

ΔΙΑΧΕΙΡΙΣΗ ΔΙΚΤΥΩΝ - NETWORK MANAGEMENT

Αρχιτεκτονική & Δρομολόγηση στο Internet

Internet Architecture & Routing

Το Εμπορικό Παγκόσμιο Internet - The Commercial Internet

Tier 1, Tier 2 Internet Service Providers (ISP), Internet eXchange Points (IXP)

Ορισμοί Μεθόδων Δρομολόγησης - Routing Definitions

Δρομολόγηση Επιπέδου Δικτύου - Network Level Routing (IP)

Πρωτόκολλα Δρομολόγησης IP: Intra-AS & Inter-AS Routing Protocols

B. Μάγκλαρης

maglaris@netmode.ntua.gr

www.netmode.ntua.gr

Νέα Κτίρια ΣΗΜΜΥ - Αίθουσα 013

24/10/2022

ΔΙΚΤΥΑ ΜΕ ΓΝΩΣΤΑ IP & ΑΥΤΟΝΟΜΕΣ ΠΕΡΙΟΧΕΣ (επανάληψη)

Announced Public IP Networks, Autonomous Domains

Autonomous System Numbers - ASN, Border Gateway Protocol - BGP

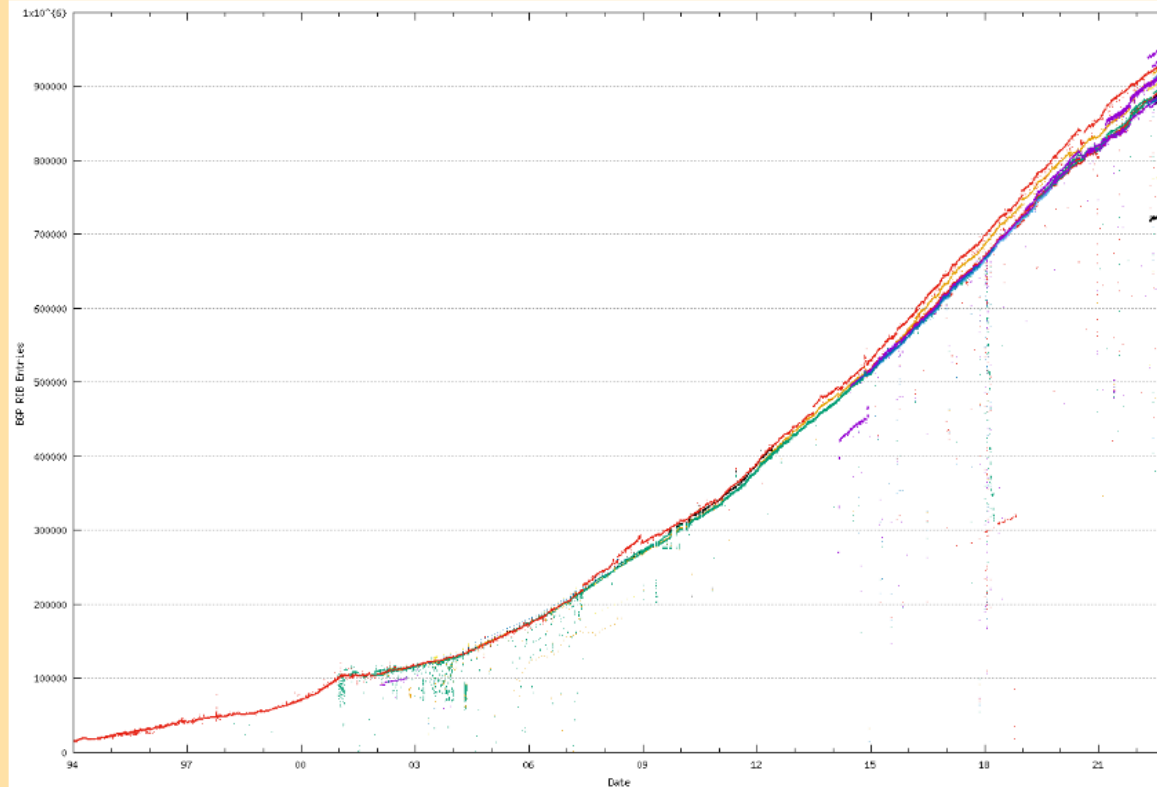
- Το Internet σήμερα (εκτίμηση **31/7/2022**)
 - Πάνω από **5.473.055.736** τελικοί χρήστες (συνδέσεις) σε συνολικό πληθυσμό **7.934.462.631**, διείσδυση **69%**
 - Γύρω στα **1.100.000** ανακοινώσιμα δίκτυα – γνωστοί προορισμοί (announced public IPv4 & IPv6 networks via **BGP announcements**)
 - Ιεραρχικά ταξινομημένα σε Αυτόνομες Διαχειριστικές Περιοχές **AS** (Autonomous Systems) με μοναδικό αριθμό **ASN** (Autonomous System Number, 16 bits ή 32 bits)
- Καταχώρηση IP & ASN σε blocks των 1024 AS's με διεθνή συντονισμό από **ICANN** (Internet Corporation for Assigned Names & Numbers) - **IANA** (Internet Assigned Number Authority) μέσω **RIR's** (Regional Internet Registries): **ARIN** (American Registry for Internet Numbers), **RIPE NCC** (Réseaux IP Européens Network Coordination Centre), **APNIC** (Asia Pacific Network Information Centre), **AFRINIC** (African Network Information Center), **LATNIC** (Latin America and Caribbean Network Information Centre)
- **12/10/2022**: Αριθμός ανακοινώσιμων (**advertised**) AS's μέσω BGP announcements: **73.477** και μη ανακοινώσιμων (**unadvertised**) **37.876** (σύνολο **111.353** από **123.901** καταχωρημένα ASN από τον IANA)

RIR	RIR Pool	Unadv	Adv	16-bit	Unadv	Adv	32-bit	Unadv	Adv
AFRINIC	1098	520	1708	376	218	684	722	302	1024
APNIC	2299	13891	13144	599	3082	4861	1700	10809	8283
ARIN	2087	12324	19187	1340	9575	14589	747	2749	4598
RIPE NCC	4814	9383	28343	3178	5070	17495	1636	4313	10848
LACNIC	2250	1758	11095	485	574	2367	1765	1184	8728
TOTAL	12548	37876	73477	5978	18519	39996	6570	19357	33481

BGP TABLES: ΑΡΙΘΜΟΣ ΓΝΩΣΤΩΝ (PUBLIC) ΔΙΚΤΥΩΝ - ΠΡΟΟΡΙΣΜΩΝ (επανάληψη)

