

ΔΙΑΧΕΙΡΙΣΗ ΔΙΚΤΥΩΝ

Αρχιτεκτονική & Δρομολόγηση στο Internet (Τμήμα 2/2)

Ορισμοί Δρομολόγησης

Άμεση – Έμμεση Δρομολόγηση

Δρομολόγηση εντός Αυτόνομης Περιοχής (IGP)

Δρομολόγηση μεταξύ Αυτονόμων Περιοχών (BGP)

Β. Μάγκλαρης

maglaris@netmode.ntua.gr

www.netmode.ntua.gr

16/10/2017

ΤΑΞΙΝΟΜΗΣΗ ΑΛΓΟΡΙΘΜΩΝ ΔΡΟΜΟΛΟΓΗΣΗΣ

- **Διαχειριστικός Έλεγχος**
 - Συγκεντρωτικός: Δημόσια Δίκτυα Μεταγωγής Πακέτου **X.25/X.75**
 - Κατανεμημένος: **Internet**
- **Δρομολόγηση με ή χωρίς Εναλλακτικούς Δρόμους**
 - Single choice:
 - Δρομολόγηση Επιπέδου 3 στο Internet: **IGP** (OSPF, IS-IS), **BGP**
 - Δρομολόγηση Επιπέδου 2: **Bridged Ethernet LANs**
 - Alternate routing:
 - Δρομολόγηση Επιπέδου 3 στο Internet: **IGP ECMP** (Equal-Cost Multi-Path)
 - Δρομολόγηση Επιπέδου 2 ½: **MPLS/TE** (Multi-Protocol Label Switching / Traffic Engineering)
 - Τηλεφωνικά Δίκτυα **PSTN** (Public Switched Telephone Networks)
- **Διαχειριστική Δυναμική**
 - Στατική δρομολόγηση:
 - **Direct Routing** στο Internet - Default Gateway
 - Υπολογισμός εναλλακτικών δρόμων με σταθερά βάρη γραμμών (**X.25/X.75**)
 - Δυναμική δρομολόγηση:
 - Μεταβαλλόμενα βάρη γραμμών (**ARPANet**)
 - Μεταβολές ανάλογα με διαθεσιμότητα γραμμών (**Internet IGP, BGP**)

ΠΡΟΤΥΠΟ ΤΡΙΩΝ ΔΙΑΣΤΑΣΕΩΝ (επανάληψη)

