PI-Grau
(Internet Protocols)

José M. Barceló Ordinas
Departamento de Arquitectura de Computadores
(UPC)

- **Topic 1: Internet Architecture&Addressing.**
  - Objectives
    - Understand the **general architecture** of Internet
    - Identify the **main actors** in the Internet architecture
    - Identify the **main organizations** implied in Internet
    - Understand **IPv4 address depletion problem**
    - Is it **IPv6** the solution to IPv4 depletion problem ?

Topic 1: Internet Architecture & Addressing.

Information Technologies (IT): branch of engineering that deals with the use of devices (computers, smart-phones, sensors, …) and communications (Internet) to store, retrieve, transmit and manipulate data.

That means that an **IT engineer** has a good technical and practical knowledge to manage the technologies dealing with the information and development of systems.

An engineer in IT is specialized in integrating information technologies more than in the information itself.

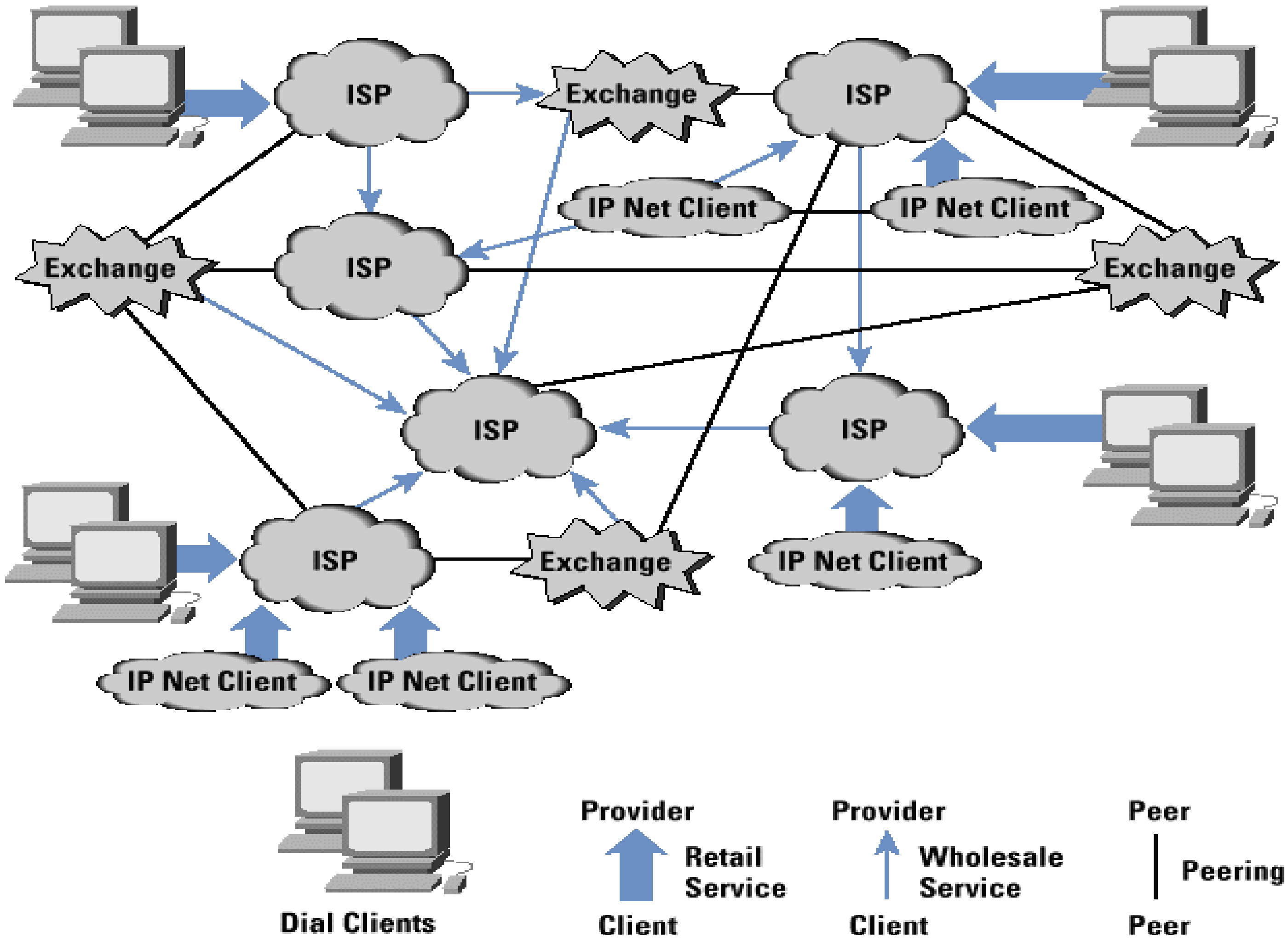
- Hardware,
- Operating Systems,
- Programming,
- Systems and networks,
- Distributed Systems,
- Security,
- System Architectures.

- **Internet:**
  - **Definition**: global system of interconnected computer networks that use the standard TCP/IP suite protocol to connect users, organizations and applications.
  - Users are inter-connected through ISP (Internet Service Provider).
- **Internet Service Provider (ISP)**
  - Provides connectivity and services to **end users** (Dial-up), **corporative networks** (IP-Net Clients) and **other ISPs.**

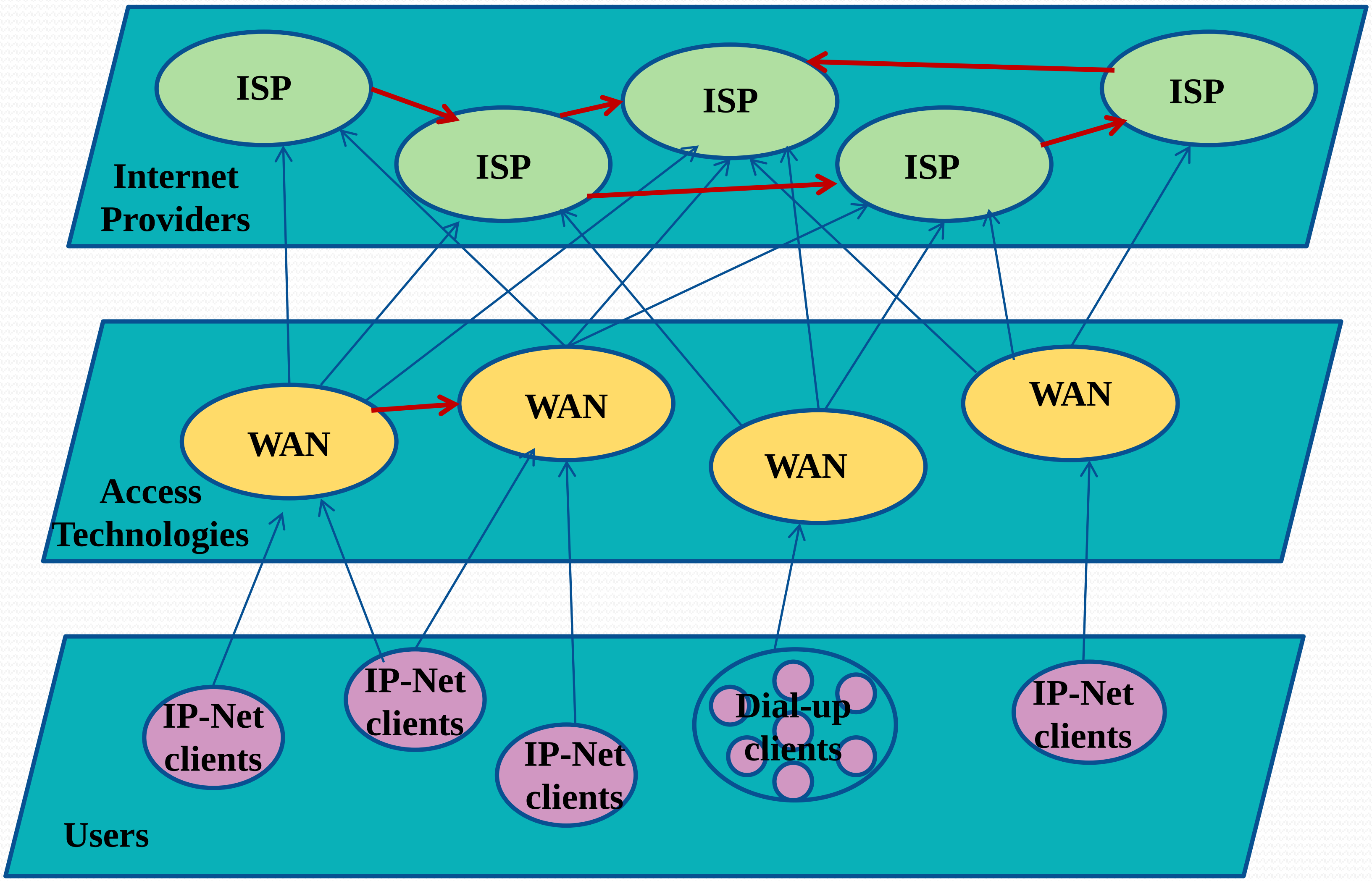
Which entities are connected on the Internet ?