<http://bgp.potaroo.net/>

Growth of the BGP Table - 1994 to Present



BGP Table Data

Report last updated at Wed, 12 Oct 2022 12:17:22 GMT

IPv4 BGP Reports

AS131072 APNIC R&D 935683
AS6447 Route-Views.Oregon-ix.net 960916

IPv4 Route-Views

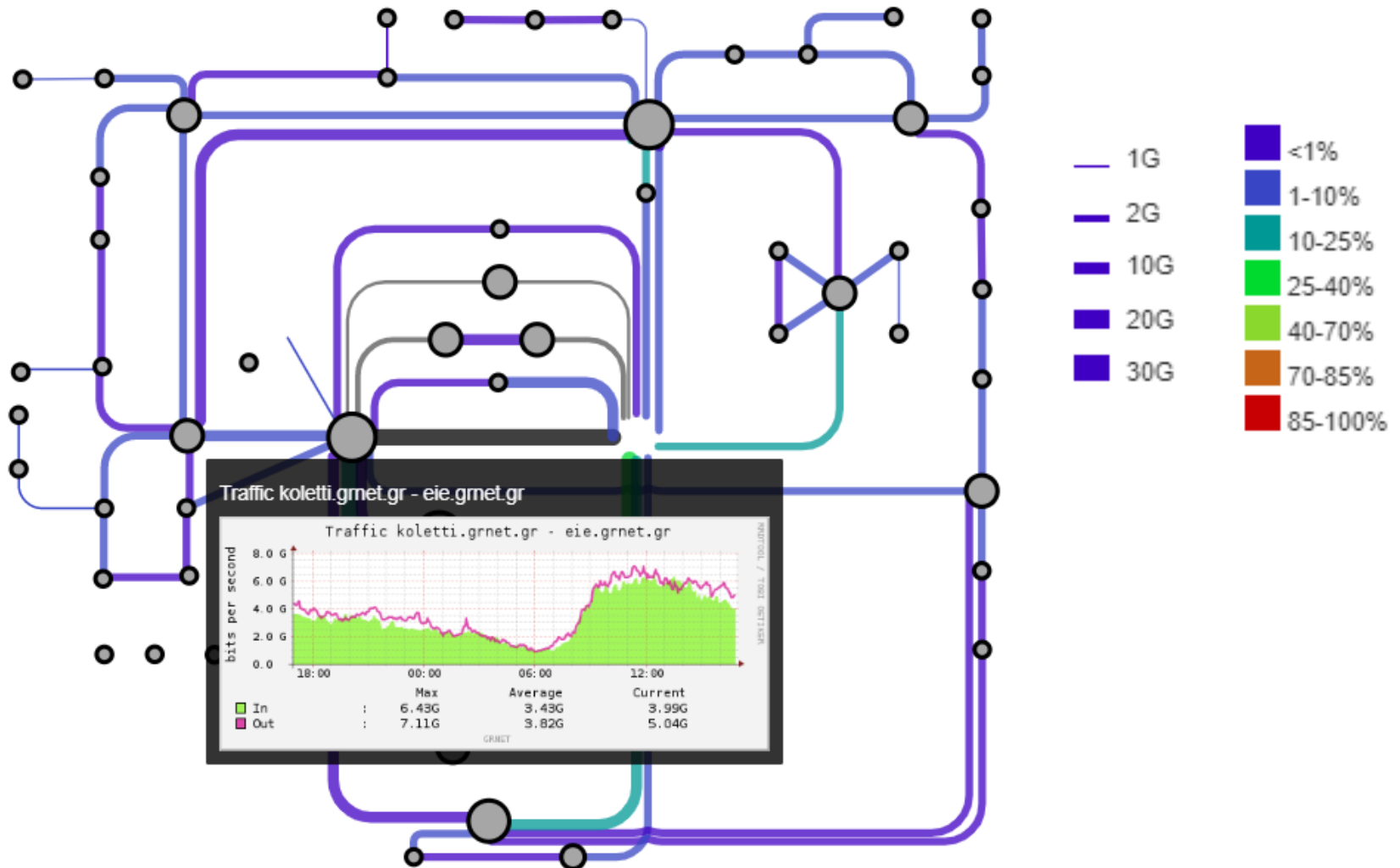
IPv6 BGP Reports

AS131072 APNIC R&D 164611
AS6447 Route-Views.Oregon-ix.net 171225

ΕΔΥΤΕ - GRNET: ΤΟ ΔΙΚΤΥΟ ΚΟΡΜΟΥ - IP Overlay (12/10/2022)

(επανάληψη)

<https://netmon.grnet.gr/networkmap/>



ΤΟ ΠΑΝΕΥΡΩΠΑΪΚΟ ΑΚΑΔΗΜΑΪΚΟ ΔΙΚΤΥΟ ΓÉANT - 2022 (επανάληψη)

ΓÉANT the pan-European
research and education network

**ΔΙΑΣΥΝΔΕΕΙ ΜΕ ΟΠΤΙΚΕΣ
ΣΥΝΔΕΣΕΙΣ ΠΟΛΛΑΠΛΩΝ
10-100 Gbps:**

42 Εθνικά Δίκτυα Έρευνας &
Εκπαίδευσης (National Research &
Education Networks – NRENs)

ΤΕΛΙΚΟΙ ΧΡΗΣΤΕΣ:

10.000 ++ Ιδρύματα

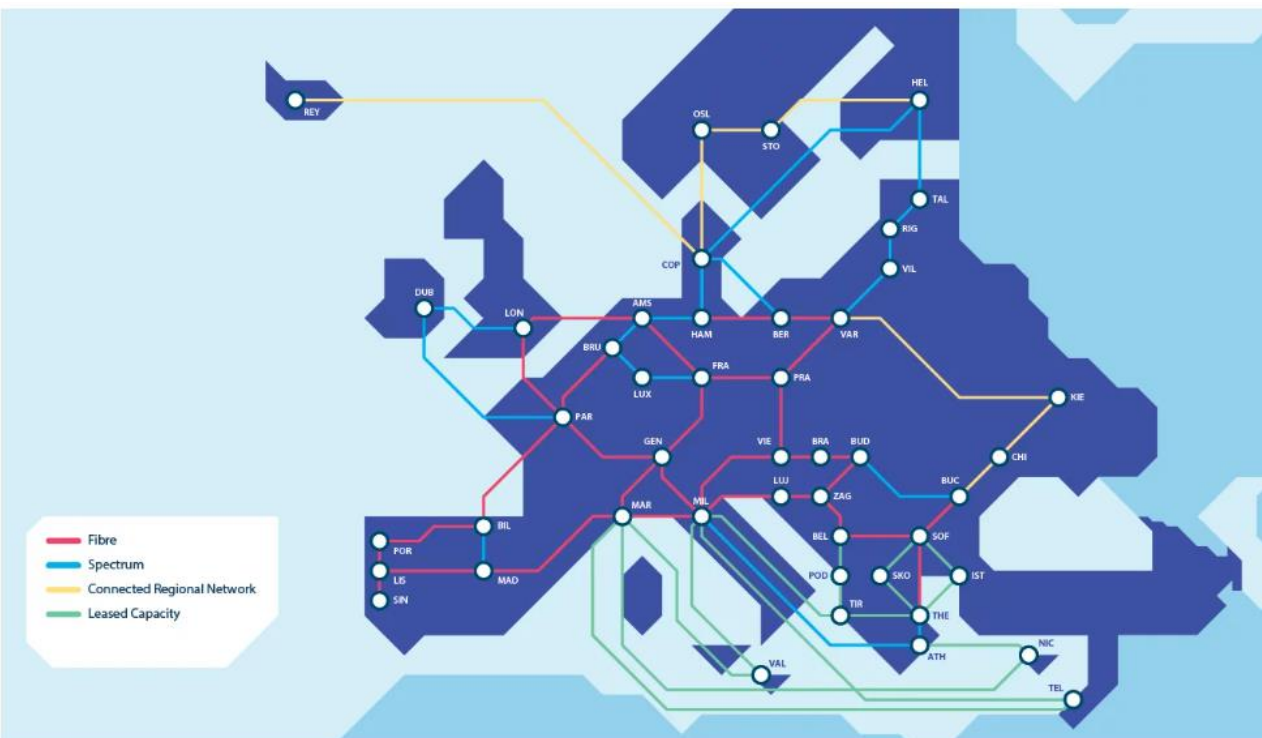
Ερευνητικές υποδομές παγκόσμιας
εμβέλειας (CERN/HEP, ITER, ESFRI...)

50 εκ. ++ φοιτητές, μαθητές,
εκπαιδευτικό προσωπικό, ερευνητές

ΔΙΑΧΕΙΡΙΣΗ:

DANTE + TERENA → ΓÉANT
Association

<http://www.geant.org/>





CAIDA'S IPV4 AND IPV6 AS CORE: VISUALIZING IPV4 AND IPV6 INTERNET TOPOLOGY AT A MACROSCOPIC SCALE IN 2020

SINCE 2000, CAIDA HAS GENERATED AS CORE GRAPHS – INTERNET TOPOLOGY MAPS ALSO REFERRED TO AS AS-LEVEL INTERNET GRAPHS – IN ORDER TO VISUALIZE THE SHIFTING TOPOLOGY OF THE INTERNET OVER TIME, AS CAN BE SEEN IN THE HISTORICAL VIEW. FOR A DETAILED EXPLANATION OF THE METHODOLOGY USED IN GENERATING THE AS CORE VISUALIZATIONS, SEE THE MAIN IPV4 & IPV6 AS CORE PAGE.

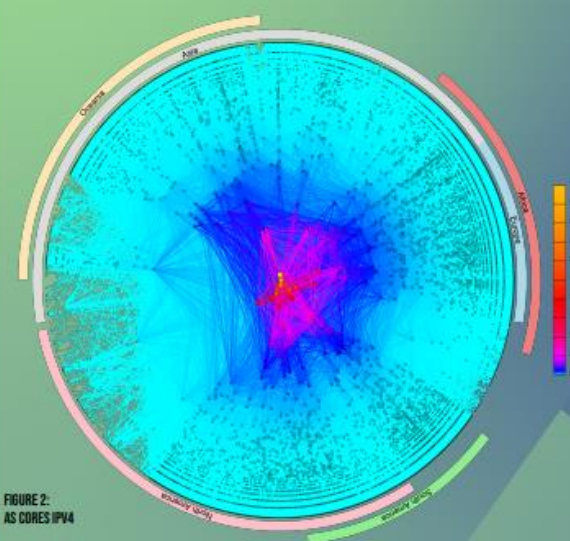


FIGURE 2:
AS CORES IPV4

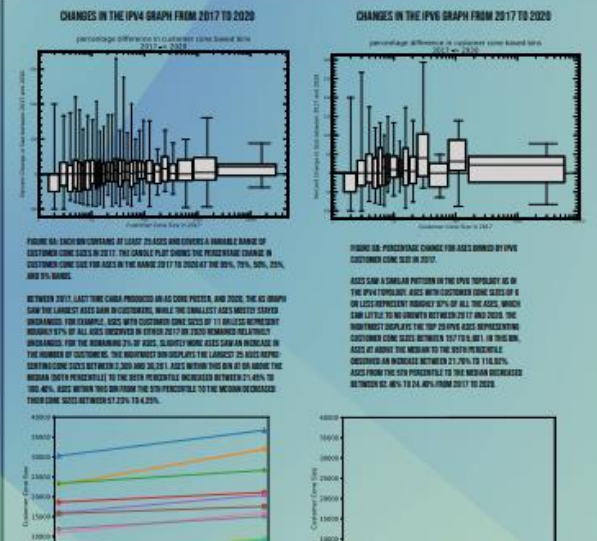
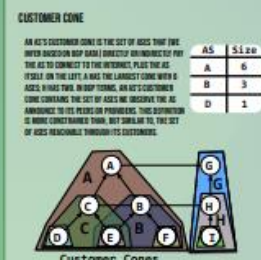


FIGURE 3:
AS CORES IPV6



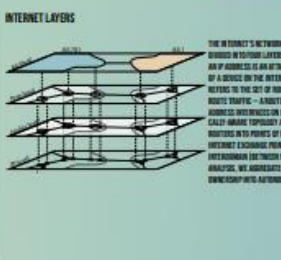
CAIDA TEAM
RONALD WALSH, DANIEL HOFFMANN, KE CLAFFY