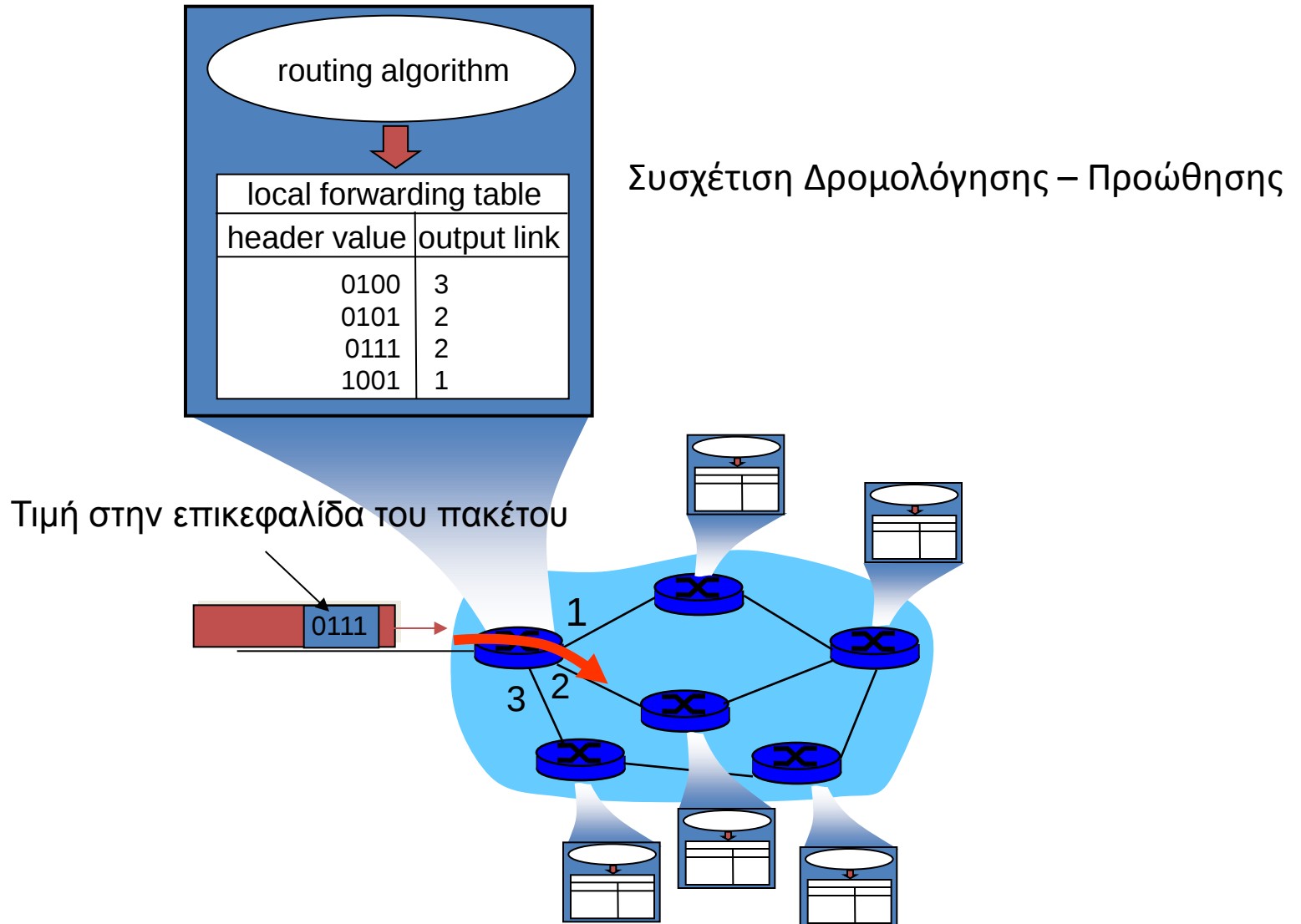
- **Διάσταση Μετάδοσης Δεδομένων - Data (forwarding) Plane**
 - Πολυπλεξία στο φυσικό επίπεδο:
 - Διαμόρφωση πλαισίων TDM: ITU-T SDH/GFP framing (από STM-1=155 Mbps → STM- n , εφεδρεία ring protection, virtual concatenation (150 Mbps VC-4, 1 Gbps VC-4-7v = 7 x VC-4))
 - Optical Digital Wrapper (ITU-T G.709: 2.5, 10, 40, 100 Gbps, Forward Error Correction - FEC)
 - Κωδικοποίηση σε πακέτα Ethernet, WiFi (IEEE 802.11), MPLS, IP
 - Προώθηση (forwarding) δεδομένων σε μεταγωγείς (switches) & δρομολογητές (routers)
- **Διάσταση Ελέγχου - Control Plane**
 - **In-band Signaling**: Σηματοδοσία ενσωματωμένη σε επικεφαλίδες πακέτων (IP headers, MPLS labels, VLAN tags)
 - Ξεχωριστά μηνύματα / πακέτα ελέγχου για σύνταξη πινάκων δρομολόγησης (Interior Gateway Protocol – IGP, Exterior/Border Gateway Protocol – EGP/BGP)
 - Πακέτα ελέγχου «υγείας» του δικτύου – ICMP/ping/traceroute
 - Μηνύματα σηματοδοσίας για αποκατάσταση μονοπατιού – path (RSVP, LDP) & αντιστοίχιση επικεφαλίδων (*labels*) σε γραμμές MPLS
 - Σηματοδοσία αντιστοίχισης time slots (ή χρώματος) σε γραμμές SDH (ή WDM)
 - Πρωτόκολλα ARP & DNS, αντιστοίχιση VLAN tags....
 - **Out-of band Signaling**: Εξαγωγή των λειτουργιών ελέγχου εκτός μηχανισμών μετάδοσης, εξωτερικές βάσεις δεδομένων και εφαρμογές **ευφυούς δικτύου**:
 - Ψηφιακή τηλεφωνία: Σηματοδοσία Common Channel Signalling CCS7, **Intelligent Networks**
 - Προγραμματιζόμενα ευφυή δίκτυα νέας γενιάς (user programmable networks): **Software Defined Networks - SDN** σε τοπικά δίκτυα **Data Centers** & αρχιτεκτονικές **Future Internet**
- **Διάσταση Διαχείρισης - Management Plane**
 - Μοντέλο Αναφοράς FCAPS: Fault, Configuration, Accounting, Performance & Security Management