- **End users** (Dial-clients):
  - Contract an access line (i.e., modem, ADLS, …). The ISP may be a telecom operator or has sub-contracts with telecom operators to offer service connectivity
- **Corporative Networks** (IP-Net-client)
  - Local Area Networks (LANs)
  - Contract switched/dedicated lines for linking their sites forming a VPN
  - Contract one or several lines (backup, redundancy, load balancing, …) to the ISP
- **Machines without human interaction**
  - Sensors (Internet of Things), robots, surveillance cameras, industrial machines, etc
- **ISP's**
  - Peering between ISPs: ISPs act with con client-to-provider, peer-to-peer and provider-to-client relationships
  - ISPs connections may be **private** (using Telecom operator lines) or **public** ("***Exchange points***")
  - **Retail services** (directly to customers) versus **Wholesail services** (in large quantities that later can be re-selled)

- Internet Architecture



- Internet Architecture



Internet Service Provider

- ISP offer different types of services
  - **Dedicated Internet Access services**
    - T1/E1 lines (1.5, 2 Mbps), T3/E3 (45, 34 Mbps), OC3 (155 Mbps), OC12 (622 Mbps), OC48 (2.5 Gbps), etc
  - **Switched Internet Access services**
    - FR (Frame Relay) or ATM
  - **Other Internet Access services**
    - Modems, ADSL, RDSI (BRI, PRI), etc
  - **Hosting/housing Services and CPD services**
    - Racks, servers (e.g.; Web), equipment, etc
  - **End user services**
    - VPNs, e-mail, news, Web, IP multicast, etc
  - **Content Provider services** (Content Distribution Networks such as Akamai)

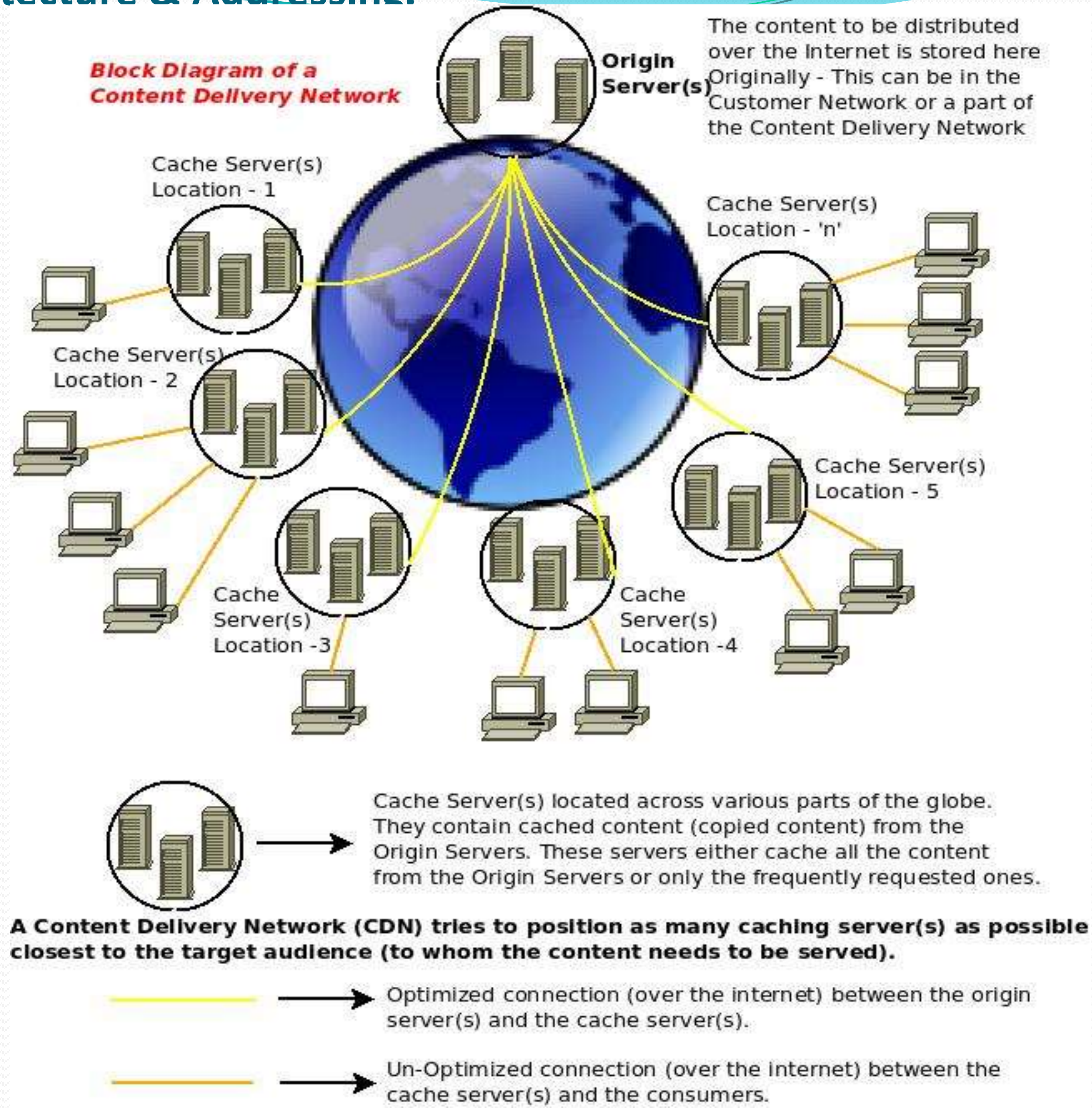
What can be done to improve the access to content?

- **From the perspective of the user:** obtain a better service
  - e.g., the user navigates and access to content and suffers delay, how to improve it ? → use the features offered by L5 protocols to improve the service like the "etag" field (Web caching) of HTTP if web access,
  - e.g., users want to share content (files) → use of User Content Networks based on P2P technologies (distributed technologies) to share contents.
- **From the perspective of the ISP:** any content that traverses an ISP can potentially be accessed by thousands-millions of users, and that has a cost,
  - E.g., some millions download the same youtube 4MB video → ISP's use Network Operator Content Networks to reduce the delay and to reduce operational costs,
  - **Proxing caches:** the operator store content in caches in order to offer better access services, **Goal:** save bandwidth/line capacity, reduce operational costs and reduce latency,
  - Examples of protocols: Internet Cache Protocol (ICP), Hypertext Caching Protocol (HTCP),

