Interdomain routing

BGP-4

Enzo Mingozzi
Professor @ University of Pisa
enzo.mingozzi@unipi.it

BGP-4

- Border Gateway Protocol v4 (rfc4271) VANILLABGP
  - **inter-AS** routing protocol (EGP)
    - inter-AS routing usually reflects **political and business relationships** between the ISP and organizations involved
    - intra-AS routing is optimized in accordance with the required technical demands
  - *How do I filter routing updates coming from a particular neighbor AS?*
  - *How do I make sure that I use this link or this provider rather than another one?*

- Exchange **network reachability information** between BGP speaking peers
- Supports **destination-based** forwarding only (CIDR)

SEGMENT ROUTING:

SI PUÒ FARE TUTTO CON SOURCE ROUTING, MA NON È MOLTO EFFICIENTE.

↓

TRADITIONAL SOURCE ROUTING

FORWARDING BASATO SOLO SU DESTINATION ADDRESS.

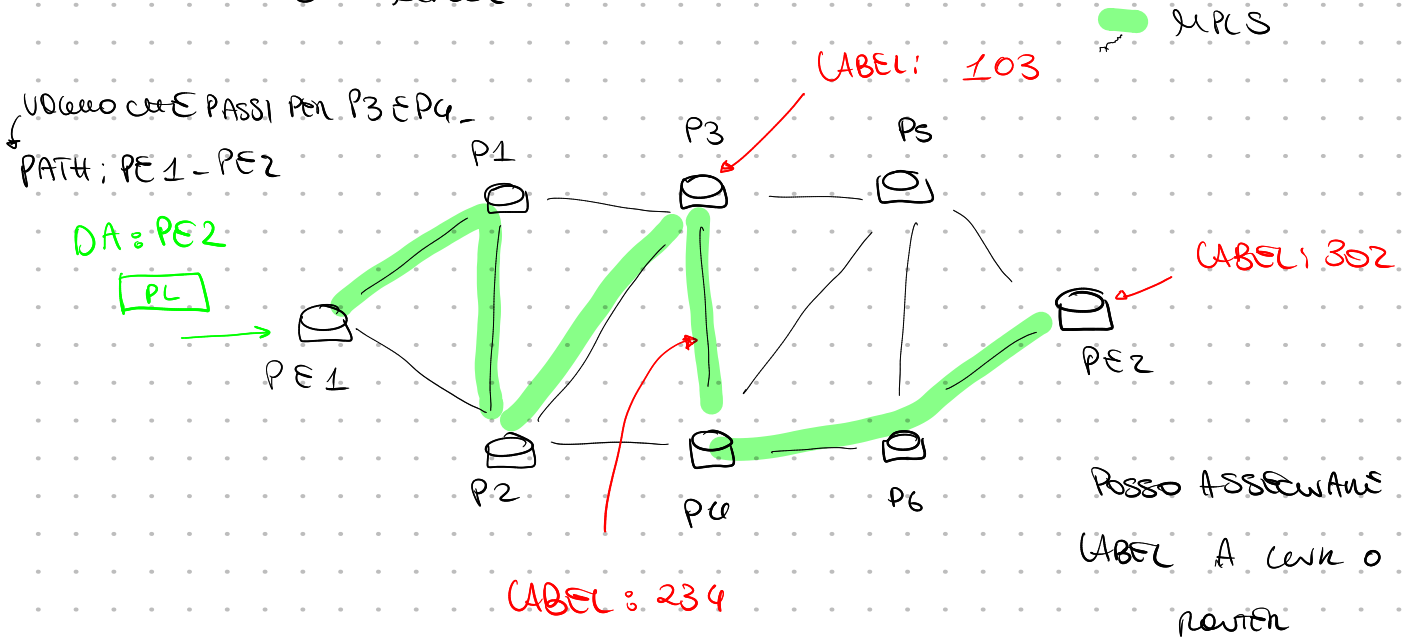
↳ QUINDI USIAMO MPLS CON TRAFFIC ENGINEERING.

↓

CONSISTE NEL FARE SOURCE ROUTING IN MODO EFFICIENTE

SEGMENT ROUTING NASCE PERCHÉ VOLEVO SOURCE ROUTING IN MODO EFFICIENTE E PUÒ ESSERE BASATO SU MPLS. → SEMBRAVA PIÙ COERENTE CHE TRATTENESSI

MPLS



POSSO ASSEGNARE LABEL A CIASCUN ROUTER

USIAMO MPLS IN MODO DIVERSO.

LE LABEL POSSONO AVERE UNO SCOPE LOCALE O GLOBALE.

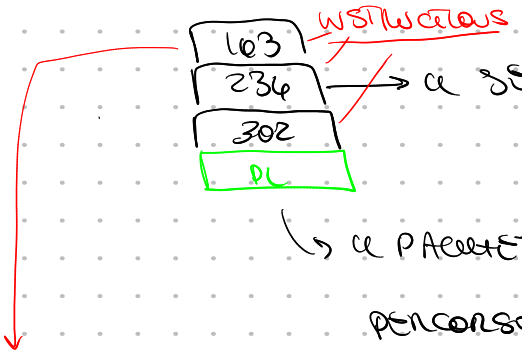
↳ TUTTI I ROUTER SANNO CHE QUEL ROUTER/LINK È UNIVOCAMENTE IDENTIFICATO DALLA LABEL.

DOBBIAMO ASSICURARCI CHE LA LABEL 103 NON SIA ASSEGNATA AD UN ALTRO ROUTER DELLA RETE.

LE LABEL POSSONO ANCHE AVERE SCOPE LOCALE, COME NEL MPLS.

CON MPLS AVREMMO CONFIGURATO UN LSP CON RSVP-TE SETTANDO UN ERO DI TIPO LOOSE.

CON SEGMENT ROUTING NON HO BISOGNO DI CONFIGURARE LSP, MA FACCIO DIRETTAMENTE A PUSH DELLE LABEL DEL PERCORSO.



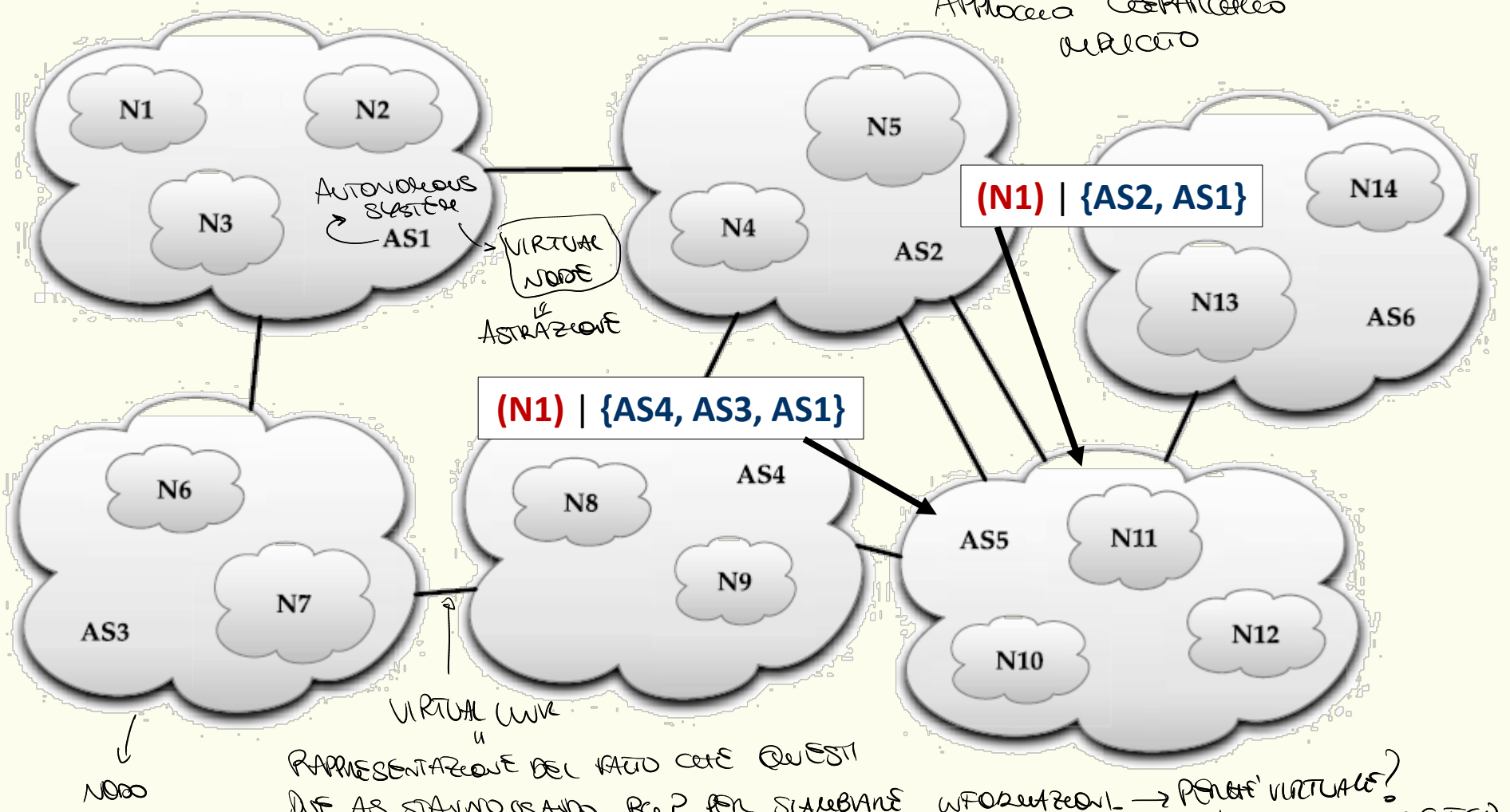
PARTO DALLA CIMA DELLO STACK, 103 IDENTIFICA P3 E ALLORA CERCO QUALE È IL MIGLIOR NEXT-HOP PER P3.

