Setup Protocol (TSP) or Tunnel Information Control (TIC) protocol used to setup tunnels





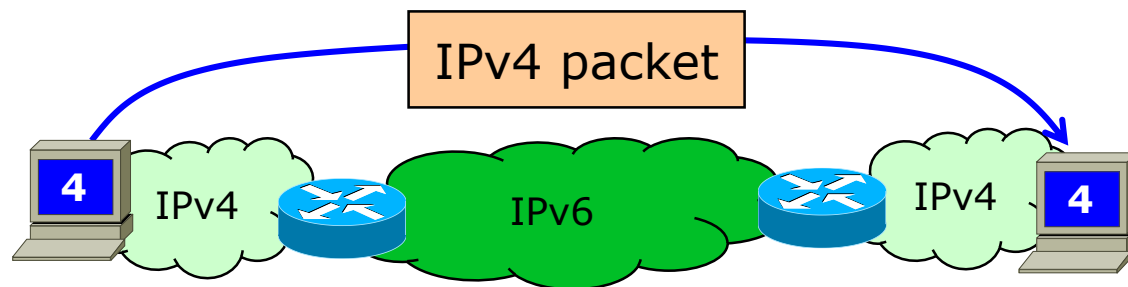
Tunnel Broker Architecture





Scalable, Carrier-grade Solutions

*Native IPv6 (IPv4) hosts
exchange IPv6 (IPv4) packets
through an IPv4 (IPv6) network*






Goals

■ Still need to support

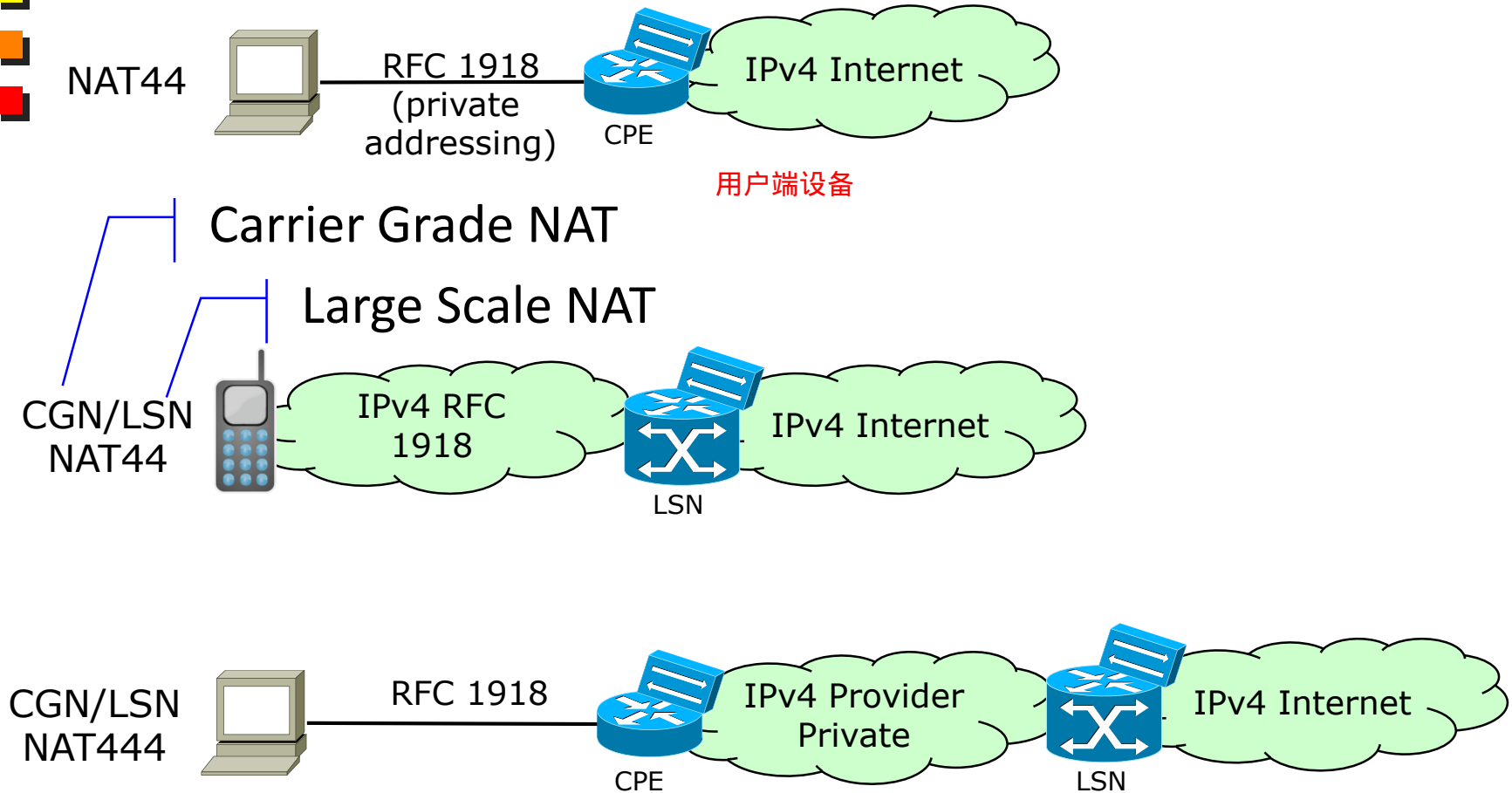
- IPv4 servers possibly communicating with IPv6 hosts
- IPv4 clients

■ Several Options

- DS-Lite
 - A+P (DS-Lite evolution)
 - MAP-T and MAP-E
 - NAT64
 - 6PE (MPLS-based)
- 



NAT is widely used today





How do we like NAT?

- Problematic with inbound sessions

- E.g., servers

- NAT + STUN/TURN may be ok for peer-to-peer sessions

- Bottleneck and single point of failure

Nevertheless

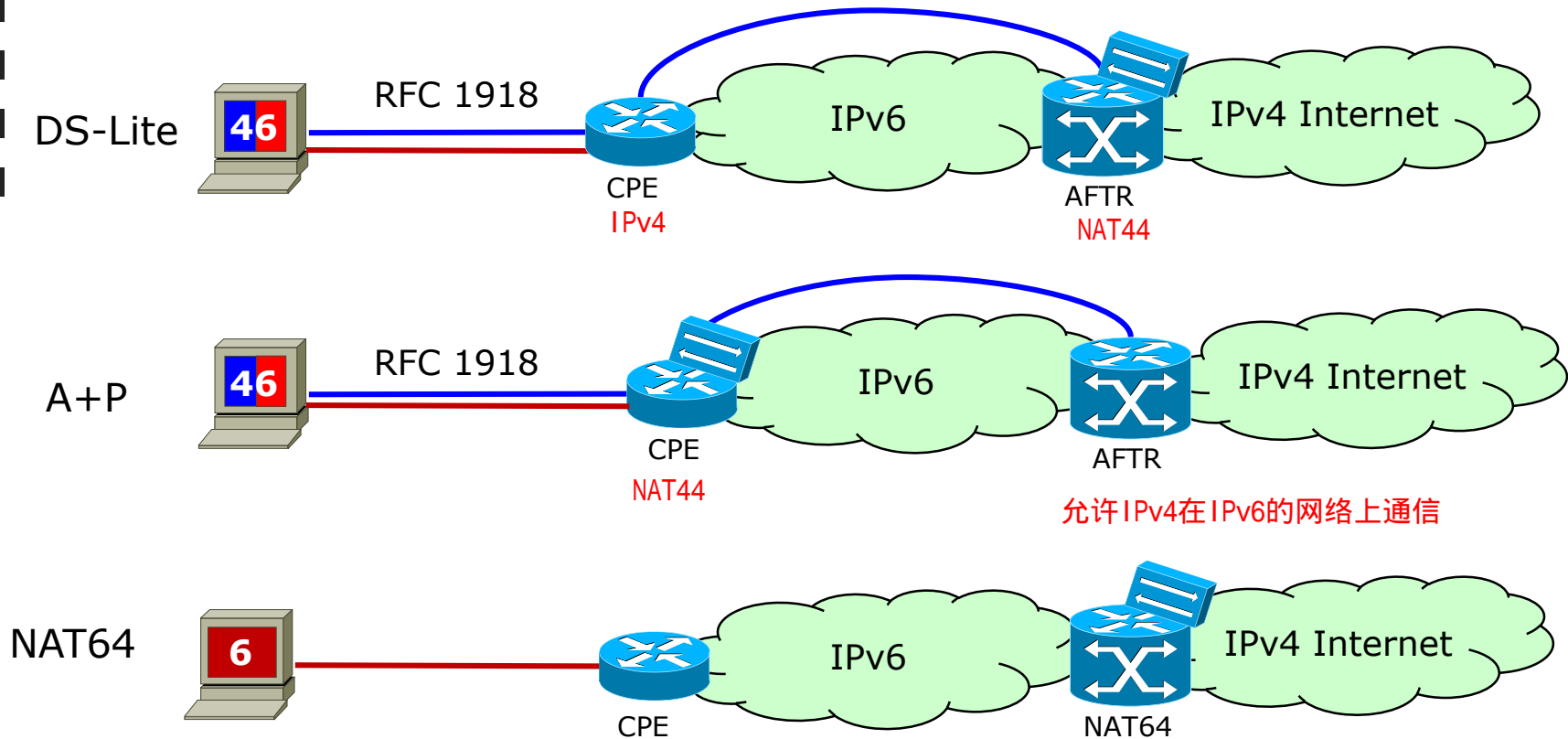
- Several (independent) cascaded instances of NAT are now very common

- Starting from virtual machines

- Difficult to do without due to scarce addresses



Same Architecture with IPv6

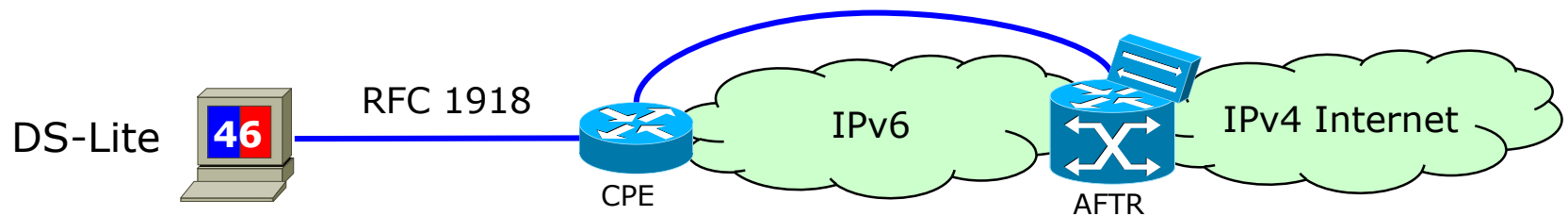


CPE: Customer Premises Equipment
AFTR: Address Family Transition Router

Note where the NAT function is located!

