

ΔΙΑΧΕΙΡΙΣΗ ΔΙΚΤΥΩΝ

Αρχιτεκτονική & Δρομολόγηση στο Internet

Επίπεδο 3:

Direct Routing, Interior Gateway Protocols (OSPF, IS-IS), Border Gateway Protocols (BGP)

Επίπεδο 2:

Ethernet Switches, Virtual Local Area Networks (VLANs), Spanning Tree Protocol (STP), Provider Backbone Bridges (PBB)

Επίπεδο 2.5:

Multi-Protocol Label Switching (MPLS)

B. Μάγκλαρης

maglaris@netmode.ntua.gr

www.netmode.ntua.gr

5/11/2018

ΤΟ ΔΙΚΤΥΟ ΤΟΥ Ε.Μ.Π. (επανάληψη)

ntua.gr (147.102.0.0/16, ASN 3323)

ΠΡΟΣΟΧΗ

Οι πίνακες δρομολόγησης στο Internet για λόγους ομοιομορφίας είναι της μορφής:

- **Prefix Δικτύου Τελικού Προορισμού ::
IP Interface Εισόδου Επόμενου Κόμβου**

ΠΑΡΑΔΕΙΓΜΑ:

Ο δρομολογητής του Ε.Μ.Π. **147.102.224.33**
βρίσκει τον δρομολογητή του ΕΚΠΑ

147.102.224.34 σαν μέλος του υποδικτύου:

- **147.102.224.32/30** (παροχή διευθύνσεων από Ε.Μ.Π.)

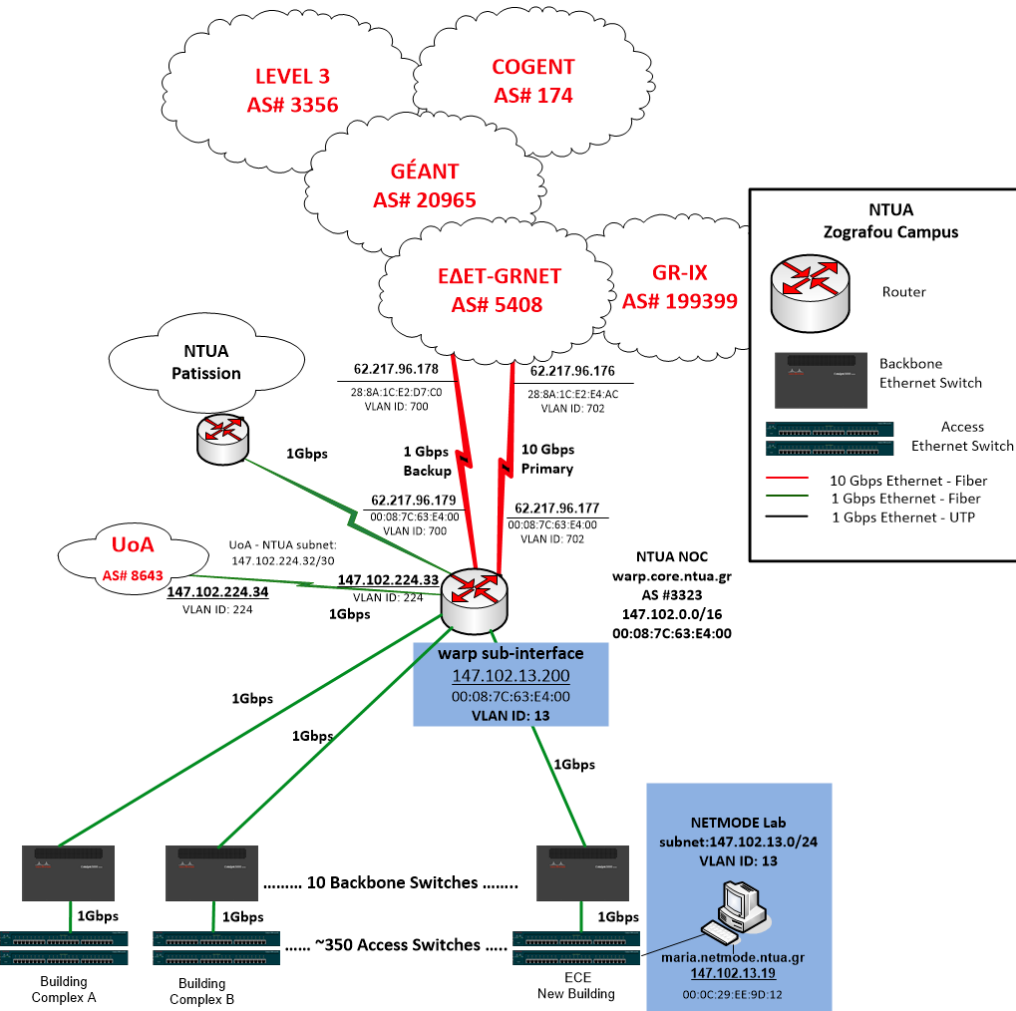
Η γραμμή Ε.Μ.Π. – ΕΚΠΑ (όπως όλες οι γραμμές σε Δίκτυα Internet) ορίζεται σαν υποδίκτυο (prefix) με 4 τουλάχιστον διευθύνσεις IP:

- Υποδίκτιο: **147.102.224.32**
- Άκρο Ε.Μ.Π.: **147.102.224.33**
- Άκρο ΕΚΠΑ: **147.102.224.34**
- Broadcast: **147.102.224.35**