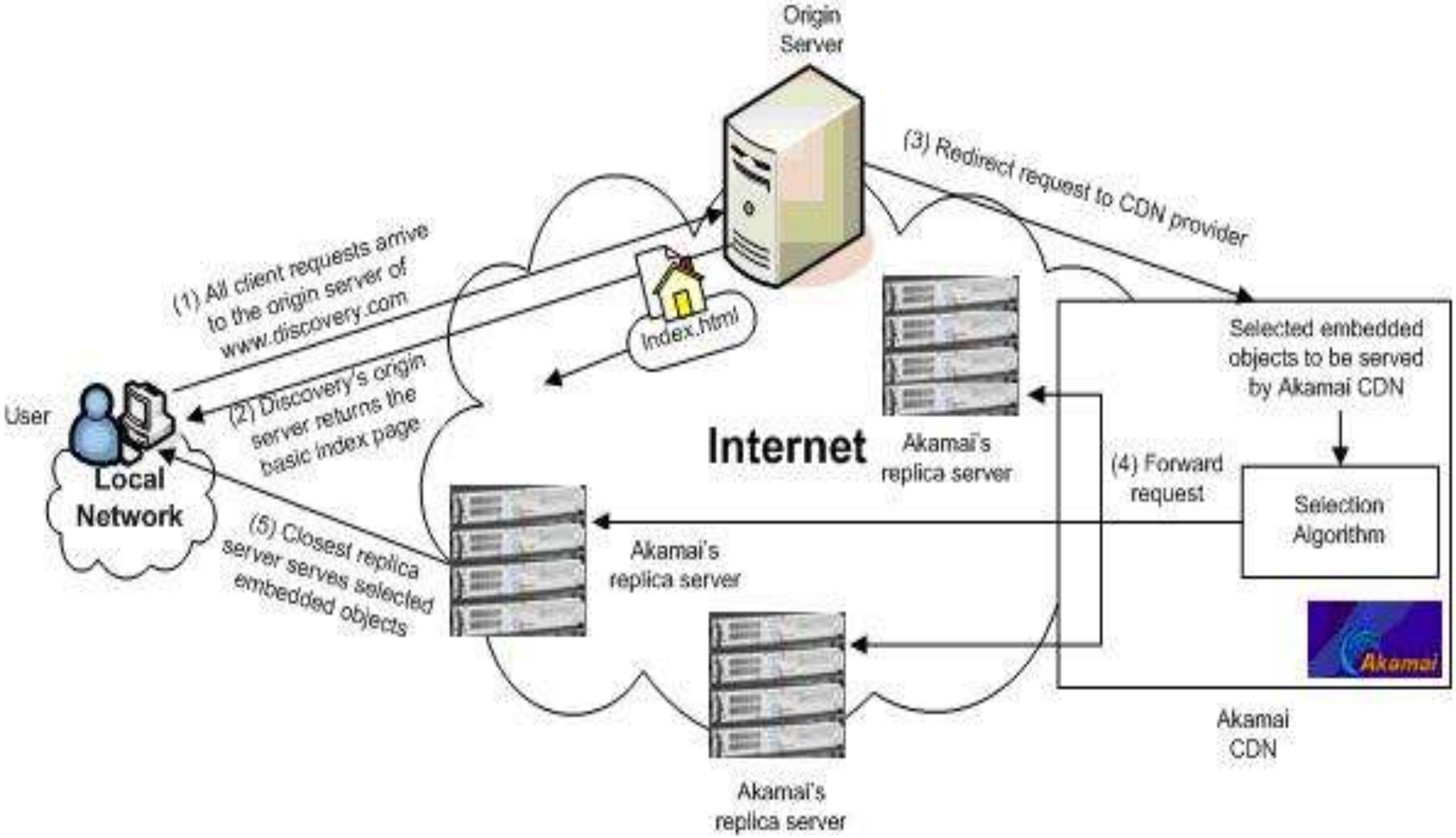
- ## What can be done to improve the access to content?
  - **From the perspective of the owner of the content:** give a good service,
    - If the owner offers a bad service in terms of delay of added value services, users will access competitors services → owners improve accessibilty to their servers and use Content Provider Content Networks such as CDN's to offer the best possible service.
    - **Farm servers**: local content networks that offer load balancing and thus better capacities and high availability of content, **Goal:** high availability & capacity,
    - **Mirrors**: distributed servers in several geographical localizations (manual access: the user has to choose the mirror server), **Goal:** reduce RTT,
    - **CDNs (Content Distribution Networks):** distributed servers in several geographical localizations (automatic: AKAMAI has more than 365.000 servers), **Goal:** reduce RTT, allow monitoring services, give value added services,

- **Content Distributions Networks (CDN):**
  - **Problem:**
    - Access to content may suffer from large RTT (Round Trip Times) when servers are located very far away from clients.
    - Access to content may suffer of peak load of demands that will impact in the QoS (Quality of Service) given to the clients.
  - **Solution:**
    - Put content as near the client as possible.
    - However, clients are around the globe, thus, a “unique” server can not be “near” all the potential clients distributed around the globe.
    - CND’s solve this problem, bringing client request to the near servers that cache the same content.

- **Content Distributions Networks (CDN):**



Topic 1: Internet Architecture & Addressing.



- **Content Distributions Networks (CDN):**

  - CDNs optimize content delivery by **putting the content closer** to the consumer and shorting the delivery path via global networks of strategically placed servers.

  - CDNs also **manage and maintain** the network elements that deliver Web content, such as text, images, and streaming audio and video, to the end user, streamlining the entire process.

  - Moreover, a CDN offers unique possibility to **provide for value added services** like customization and adaptation of content, virus scanning and ad insertion or accounting services.

- **Content Distributions Networks (CDN) functions:**

  - **Redirection of services**: to direct a request to the cache server that is the closest and most available.
  - **Distribution services**: a distributed set of surrogate servers that cache content on behalf of the origin server, mechanisms to bypass congested areas of the Internet or technologies like IP-multicast, and replication services.
  - **Accounting services**: to handle, measure, and log the usage of content.

- **Content Distributions Networks (CDN):**
  - Methodologies used in CDN’s
    - **Request Routing mechanisms (RFC 3568):** A request routing system uses a set of metrics in an attempt to direct users to surrogate that can best serve the request.
      - **DNS request routing:** a specialized DNS server is inserted in the middle of the DNS resolution process, returning a different set of A, NS or CNAME records based on user defined policies, metrics, or a combination of both
      - **Transport request routing:** the Request-Routing system inspects the information available in the first packet (client's IP address, port information, and layer 4 protocol) of the client's request to make surrogate selection decisions
      - **Application request routing:** Deeper examination of client's packets provides fine-grained Request-Routing control down to the level of individual objects. E.g. HTTP info, MIME, cookies,…

Content Distributions Networks (CDN):

- **Examples**: AKAMAI
  - 365,000 servers in 1,500 locations/networks across 135 countries, in Sept 2022
  - Uses DNS entries (CNAME)
  - If the object is cached → deliver the content
  - If the object is dynamic → handle it to the original server
- **Summary** - advantages of using CDNs
  - Reduced Latency
  - High Scalability
  - High Availability
  - Increased Offload

DNS request routing:

nslookup
> **www.microsoft.com**
Server: 147.83.32.3
Address: 147.83.32.3#53

Non-authoritative answer:
www.microsoft.com canonical name = **toggle.www.ms.akadns.net**.
toggle.www.ms.akadns.net canonical name = g.www.ms.akadns.net.
g.www.ms.akadns.net canonical name = lb1.www.ms.akadns.net.
Name: lb1.www.ms.akadns.net
Address: 65.55.12.249

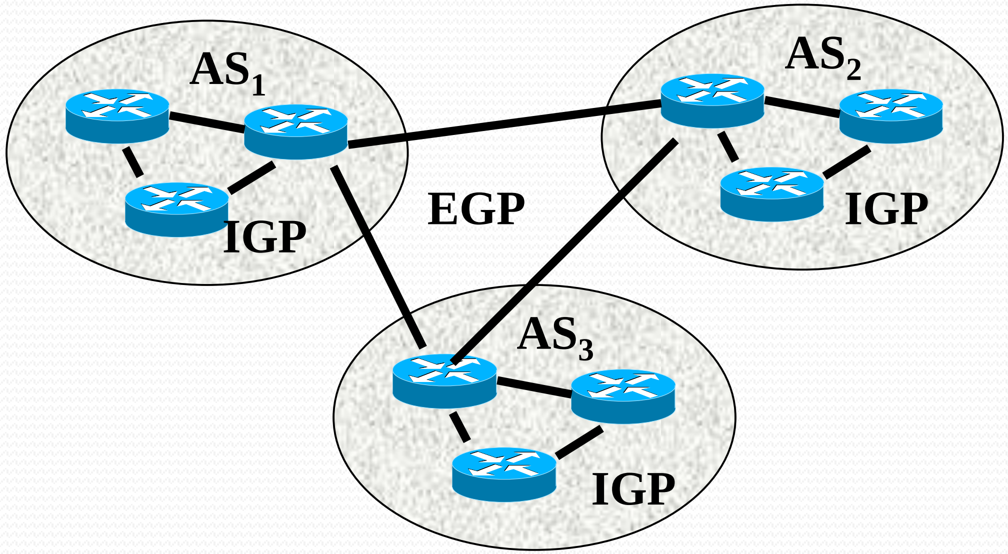
> **www.cisco.com**
Server: 147.83.32.3
Address: 147.83.32.3#53