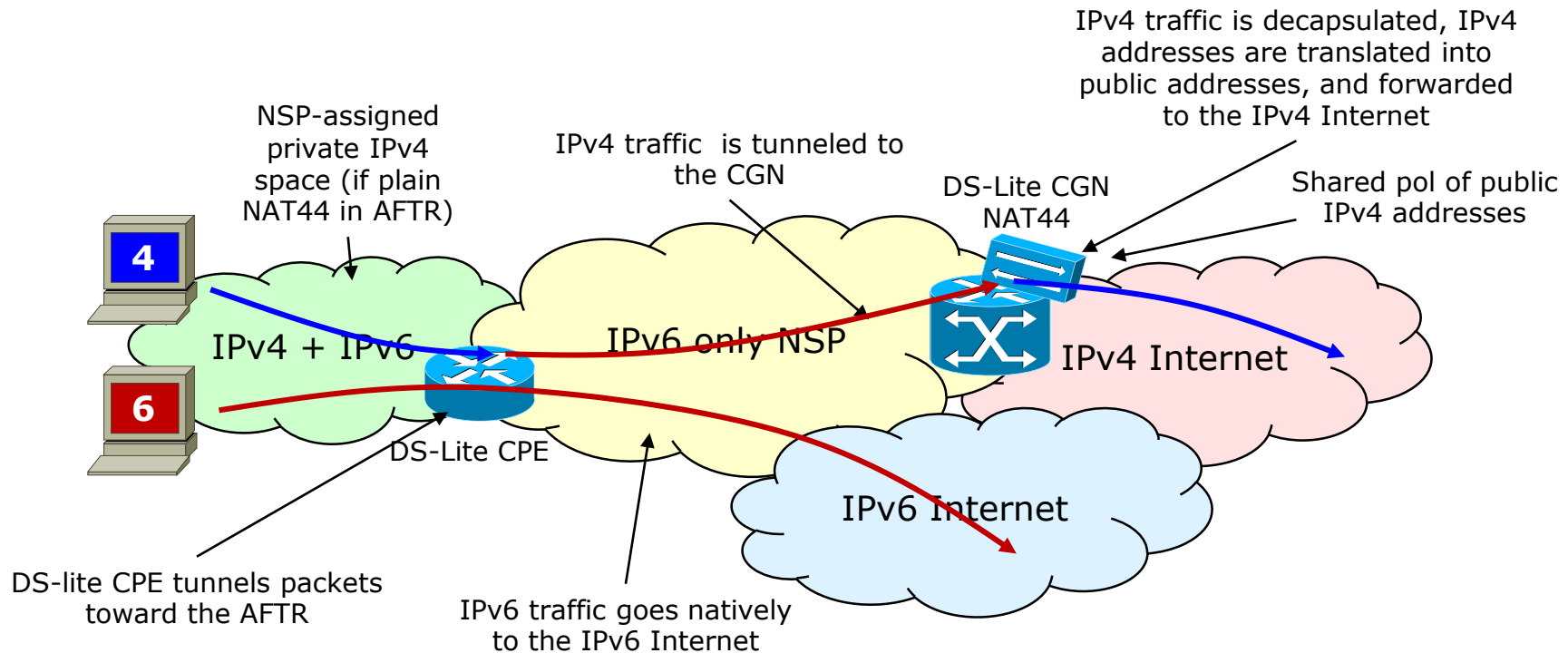
AFTR: Address Family Transition Router

- Allows IPv4 host to communicate with IPv4 hosts over an IPv6 carrier infrastructure
 - Residential host and current providers
- Features an IPv6 tunnel concentrator and possibly a large-scale NAT
- In use in DS-Lite and A+P



DS-Lite (Dual-Stack Lite)

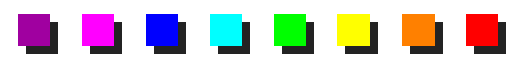
- Dual-stack at the edge
- IPv6-only Service Provider backbone





Properties

- Reduces requirement for IPv4 addresses compared to dual-stack approach
 - Dual-stack requires public IPv4 address per host
- Extended NAT enables customer assigned (i.e., overlapping) addressing
 - IPv6 address of CPE in NAT table





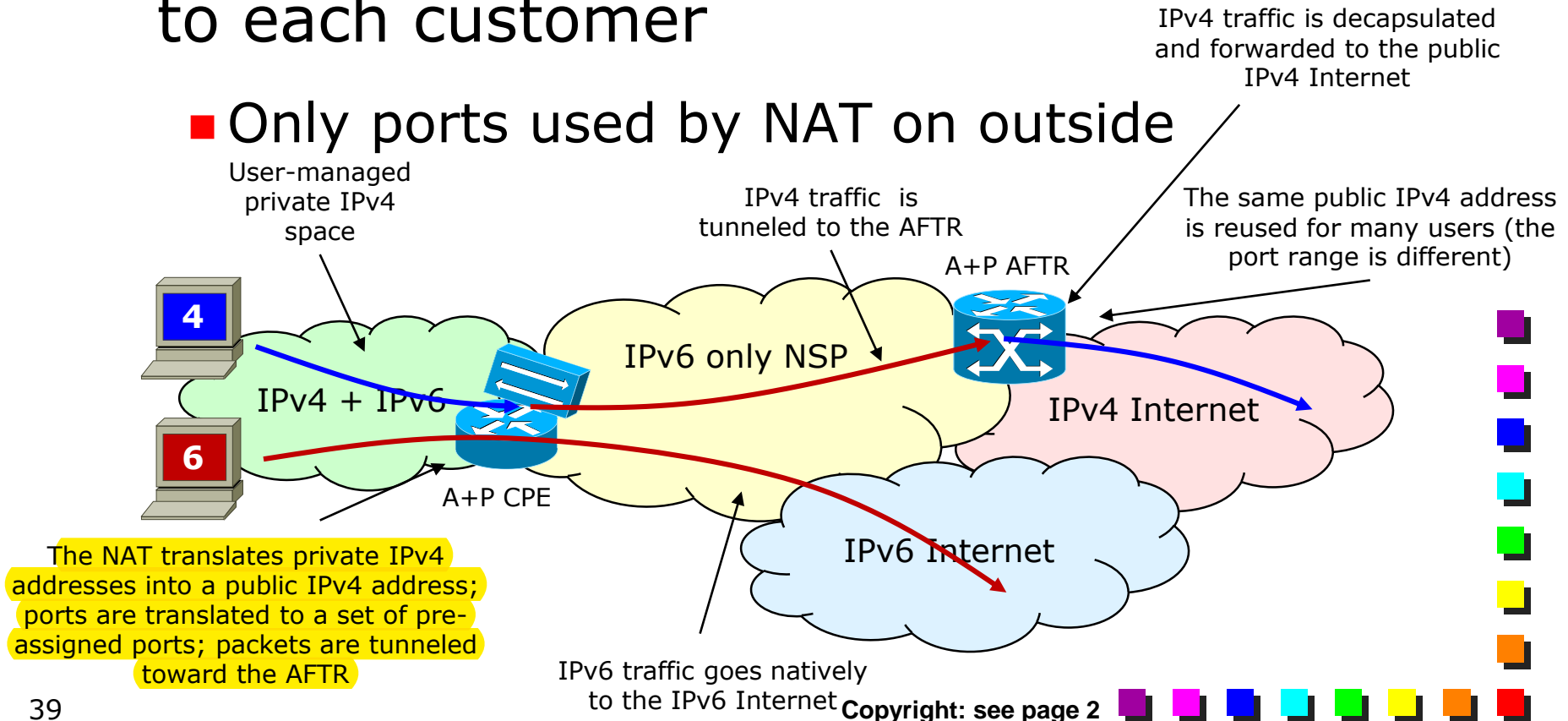
Limitations

- NAT is not under control of customer
 - Same problem as with CGN
- Problematic with servers
 - Static mapping and port forwarding cannot be configured



A+P (Address plus Port)

- NAT is under control of customer
- Ranges of TCP/UDP ports are assigned to each customer
- Only ports used by NAT on outside