P3 RICONOSCE CHE FA PARTE DEL SECONDO SEGMENTO, FA LA POP E CONTINUA.

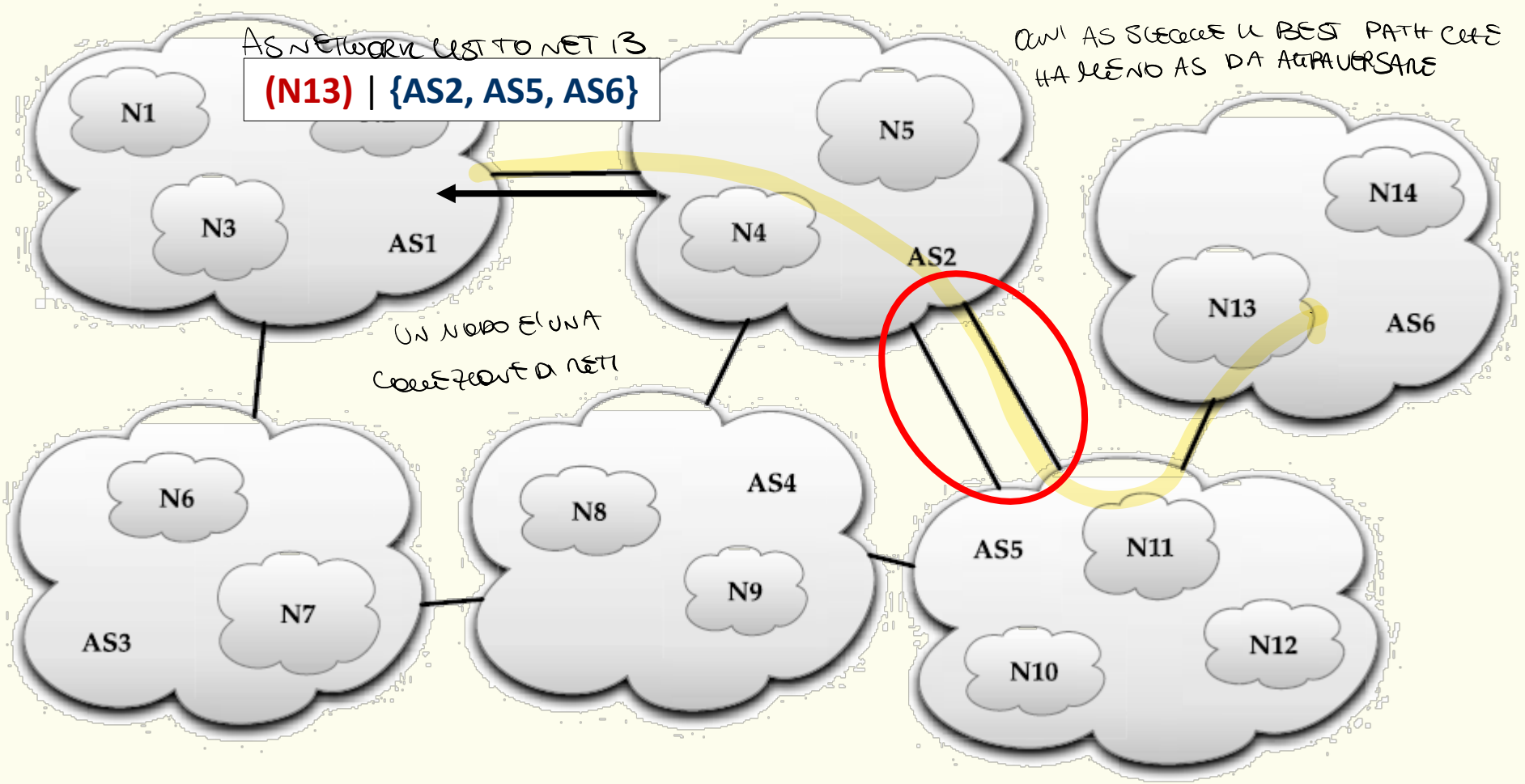
NON ABBIAMO BANDWIDTH RESERVATION!!!