Non-authoritative answer:
www.cisco.com canonical name = **www.cisco.com.akadns.net**.
www.cisco.com.akadns.net canonical name = geoprod.cisco.com.akadns.net.
geoprod.cisco.com.akadns.net canonical name = www.cisco.com.edgekey.net.
www.cisco.com.edgekey.net canonical name = www.cisco.com.edgekey.net.globalredir.akadns.net.
www.cisco.com.edgekey.net.globalredir.akadns.net canonical name = e144.cd.akamaiedge.net.
Name: e144.cd.akamaiedge.net
Address: 88.221.32.170

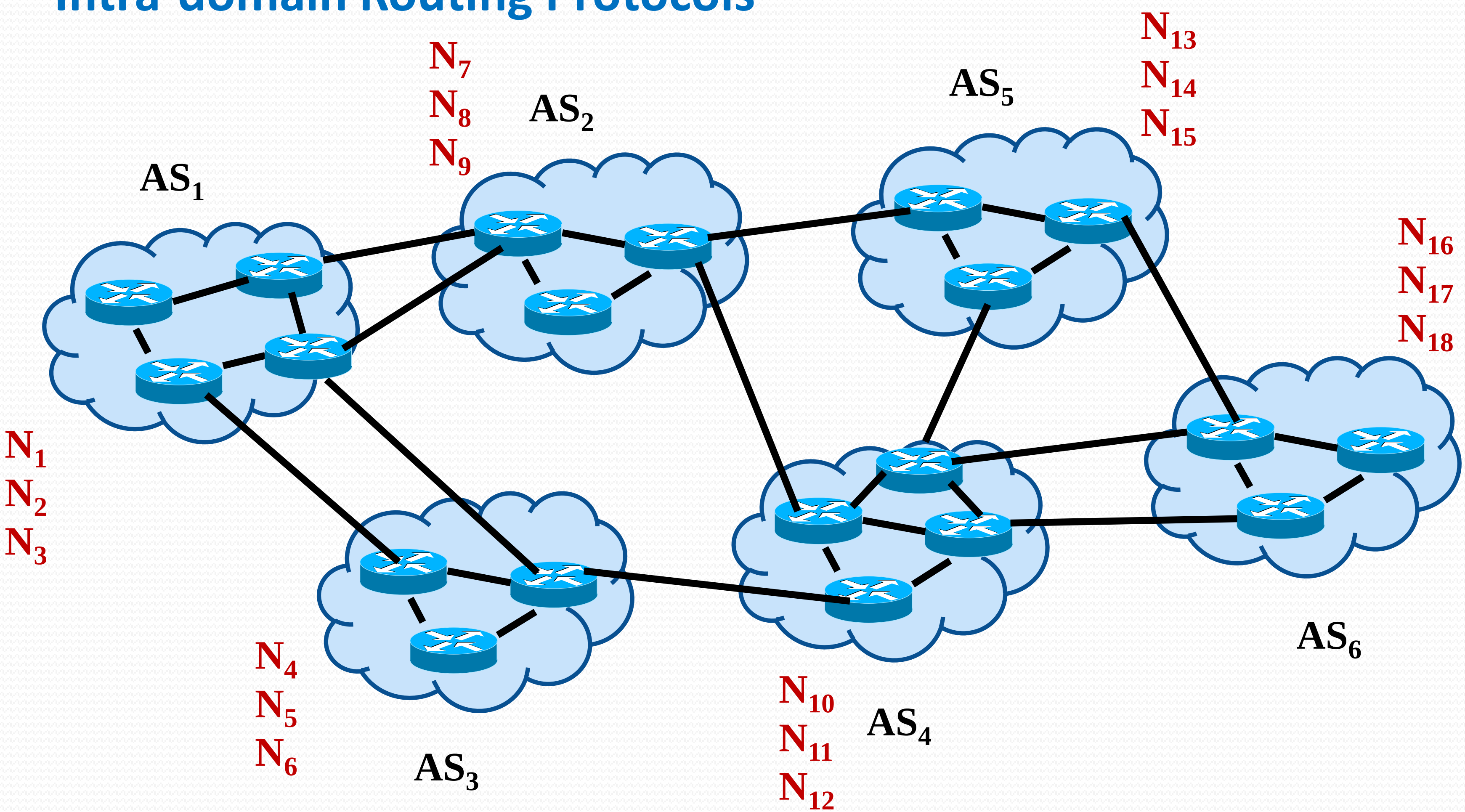
- **SLA (Service Level Agreement):** agreement that guarantees the contract offered by an ISP and that penalize the ISP if this one does not comply the contract.
  - **Customer-based SLA**: An agreement with an individual customer group, covering all the services they use.
  - **Service-based SLA**: An agreement for all customers using the services being delivered by the service provider.
    - An email system for the entire organization.
  - **Multi-level SLA**: The SLA is split into the different levels, each addressing different set of customers for the same services, in the same SLA:
    - **Corporate-level SLA**: Covering all the generic **service level management** (**SLM**) issues appropriate to every customer throughout the organization.
    - **Customer-level SLA**: covering all SLM issues relevant to the particular customer group, regardless of the services being used.
    - **Service-level SLA**: covering all SLM issue relevant to the specific services, in relation to this specific customer group.

- **SLA (Service Level Agreement)** - aspects a SLA address:
  - **Availability level** (% of time that the customer agrees in losing connectivity) → measured as the number of delivered packets respect to the transmitted
  - **Bandwidth contracted** (Mb/s)
  - **Throughput** (Mb/s) in high availability (rushy hours)
  - **Response time against connectivity failures** (e.g. 24/7 service), penalize with respect the amount of time without connectivity
  - **Redundancy** (multi-homing)
  - **Security**
  - **Monitoring services**
  - **Quality of Service (QoS):** service levels (e.g. Gold, Silver, …)
    - Packet marking, dropping, end-to-end delays, ….
- **Service credits:** something the Service Provider may offer in case your SLA is not achieved.
- The SLA is a customer support service
  - http://aws.amazon.com/es/s3-sla/
  - https://www.bmc.com/blogs/sla-template-examples/

- **Autonomous Systems (AS) or Routing Domain: is a group of IP networks run by one or more network operators with a single, clearly defined routing policy.**
  - AS are identified with 16 bits (65535 AS's)
  - AS's exchange routes (IP subnets) using **Inter-domain Routing Protocols** or **External Gateway Protocols (EGP's)** such as BGPv4
  - Internally, AS's exchange routes (IP subnets) using **Intra-domain Routing Protocols** or **Internal Gateway Protocols (IGP's)** such as OSPF, IS-IS, EIGRP, …



- **Autonomous Systems (AS) Inter-domain Routing Protocols and Intra-domain Routing Protocols**



Internet Service Provider (ISP) and Autonomous Systems (AS):

- An ISP is an administrative entity that may have one or more AS numbers assigned depending of its architecture and geographical situation
- In general an AS number may be assigned to an ISP or to a Corporative Network,
- Thus, not all AS are ISP, however all ISPs have one or more AS number assigned
- We will clarify ISP in the next sections

Regional Internet Registries (RIRs):

- Independent organizations that supports Internet resources coordination in a geographical region and develop consistent policies and promote best current practice for the Internet.