Features

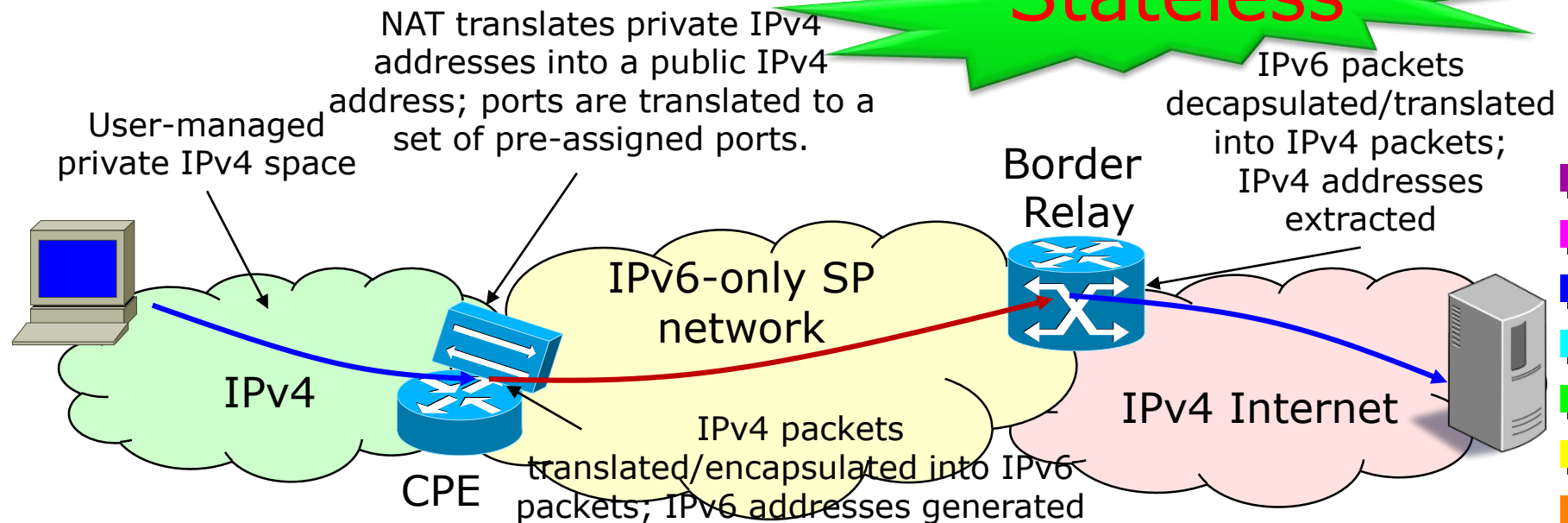
- No problem with overlapping private address spaces at customers'
- Ports can be assigned automatically to CPE using the Port Control Protocol (PCP)
 - CPE can negotiate more ports any time
- AFTR is just a IPv4-in-IPv6 (proto-41) tunnel terminator
 - NAT44 is no longer needed in the AFTR



Mapping Address and Port (MAP)


- Multiple CPEs use same public IPv4 address
- Set (not range) of ports assigned to CPE

Stateless





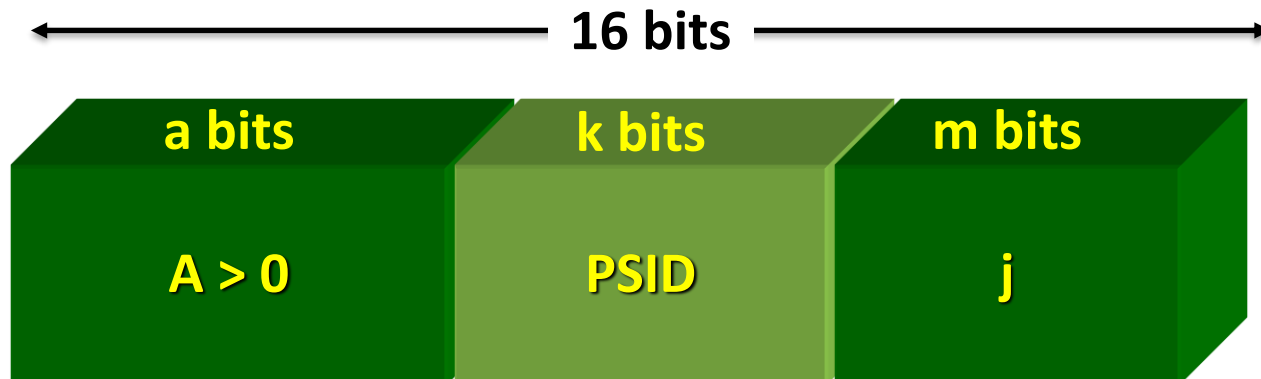
Mapping Address and Port (MAP)

- Client IPv4 address and port set mapped to unique IPv6 address
 - Prefix routed to CPE
 - IPv4 public server address mapped to unique IPv6 address
 - Prefix routed to Border Relay
 - MAP-E: MAP with Encapsulation
 - IPv4 packets are tunneled
 - MAP-T: MAP with Translation
 - IPv4 packets are translated into IPv6 packets and then back to IPv4
- 



Port Set

- Each CPE is assigned a unique PSID (Port Set Identifier) and public IPv4 address pair
- CPE uses port numbers with that PSID value, varying A and j
 - $A > 0$ to avoid static port numbers

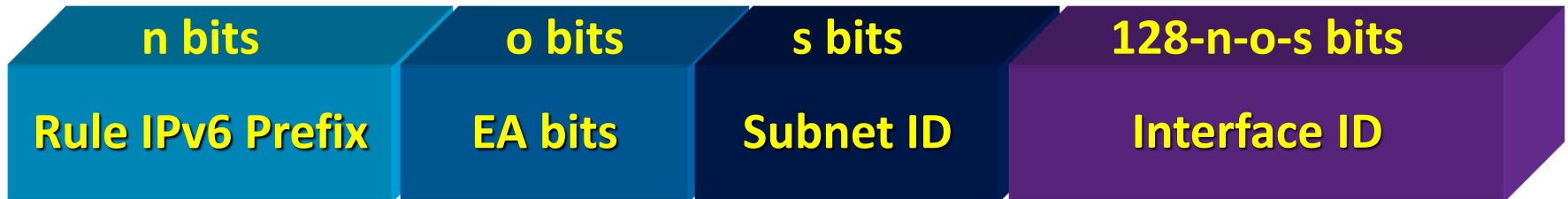




CPE IPv6 Address

- EA (Embedded Address) bits contain PSID and (partial) IPv4 address
 - Uniquely identifies the CPE