ANTI-ΠΑΡΑΔΕΙΓΜΑ:

Ο δρομολογητή του Ε.Μ.Π. **62.217.96.177** βρίσκει τον δρομολογητή του ΕΔΕΤ **62.217.96.176** σαν μέλος του υποδικτύου:

- **62.217.96.176/31** (παροχή διευθύνσεων από ΕΔΕΤ)



ΣΥΝΔΕΣΕΙΣ ΜΕΤΑΞΥ ΔΡΟΜΟΛΟΓΗΤΩΝ (Links between Routers) *(επανάληψη)*

- Για ομοιομορφία της δρομολόγησης, κάθε γραμμή ορίζεται (συνήθως) σαν δίκτυο με 4 τουλάχιστον διευθύνσεις (/30)
- Παράδειγμα: Μεταξύ ΕΜΠ **147.102.0.0/16** & Παν. Αθηνών **195.134.64.0/18** ορίζεται το «δίκτυο» **147.102.224.32/30**
 - Υποδίκτυο: **147.102.224.32/30**
 - Άκρο ΕΜΠ: **147.102.224.33/30**
 - Άκρο Παν. Αθηνών: **147.102.224.34/30**
 - Broadcast: **147.102.224.35/30**

ΑΛΓΟΡΙΘΜΟΙ ΕΥΡΕΣΗΣ ΔΡΟΜΩΝ ΣΤΟ ΕΠΙΠΕΔΟ 3 ΤΟΥ INTERNET (επανάληψη) LS (Link State) - Αλγόριθμος Dijkstra

- Παράδειγμα εφαρμογής: **IGP - OSPF** (Open Shortest Path First)
 - **Link State Data Base** + αλγόριθμος **Dijkstra** στον κορμό Αυτόνομου Δικτύου (Core of an Autonomous System, OSPF Area 0):
 - Τρέχει σε όλους τους δρομολογητές κορμού που πρέπει να έχουν πλήρη & ενιαία εικόνα της κατάστασης - τοπολογίας του δικτύου για **συμβατότητα υπολογισμού πινάκων δρομολόγησης**
 - **Κόστος γραμμών** ανάλογο με την ταχύτητα ή οριζόμενα από τον Διαχειριστή
 - Ανακοινώσεις κατάστασης δρομολογητών κορμού (OSPF Area 0) και συνδέσεων: μέσω **LSA** (Link State Advertisements) **v2** για IPv4 ή **v3** για IPv6 (**IP signals** χωρίς TCP/UDP transport layer)
 - **Ανανέωση LSA**: Ανά 30 min ή με αλλαγή κατάστασης ή αν εξαντληθούν 60 min
 - Στα περιφερειακά υποδίκτυα, **OSPF Stub Areas**: Στατική επιλογή **Default Gateway**

ΑΛΓΟΡΙΘΜΟΙ ΕΥΡΕΣΗΣ ΔΡΟΜΩΝ ΣΤΟ ΕΠΙΠΕΔΟ 3 ΤΟΥ INTERNET

DV (Distance Vector) - Αλγόριθμος Bellman - Ford

- Παράδειγμα εφαρμογής: **EGP - BGP** (Border Gateway Protocol)
 - **e-BGP**: External BGP → Πίνακες σε Border Gateways με εκτιμήσεις ενδιαμέσων public AS's (έως 62.000) προς 750.000 prefixes (public δίκτυα - προορισμοί)
 - **i-BGP**: Internal BGP (μεταξύ δρομολογητών κορμού ενός AS)
 - Για προορισμούς ενθυλακωμένους σε πολλαπλά prefixes: Προτίμηση βάσει **longest prefix match**
 - Υπολογισμός **reachability & AS paths** ανά prefix σε Border Gateways: Με βάση **advertisements** (**TCP signals**) από γειτονικά AS's και αλγόριθμο δρομολόγησης **Bellman-Ford**
 - Επιλογή μεταξύ εναλλακτικών δρόμων για προώθηση πακέτων στον forwarding table των Border Gateways με βάση διαχειριστικές πολιτικές (**weight, preferences...**)

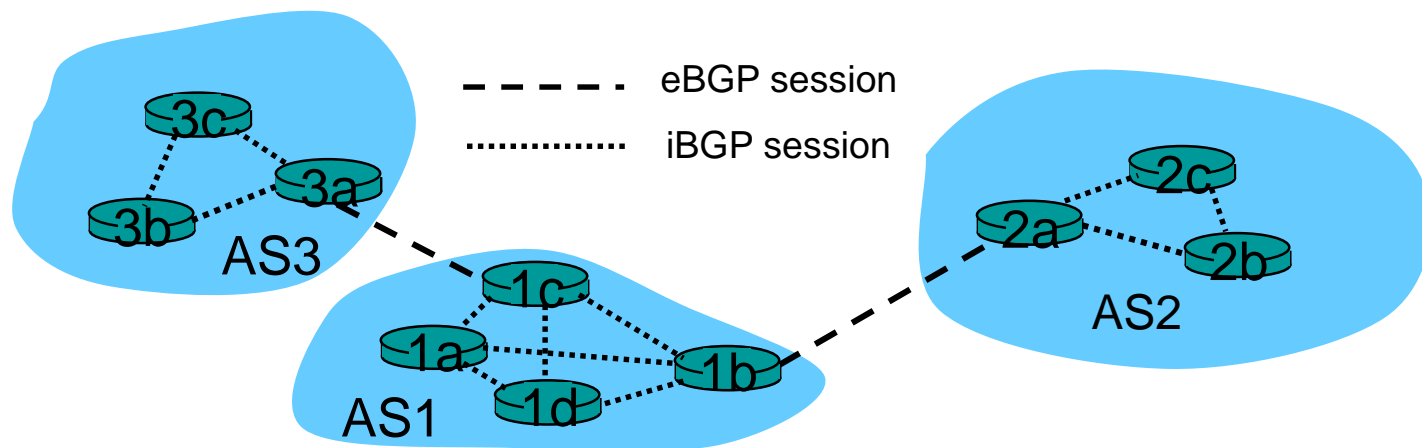
ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Inter-AS Routing, Σηματοδοσία BGP