SPONSORING INSTITUTIONS
NSF, DARPA, AND OTHERS

ADVISORY BOARD
DAVID BRIN, JEFFREY HEINRICH, AND OTHERS

CONTRIBUTORS
ANDREW CHEN, AND OTHERS

CONTACT
CAIDA@CAIDA.ORG



CAIDA TEAM
RONALD WALSH, DANIEL HOFFMANN, KE CLAFFY

SPONSORING INSTITUTIONS
NSF, DARPA, AND OTHERS

ADVISORY BOARD
DAVID BRIN, JEFFREY HEINRICH, AND OTHERS

CONTRIBUTORS
ANDREW CHEN, AND OTHERS

CONTACT
CAIDA@CAIDA.ORG



CAIDA TEAM
RONALD WALSH, DANIEL HOFFMANN, KE CLAFFY

SPONSORING INSTITUTIONS
NSF, DARPA, AND OTHERS

ADVISORY BOARD
DAVID BRIN, JEFFREY HEINRICH, AND OTHERS

CONTRIBUTORS
ANDREW CHEN, AND OTHERS

CONTACT
CAIDA@CAIDA.ORG



CAIDA TEAM
RONALD WALSH, DANIEL HOFFMANN, KE CLAFFY

SPONSORING INSTITUTIONS
NSF, DARPA, AND OTHERS

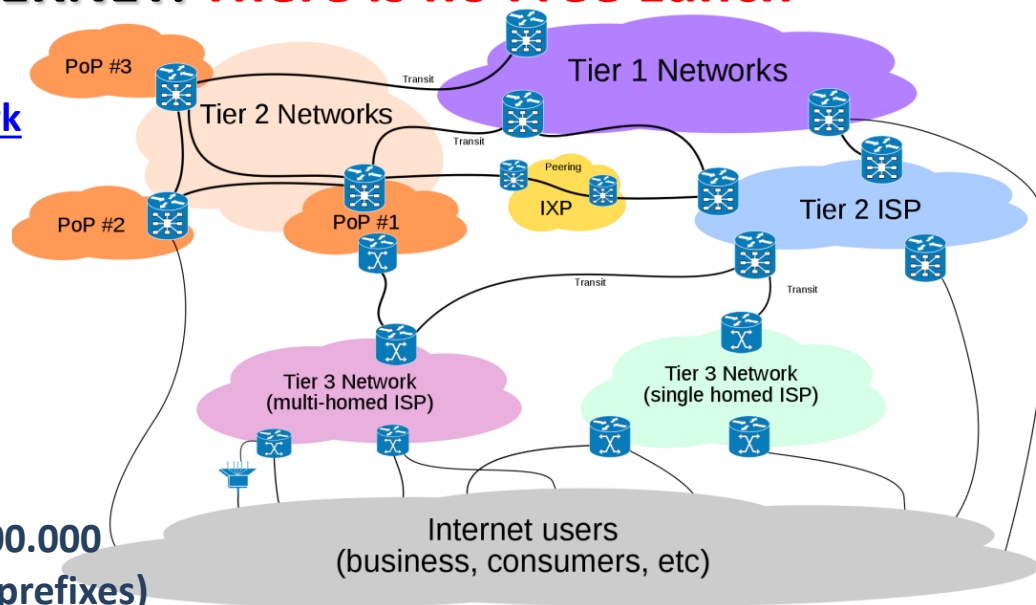
ADVISORY BOARD
DAVID BRIN, JEFFREY HEINRICH, AND OTHERS

CONTRIBUTORS
ANDREW CHEN, AND OTHERS

CONTACT
CAIDA@CAIDA.ORG

ΤΟ «ΕΜΠΟΡΙΚΟ» INTERNET: There is no Free Lunch

http://en.wikipedia.org/wiki/Tier_1_network



Οι 15 **Tier 1** ISP's με πρόσβαση και στα 1.100.000 δίκτυα - γνωστούς προορισμούς (IPv4 & IPv6 prefixes)

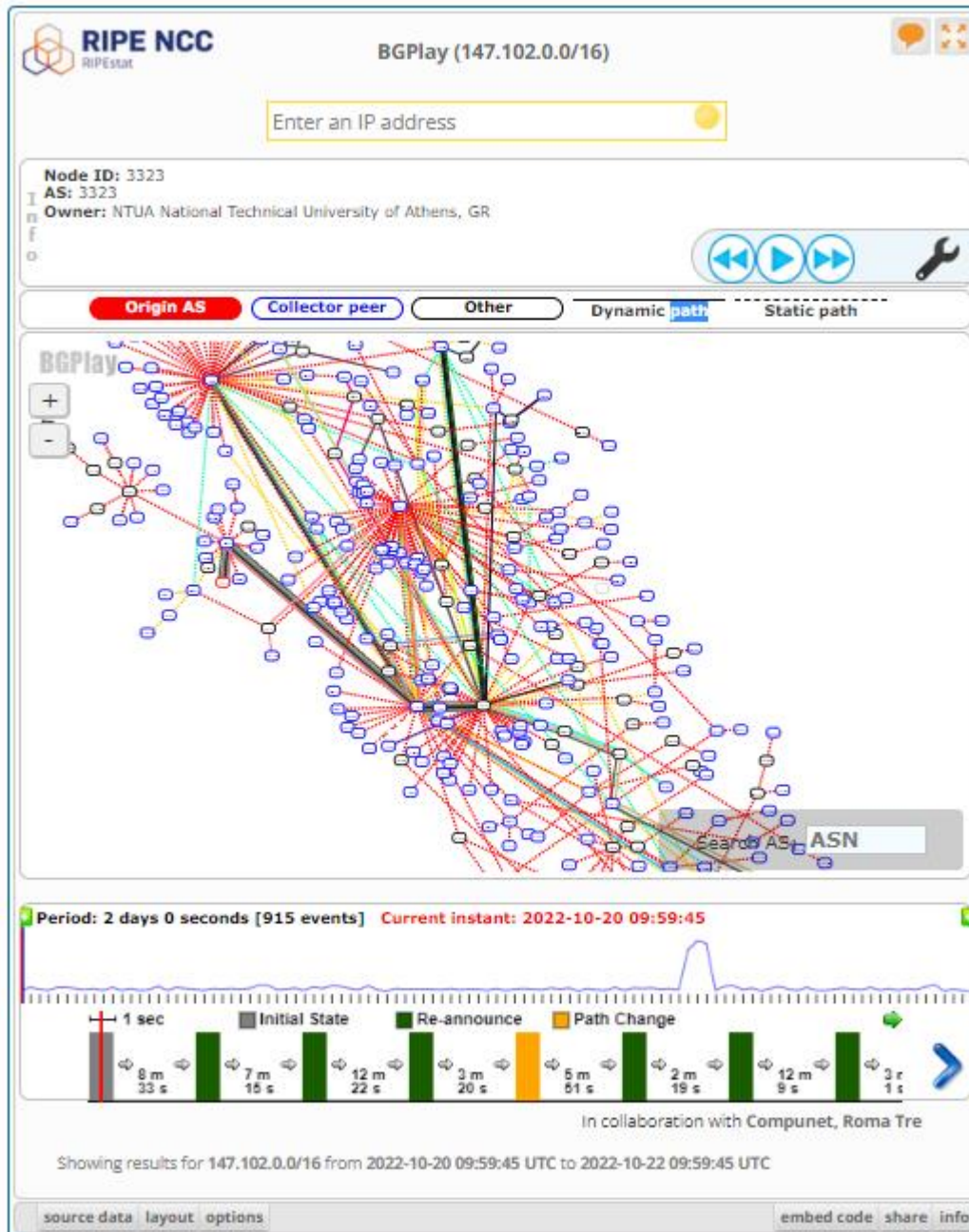
Name	Headquarters	AS number
AT&T ^[11]	United States	7018
Deutsche Telekom Global Carrier ^[13]	Germany	3320
GTT Communications	United States	3257
Liberty Global ^{[17][18]}	United Kingdom ^[19]	6830
Lumen Technologies (formerly CenturyLink, formerly Level 3) ^{[21][22][23]}	United States	3356
NTT Communications (formerly Verio) ^[26]	Japan	2914
Orange ^[27]	France	5511
PCCW Global	Hong Kong	3491
T-Mobile US (formerly Sprint) ^[29]	United States	1239
Tata Communications (formerly Teleglobe) ^[31]	India	6453
Telecom Italia Sparkle (Seabone) ^[33]	Italy	6762
Arelion (formerly Telia Carrier) ^[34]	Sweden	1299
Telxius (Subsidiary of Telefónica) ^[36]	Spain	12956
Verizon Enterprise Solutions (formerly UUNET) ^[42]	United States	701
Zayo Group (formerly AboveNet) ^[44]	United States	6461