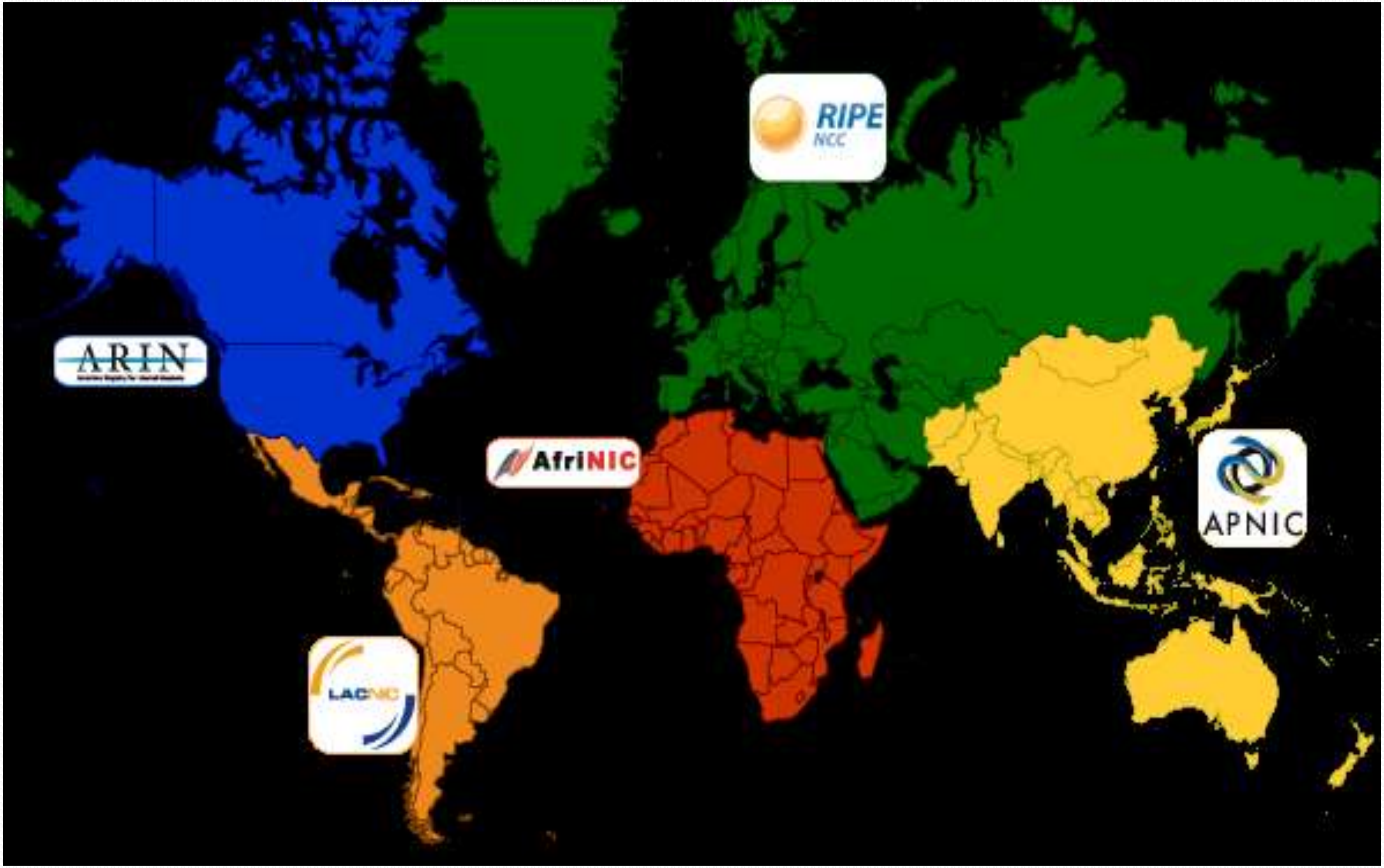
- Each RIR **manage**:

  - **IPv4 and IPv6 Address blocks** and **AS number** assignments

  - **DO NOT** manage Name Domains (done by ICANN: Internet Corporation for Assigned Names and Numbers)

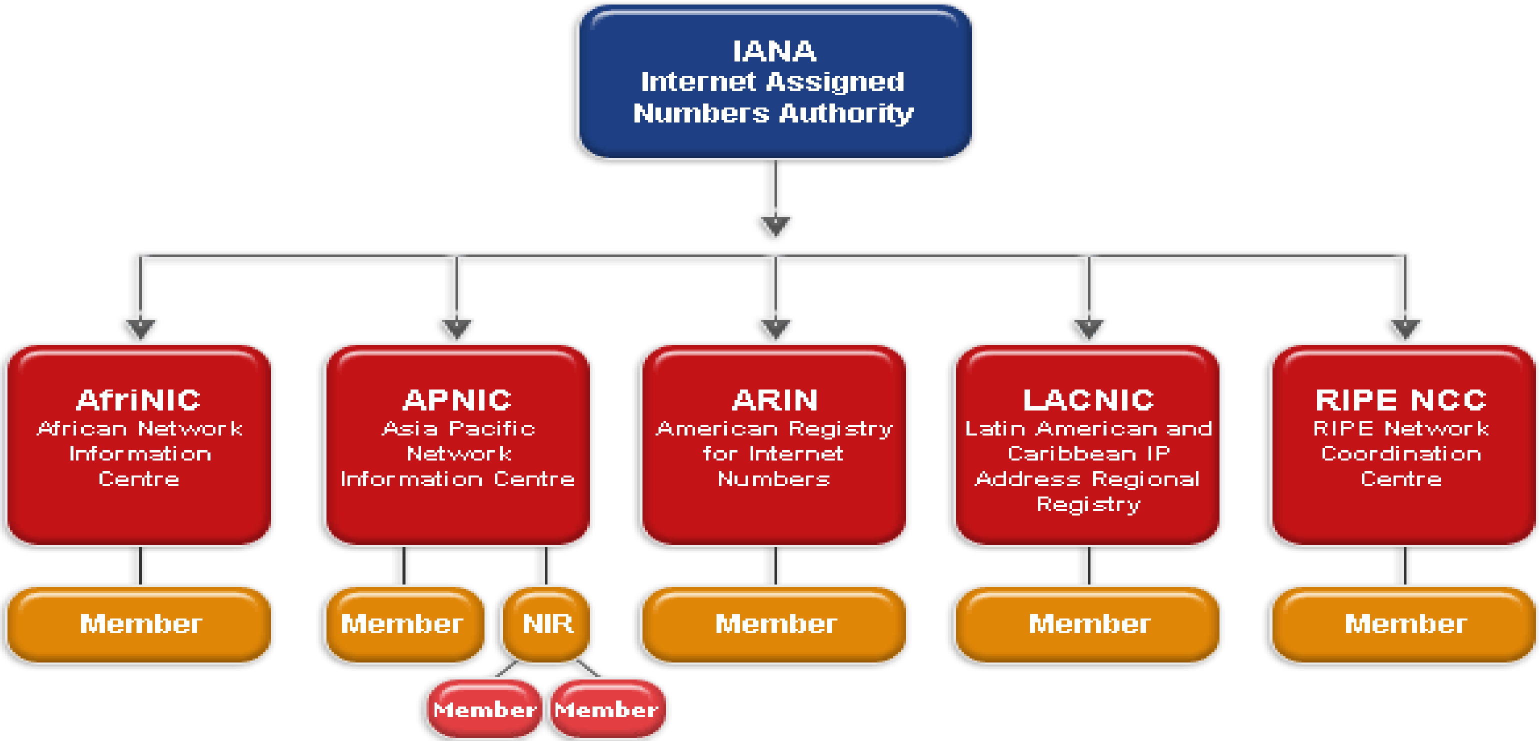
- There are 5 Regions
  - ARIN, RIPE, AFRINIC, LACNIC, APNIC

- **Regional Internet Registries (RIRs):**



Regional Internet Registries (RIRs):