- Οι πίνακες δρομολόγησης **RIB** (Routing Information Base) προς τα 750,000 δημόσια **prefixes** (δίκτυα) στο Internet, τηρούνται στην ηλεκτρονική μνήμη των συννοριακών δρομολογητών (**border routers**) των 62,000 Αυτόνομων Συστημάτων (**AS**)
- Ένας border router μπορεί να μην καταγράφει πληροφορία για προορισμούς τους οποίους «βλέπει» **αποκλειστικά** μέσω άλλου AS υψηλότερης ιεραρχίας (π.χ. Ε.Μ.Π. – GRNET) ή να συναθροίζει (**aggregate**) διαδοχικά prefixes (δίκτυα προορισμούς)
- Η πληροφορία για τα δίκτυα (prefixes) που «βλέπει» ένας border router από τις επιλογές εξόδου του, τηρείται σε πίνακα **NLRI** (Network Layer Reachability Information) που ανανεώνεται μέσω της σηματοδοσίας BGP
- Η σηματοδοσία (**signaling επιπέδου ελέγχου**) του **BGP** υλοποιείται από Control Messages που διακινούνται με πρωτόκολλο TCP για αξιόπιστο έλεγχο κυκλοφορίας (flow control). Οι εντολές ελέγχου του BGP είναι:
 - **OPEN**: ανοίγει TCP σύνδεση στο γείτονα (peer) και προαιρετικά ταυτοποιεί το απέναντι άκρο
 - **UPDATE**: ανακοινώνει νέα path ή αποσύρει (withdraws) παλαιότερα
 - **KEEPALIVE**: κρατάει την σύνδεση ανοιχτή σε περίπτωση που δεν υπάρχουν UPDATES ή ACK σε αίτηση OPEN
 - **NOTIFICATION**: ανακοίνωση σφαλμάτων σε προηγούμενα μηνύματα ή για να κλείσει η σύνδεση

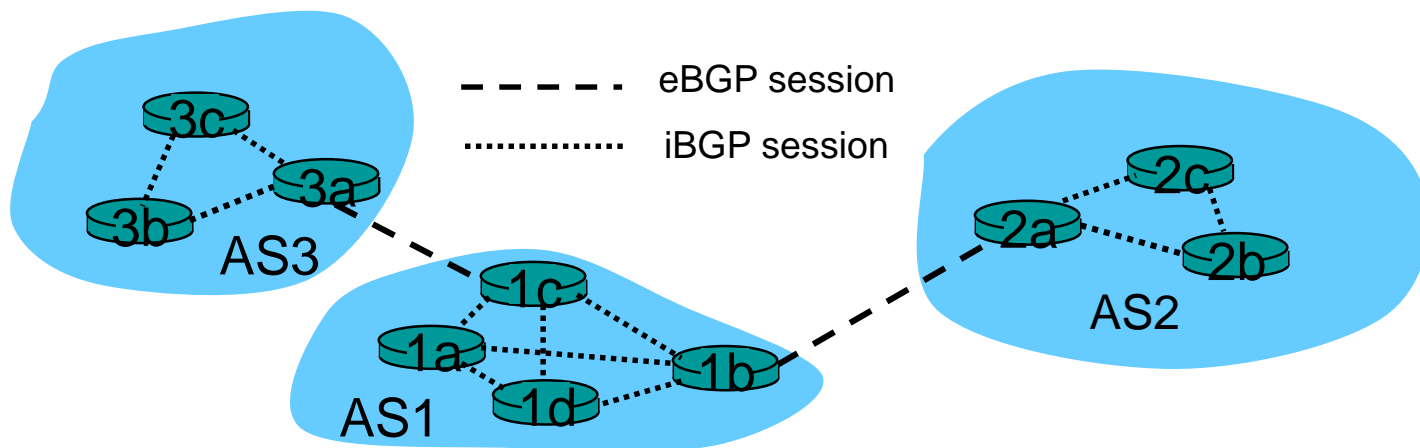
ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΟΥ BGP

- Τα ζεύγη από συνοριακούς δρομολογητές (BGP peers) ανταλλάσσουν πληροφορίες δρομολόγησης (routing info) πάνω από ημι-σταθερές συνδέσεις TCP: **BGP sessions**
 - Οι BGP sessions δεν χρειάζεται να αντιστοιχίζονται σε φυσικές συνδέσεις links
- Όταν το AS2 ανακοινώνει ένα πρόθεμα (prefix υποδικτύου προορισμού) προς AS1:
 - Το AS2 **υπόσχεται** ότι θα προωθεί πακέτα με διεύθυνση προορισμού που να ανήκει στο δεδομένο prefix
 - Το AS2 μπορεί να συναθροίσει (aggregate) prefixes υποδικτύων στις ανακοινώσεις του