Tier 2 ISP's με ελαφρά μειωμένη πρόσβαση (συχνά θεωρούμενοι **Tier 1**)

Name	Headquarters	AS Number
China Telecom	China	4134 4809
Singtel ^[48]	Singapore	7473
Cogent Communications (formerly PSINet) ^[49]	United States	174
Comcast ^[52]	United States	7922
Hurricane Electric ^[53]	United States	6939
RETN ^[56]	United Kingdom	9002
Vodafone Carrier Services (formerly Cable & Wireless) ^[57]	United Kingdom	1273
Verizon Enterprise Solutions (formerly XO Communications) ^{[59][60]}	United States	2828
Telstra ^[61]	Australia	4637

ΠΑΡΟΧΗ INTERNET ΣΤΟ Ε.Μ.Π. (NTUA - ASN 3323)

<https://stat.ripe.net/special/bgplay> (22/10/2022)



NTUA (3323)

GRNET (5408)

GÉANT (21320)

GÉANT Internet Feeds

- **LEVEL3 (3356)**
- **COGENT 174 (174)**
- **HURRICANE US (6939)**
- **NORDUnet (2603)**

TO ATHENS INTERNET EXCHANGE (AIX)

Πρωτοβουλία του Ε.Μ.Π. & του GRNET, 1996

ΜΕΛΗ: **Tier 2** ISP's της Ελλάδας & ΕΔΕΤ για Εθνικό Peering

- GRNET (ΕΔΕΤ)
- Forthnet
- Hellas On Line
- Altec Telecoms
- NetOne
- Vivodi
- Verizon Hellas
- ON Telecoms
- OTENET
- AT&T Global Network Services Hellas
- ORANGE BUSINESS SERVICES
- Vodafone NET
- WIND
- Tellas
- Lannet

Το BGP δεν ανακοίνωνε Εθνικούς Προορισμούς συνδρομητών εκτός Ελλάδος μέσω AIX, μόνο από *Tier 1 – Tier 2 feeds* των παρόχων τους (πιθανή συνεργασία μόνο σε καταστάσεις εκτάκτου ανάγκης)

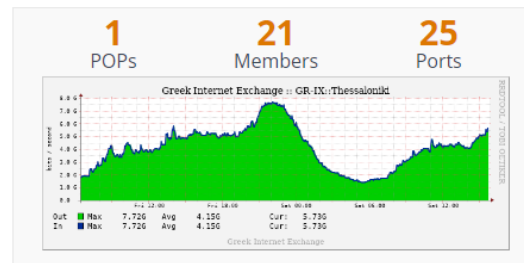
<https://www.gr-ix.gr/members/>

<https://www.gr-ix.gr/>

3 σημεία στέγασης (Points-of-Presence, PoP's)

1 σημείο στέγασης (Point-of-Presence, PoP)

GR-IX::THESSALONIKI



Greek Internet Exchange :: Greek Internet Exchange

bits / second

Nov 2021 Jan 2022 Mar 2022 May 2022 Jul 2022 Sep 2022

Out	Peak	Avg
In	Peak	Avg

Cur: 282.32G

Cur: 282.32G

Greek Internet Exchange

<https://portal.gr-ix.gr/statistics/ixp>

GR-DC-Athens					GR-DC-Thessaloniki				
Member	AG	ATH1	ATH2	ATH3	Member	AG	THESS1		
Arctura	1909		✓		CloudEon	1009	✓		
Artemis	4045		✓		Cloudops	1968			
AthenaExchange Group	1008		✓	✓	Devotec Solutions	20060			
CI	20740		✓		National Infrastructure for Research and Technology (NIKIT)	949			
City of Athens	20068		✓		HEX	6796	✓		
CloudEon	1009		✓		Inter Telecom	6870			
Connecticut	10700			✓	INICT	0087			
CNET	006			✓	Microsoft	200479	✓		
CyberGigawatt	686			✓	NETS	100	✓		
Diagrams	1968		✓		OSI	6796	✓		
Edgex Networks	4707				SECCompany	6890			
EIS	0000	✓			Subgroup	0000	✓		
Euro Telecom	27770		✓		Synapsion	0000	✓		
Eurolink	0607			✓	Telteknet	6796	✓		
FSN	20070			✓	TeleNet	19000			
Fullercom	0000			✓	Unidat	20000	✓		
Ghana	0000			✓	Vision	0000			
Google Ireland	20060			✓	VoilaNet	0000			
Greek Internet Exchange (GIG)	1909				WNET	2600	✓		
National Infrastructure for Research and Technology (NIKIT)	949		✓	✓	DND services (per user name server)				
HEX	6796			✓	AT12 (per user DND lookup for IP: 1918 address)	10	✓		
Hellas Telecommunications and Post Corporation	20000		✓		ICP (ip)	0000	✓		
HotMail	0000		✓						
HyperionElastic	0000		✓						
India	20070			✓					
Information Society IS	0000		✓	✓					
Interconnect	20000			✓					
Inter Telecom	6870			✓					
Internet Hellas	20000			✓					
KappaCloud	10070			✓					
Latam	19000			✓					
LatviaHellas	0000		✓	✓					
Microsoft	200079			✓					
Microsoft	19000			✓					
Microsoft	0000			✓					
Microsoft	20000		✓	✓					
Nets	0000			✓					
NETS	100		✓	✓					
OneGreece Global Communications	0000			✓					
OSI	6796			✓					
Packet Clearing House/Route Collector	0000		✓						
PC200 Global	0000			✓					
Proton	20070			✓					
Real Time Data	20000		✓						
Red Beans Ltd	0000			✓					
SE Global	6890			✓					
Servid	20000			✓					
Sello Connect	4000			✓					
Subgroup	0000			✓					
Synapsion	27000			✓					
Synapsion	0000			✓					
TS Quantum Service	19000			✓					
TopHost	19000			✓					
Unidat	000		✓						
Unidat	0000			✓					
WNET	2600		✓	✓					
Xelate	20070			✓					
DND services (per user name server)									
AT12 (per user DND lookup for IP: 1918 address)					10	✓			
ICP (ip)					0000	✓			
Packet Clearing House (Real Servers and about 4000 IPv6 DNs)					00	✓			
RIP (v6 only)					2070	✓			
Message Internet Resolution Service					2600	✓			