VANTAGGIO: NON ABBIAMO BISOGNO DI RSVP-TE

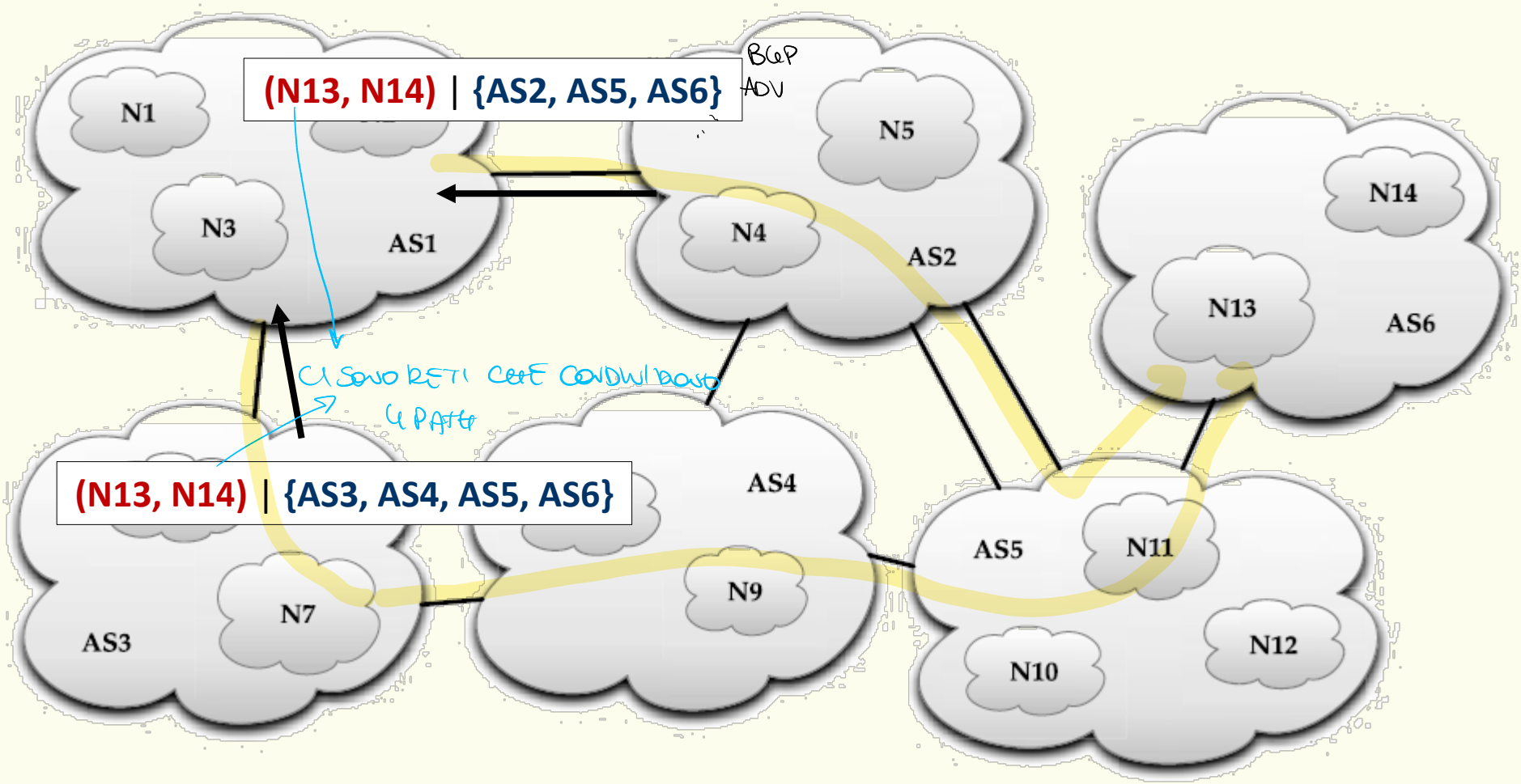
Path vector routing



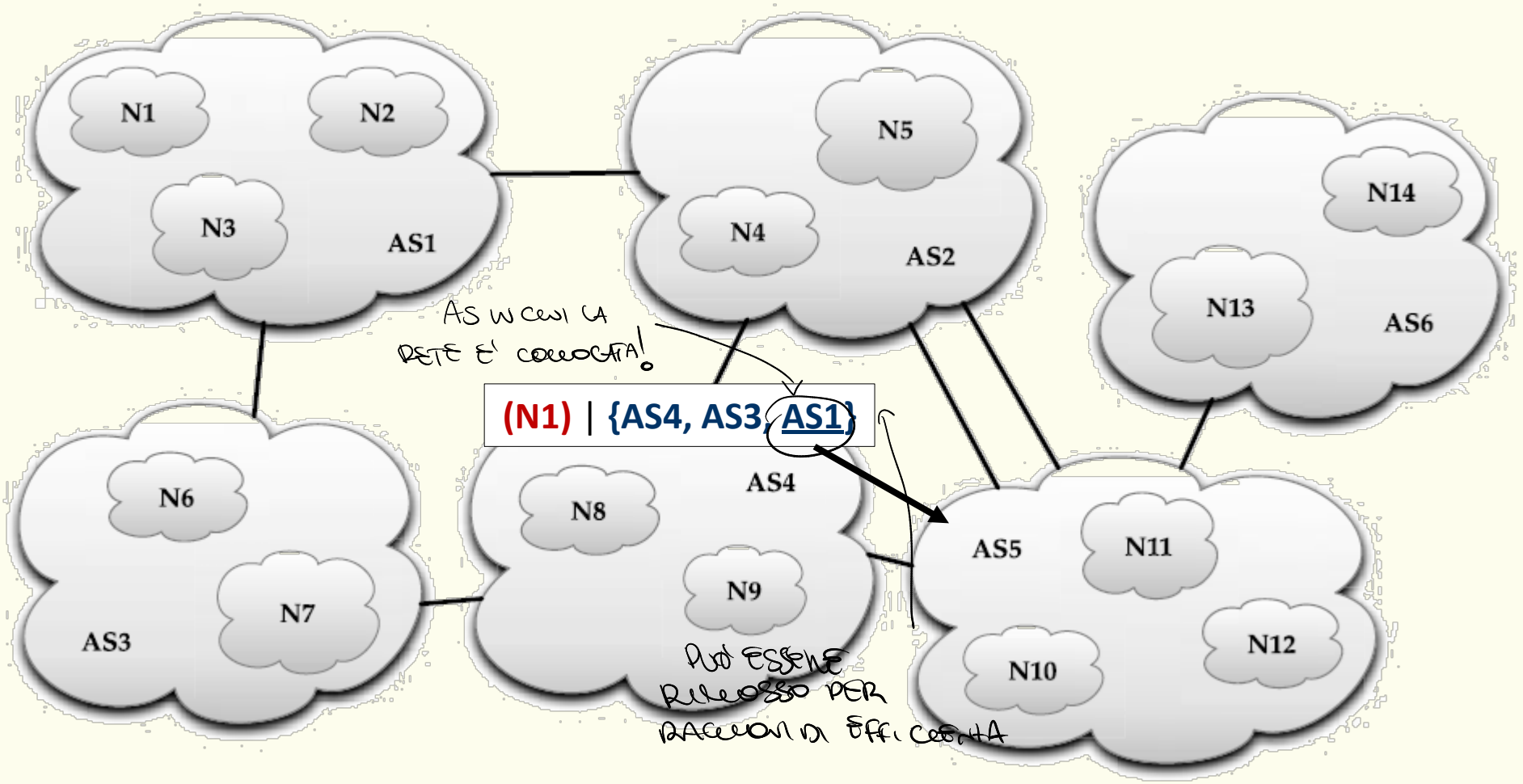
AS hop count metric



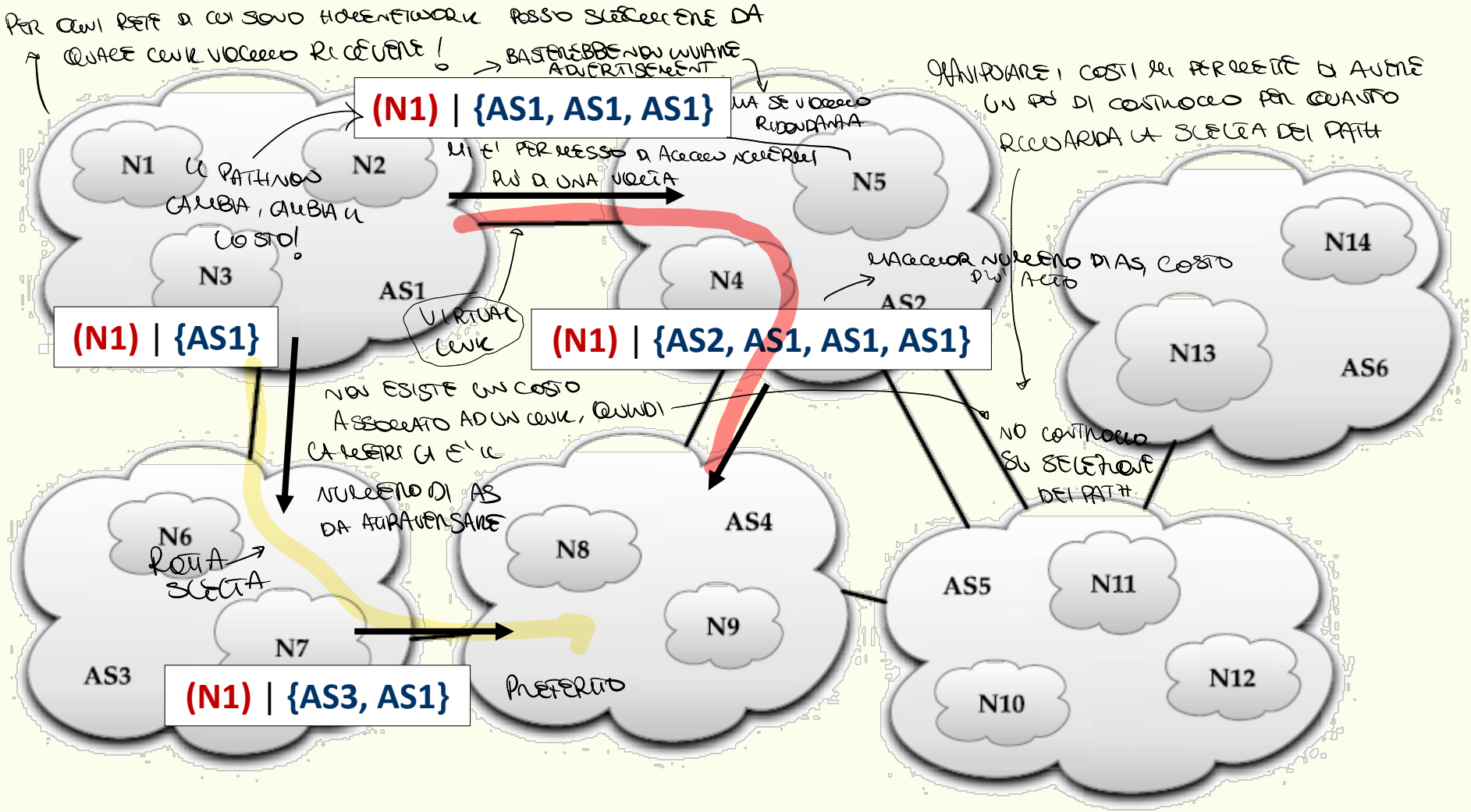
Set of networks advertisement



Home AS advertisement



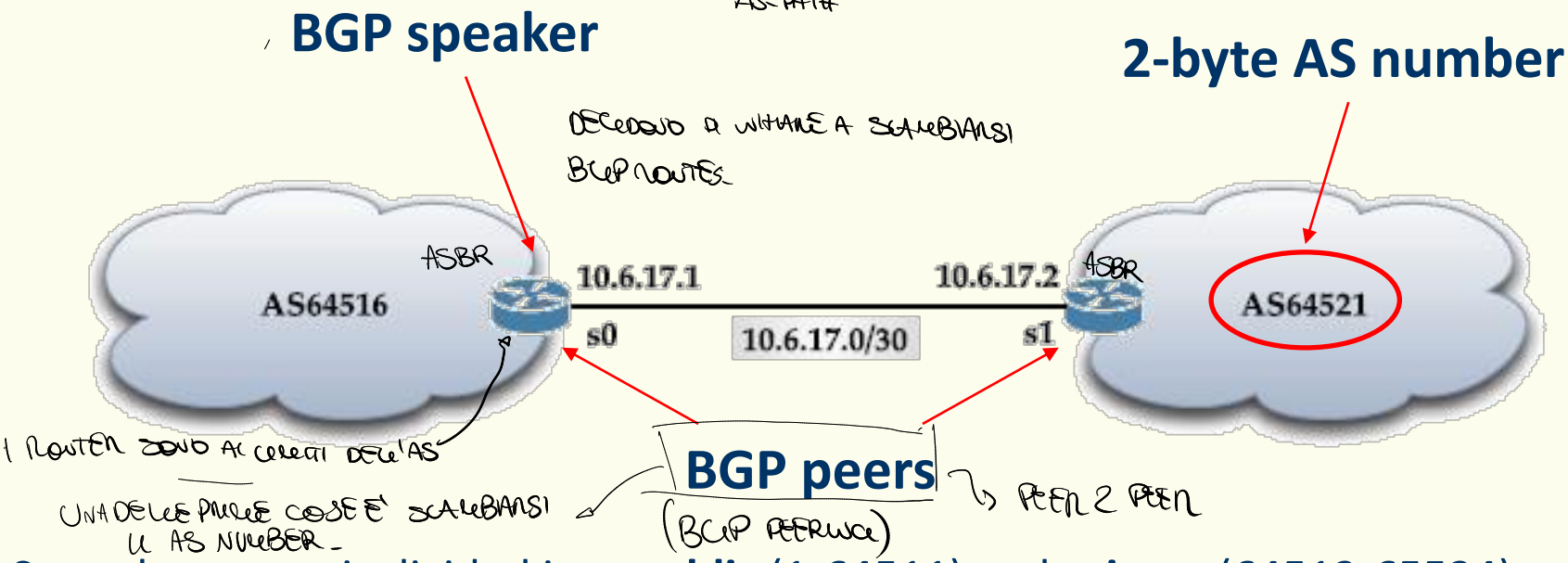
Path manipulation



Terminology

Network = IP prefix (A.B.C.D/n) → **network layer reachability information (NLRI)**