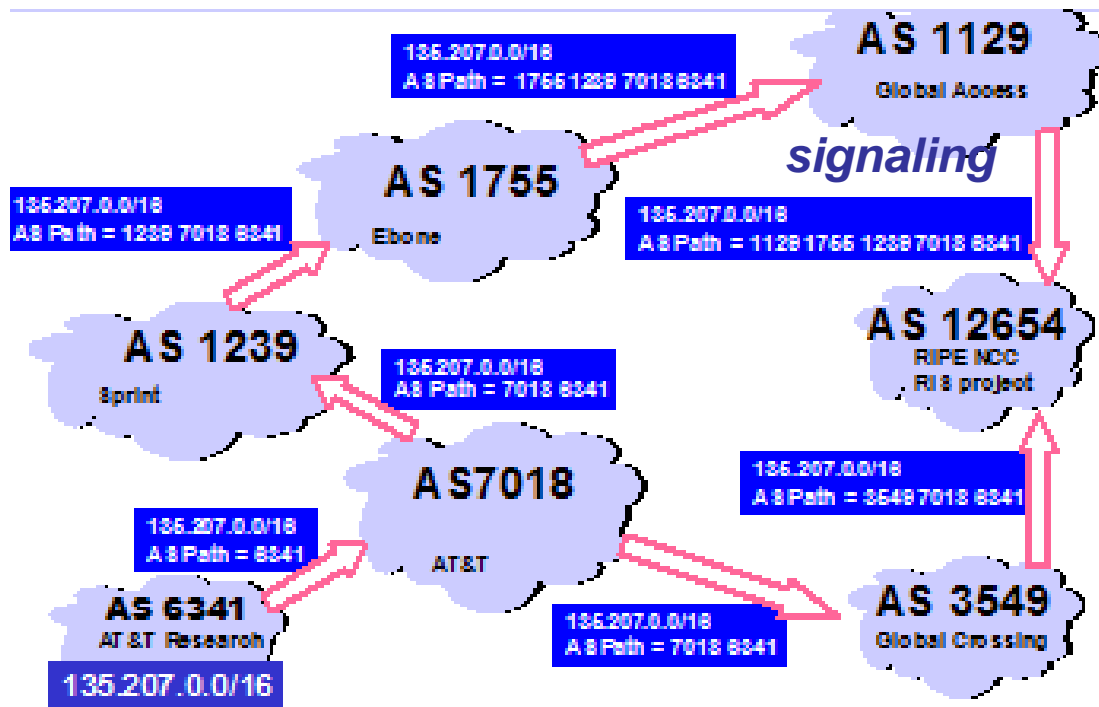
ΔΙΑΝΟΜΗ BGP REACHABILITY INFO

- Με χρήση σύνδεσης TCP, το πρωτόκολλο **eBGP (external BGP)** μεταξύ των border gateways 3a και 1c στέλνει **prefix reachability info** της AS3 στην AS1
 - 1c μπορεί να χρησιμοποιήσει **iBGP (internal BGP)** για διανομή νέων **prefix reachability info** σε όλους τους δρομολογητές κορμού της AS1
 - 1b μπορεί να ξανα-ανακοινώσει νέο **prefix reachability info** στο AS2 πάνω από σύνδεση eBGP μεταξύ 1b-to-2a
- Ένας δρομολογητής όταν μαθαίνει νέο **network prefix**, δημιουργεί routing entry στο πίνακα προώθησης (**forwarding table**)
- Οι δρομολογητές που μετέχουν στο iBGP μέσα σε μια AS πρέπει να είναι απ' ευθείας διασυνδεδεμένοι (**fully connected iBGP routers**)



ΑΝΑΚΟΙΝΩΣΗ ΔΙΚΤΥΟΥ 135.207.0.0/16 ΜΕΣΩ BGP

(από παρουσίαση του Timothy G. Griffin,
AT&T Research, Paris 2002)