ΤΑΞΙΝΟΜΗΣΗ ΑΛΓΟΡΙΘΜΩΝ ΔΡΟΜΟΛΟΓΗΣΗΣ

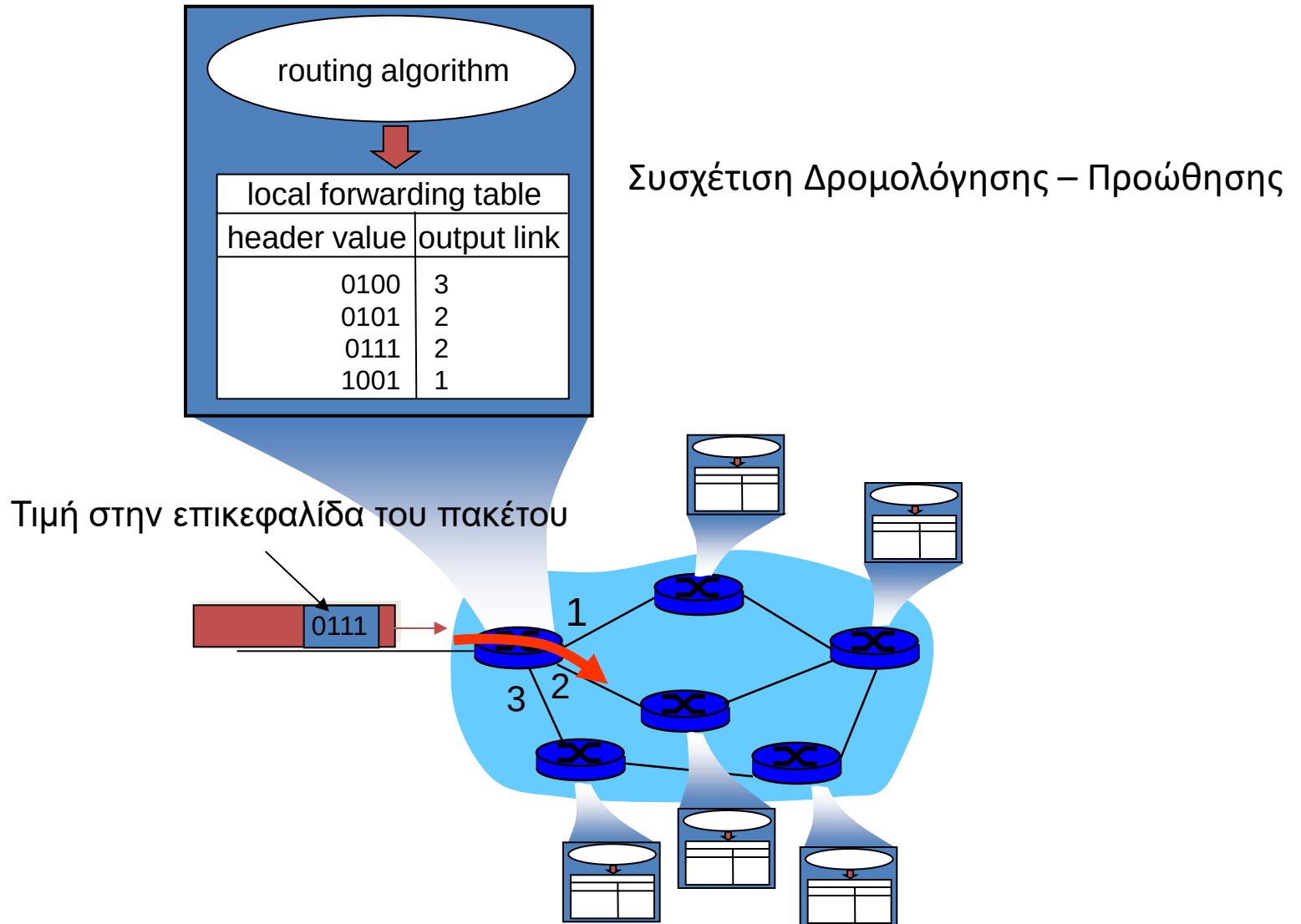
- **Διαχειριστικός Έλεγχος**
 - Συγκεντρωτικός: Δημόσια Δίκτυα Μεταγωγής Πακέτου **X.25/X.75**
 - Κατανεμημένος: **Internet**
 - Interior Gateway Protocols - **IGP**
 - Exterior (Border) Gateway Protocols - **EGP/BGP**
- **Δρομολόγηση με ή χωρίς Εναλλακτικούς Δρόμους**
 - Single choice:
 - Δρομολόγηση Επιπέδου 3 στο **Internet**: **Host Routing**, **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:
 - **Host 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**)

ΔΡΟΜΟΛΟΓΗΣΗ ΣΤΟ INTERNET

- **Πρωώθηση (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

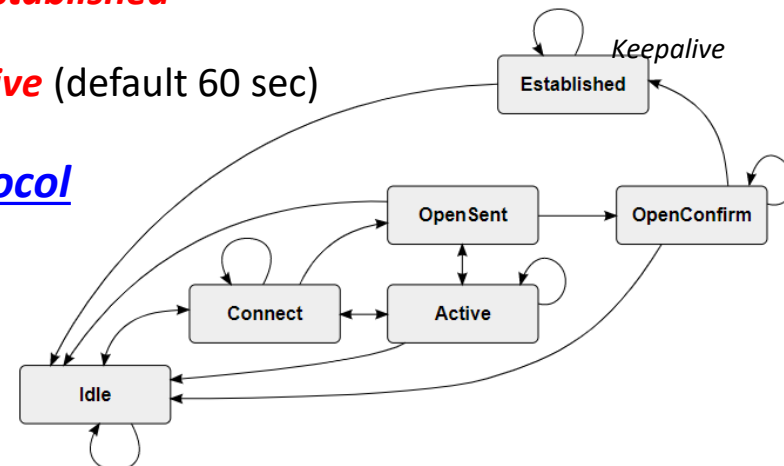
- **Intra-AS Routing, Interior Gateway Protocols (IGP):** Μια έξοδος προς επόμενο Interface για κάθε υποδίκτυο (subnet) τελικό προορισμό εντός ενός AS
 - RIP (Routing Internet Protocol): Πρωτόκολλο παλαιάς γενιάς, βασισμένο σε αλγόριθμους **distance vector (Bellman Ford)**
 - OSPF (Open Shortest Path 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