BGP route: IP prefix destination(s) → attributes of an AS–path by a receiving AS through an UPDATE message



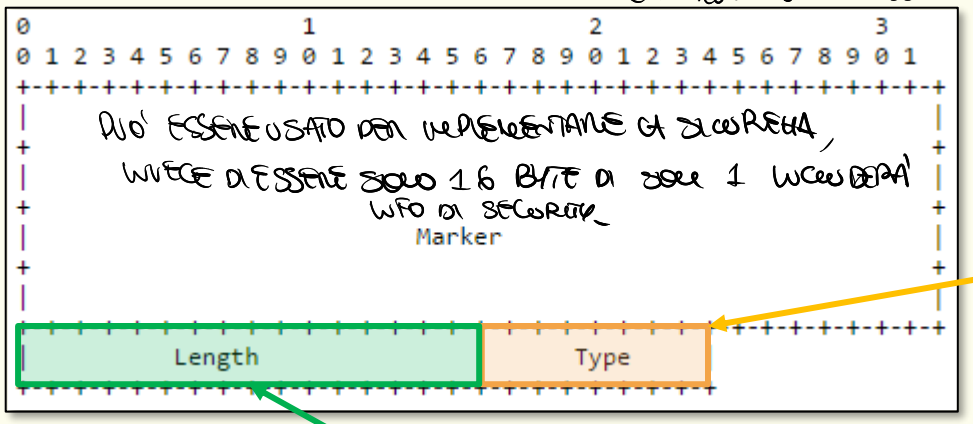
AS number space is divided into **public** (1-64511) and **private** (64512-65534)

4-byte AS number space is also defined → FLAT SPACE, SOLO IDENTIFICATORI

BGP operations

- Four BGP message types (+ 1) exchanged over TCP (port 179)

NON VOGLIO INVIARE UNO STREAM DI BIT QUINDI HO BISOGNO DI UN MARKER PER CAPIRE L'INIZIO DI UN MESSAGGIO



- OPEN
- UPDATE
- NOTIFICATION
- KEEPALIVE

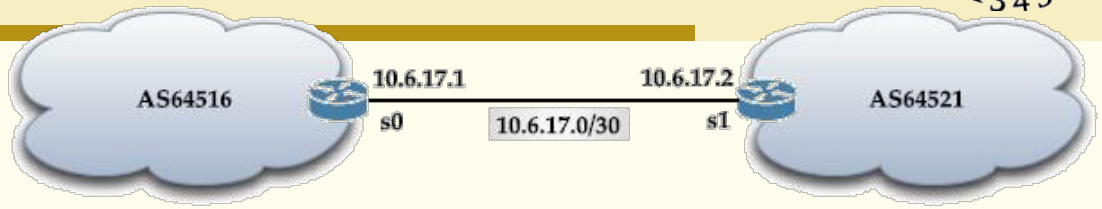
OBBLIGATORI

- ROUTE-REFRESH (optional)

NON OBBLIGATORIO DA SUPPORTARE PER I BGP SPEAKER.

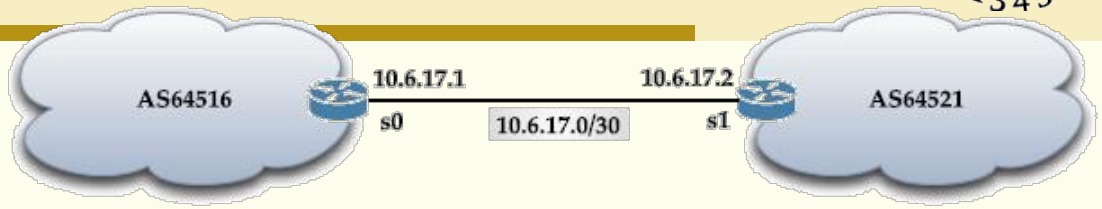
≤ 4096

BGP operations



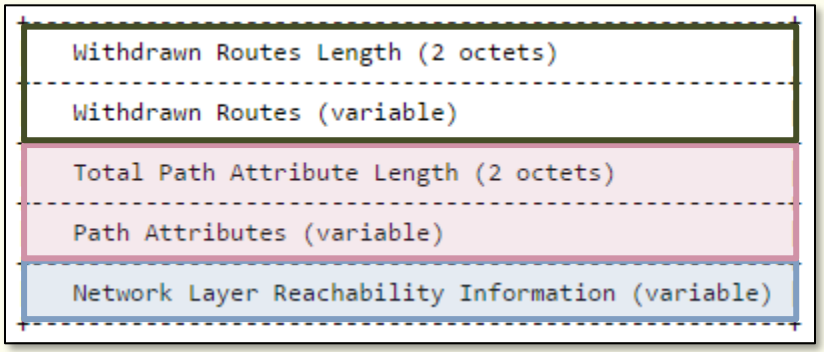
- **OPEN** message
  - First message sent to configured BGP peer (contains AS number)
- **KEEPALIVE**
  - Exchanged periodically, approximately three times every *hold time* at most
- **NOTIFICATION**
  - Closes a BGP session

BGP operations



- **UPDATE** message
  - Sent to exchange information about *networks*
  - Works in PUSH mode
  - Announced routes must be explicitly withdrawn

TRASPORTA BGP ROUTE → SET DI PREFISSI IP CON ATTRIBUTI

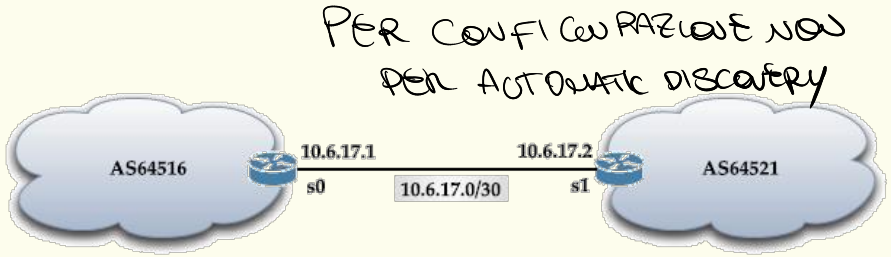


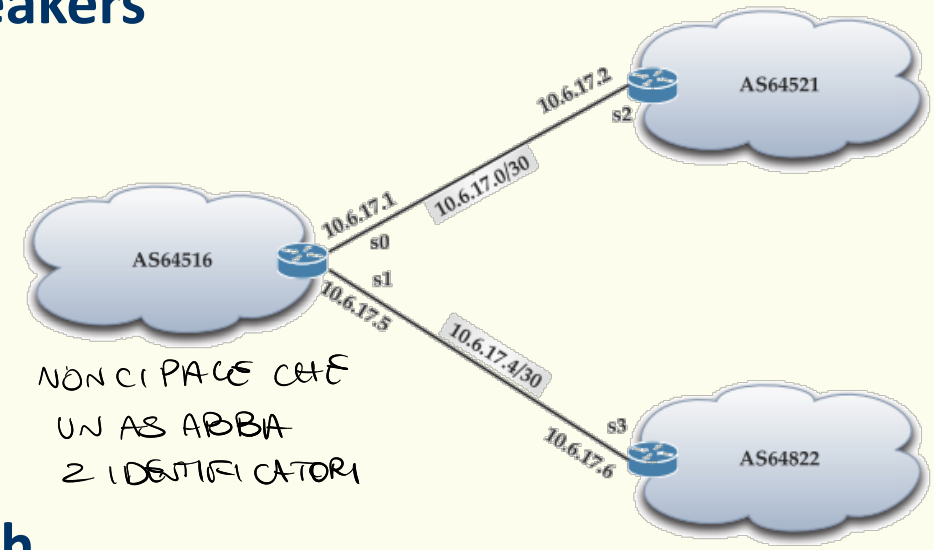
A sequence of path attributes (TLV)
- ORIGIN
- AS_PATH
- NEXT_HOP
- LOCAL_PREF
- …