BGP 4: RFC 4271

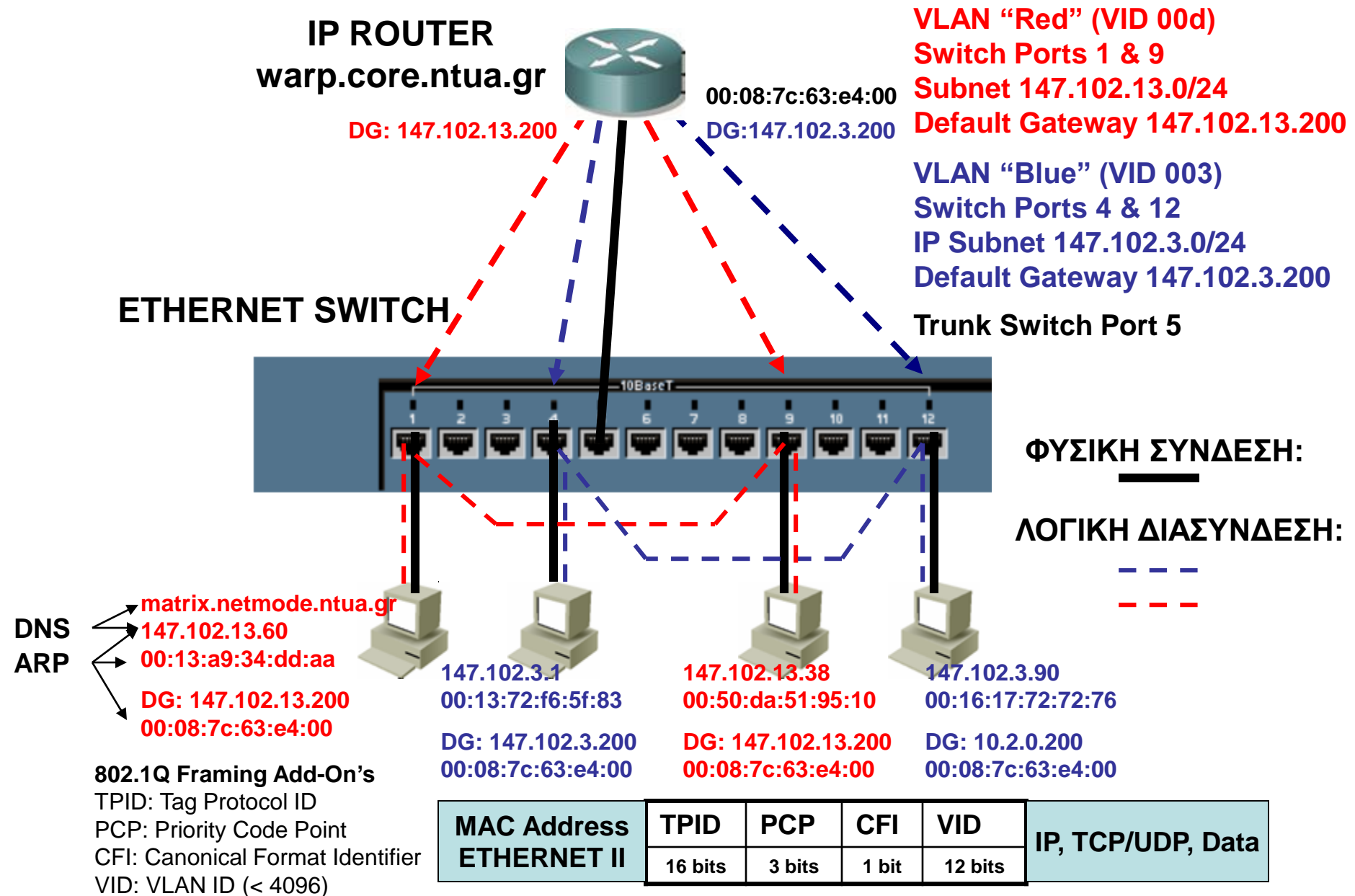
Signaling over
TCP Port 179

iBGP: Internal BGP
(pass inter-AS peering
to intra-AS fully connected
routers)

eBGP: External BGP
(between AS's over direct
router inter-AS links)

ΔΡΟΜΟΛΟΓΙΣΗ ΕΠΙΠΕΔΟΥ 2 - LINK LAYER

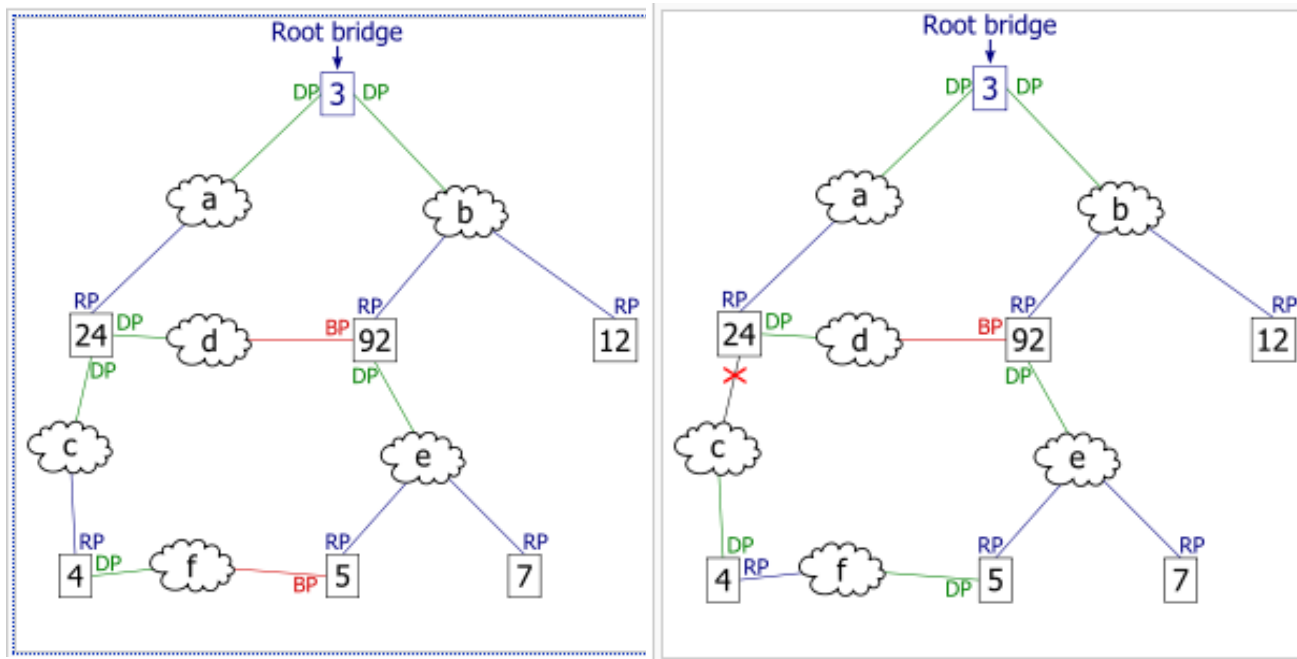
ETHERNET & ΕΙΚΟΝΙΚΑ ΤΟΠΙΚΑ ΔΙΚΤΥΑ VLAN (IEEE 802.1Q)