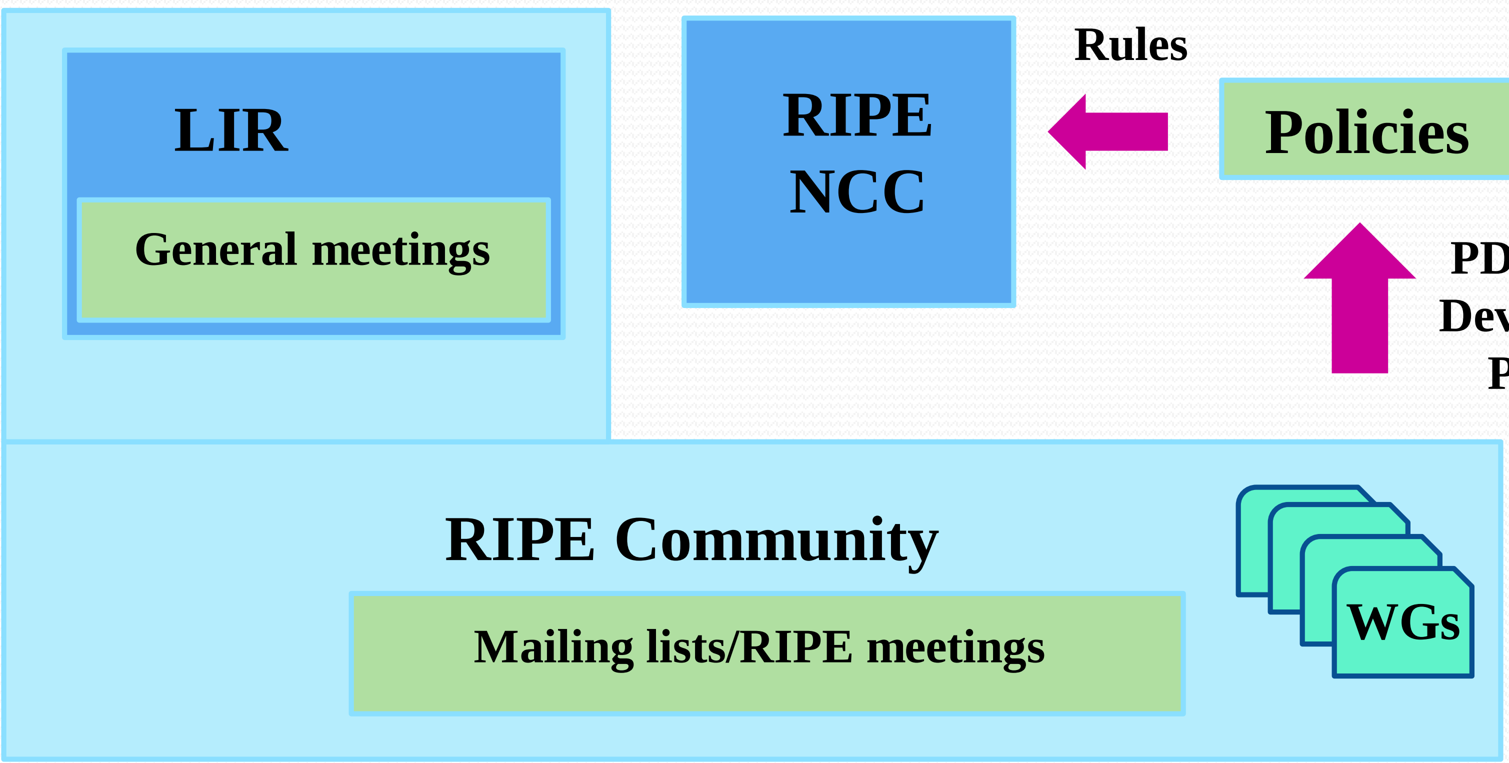
- LIR (Local Internet Registry)→ Members of RIR
- NIR (National Internet Registry) → coordinate IP allocations at national level. There are no NIR in Europe, but APNIC and LACNIC have them.
- **RIRs allocate IP address space and AS Numbers to Local Internet Registries (LIRs) that assign these resources to End Users**



- **RIPE NCC (RIPE Network Coordination Centre, http://www.ripe.net/)**
  - Provides Internet resource allocations, registration services and coordination activities that support the operation of the Internet globally.
    - **RIPE NCC** engages in a range of activities that can be defined as "**Internet Governance**". These activities include working with the technical community, governments, regulators, civil society and law enforcement agencies.
    - **LIR (Local Internet Register) Services:** members of a RIR. Organization that has been allocated a block of IP addresses by a RIR, and that assigns most parts of this block to its own customers
    - **Data & Tools:** provides databases and monitoring tools that support stable, reliable and secure Internet operations.
      - RIPE Database support: contains registration details of IP addresses and AS Numbers originally allocated by the RIPE NCC
      - Operates 1 of the 13 K-root Name Servers
      - Provides high-quality measurements and analysis that can be used for a variety of operational, media, governmental and law enforcement activities.

Topic 1: Internet Architecture & Addressing.

- **RIPE NCC (RIPE Network Coordination Centre, http://www.ripe.net/)**
  - **RIPE NCC** is the RIR**,**
  - **RIPE** (Réseaux IP Européens) is a forum open to all parties with an interest in the technical development of the Internet. The RIPE community's objective is to ensure that the administrative and technical coordination necessary to maintain and develop the Internet,
    - RIPE is not a legal entity and has no formal membership, anybody can participate.



- **RIPE NCC (RIPE Network Coordination Centre, http://www.ripe.net/):**
  - **Working Groups (WG's):** Address policy, routing, database, anti-abuse, cooperation, DNS, IPv6, RIPE NCC Services, connect, open source, measurement, analysis and tolos, IoT
  - **Policy Developmetn Process:**
    - **Open:** anybody can participate and done at mailing lists and at meetings,
    - **Transparent:** list discussions archived and public, meetings trasncripted,
    - **Developed bottom-up:** from you to them. You make the policies and RIPE NCC implements them

Criteria to obtain an AS number

- Only when the routing policy is different from your ISP provider routing policy (RFC 1930)
  - That implies that the customer has to be connected to at least two ISP (multi-homing) with different routing policies
- ¿What is a **LIR (Local Internet Registry)** ?
  - A LIR is an organization that has been allocated a block of IP addresses by a RIR, and that assigns most parts of this block to its own customers.
  - Most LIRs are ISP's, enterprises, or academic institutions.
  - Membership in an RIR is required to become an LIR, so any member of a RIR is a LIR. Not all AS's are LIR's, since you can operate with an AS number obtaining your IP blocks from a LIR,
  - See http://www.ripe.net/membership/maps/ in order to see ISP and LIR in any geographical zone covered by RIPE

Topic 1: Internet Architecture & Addressing.

- **Look at some Statistics from RIPE:**
  - https://www-public.imtbs-tsp.eu/~maigron/RIR_Stats/RIR_Delegations/World/ASN-ByNb.html
  - https://labs.ripe.net/statistics/
  - **See number of LIRs**
  - **% of membership per country**

ICANN, IANA and IETF

- **ICANN (Internet Corporation for Assigned Names and Numbers, https://www.icann.org):** in charge of making policy decisions about how the domain name system is run.
  - ICANN helps coordinate the Internet Assigned Numbers Authority (IANA) functions
  - ICANN performs the actual technical maintenance work of the Central Internet Address pools and DNS root zone registries pursuant to the Internet Assigned Numbers Authority (IANA) function contract.