List of networks {address/prefix length}

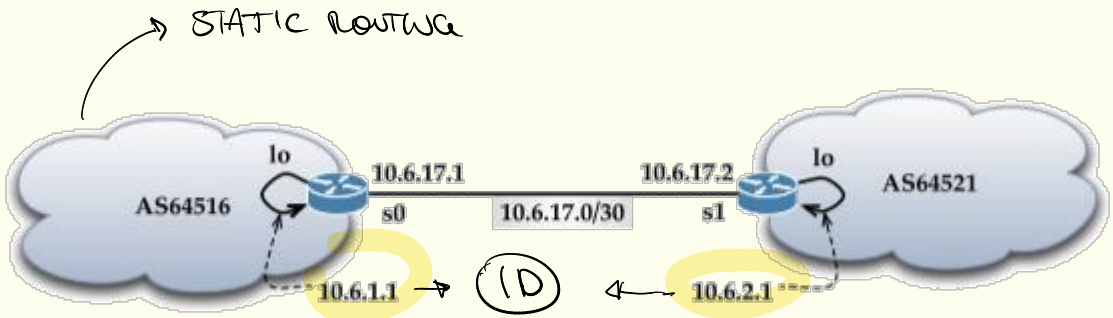
BGP initialization

p2p connections between BGP speakers





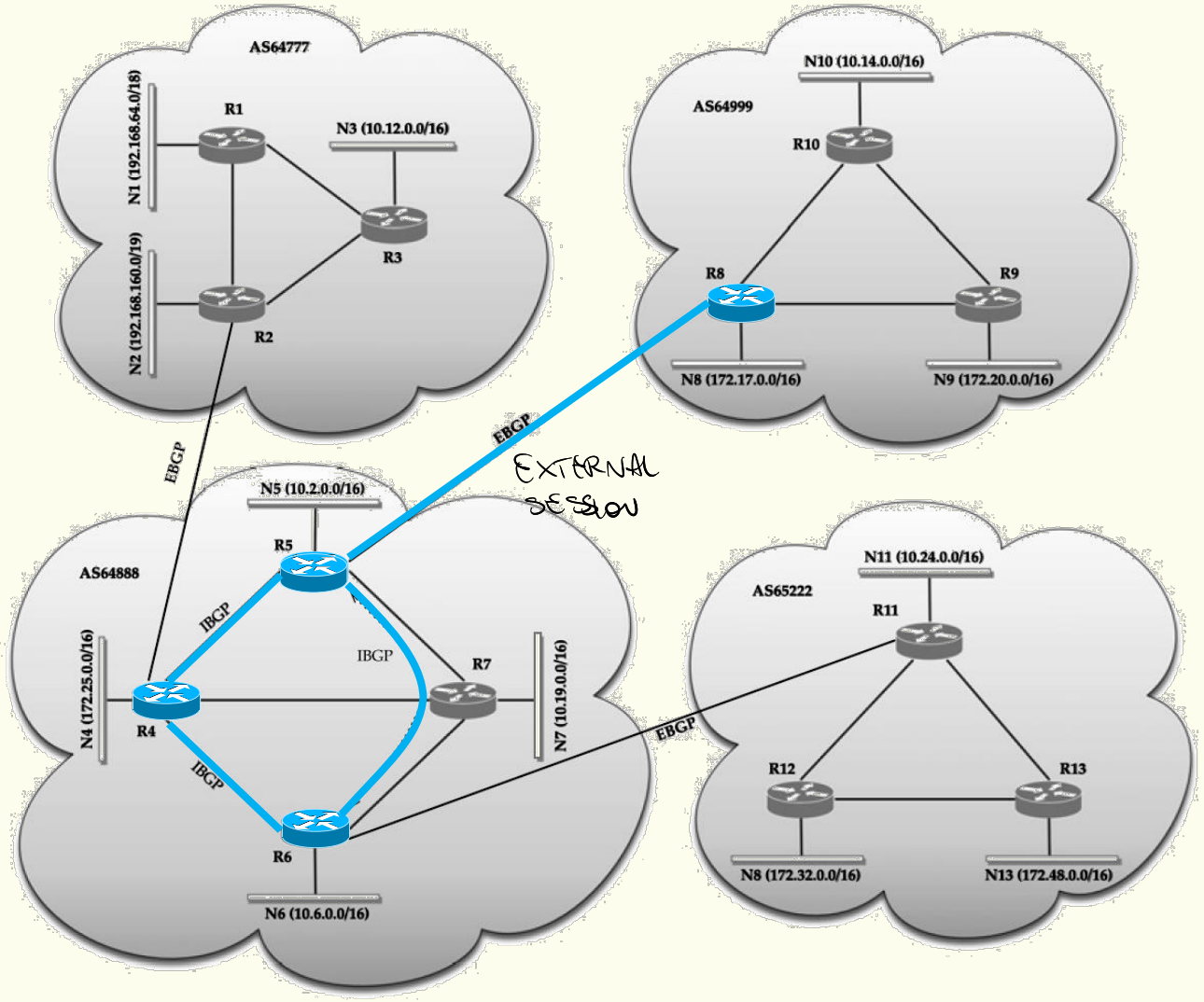
Loopback interface-based approach



(external) eBGP vs. (internal) iBGP

eBGP: BGP speakers are in different ASes, e.g. R5 and R8

iBGP: BGP speakers are in the same AS, e.g. R5 and R6

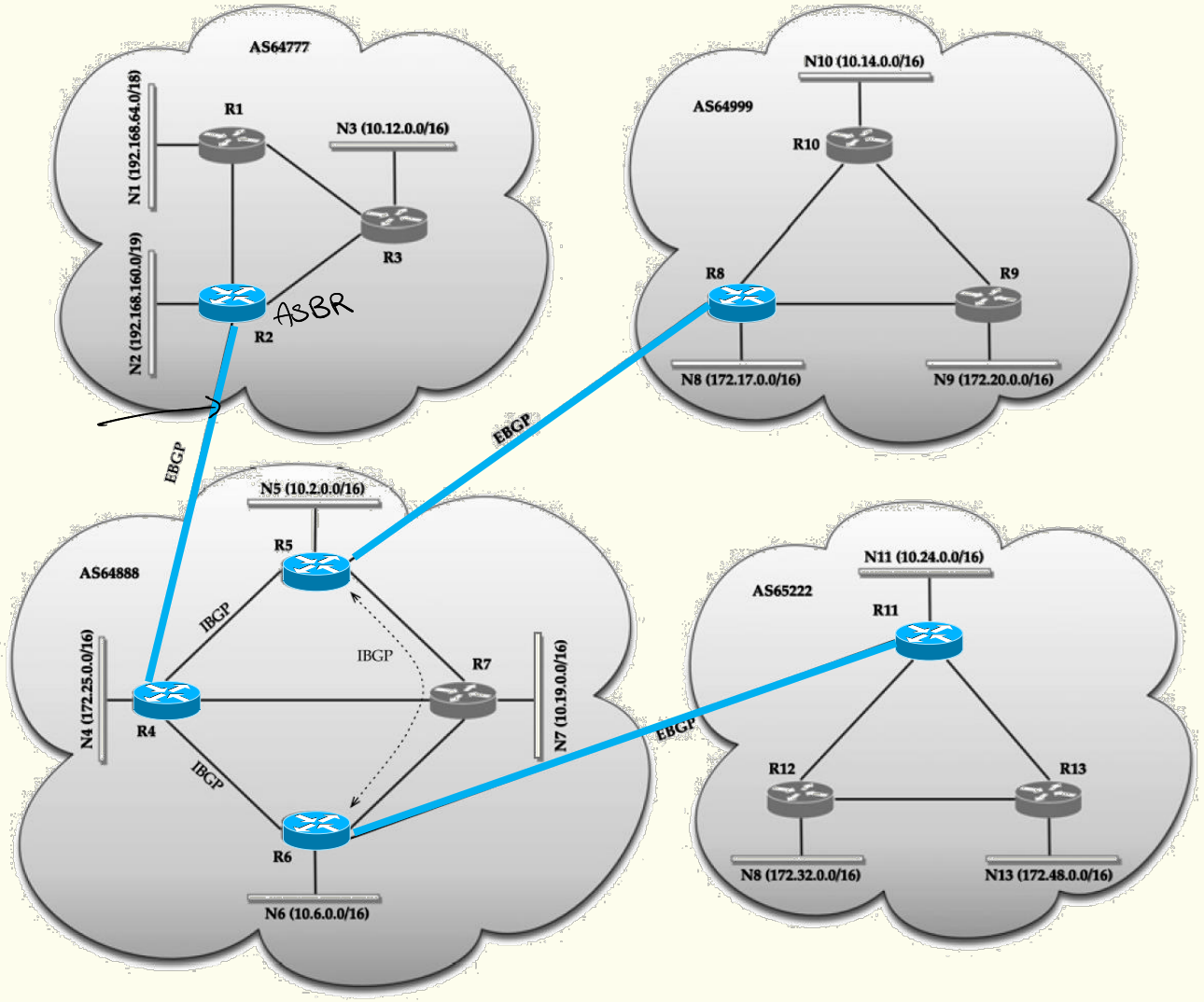


Why iBGP?