ΠΡΩΤΟΚΟΛΛΟ ΔΙΑΜΟΡΦΩΣΗΣ ΔΕΝΔΡΙΚΗΣ ΤΟΠΟΛΟΓΙΑΣ ΜΕΤΑΓΩΓΕΩΝ ETHERNET (1/2)

Spanning Tree Protocol - STP, IEEE 802.1D

- Εξέλιξη των Αλγορίθμων Διάρθρωσης Διαφανών Γεφυρών **Spanning Tree Protocol (STP) for Transparent Ethernet Bridges** → **STP Ethernet Switches**
- Radia Perlman**, DEC & MIT 1985 <http://www1.cs.columbia.edu/~ji/F02/ir02/p44-perlman.pdf>
- Αναδιαμόρφωση Spanning Tree http://en.wikipedia.org/wiki/Spanning_tree_protocol
- Χρόνος Αντίδρασης σε Βλάβη: ~ **60 sec**



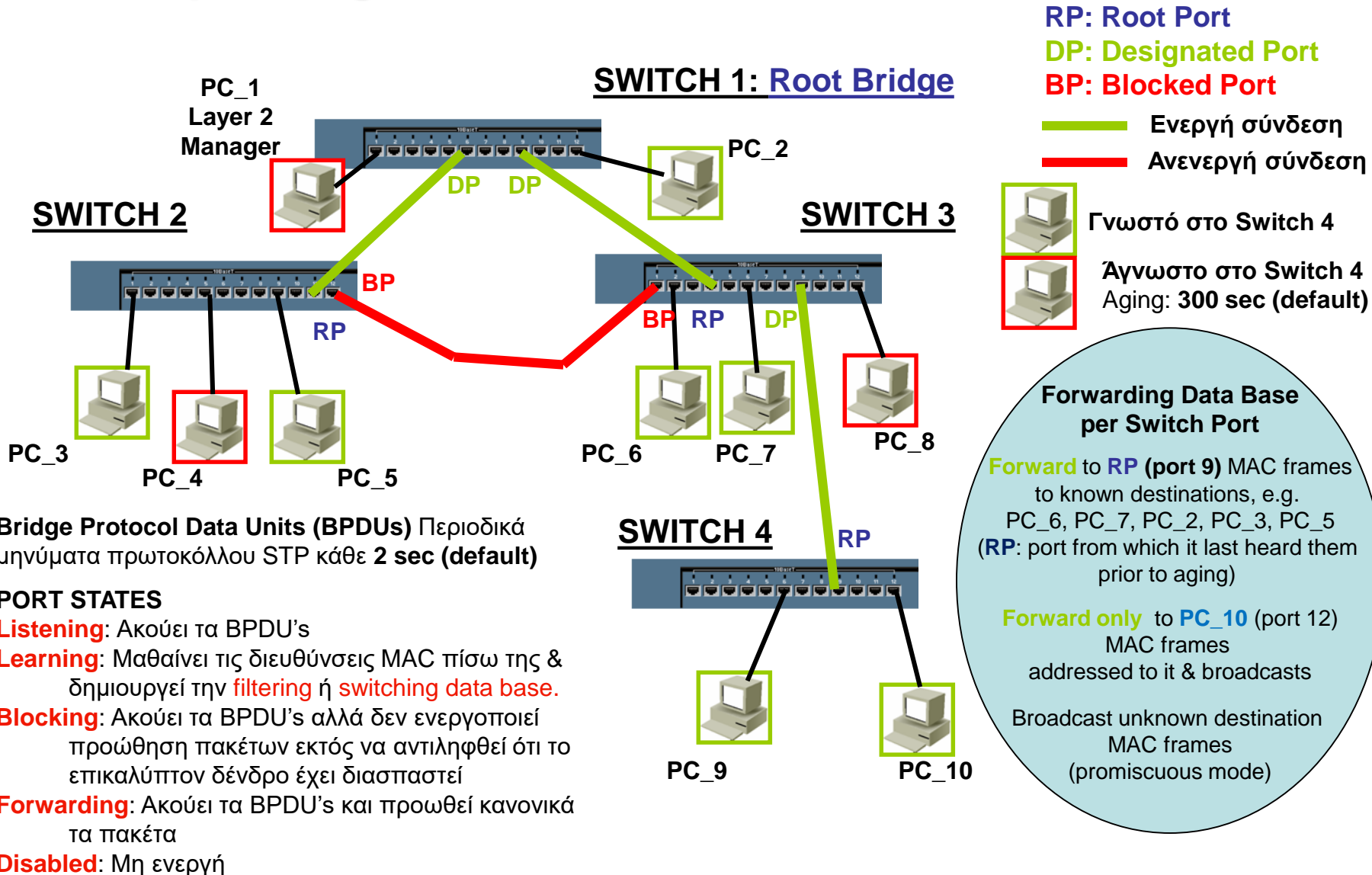
Γέφυρες (Bridges, Switches):
3 (**Root**), 24, 92, 4, 5, 7, 12