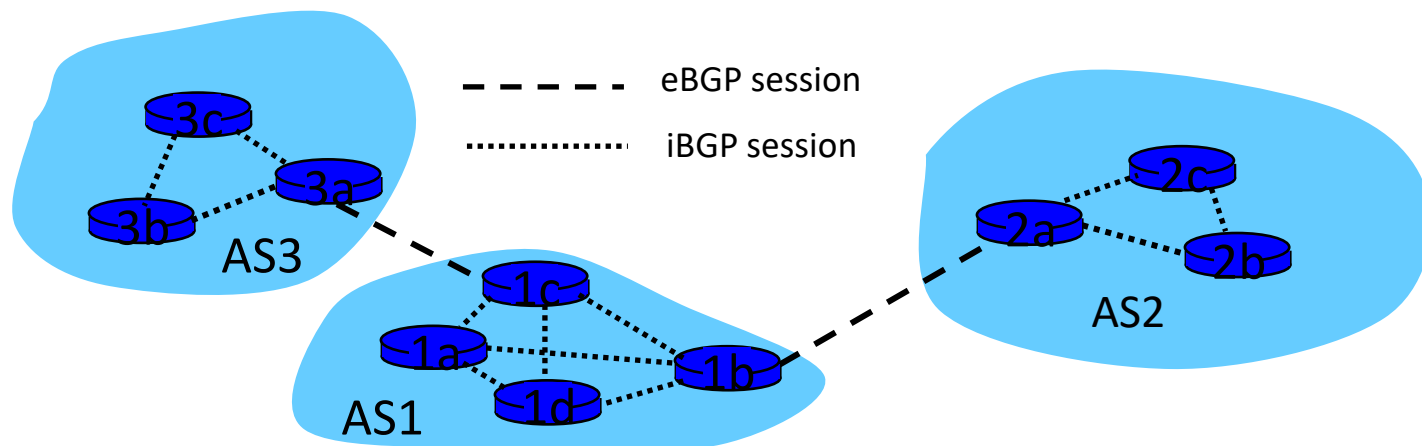
- **Inter-AS Routing, Exterior (Border) Gateway Protocols (EGP/BGP):** Πολλές εναλλακτικές διαδρομές με βάρη προς όλα τα γνωστά δίκτυα (περίπου **1.100.000** σήμερα) μεταξύ ακραίων (border) routers αυτονόμων συστημάτων (Autonomous Systems – **AS's**, περίπου **111.353** σήμερα - **73.477** ανακοινώσιμα)
 - 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
 - BGP (RFC 4271): Control plane packets (signals) μεταξύ BGP peers σε TCP port 149
 - Finite State Machine:
Idle, Connect, Active, OpenSent, OpenConfirm, Established
 - BGP packet types:
Open, Update, Notification (error signal), **Keepalive** (default 60 sec)