Stub AS (e.g. AS64777):
no need for eBGP speakers to redistribute BGP routes within the AS

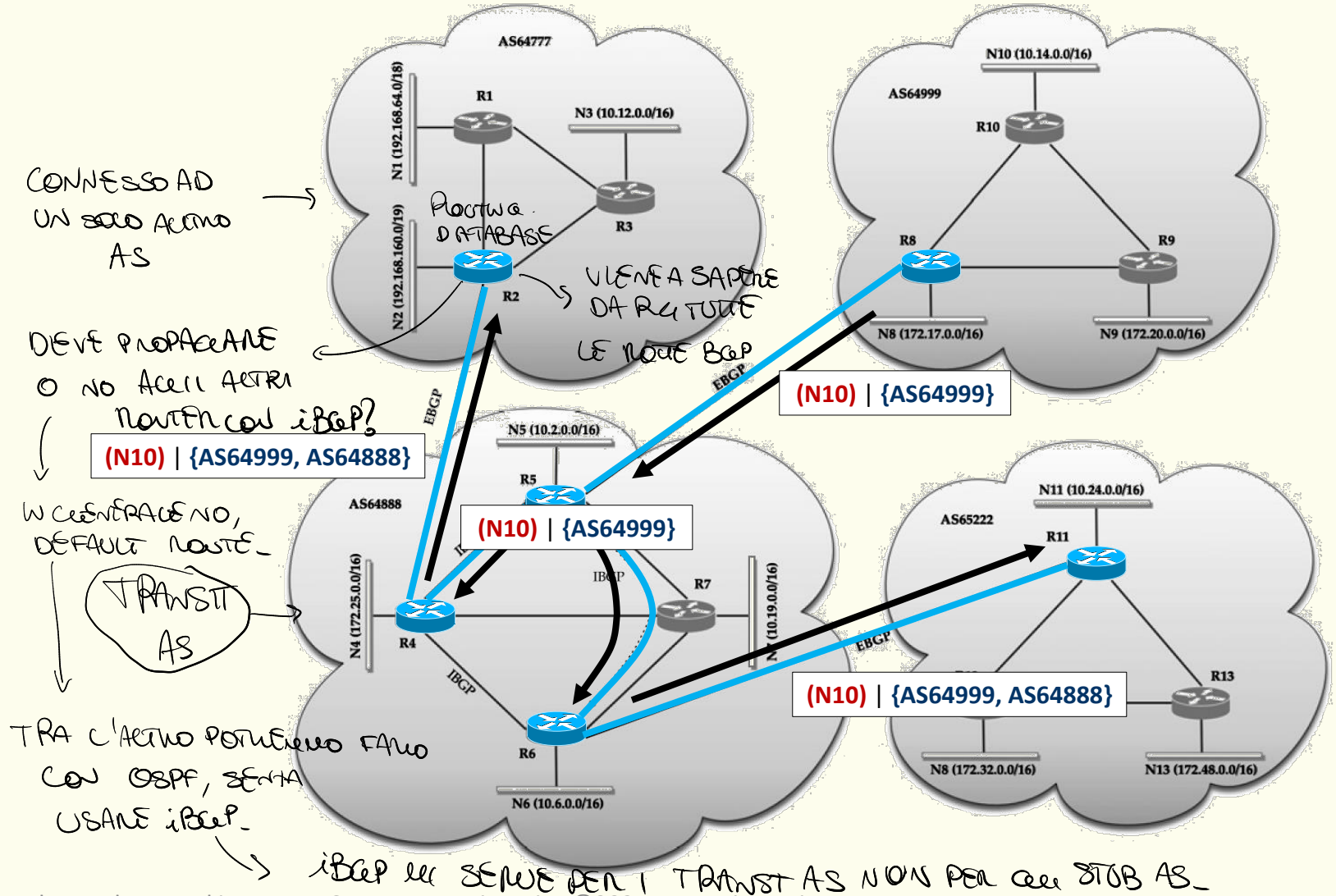
PER SAPERE QUALI ALTRI PREFISSI RACCOGLIERE

Transit AS (e.g. AS64888):
BGP routes learned by one eBGP speaker need be redistributed to the other eBGP speakers



Why iBGP?

PERCHÉ DOVREMMO AVERE BGP SPEAKERS ALL'INTERNO DELLO STESSO AS?



CONNESSO AD UN SOLO ALTRO AS

DEVE PROPAGARE O NO AGLI ALTRI ROUTER CON iBGP?

IN GENERALE NO, DEFAULT ROUTE.

TRANSIT AS

TRA L'ALTRO POTREMMO FARLO CON OSPF, SENZA USARE iBGP.

PUNTO DI DATABASE

VIENE A SAPERE DA TUTTE LE ROUTE BGP

iBGP SERVE PER I TRANSIT AS NON PER GLI STUB AS.

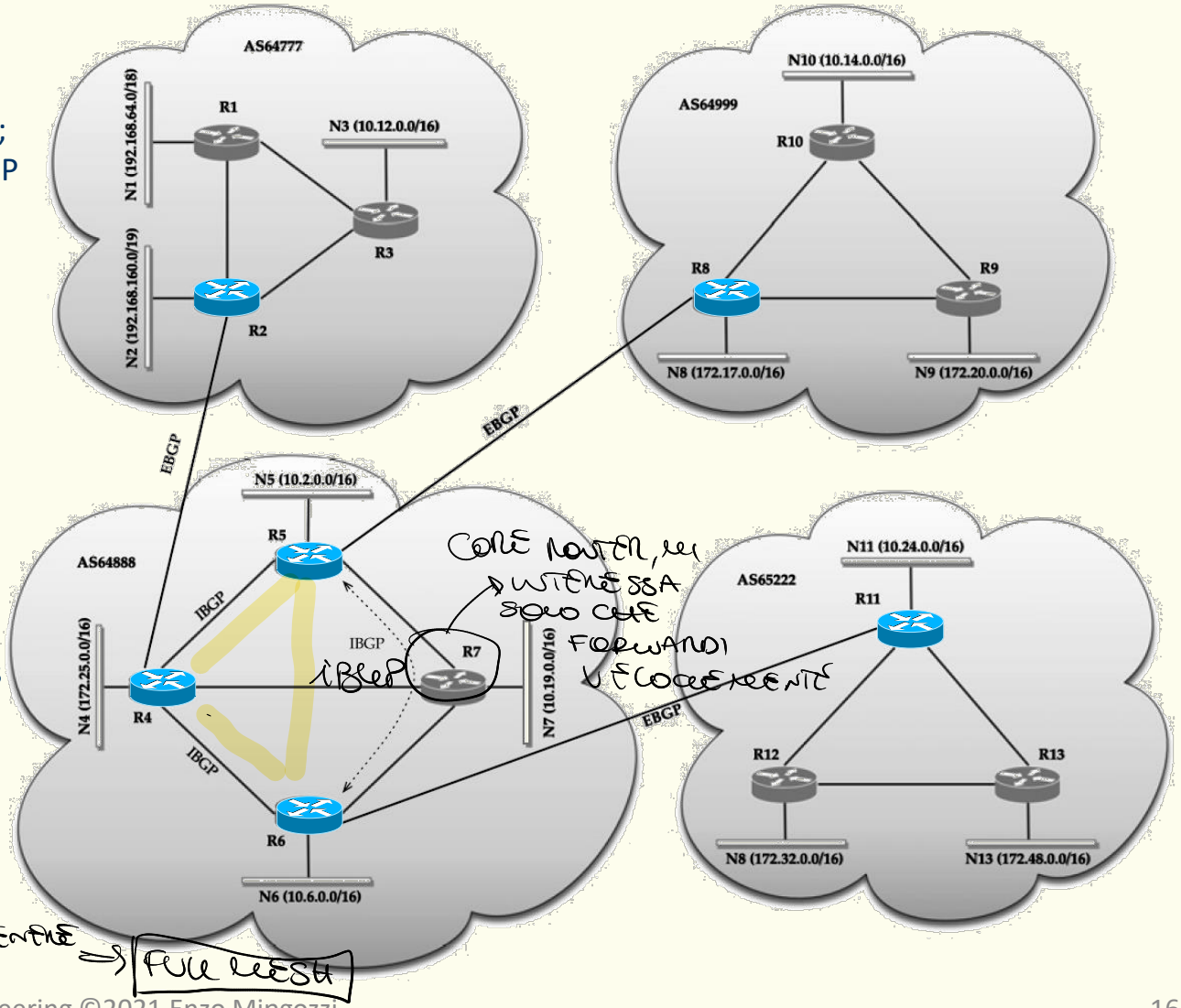
Rules

Rule 1 A BGP speaker can advertise IP prefixes it has learned from an eBGP speaker to a neighboring iBGP speaker; similarly, a BGP speaker can advertise IP prefixes it has learned from an iBGP speaker to an eBGP speaker

Rule 2 An iBGP speaker cannot advertise IP prefixes it has learned from an iBGP speaker to another peer iBGP speaker

Two reasons:

1. Avoid looping of BGP route updates within the AS
2. No need to advertise internal routes



Rules

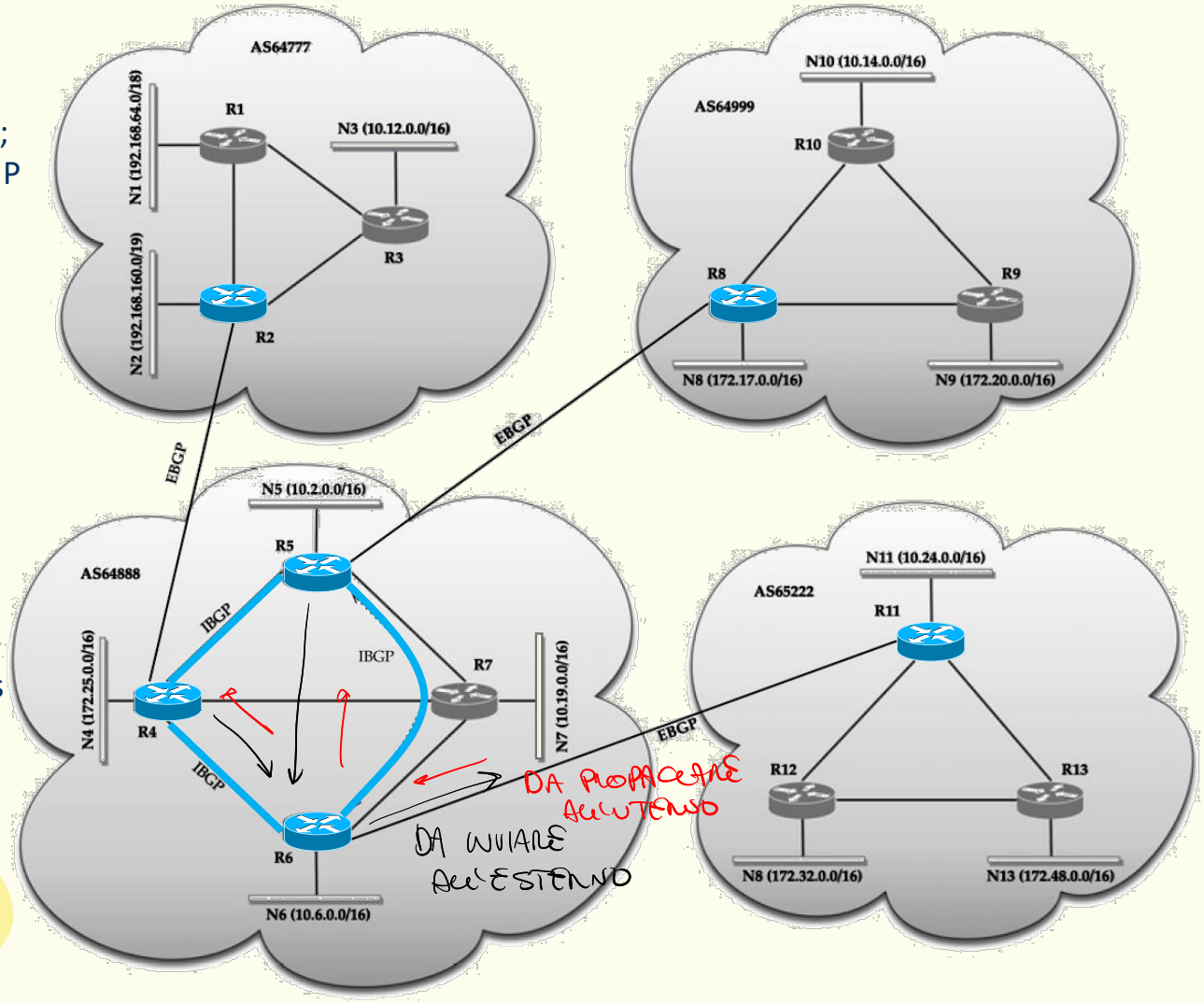
Rule 1 A BGP speaker can advertise IP prefixes it has learned from an eBGP speaker to a neighboring iBGP speaker; similarly, a BGP speaker can advertise IP prefixes it has learned from an iBGP speaker to an eBGP speaker

Rule 2 An iBGP speaker cannot advertise IP prefixes it has learned from an iBGP speaker to another peer iBGP speaker

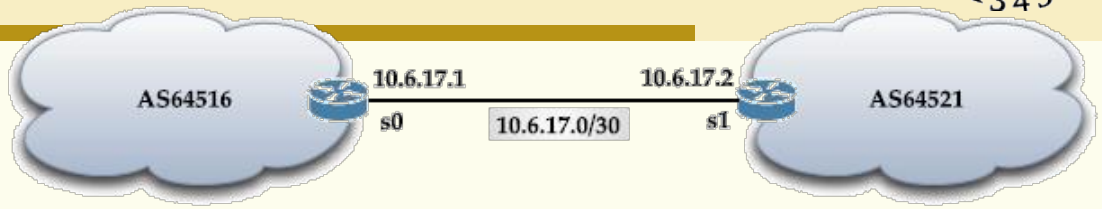
Two reasons:

1. Avoid looping of BGP route updates within the AS
2. No need to advertise internal routes

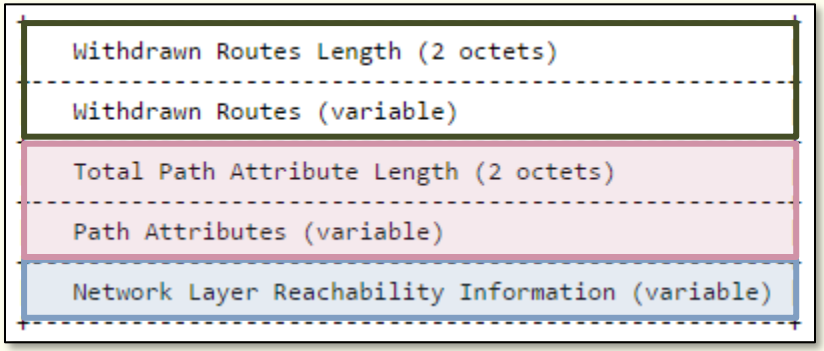
A full mesh iBGP connectivity is needed



PATH attributes



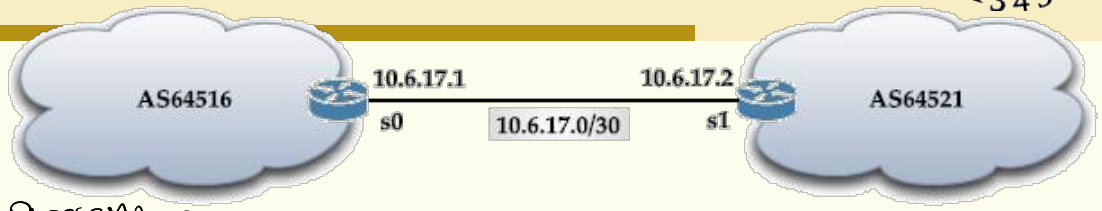
- **UPDATE** message
  - Path attributes advertised for a set of routes
  - Keep track of route-specific information such as path information, and degree of preference of a route
  - Used in the BGP filtering and route decision process



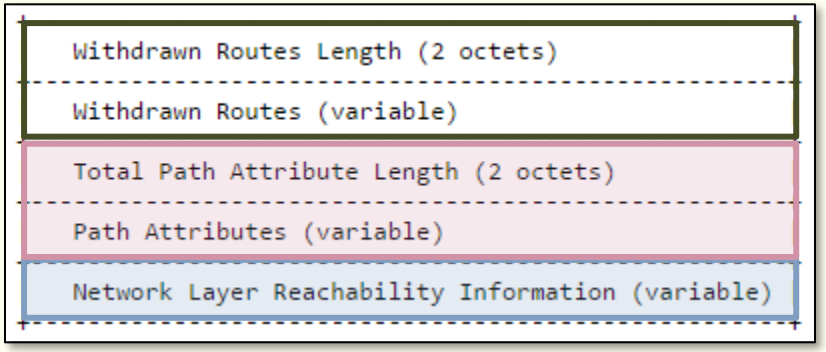
A sequence of path attributes (TLV)
- ORIGIN
- AS_PATH
- NEXT_HOP
- LOCAL_PREF
- …

List of networks {address/prefix length}

PATH attributes



- **UPDATE** message
  - **Well-known mandatory**: Must be present in all UPDATE messages
  - **Well-known discretionary**: Could be present in UPDATE messages
  - **Optional transitive**: If not recognized, are propagated to other neighbors
  - **Optional non-transitive**: Discarded if not recognized