Τοπικά δίκτυα Ethernet:
a, b, c, d, e, f

RP: Root Port
DP: Designated Port
BP: Blocking Port

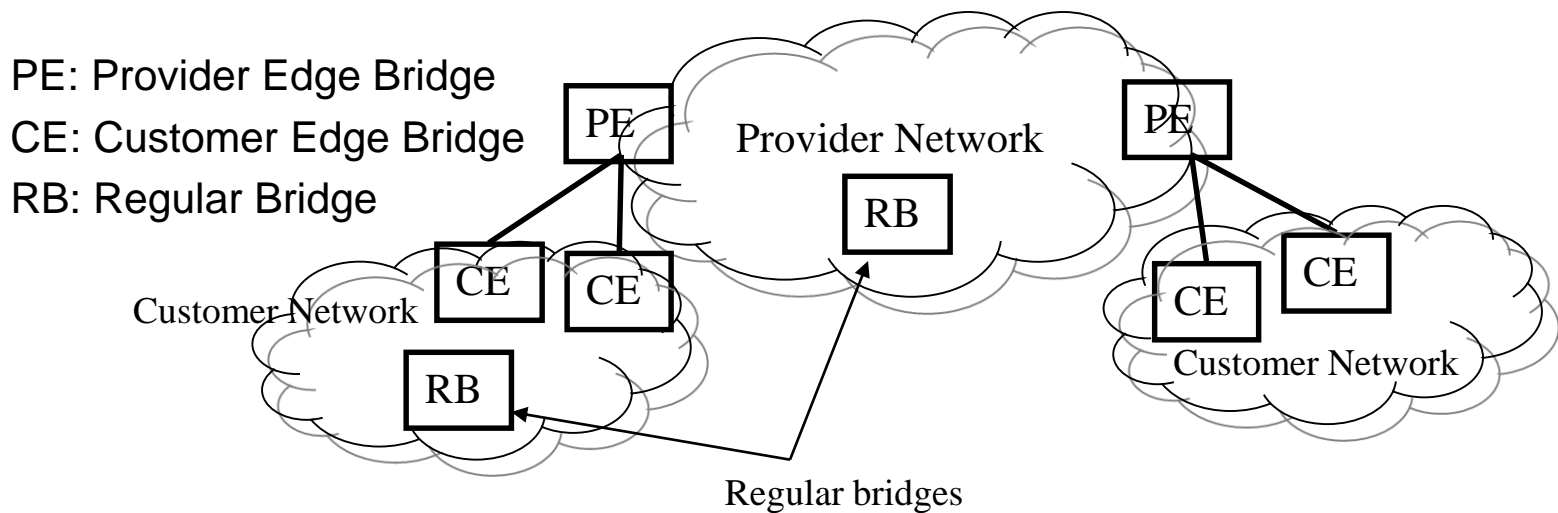
ΠΡΩΤΟΚΟΛΛΟ ΔΙΑΜΟΡΦΩΣΗΣ ΔΕΝΔΡΙΚΗΣ ΤΟΠΟΛΟΓΙΑΣ ΜΕΤΑΓΩΓΕΩΝ ETHERNET (2/2)

Spanning Tree Protocol - STP, IEEE 802.1D



ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ 2 ΣΕ ΔΙΚΤΥΑ ΠΑΡΟΧΩΝ

Provider Backbone Bridges – PBB (mac-in-mac, IEEE 802.1ah)



IEEE 802.1ah (2008): Επέκταση Ethernet **LAN's** (**IEEE 802.3z: 1 GigE, IEEE 802.3ae: 10 GigE**) σε Μητροπολιτικά Δίκτυα **MAN** & Δίκτυα Κορμού Ευρείας Περιοχής **WAN** (**1-10-40-100 GigE**)