ΔΡΟΜΟΛΟΓΗΣΗ ΣΤΟ INTERNET

- **Προώθηση (forwarding):** Μετακίνηση πακέτων από την είσοδο δρομολογητή σε κατάλληλη έξοδο (*λειτουργία data plane*)
- **Δρομολόγηση (routing):** καθορισμός διαδρομής πακέτων από πηγή προς προορισμό, *routing algorithms* (*λειτουργία control plane*)

Σημείωση: Στις επόμενες διαφάνειες χρησιμοποιήθηκε υλικό υποστήριξης του βιβλίου των Kurose & Ross “**Computer Networking: A Top-Down Approach,**” Pearson Publisher, 6th Edition.

ΔΡΟΜΟΛΟΓΗΣΗ ΣΤΟ INTERNET



ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Layer 3 Intra-AS Routing

- **Intra- AS Routing, Interior Gateway Protocols (IGP):** Μια έξοδος προς επόμενο Interface για κάθε υποδίκτυο (subnet) τελικό προορισμό
 - RIP (Routing Internet Protocol): Πρωτόκολλο παλαιάς γενιάς, βασισμένο σε αλγόριθμους **distance vector (Bellman Ford)**
 - OSPF (Open Shortest Path First): Το πιο διαδομένο σήμερα, βασισμένο σε αλγόριθμους **link state (Dijkstra)**. Για λόγους κλιμάκωσης, μπορεί να υλοποιηθεί ιεραρχικά γύρω από την περιοχή κορμού (**backbone area 0** ή 0.0.0.0) και με συνδεόμενες περιφερειακές περιοχές (**stub areas**). Για δρομολόγηση πακέτων IPv4, ισχύει η OSPF Version 2: RFC 2328, 1998
 - IS-IS
 - Δυνατότητα πολλαπλών εναλλακτικών δρόμων ίσου κόστους (ECMP): Direct routing μεταξύ γειτονικών δρομολογητών με παράλληλες συνδέσεις και επιλογή εξόδου με proprietary αλγορίθμους (π.χ. Per-Packet Round-Robin, Per-Flow Load Balancing, Per Source -Destination Traffic Engineering...) Για αυτοματισμό του OSPF απαιτείται τροποποίηση του αλγορίθμου **Dijkstra για k-Shortest paths**

ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Layer 3 Inter-AS Routing

- **Inter-AS Routing, Exterior (Border) Gateway Protocols (EGP/BGP):** Πολλές εναλλακτικές διαδρομές με βάρη προς όλα τα γνωστά δίκτυα (περίπου 700.000 σήμερα) μεταξύ ακραίων (border) routers αυτονόμων συστημάτων (Autonomous Systems, AS, περίπου 60.000 σήμερα)
 - Η διαδρομή καταγράφεται στον BGP Table των ακραίων δρομολογητών (border gateways) ενός AS ανά prefix (υποδίκτυο προορισμού) και την σειρά των AS's της προτεινόμενης διαδρομής (μαζί με το βάρος της)
 - Οι πίνακες BGP φυλάσσονται στην ηλεκτρονική μνήμη των border gateways και ανανεώνονται δυναμικά όποτε υπάρχουν αλλαγές στο Internet με ευθύνη των γειτονικών δρομολογητών (border gateways) που ανακοινώνουν τα δίκτυα των αυτονόμων κοινοτήτων (AS's) που γνωρίζουν (*advertising*)
 - Ο υπολογισμός των «βέλτιστων» Inter-AS δρόμων γίνεται κατανεμημένα, βασισμένος σε αλγόριθμους *distance vector (Bellman Ford)* με κόστη τα βάρη των συνδέσεων μεταξύ border gateways. Οι δρόμοι Inter-AS μπορεί να αλλάξουν στη πορεία ενός πακέτου προς τον τελικό του προορισμό

ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Είδη Intra-AS Routing

- Άμεση δρομολόγηση (**direct routing**)
 - Κάθε κόμβος (PC, router) στέλνει πακέτα IP σε **interface κόμβου** του ίδιου υποδικτύου
- Έμμεση δρομολόγηση (**indirect routing**)
 - Ο τελικός κόμβος στέλνει πακέτα IP σε κόμβο του ίδιου δικτύου – υποδικτύου, χρησιμοποιώντας ενδιάμεσους δρομολογητές (routers). Για την προώθηση πακέτων σε Επίπεδο 2 (L2, Ethernet) μαθαίνει την αντιστοίχιση του **interface δρομολογητή (gateway) & την διεύθυνση L2 (MAC) μέσω ARP (Address Resolution Protocol)**
 - Ο δρομολογητής πρέπει να γνωρίζει τη διαδρομή (επόμενο interface δρομολογητή) προς το δίκτυο – υποδίκτυο προορισμού
 - Μεταξύ δρομολογήτων κορμού εντός AS υπάρχουν πίνακες δρομολόγησης (routing tables) προς όλα τα εσωτερικά υποδίκτυα (και τα άμεσα συνδεδεμένα σε γειτονικά AS) που ανανεώνονται με πρωτόκολλα **IGP** (συνήθως **OSPF** με core και stub areas)
 - Οι τελικοί κόμβοι στέλνουν πακέτα με διεύθυνση προορισμού εκτός του δικτύου – υποδικτύου τους σε **default gateway** (π.χ. 147.102.13.200)

ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Intra-AS Routing, Host Routing Table

- Εγγραφές του τύπου (N, R)

- N: Δίκτυο προορισμού
- R: Επόμενο interface δρομολογητή (gateway)

- Host routing table σε λειτουργικό Windows από το μηχάνημα με IP **147.102.13.32**

> netstat -nr

Routing Table:

Network	Destination	Netmask	Gateway	Interface	Metric
	0.0.0.0	0.0.0.0	147.102.13.200	147.102.13.32	20
	127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
	147.102.13.0	255.255.255.0	147.102.13.32	147.102.13.32	20
	147.102.13.32	255.255.255.255	127.0.0.1	127.0.0.1	20
	147.102.255.255	255.255.255.255	147.102.13.32	147.102.13.32	20
	224.0.0.0	240.0.0.0	147.102.13.32	147.102.13.32	20

- Προς το ίδιο τοπικό υποδίκτυο **147.102.13.0/24** σαν gateway ορίζεται κατευθείαν (**direct**) το τοπικό interface **147.102.13.32**
 - Προς destination `dolly.netmode.ntua.gr` (**147.102.13.10**) gateway θα είναι το τοπικό interface **147.102.13.32**
- Προς όλα τα άλλα δίκτυα **0.0.0.0** σαν gateway ορίζεται το **147.102.13.200** (**default gateway: router.netmode.ntua.gr**)
- Προς **local host 127.0.0.0/8** (π.χ. για δοκιμή δικτυακών εφαρμογών τοπικά) ή προς το ίδιο το **147.102.13.32/32** «προωθούνται» στο «interface» **127.0.0.1**
- Προς διεύθυνση broadcast **147.102.255.255/32** σαν gateway ορίζεται κατευθείαν (direct) το τοπικό interface **147.102.13.32** (η διεύθυνση δεν ισχύει στο δίκτυο του ΕΜΠ)
- Προς διευθύνσεις multicast **224.0.0.0/4** σαν gateway ορίζεται κατευθείαν (direct) το τοπικό interface **147.102.13.32**

ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Intra-AS Routing, Links between Routers

- 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**

ΠΑΡΑΔΕΙΓΜΑ ΕΣΩΤΕΡΙΚΗΣ ΔΡΟΜΟΛΟΓΗΣΗΣ: ΤΟ ΔΙΚΤΥΟ ΤΟΥ Ε.Μ.Π.