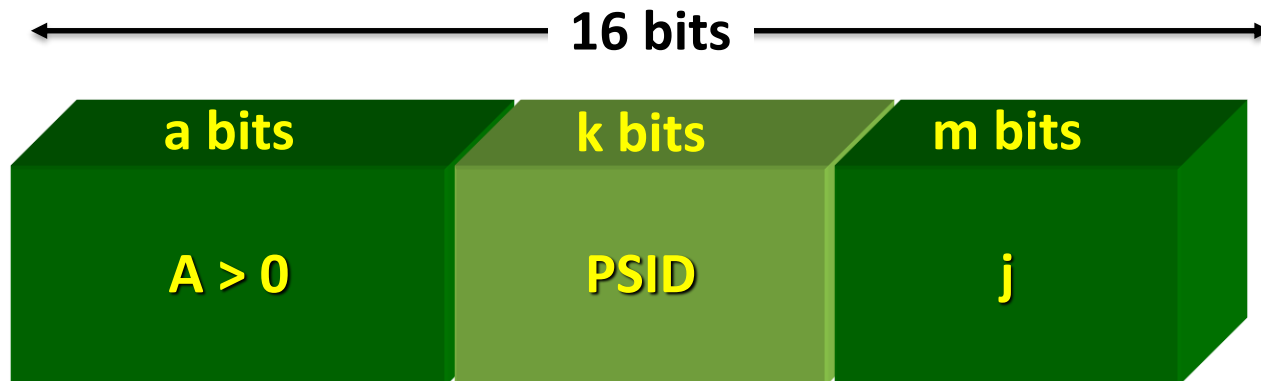
← End-user IPv6 prefix →



Mapping Rule

- Rule IPv6 prefix
- Rule IPv4 prefix
- EA bits length

Moreover, a PSID offset (value of a) is set for the whole MAP domain



Sample CPE IPv6 Address

■ Rule IPv6 prefix: 2001:1:1100::/40

■ Rule IPv4 prefix: 195.2.2.0/24

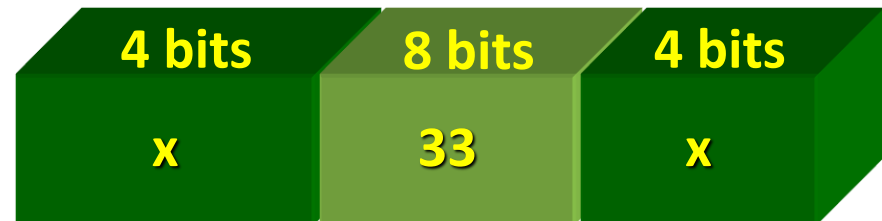
■ EA bit length: 16

■ PSID offset: 4

■ PSID: 0x33


■ Address used by CPE: 195.2.2.4 PSID offset

偏移地址即为子网中的地址



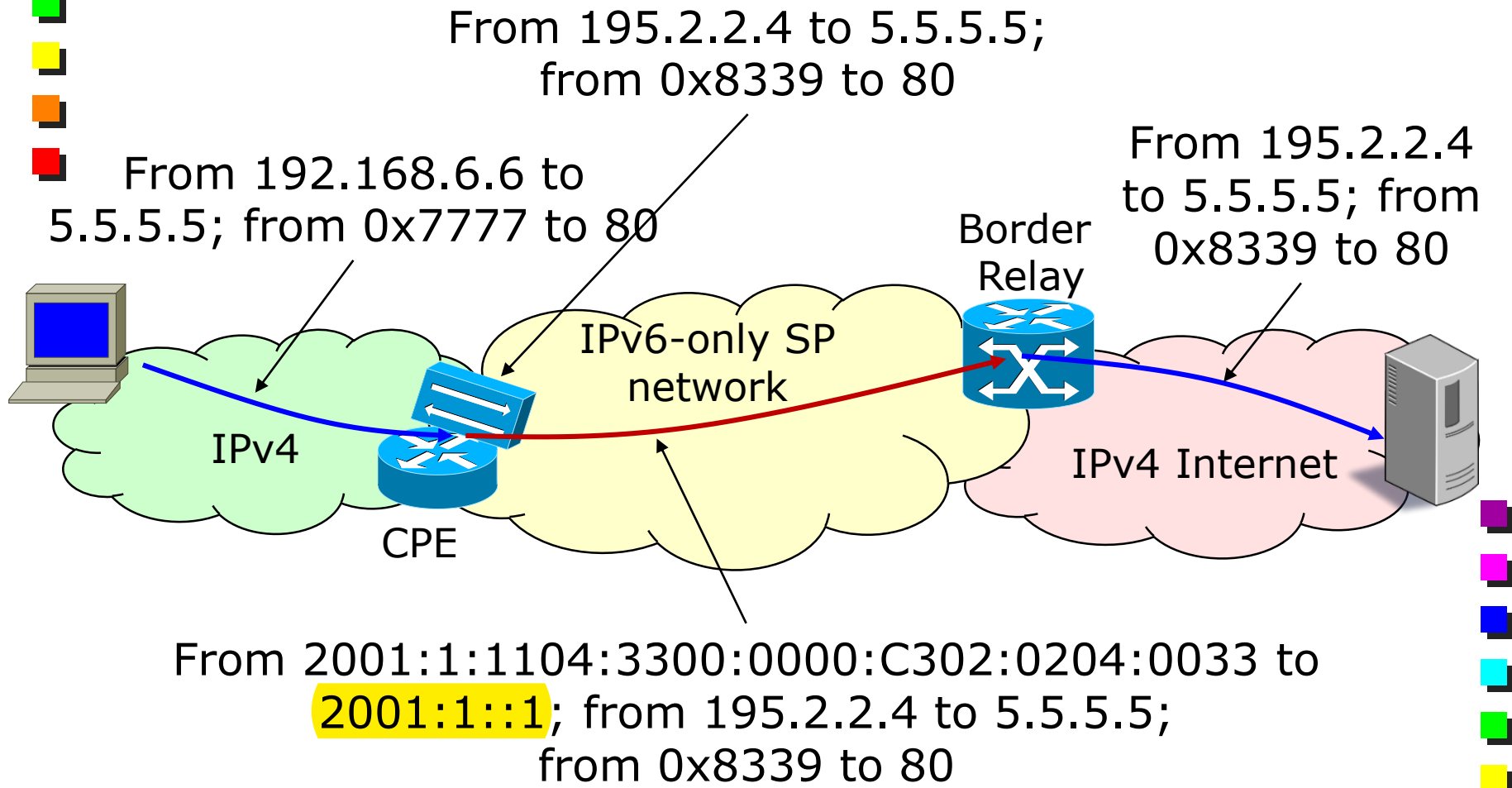


Border Relay (BR)

- BR address must be known to CPEs
 - Multiple BRs might have same address
 - Anycasting
 - MAP-E: BR address terminates tunnel
 - MAP-T: prefix associated to BR used for translation of outside IPv4 addresses
 - BR prefix is advertised on the backbone
 - Might be advertised by multiple BRs
- 