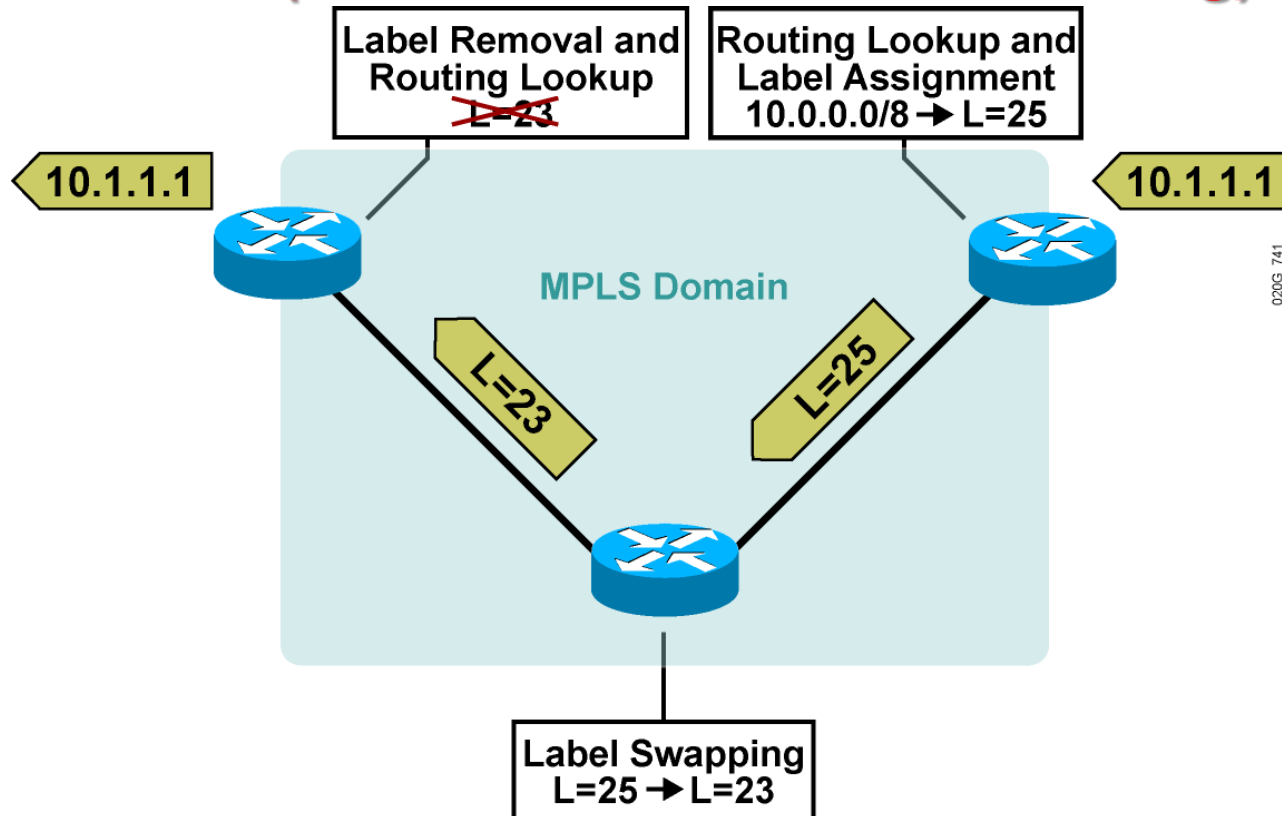
- Τυποποίηση πρωτοκόλλων **VPLS**, **mac-in-mac** και **QinQ** tunnels για επέκταση VLAN's μεταξύ τοπικών δικτύων LAN's μέσω Layer 2 VPNs
- Προς συρρίκνωση τοπολογίας επιπέδου 3 → collapsed backbone με μηχανισμούς μεταφοράς επιπέδου 2: **10-100 Gig point-to-point Ethernet transport**

ΠΑΡΑΔΟΣΙΑΚΗ ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ 3

<http://labjarkom.ilkom.unsri.ac.id/userfiles/MPLS-1.ppt>

ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ 2.5:

MPLS (Multi-Protocol Label Switching)

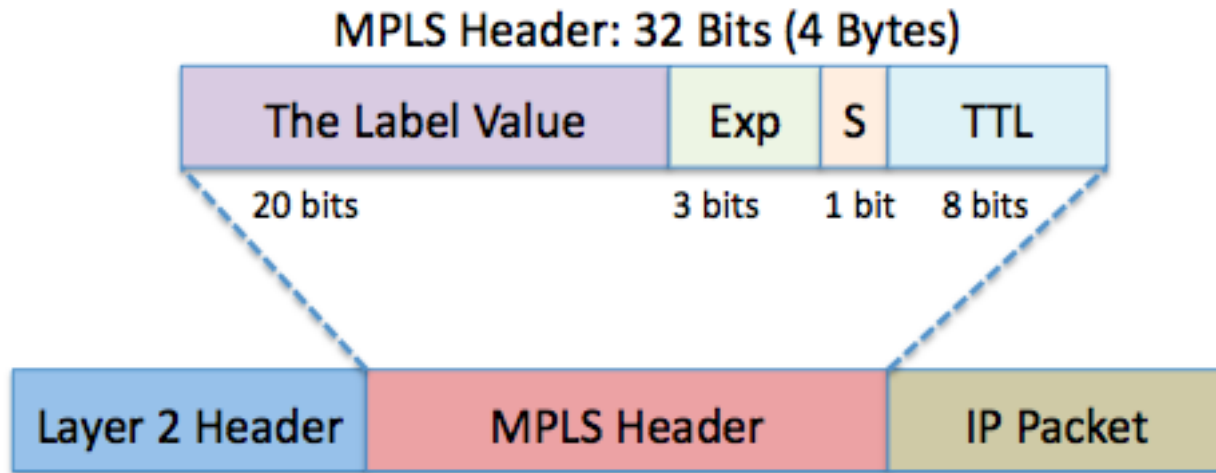


MPLS core routers : Label Switch Router – LSR
Αντικαθιστούν (swap) Labels
Πρωθούν τα πακέτα με βάση πίνακες δρομολόγησης ανά Label

MPLS edge routers: Edge LSR, Label Edge Router – LER
Εισάγουν/διαγράφουν επικεφαλίδες MPLS
Δρομολογούν με βάση πίνακες δρομολόγησης IP και Labels

MPLS HEADER

<http://blog.ine.com/2010/02/21/the-mpls-forwarding-plane/>



- **Label :** Label value (0 to 15 are reserved for special use) assigned to **source & destination IP (flows) (traffic engineering option)**
- **Exp :** Experimental Use
- **S :** Bottom of Stack (set to 1 for the last entry in the label)
- **TTL :** Time To Live