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 δρομολογητή του Ε.Μ.Π. **147.102.224.33** βρίσκεται τον δρομολογητή του ΕΚΠΑ **147.102.224.34** σαν μέλος του υποδικτύου:

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

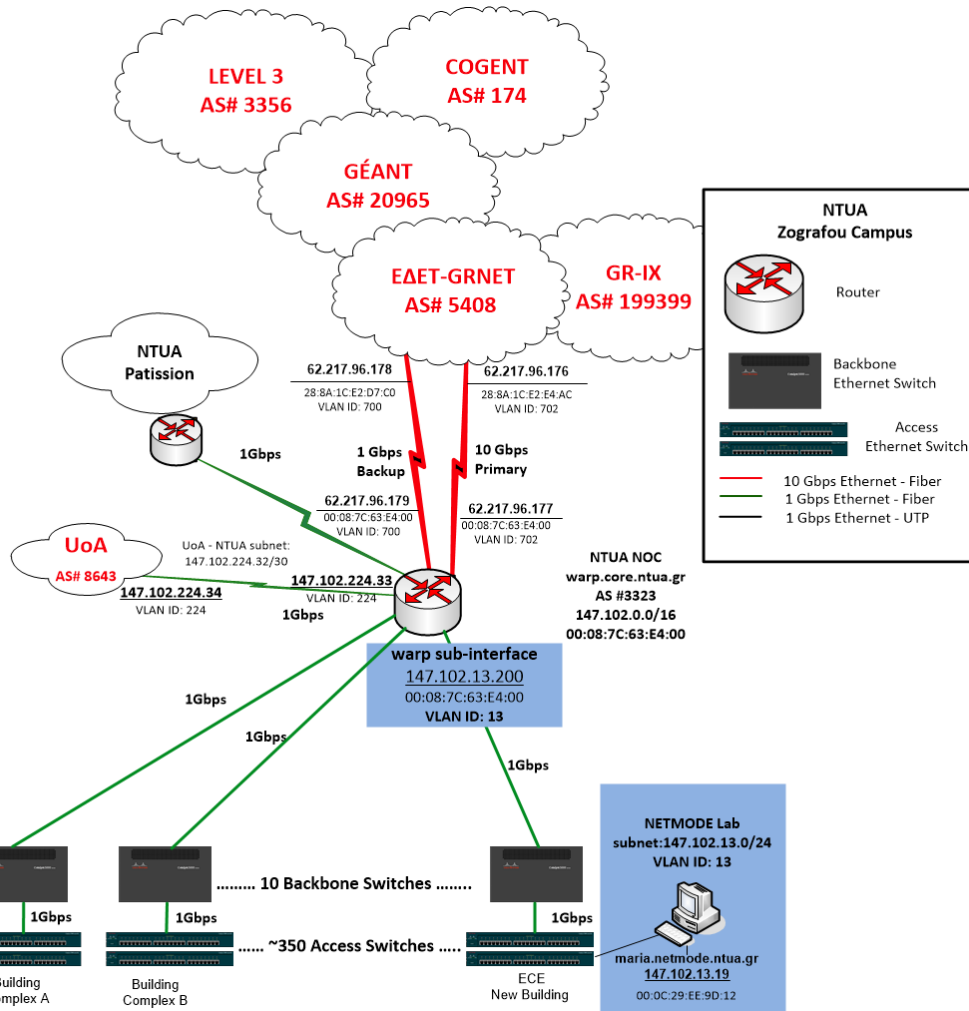
Η γραμμή Ε.Μ.Π. – ΕΚΠΑ (όπως όλες οι γραμμές σε Δίκτυα Internet) ορίζεται σαν υποδίκτυο (prefix) με 4 τουλάχιστον διευθύνσεις IP:

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

ΑΝΤΙ-ΠΑΡΑΔΕΙΓΜΑ:

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

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



ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Inter-AS Routing, Border Gateway Protocols - BGP