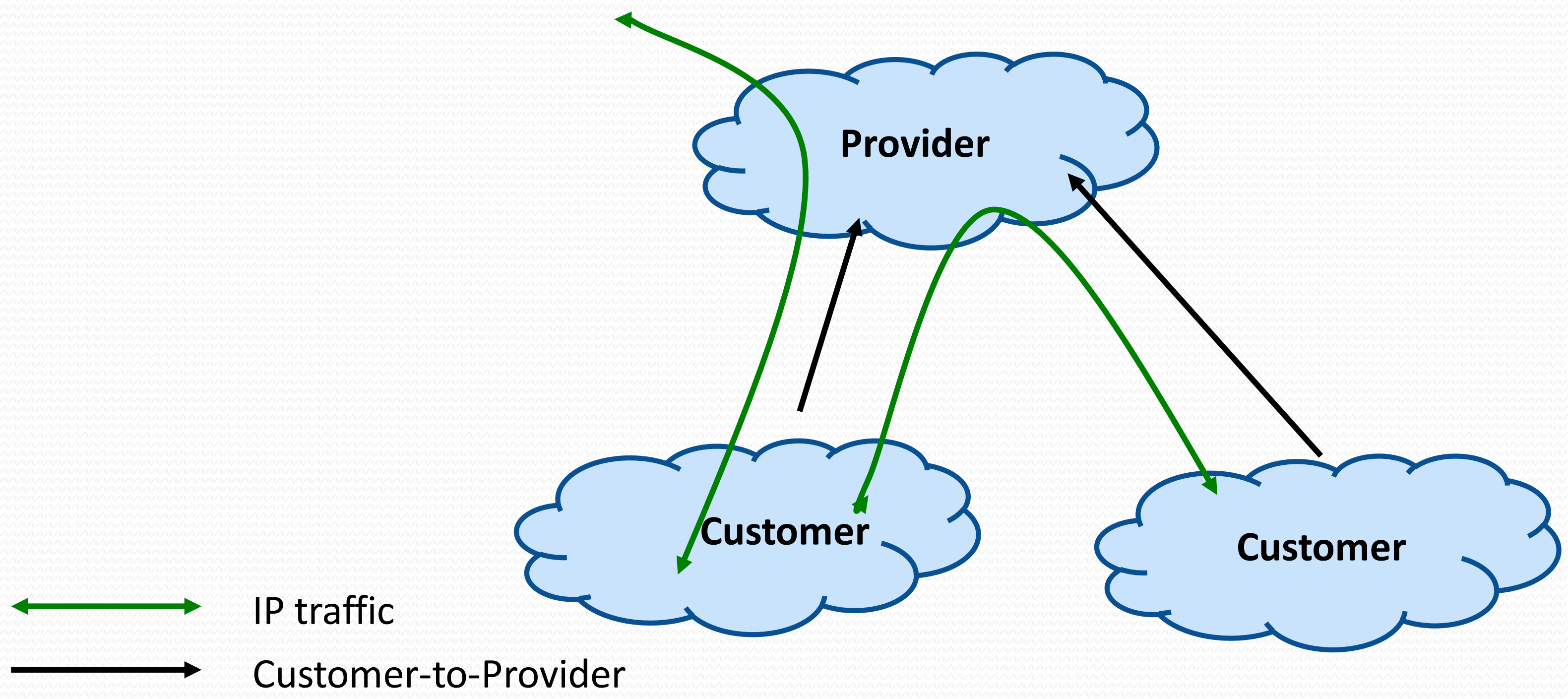
- **IANA (Internet Assigned Numbers Authority):** is responsible for managing the operational aspects of all of the various number spaces that make the various protocols work.
  - Oversees global IP address allocation, autonomous system number allocation, root zone management in the Domain Name System (DNS), media types, and other Internet Protocol-related symbols and Internet numbers.
  - Delivers IP@ and ASN to RIRs
  - IANA depends on ICANN.
  - https://icannwiki.org/Internet_Assigned_Numbers_Authority (ICANN vs IANA)

- **IETF (Internet Engineering Task Force):** is in charge of the engineering activities.
  - Develops and promotes voluntary Internet standards (TCP/IP suite, not L1/L2 that are on IEEE or ITU standards).

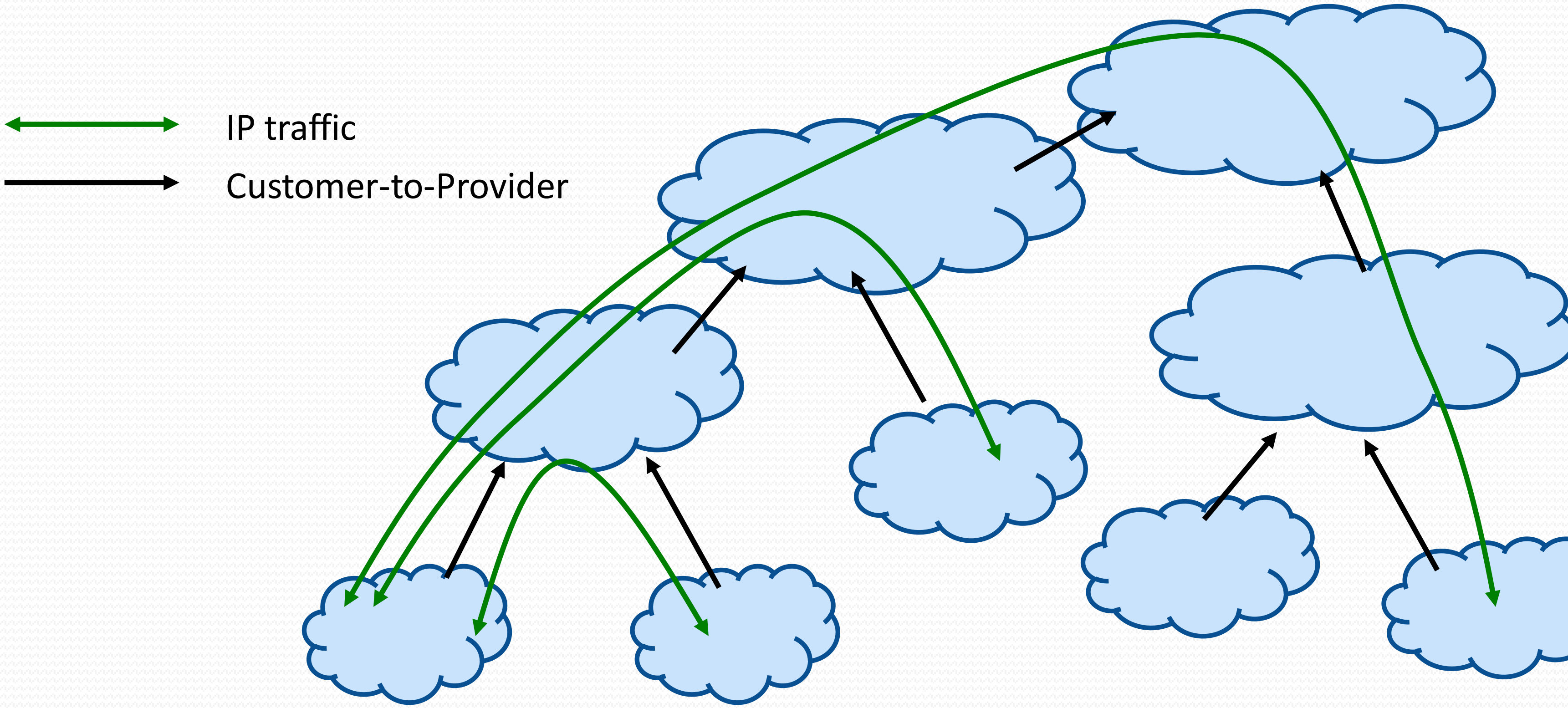
- **IANA (Internet Assigned Numbers Authority)**
  - Organization that dates from 70’s, originally, until 2016 hold by USA organization, now under International organizations called **Public Technical Identifiers (PTI).**
  - **Main tasks:** allocate and maintain unique codes and numbering systems that are used in the technical standards (“protocols”) that drive the Internet.
  - Activities can be broadly grouped in to **three categories (see https://pti.icann.org)**:
    - Domain Names**:** Management of the DNS Root, the .int and .arpa domains, and an IDN practices resource.
    - Number Resources**:** Co-ordination of the global pool of IP and AS numbers, primarily providing them to Regional Internet Registries (RIRs).
    - Protocol Assignments**:** Internet protocols’ numbering systems are managed in conjunction with standards bodies.