https://en.wikipedia.org/wiki/Border_Gateway_Protocol



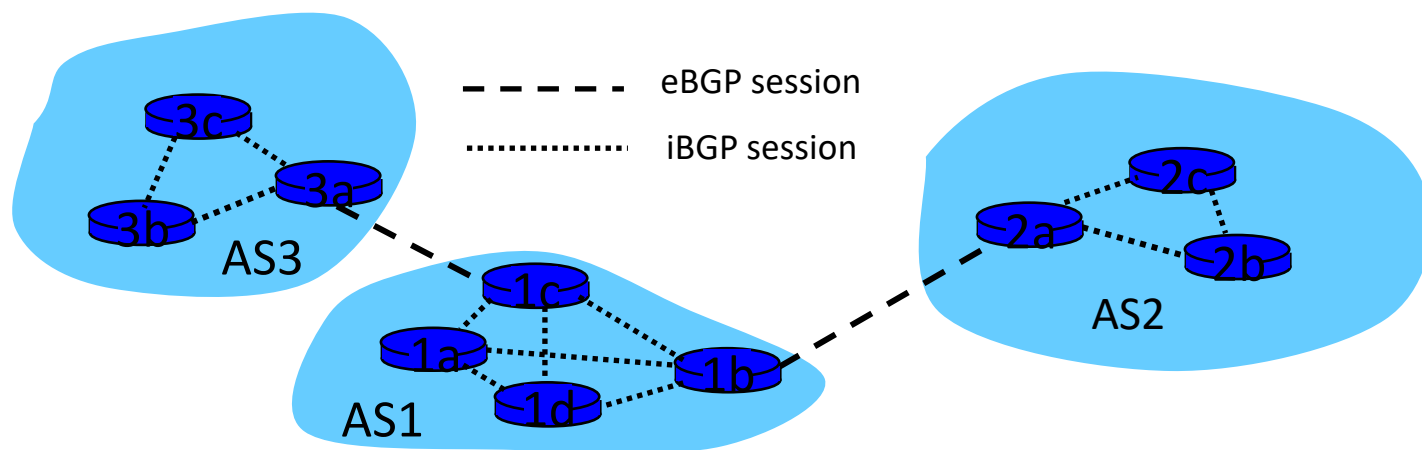
ΒΑΣΙΚΗ ΛΕΙΤΟΥΡΓΙΑ ΤΟΥ BGP

- 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

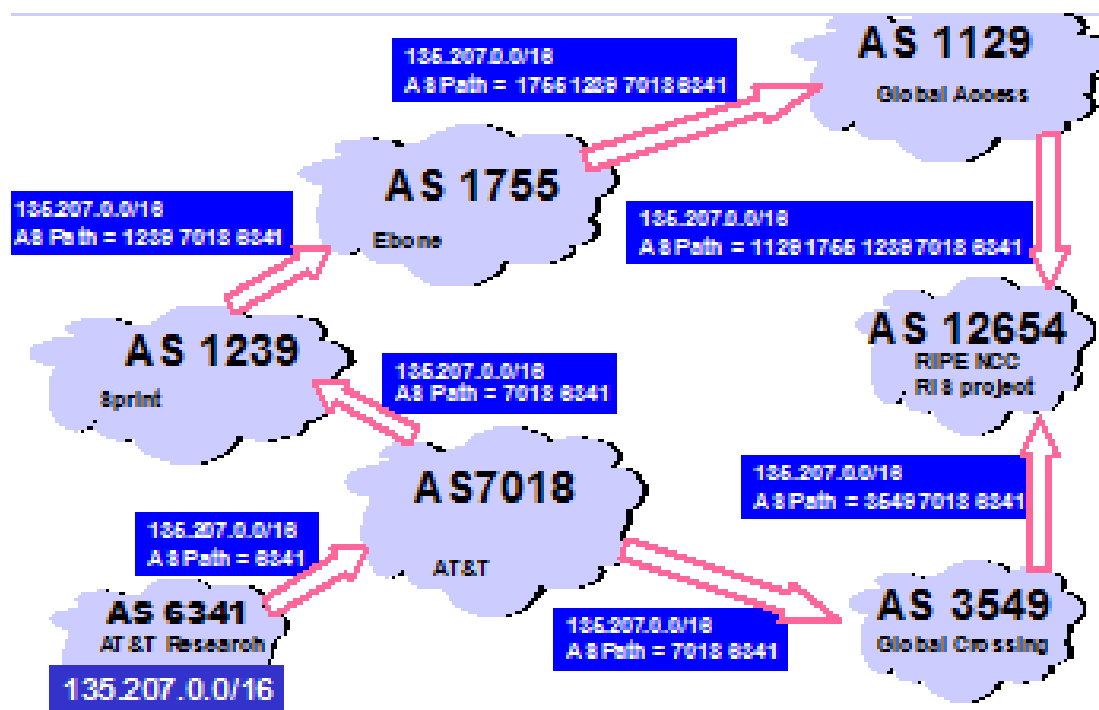


ΔΙΑΝΟΜΗ BGP REACHABILITY INFO

- 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**)

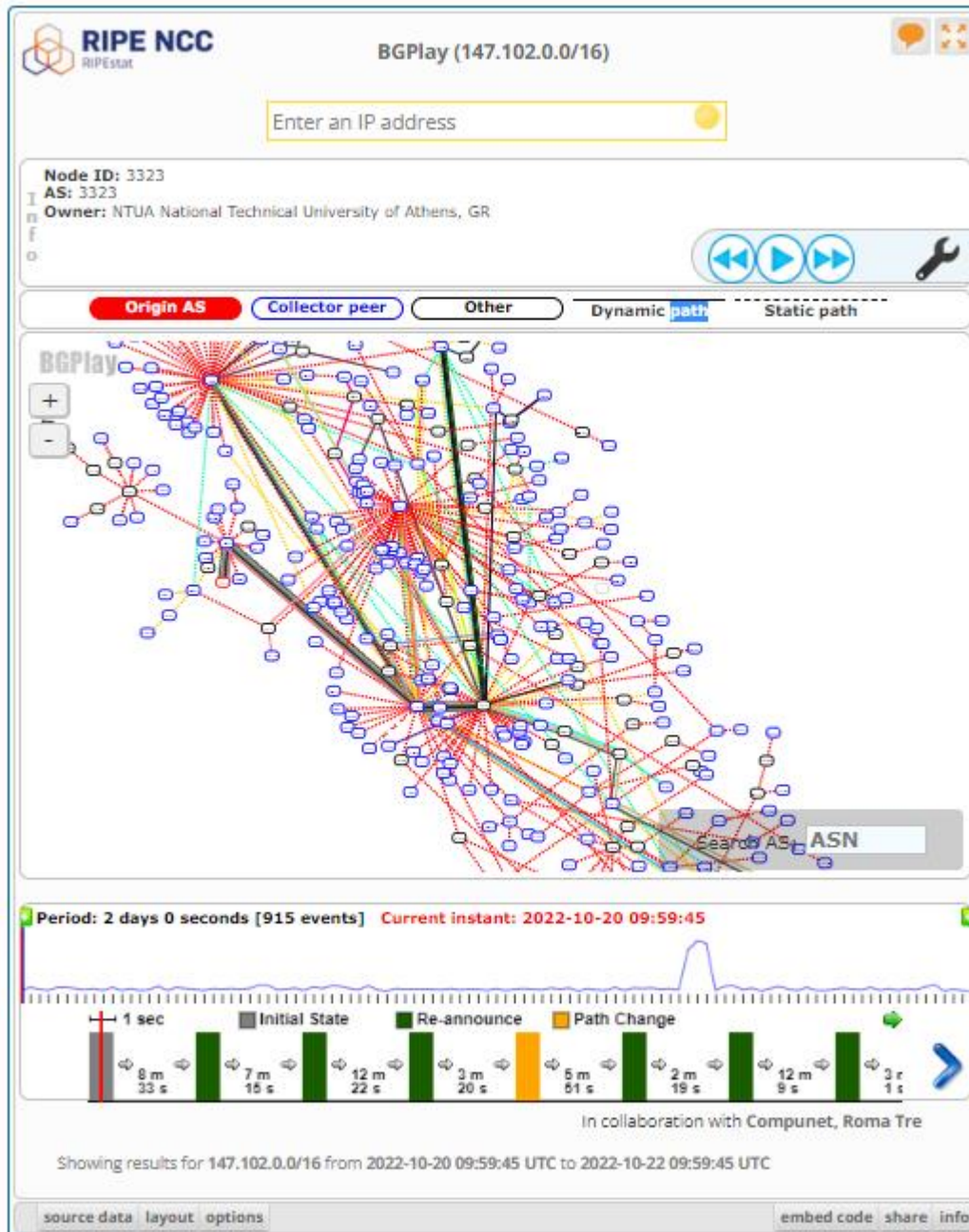


ΠΑΡΑΔΕΙΓΜΑ ΑΝΑΚΟΙΝΩΣΗΣ
ΔΙΚΤΥΟΥ 135.207.0.0/16 ΜΕΣΩ BGP (επανάληψη)
(από παρουσίαση του Timothy G. Griffin, AT&T Research, Paris 2002)



ΠΑΡΟΧΗ INTERNET ΣΤΟ Ε.Μ.Π. (NTUA - ASN 3323)

<https://stat.ripe.net/special/bgplay> (22/10/2022)



NTUA (3323)

GRNET (5408)

GÉANT (21320)

GÉANT Internet Feeds

- **LEVEL3 (3356)**
- **COGENT 174 (174)**
- **HURRICANE US (6939)**
- **NORDUnet (2603)**