Life of a Packet with MAP-E



Outside IPv4 Destinations

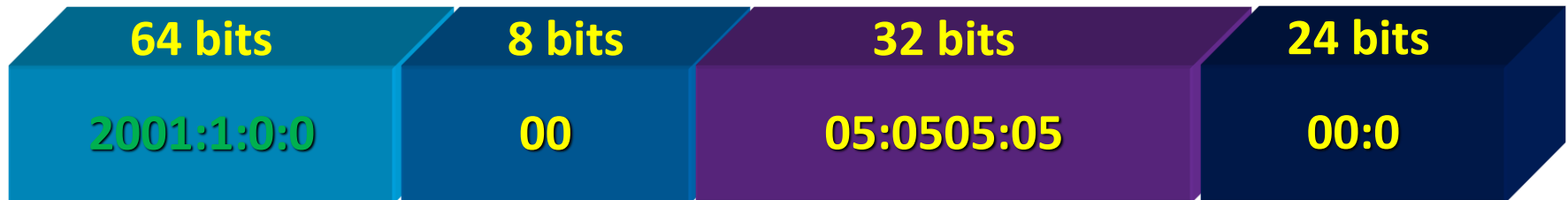
■ IPv4-embedded IPv6 address

■ RFC 6052



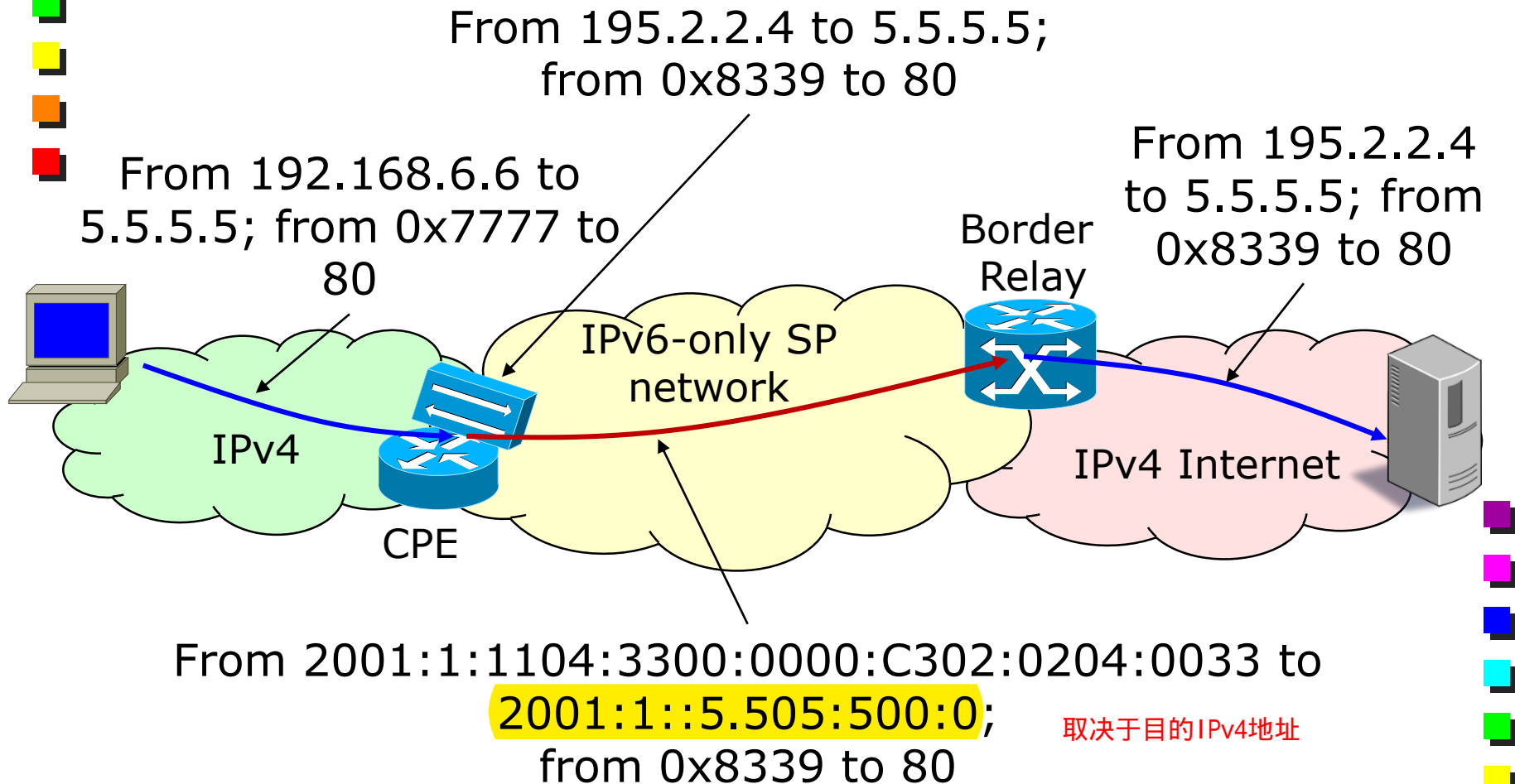
■ BR prefix: 2001:1::/64

■ Outside destination: 5.5.5.5

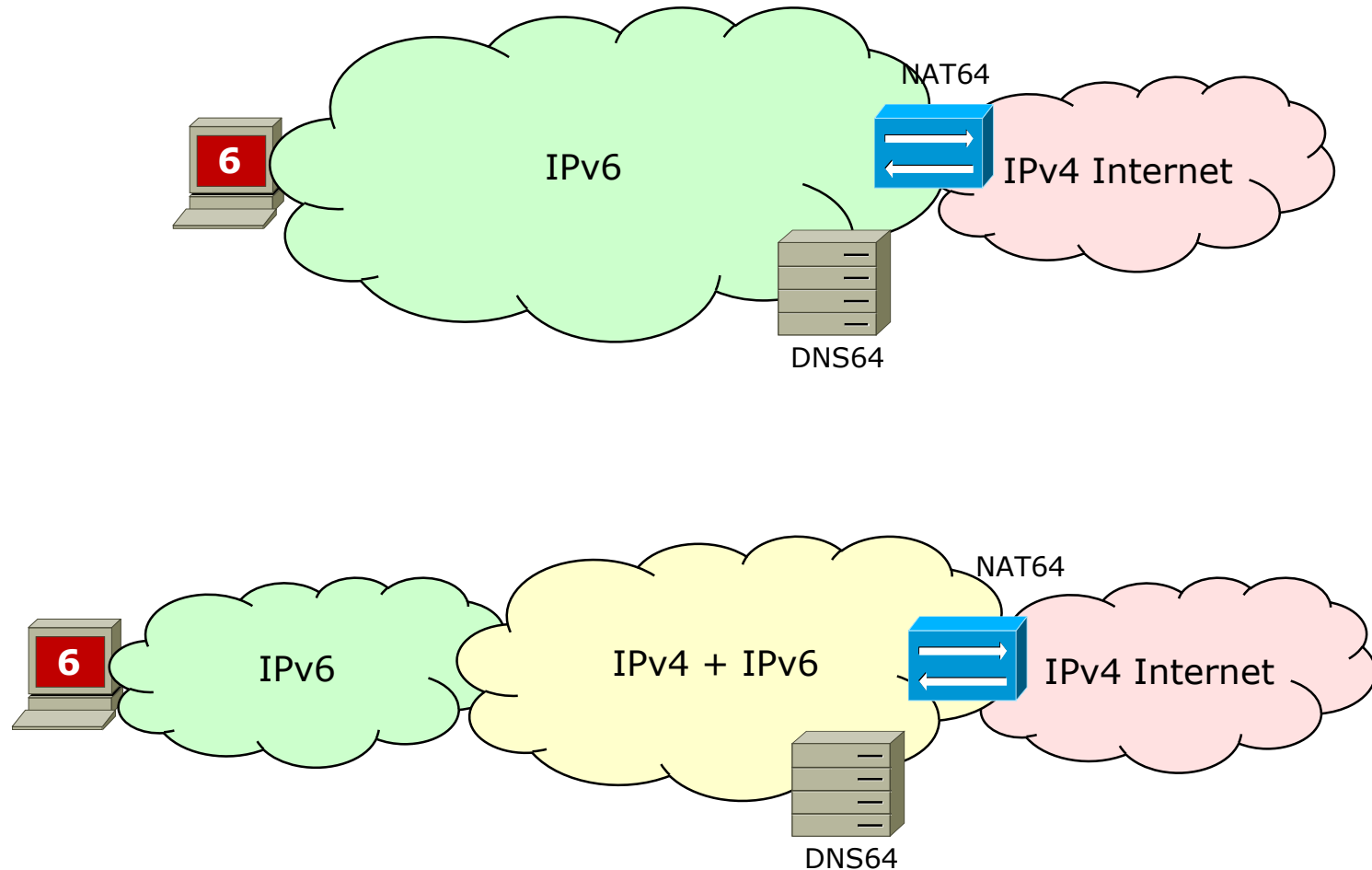


■ 2001:1::5.505:500:0

Life of a Packet with MAP-T




NAT64 + DNS64: Deployment Scenarios



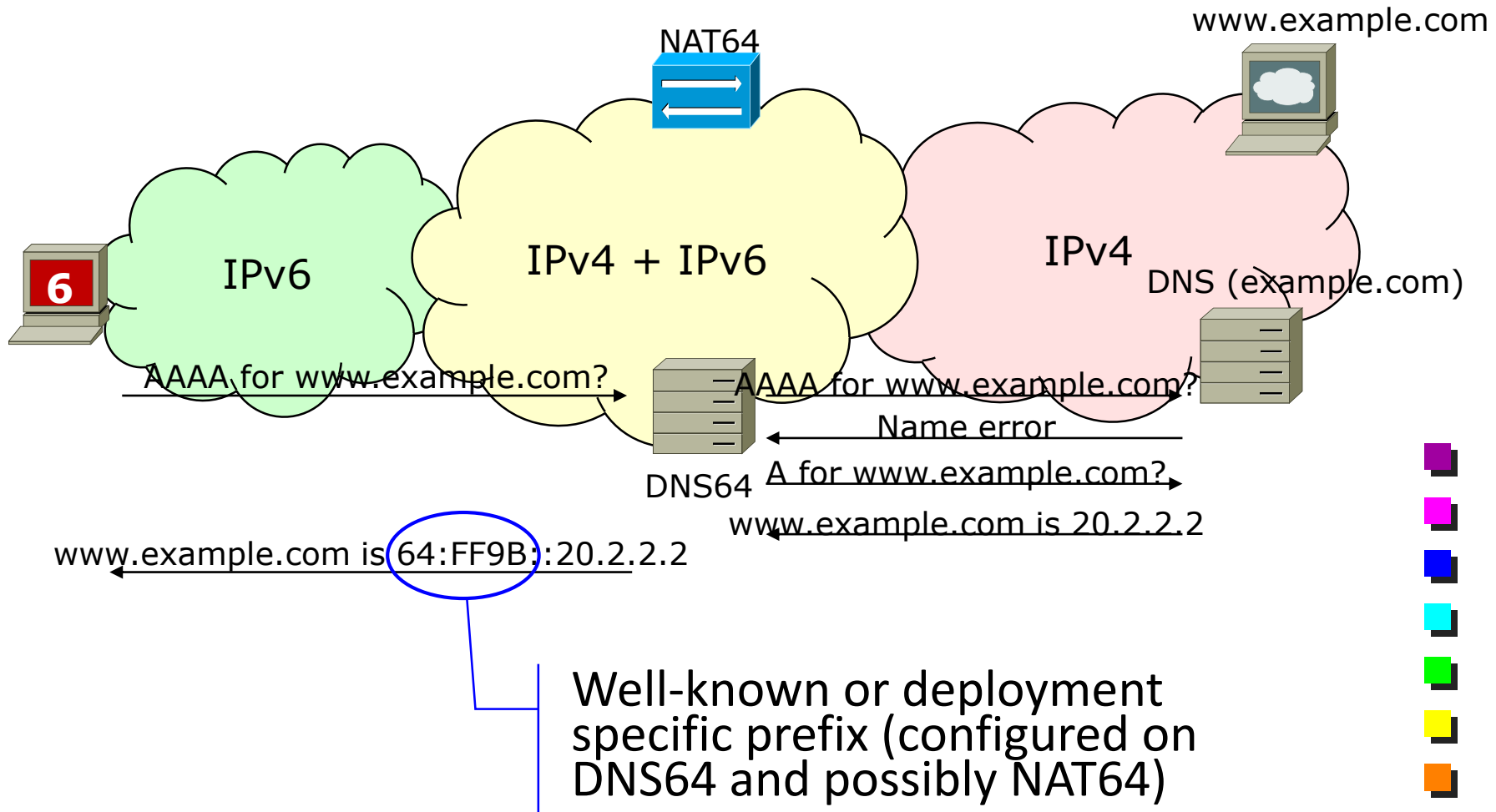


NAT64 + DNS64: Principles

- An IPv6 prefix is dedicated to mapped IPv4 addresses
 - Either well-known or network-specific
 - RFC 6052
 - DNS64 maps A records into AAAA using NAT64 prefix, then serves A and AAAA records to the client
 - NAT64 router advertises NAT64 prefix into IPv6 network to attract traffic toward IPv4 hosts
- 



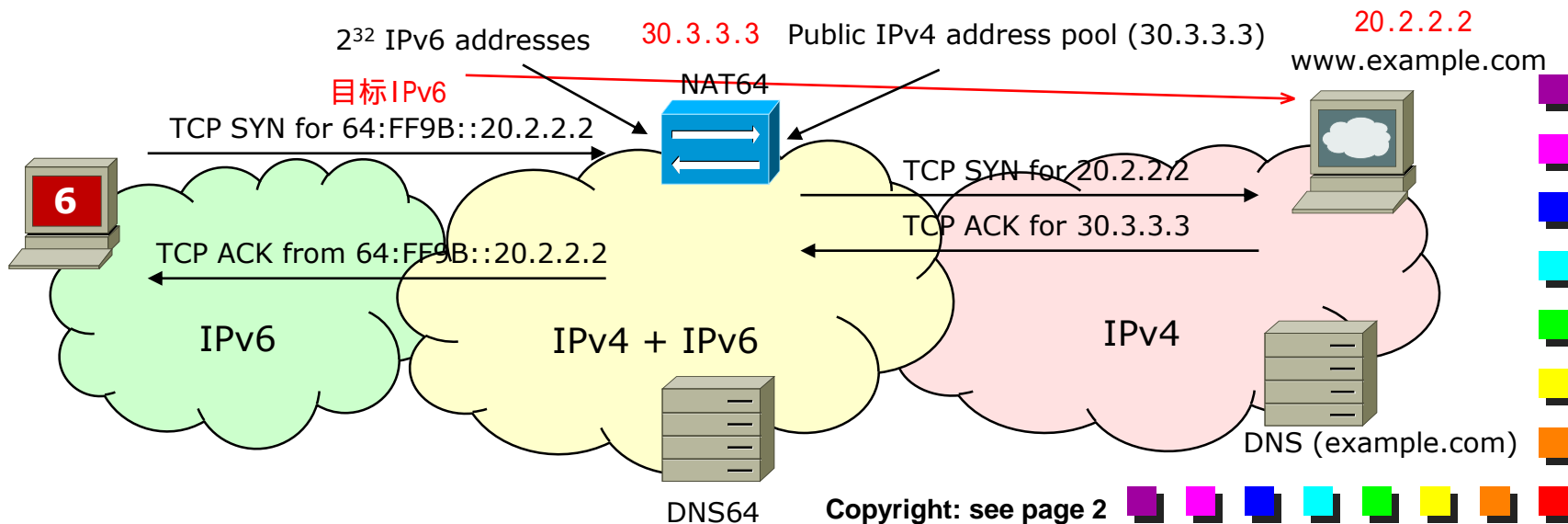
DNS64: Name Resolution



NAT64: Packet Forwarding

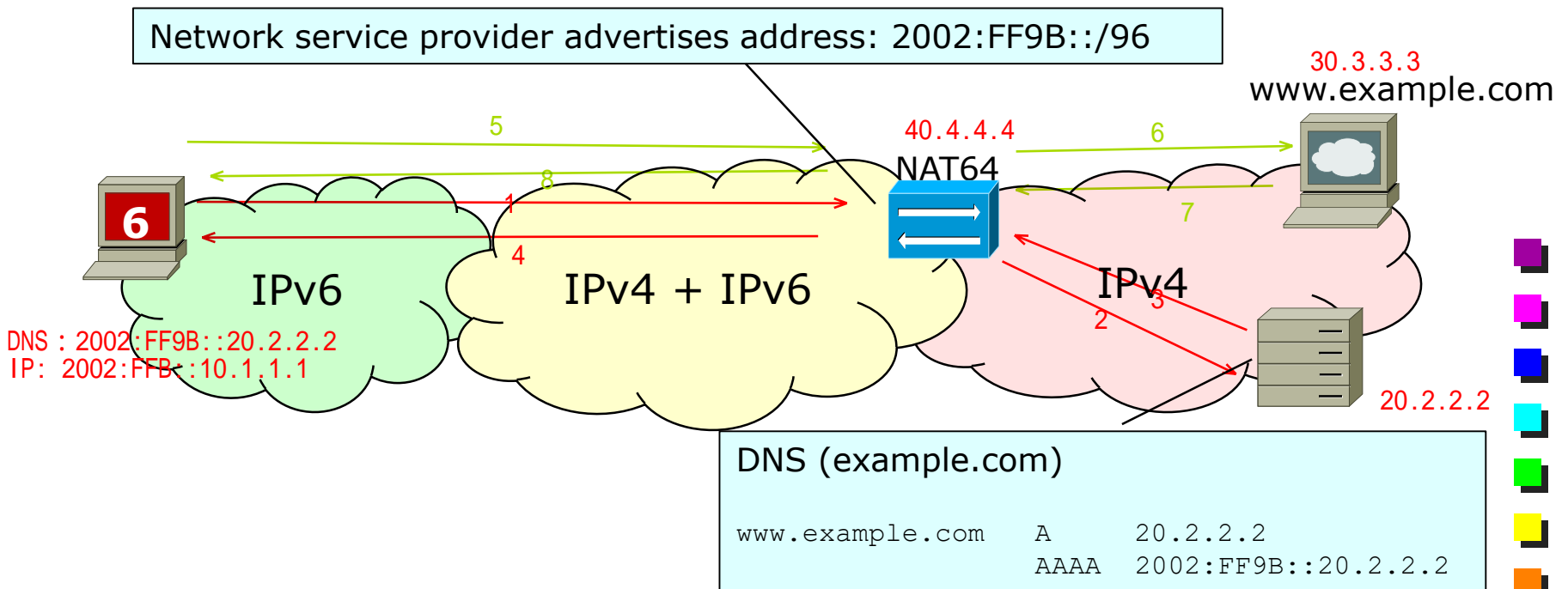
■ NAT64 (outbound)