- Το BGP υλοποιείται μεταξύ συνοριακών δρομολογητών (border routers) Αυτονόμων Κοινοτήτων (AS) ώστε να δίνει σε κάθε AS τον τρόπο για να:
 - Μαθαίνει πληροφορίες πρόσβασης (subnet reachability information) από γειτονικά AS/δρομολογητές
 - Προωθεί reachability information σε όλους τους εσωτερικούς (AS-internal) δρομολογητές κορμού
 - Επιτρέπει σε υποδίκτυα με δημόσιες διευθύνσεις IP (**prefixes**) να ανακοινώσουν την ύπαρξη τους στο υπόλοιπο Internet: **“Είμαι εδώ”**
 - Καθορίζει “καλούς” routes προς όλα τα δημόσια δίκτυα (prefixes) με βάση **reachability information** και **policy**. Οι συνοριακοί δρομολογητές μπορεί να γνωρίζουν όλα τα 700.000 δίκτυα-προορισμούς του global Internet ή μέρος τους (αν δεν υπάρχουν εναλλακτικοί δρόμοι για κάποιους προορισμούς)
- Πέρα του BGP μεταξύ AS border routers (external BGP - **eBGP**) απαιτείται και διάχυση της πληροφορίας εντός AS (interior BGP – **iBGP**)

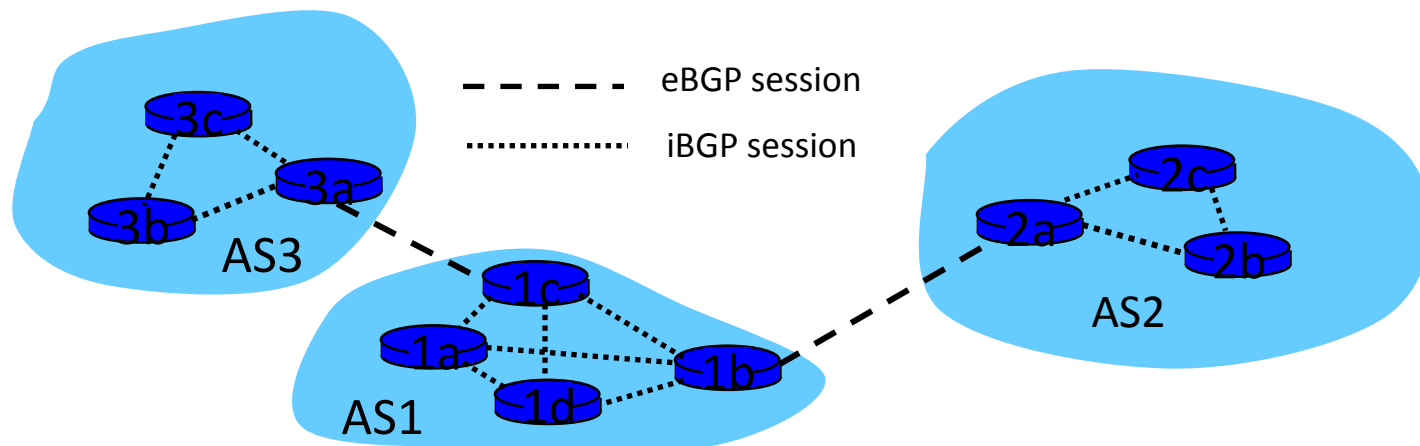
ΔΡΟΜΟΛΟΓΗΣΗ ΕΠΙΠΕΔΟΥ ΔΙΚΤΥΟΥ

Inter-AS Routing, Σηματοδοσία 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

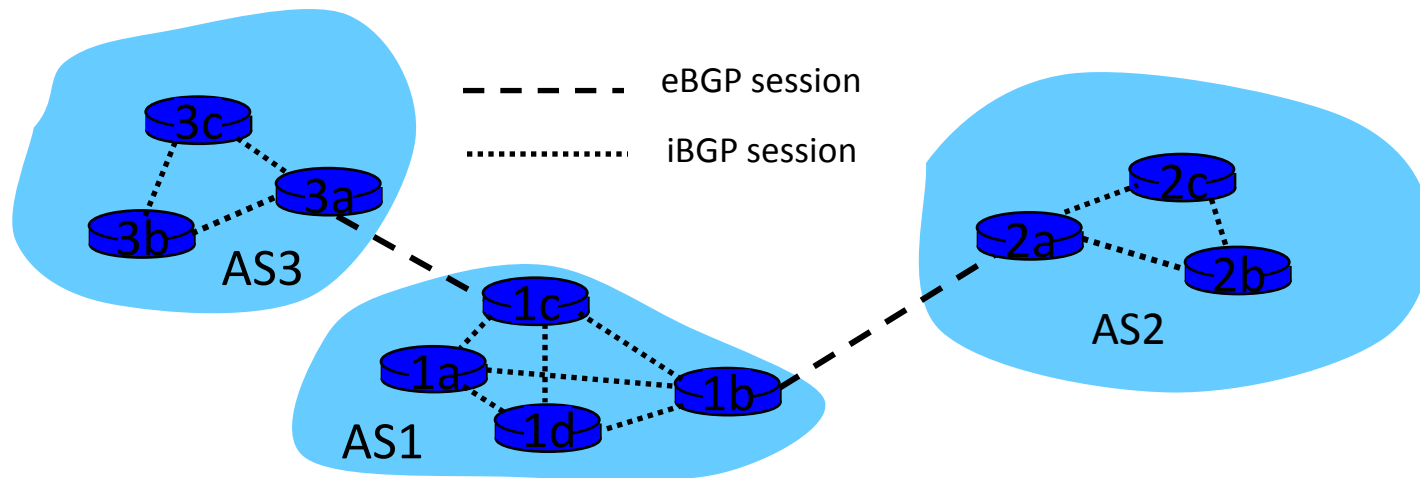
ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΤΟΥ BGP