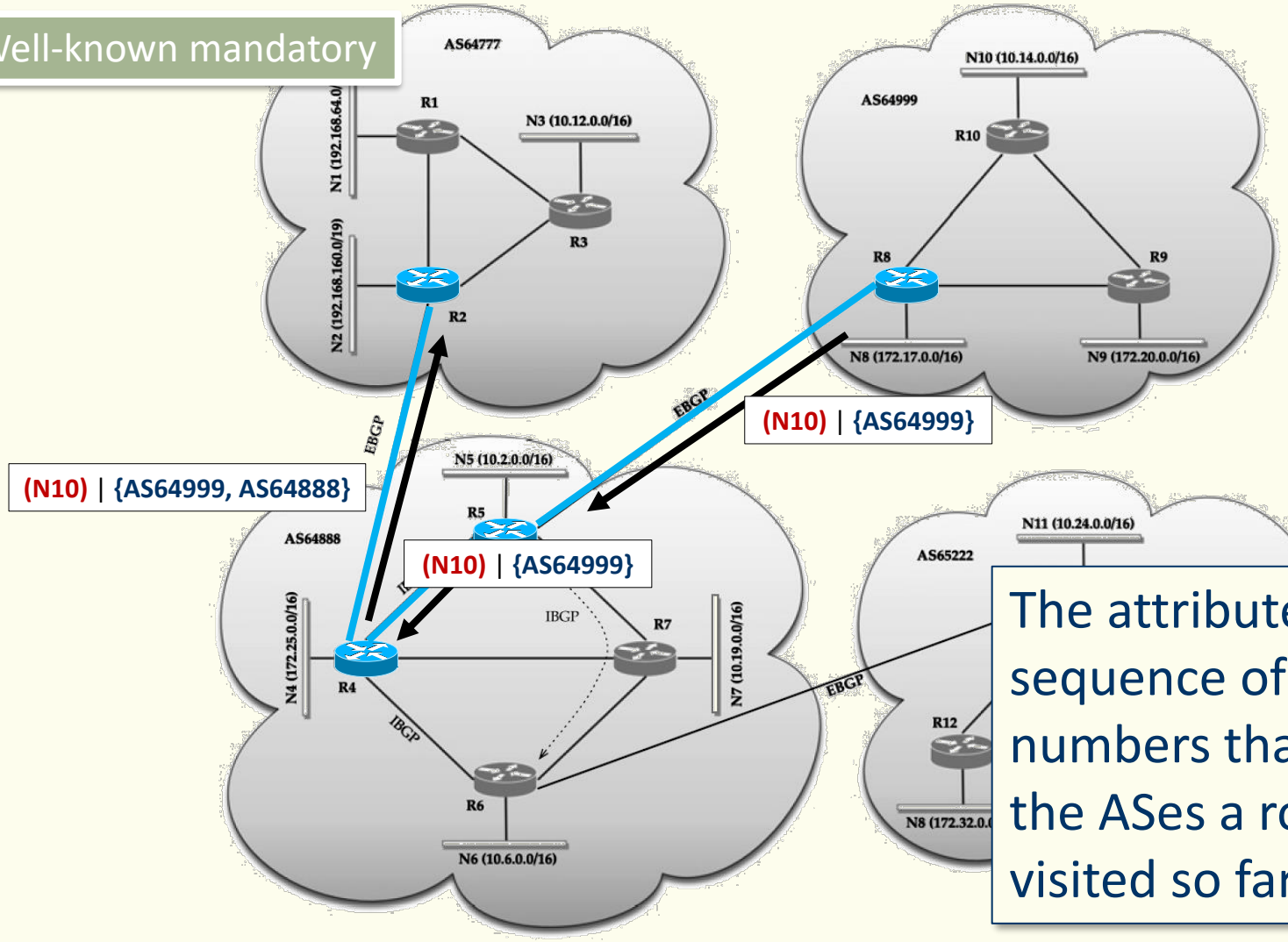


A sequence of path attributes (TLV)
- ORIGIN
- AS_PATH
- NEXT_HOP
- LOCAL_PREF
- …

List of networks {address/prefix length}

AS_PATH

Well-known mandatory

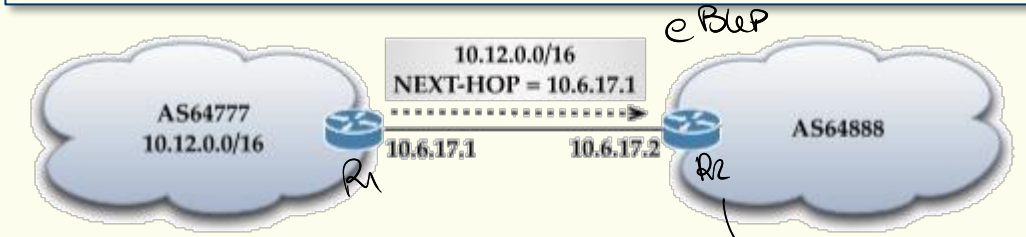


The attribute stores a sequence of AS numbers that identifies the ASes a route has visited so far

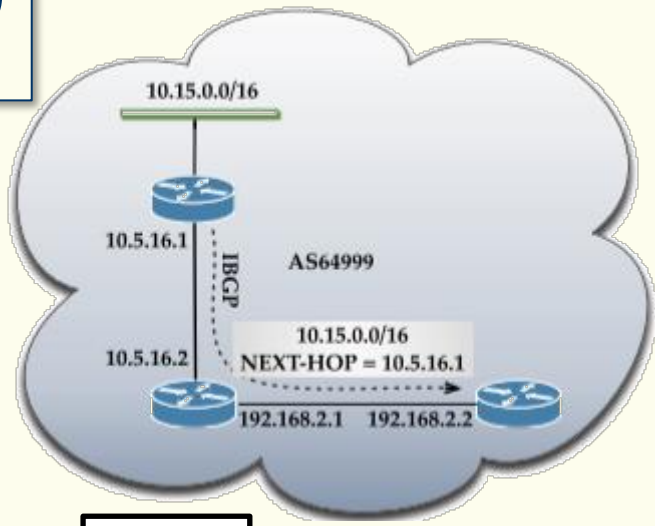
NEXT_HOP

The attribute defines the IP address of the router that SHOULD be used as the next hop (not necessarily 1-hop) to the destinations listed in the UPDATE message

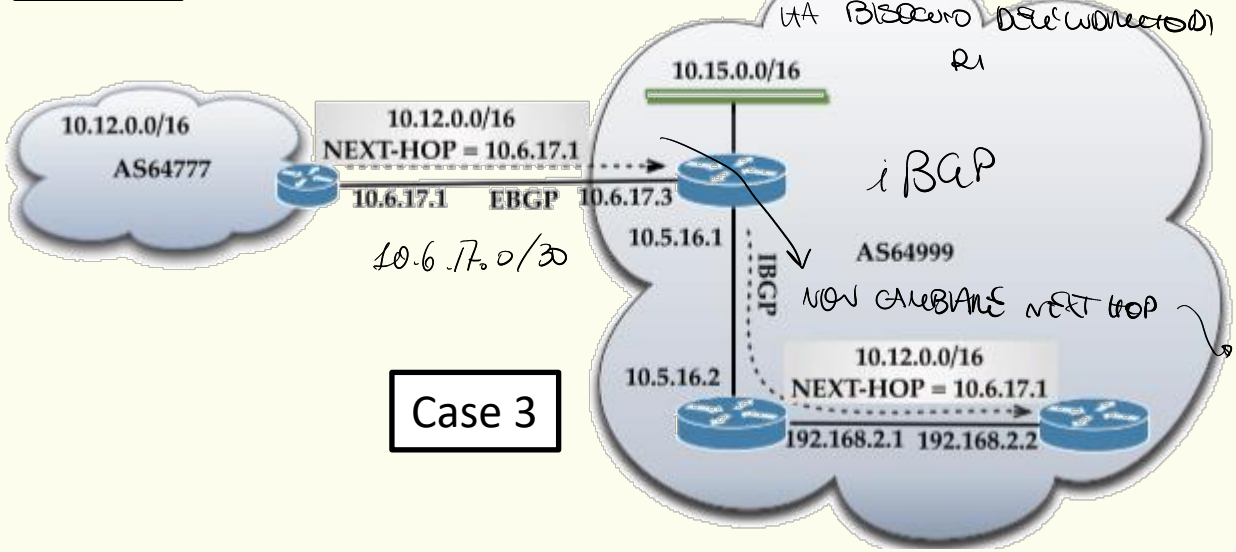
Well-known mandatory



Case 1



Case 2

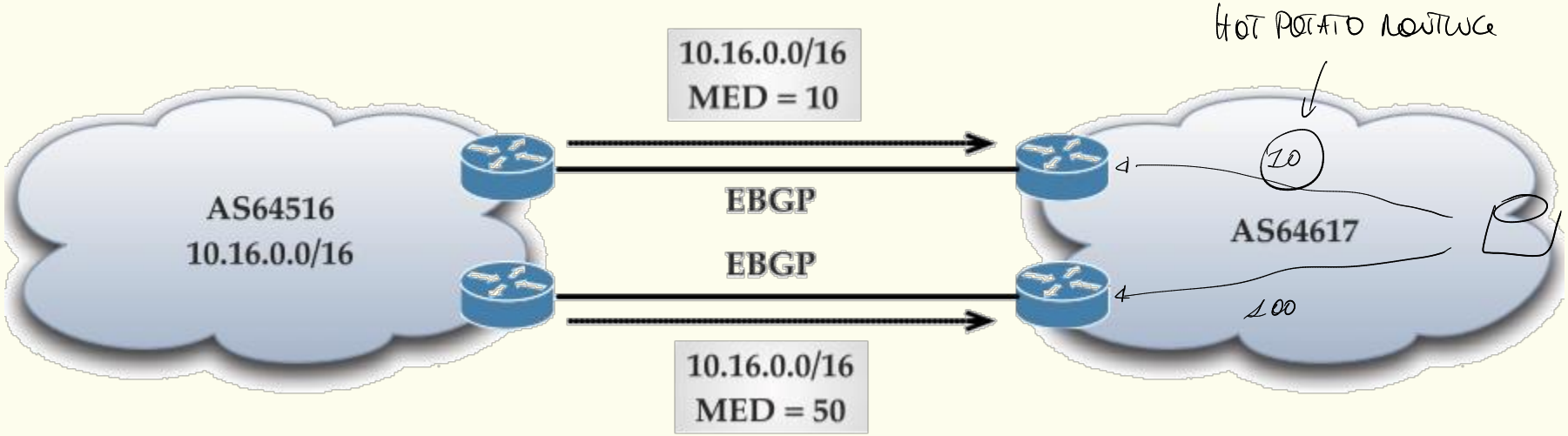


Case 3

MED (MULTI–EXIT–DISCRIMINATOR)

The attribute is a metric meant for use when there are multiple external links to a neighboring AS

Optional non-transitive



LOCAL_PREF

The attribute defines a metric used internally within an AS between BGP speakers, helpful in selection when the AS has connectivity to multiple ASes

Well-known discretionary