- Translates IPv6 address and packet into IPv4
- Picks a free IPv4 address/port from its pool
- Builds NAT session entry



A Specific Use Case: Individual IPv4 Server Reachability

No need for DNS64: a regular DNS is enough





NAT64 Deeper Dive


■ NAT64 prefix

- Any /32, /40, /48, /56, /64 or /96 prefix
- Well-Known Prefix (WKP): 64:FF9B::/96
- /64 recommended for Network Service Providers

■ Stateful NAT64

- Very similar to PAT (stateful NAT44)
- Individual TCP and UDP sessions + ICMP replies are translated
- Source IPv6 address + port number used in the lookup

■ Stateless NAT64

- Each IPv6 address is translated into one IPv4 address
 - Either static mapping or IPv6 hosts have NAT64 prefix
 - Only ICMP packets and IP headers are translated
 - Limited use: few IPv6 hosts
 - E.g., IPv6 only servers, which requires static mapping
- 



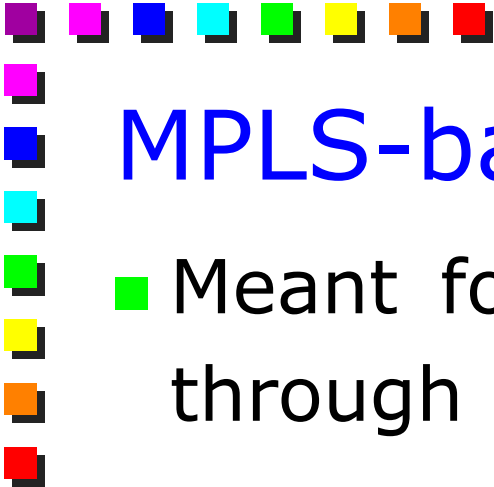
NAT64 + DNS64 Limitations

- Only when the DNS is involved
 - I.e., hostnames are used
 - E.g., it does not work in case the user directly specifies an IPv4 address
 - E.g., `ping 1.2.3.4`
- No DNSSEC
 - In DNSSEC authoritative DNS signs record
 - But DNS64 modifies records

MPLS-based

■ Meant for through

- MPLS-based
- Meant for through





IPv6 and MPLS

不可知，也就是对于所传输的协议类型care, MPLS只是转发数据到 Edge MPLS router就好

- Data plane is agnostic to IPv4/IPv6
 - Forwarding is based on labels
- Control Plane is not
 - Destinations are identified with an IP address



Native IPv6 over MPLS

Full Control Plane upgrade

IPv6 support in


- Routing
- Label Distribution Protocol
- Management

Possibly dual (IPv4 and IPv6) support






IPv6 over Circuit Transport

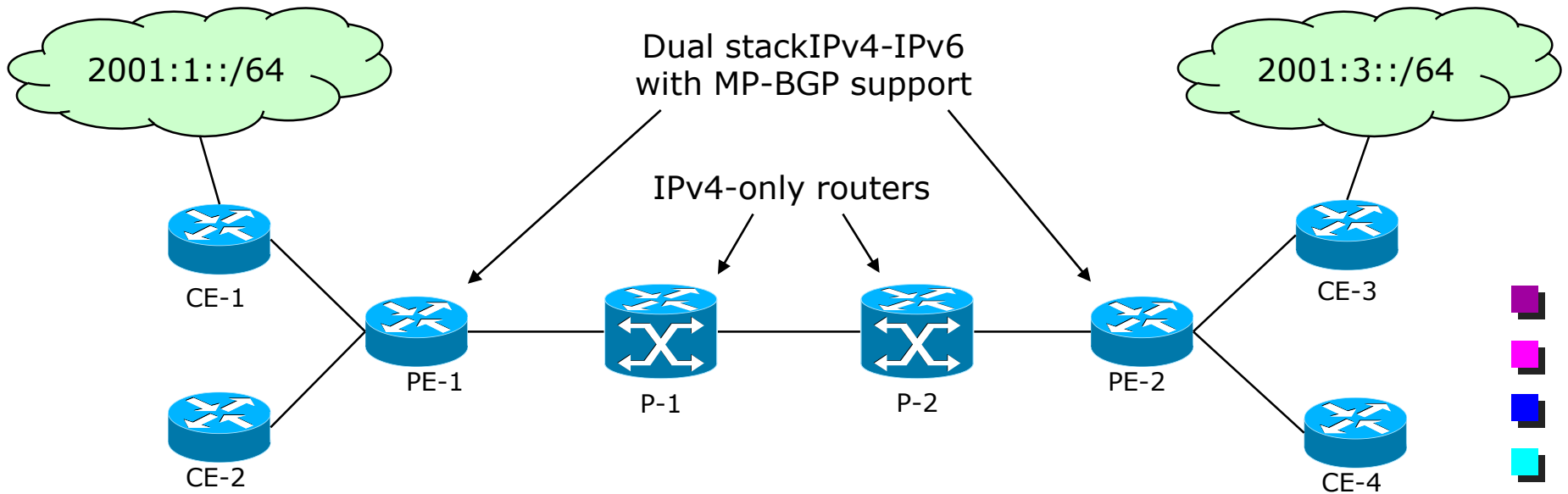
- MPLS as layer 2 connectivity
 - LSP -> L2 tunnel
 - PEs (i.e., label edge routers) are IPv6 aware
 - Static routes to IPv6 destinations
 - Routing with remote PEs
 - No changes needed to P routers
 - Scalability problems
 - L2 tunnels and routes configured manually
 - Possibly mesh topologies of L2 tunnels
- 



6PE

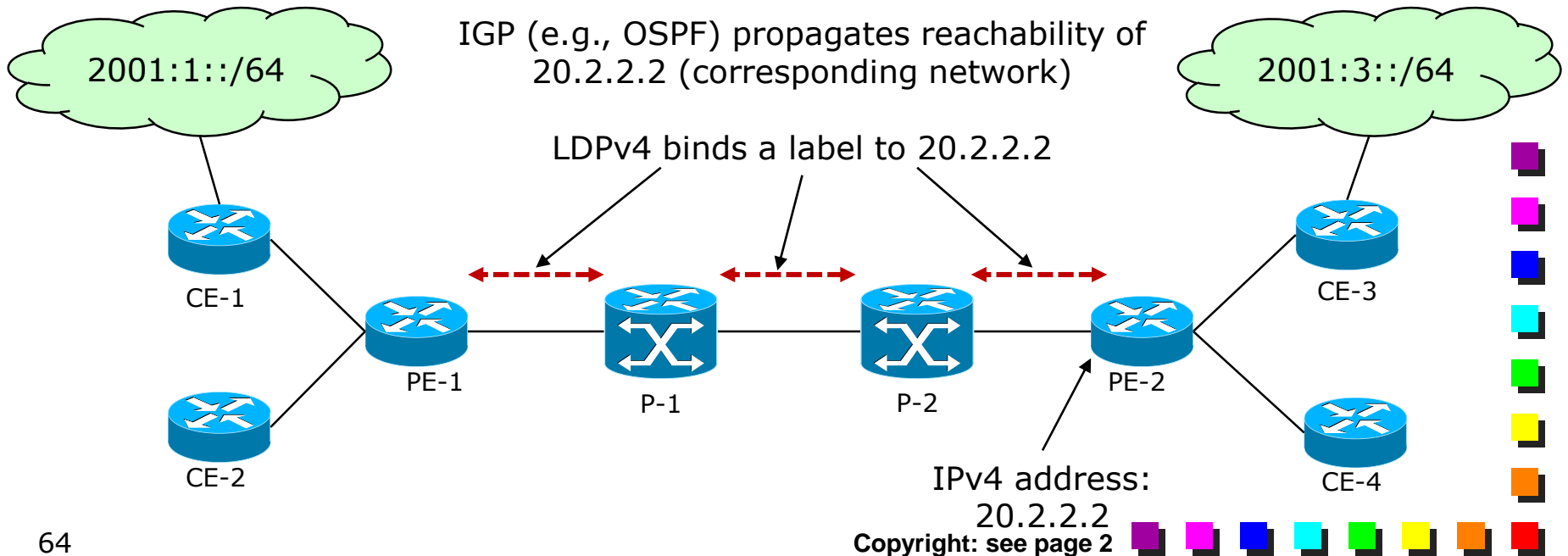
- Only PE routers need to be changed
 - Same mechanisms as MPLS-based VPN
 - VPN services can be offered on the same backbone
 - Very scalable
 - Minimum configuration required
 - As more customers requires IPv6 connectivity, provider might consider migrating to native support
- 