- Τα ζεύγη από συντοριακούς 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



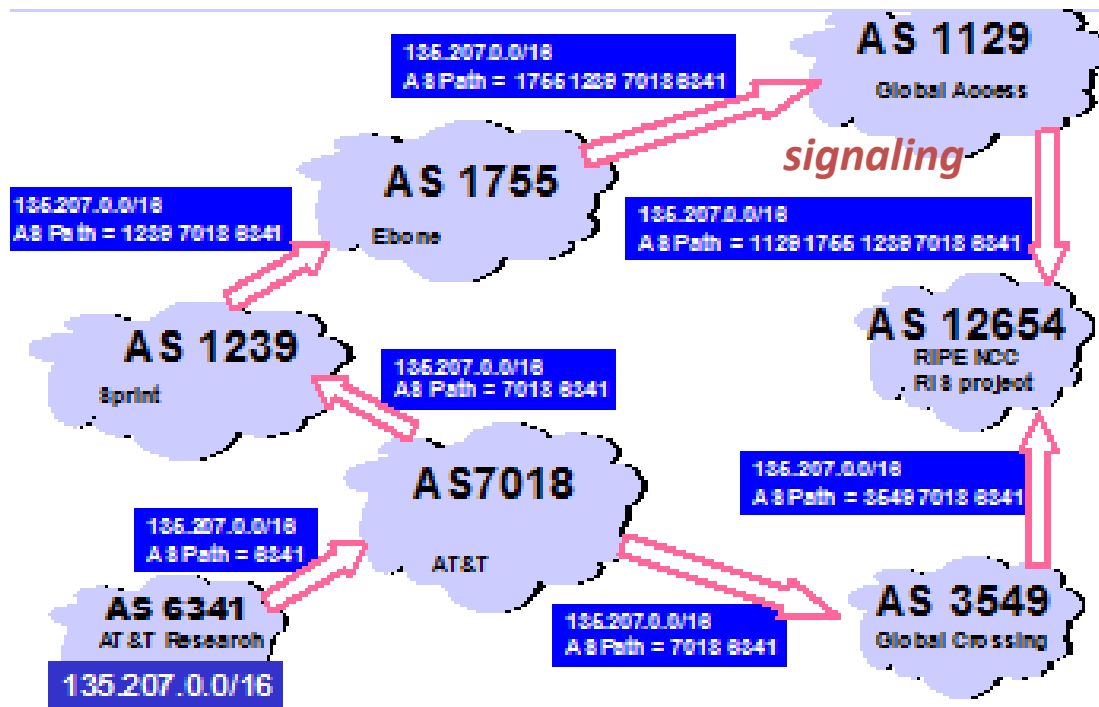
ΔΙΑΝΟΜΗ BGP REACHABILITY INFO

- Με χρήση σύνδεσης TCP, το πρωτόκολλο **eBGP** μεταξύ των border gateways 3a και 1c στέλνει *prefix reachability info* της AS3 στην AS1
 - 1c μπορεί να χρησιμοποιήσει **iBGP** 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



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

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

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