Architecture



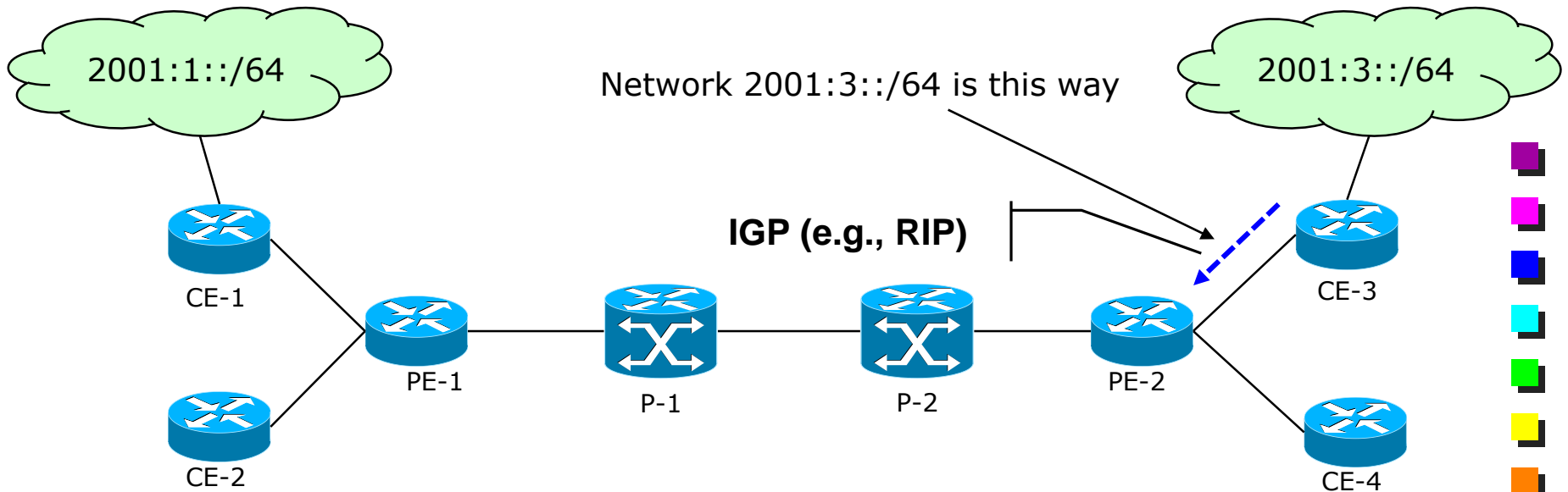
Internal Routing

- 6PE and P routers advertise their IPv4 prefixes with
- Labels are bound to each PE
- Topology-based label binding



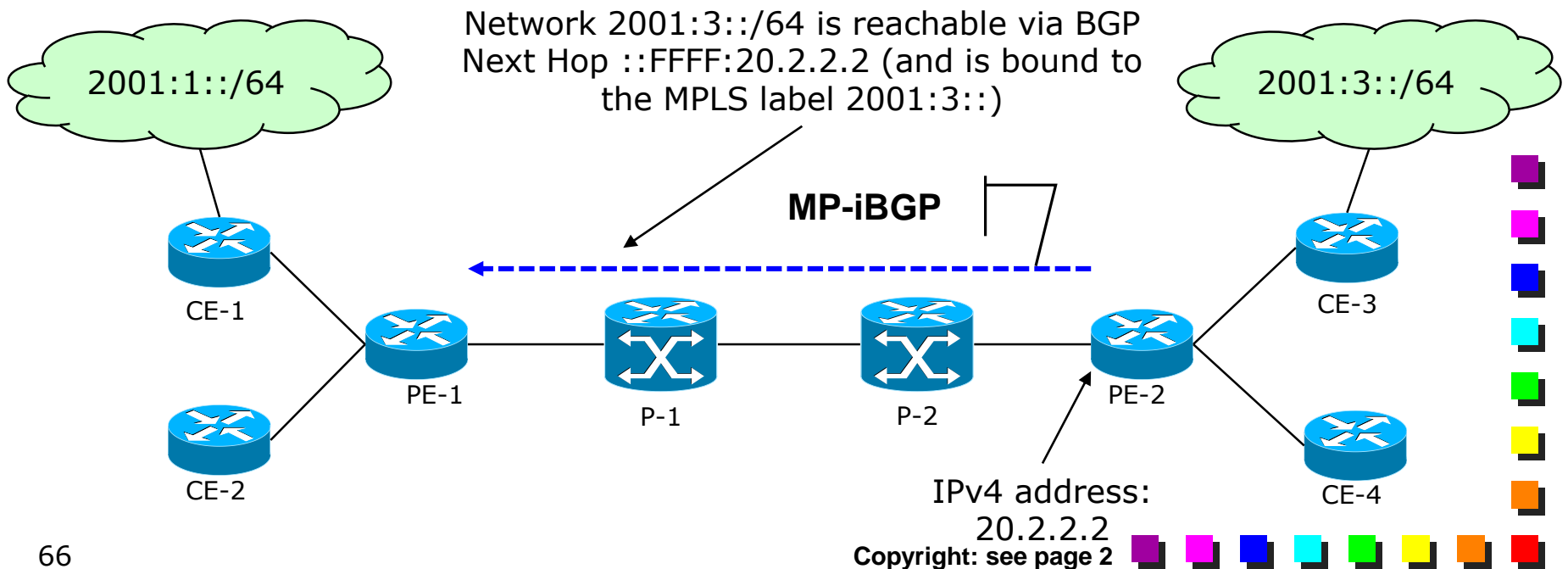
Announcing IPv6 Networks

- CE and 6PE connect through **native** IPv6 interfaces
- Routing information is exchanged
 - Any routing protocol (RIP, OSPF, BGP)
 - Static routes



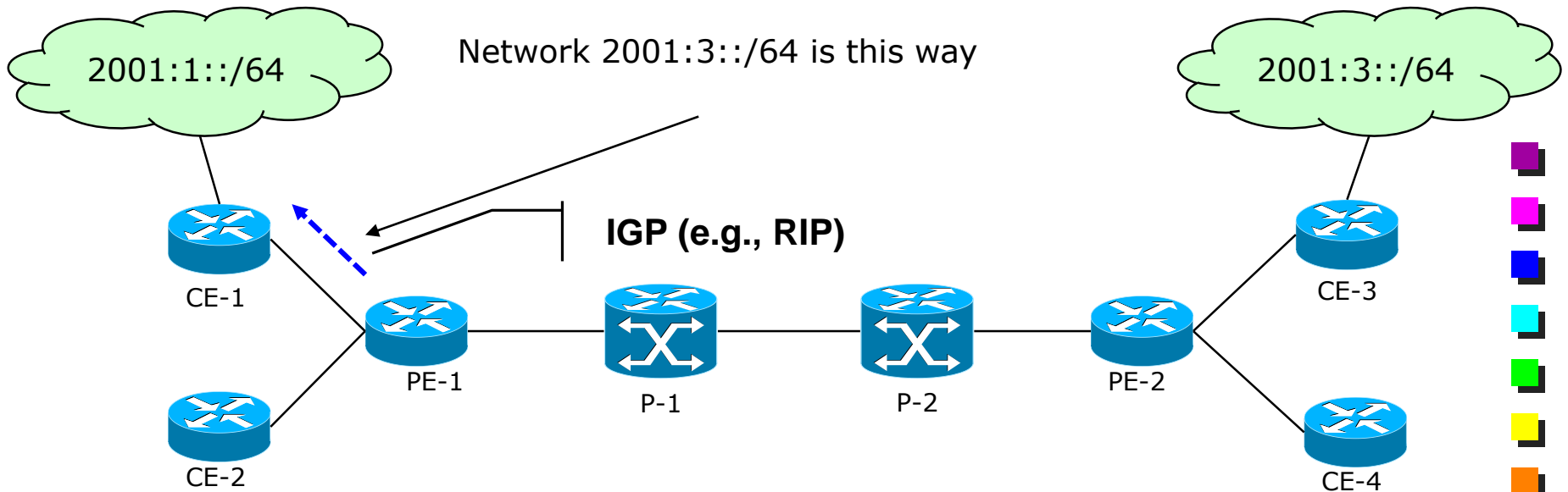
Propagating IPv6 Routes

- 6PEs exchange routing information through MP-BGP sessions over IPv4
- 6PE IPv4-mapped address as the BGP Next-Hop



Redistributing IPv6 Routes

- PEs propagate advertisements to IPv6 networks through IGP
- This may not be needed is an IPv6 a default route is used in CE-1





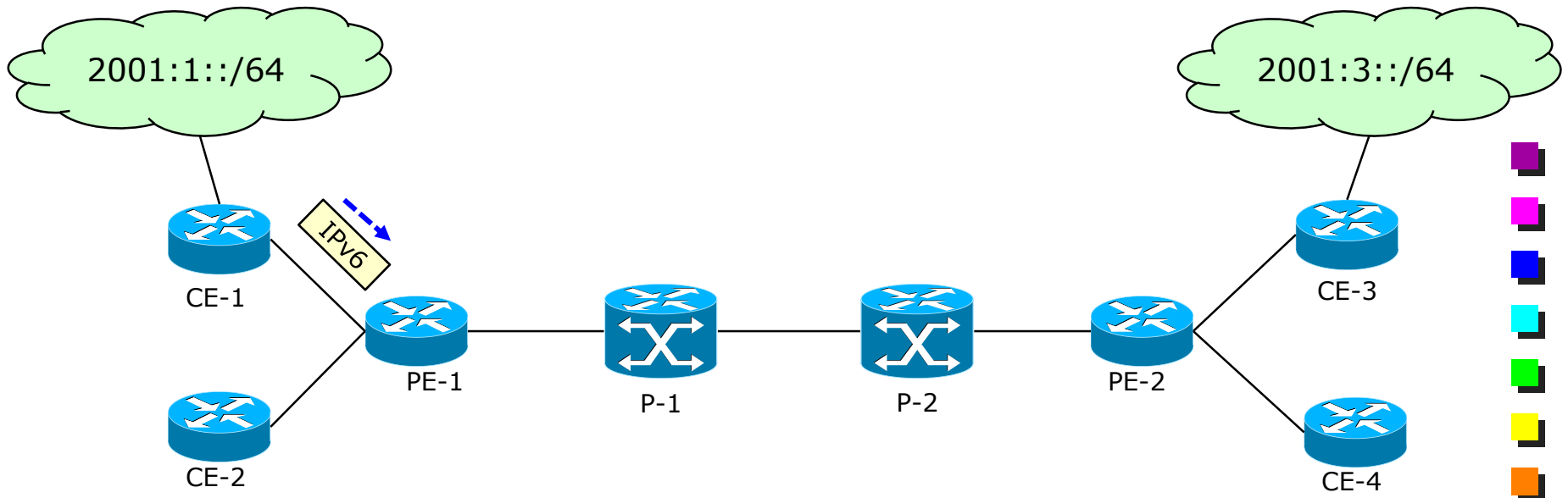
Label Distribution

- For IPv4 destinations among P and 6PE routers
 - LDP or RSVP
- For IPv6 destination directly among 6PEs using MP-iBGPv4



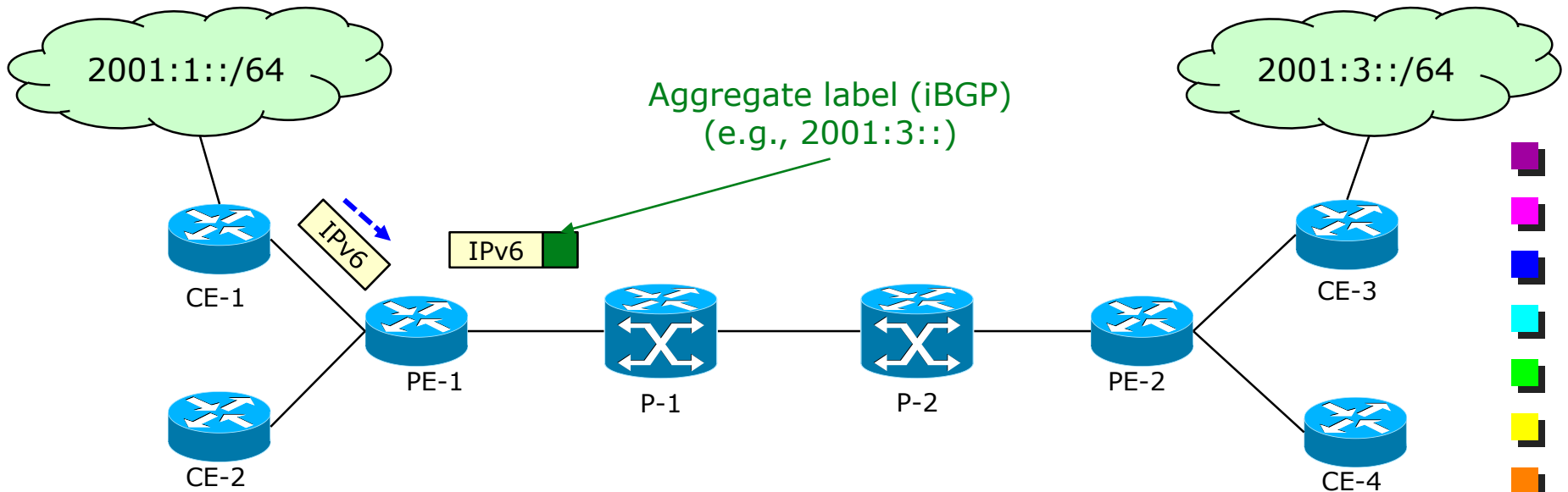
IPv6 Packet Forwarding

- CE routers send IPv6 packets to 6PE routers
- Static (default) route or dynamic route from IGP



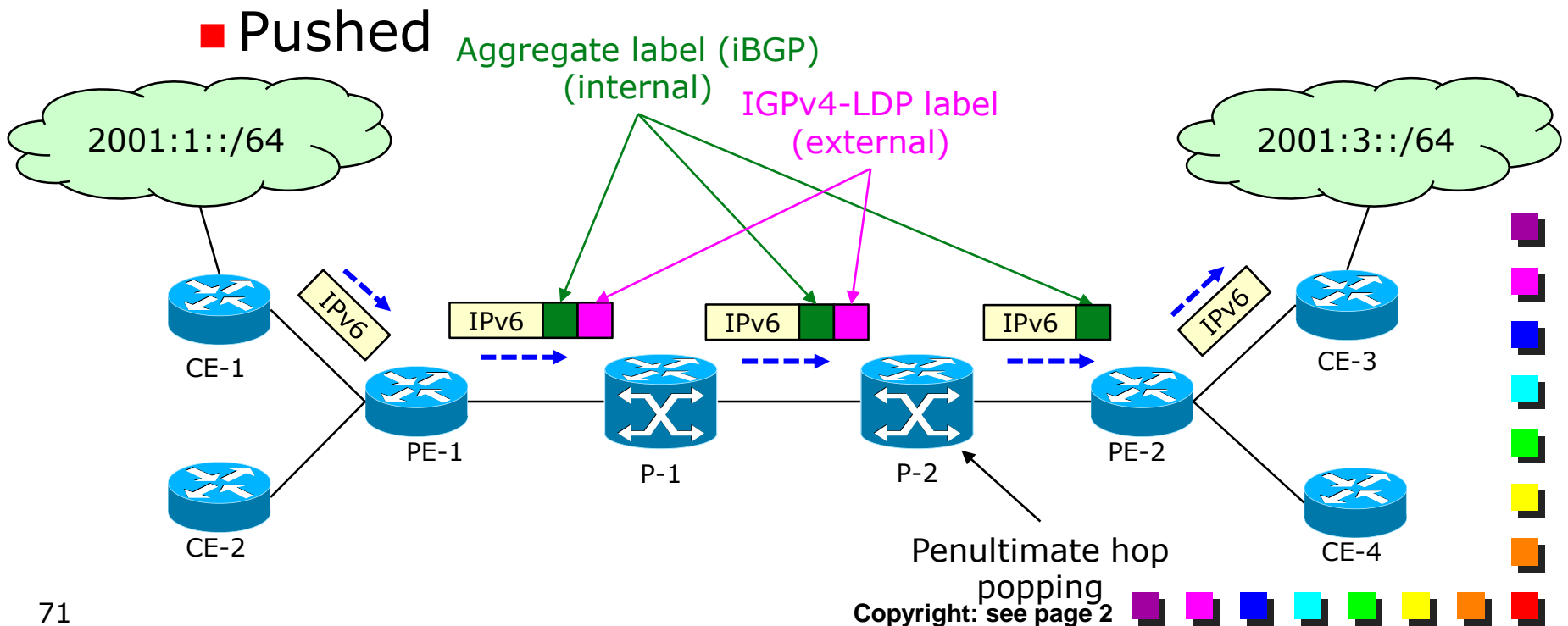
IPv6 Packet Forwarding

- 6PE has label mapping for destination
 - Label as distributed by iBGP (e.g., 2001:3::)
 - Next hop: 6PE identified the BGP Next Hop
 - IPv4-mapped IPv6 address (e.g., ::FFFF:20.2.2.2)



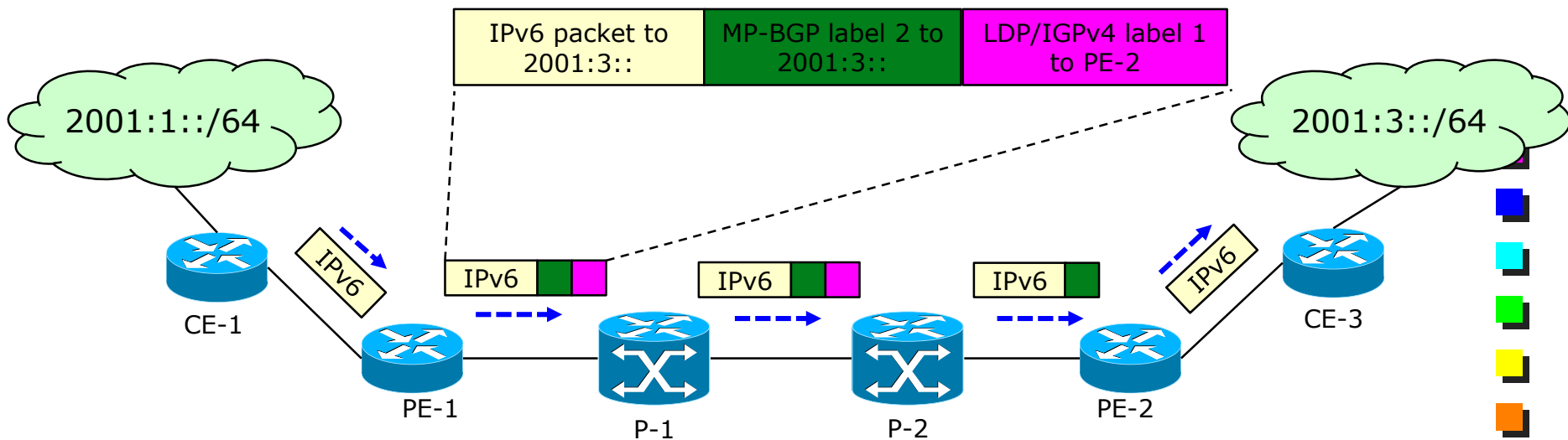
IPv6 Packet Forwarding

- Next hop is reachable through IPv4
 - Advertised by IGP
- Label mapping distributed with LDP/RSVP
 - Pushed



A closer look to the labels

- Inner label in principle not required for
 - It keeps the solution the same as for VPN
 - Penultimate hop should be able to forward IPv6 packet
 - Or PHP should not be performed





References

- Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, August 2011.
- R. Bush, "The Address plus Port (A+P) Approach to the IPv4 Address Shortage," RFC6346, August 2011.
- J. De Clercq , D. Ooms, S. Prevost, F. Le Faucheur, "Connecting IPv6 Islands over IPv4 MPLS Using IPv6 Provider Edge Routers (6PE)," RFC 4798, February 2007.