- **Peering (exchange of routes and traffic)**
  - Each AS may establish several types of relationships (peering) with several policies
  - **AS relationships:**
    - **Provider-to-Customer:** provider offer transit services to customers
    - **Customer-to-Provider:** a customer needs at least one provider
    - **Peer-to-Peer (non-transit):** two AS's agree in exchange their routes and their customer routes but do not transit others
    - **Peer-to-Peer transit (Siblings):** two AS's agree in exchange their routes and transit any other route

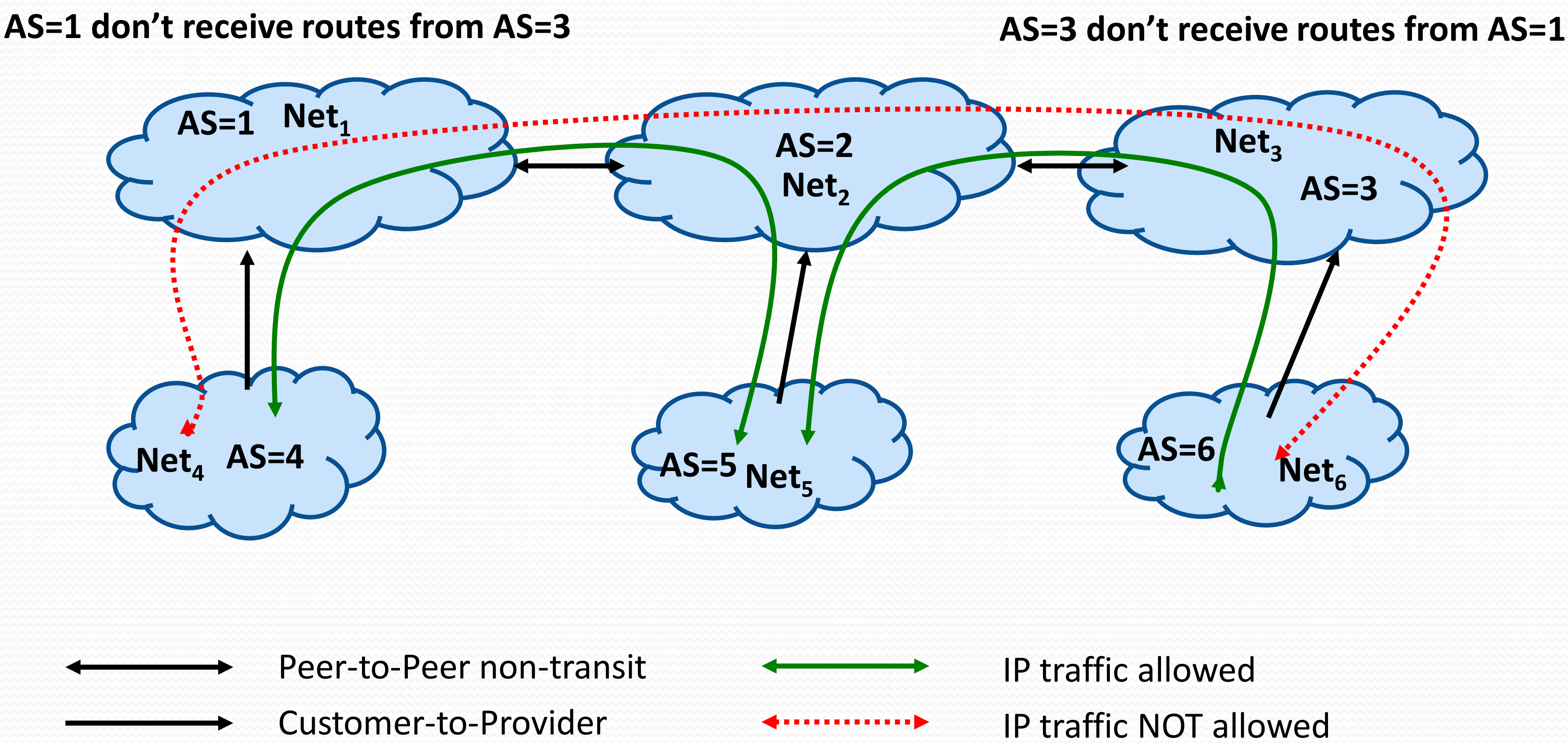
- **Provider-to-Customer:** provider offer transit to customer



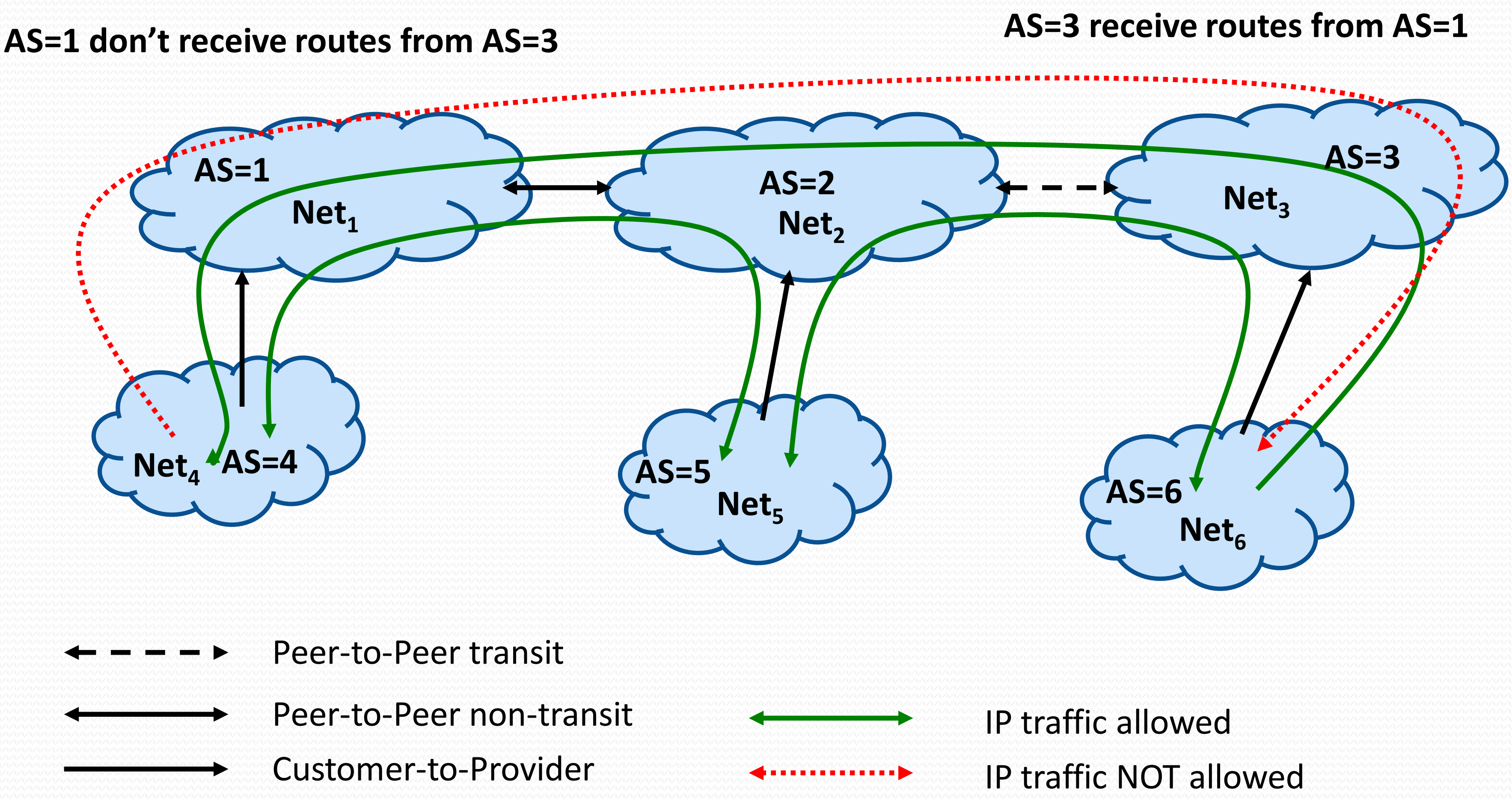
- **Providers form a hierarchy**



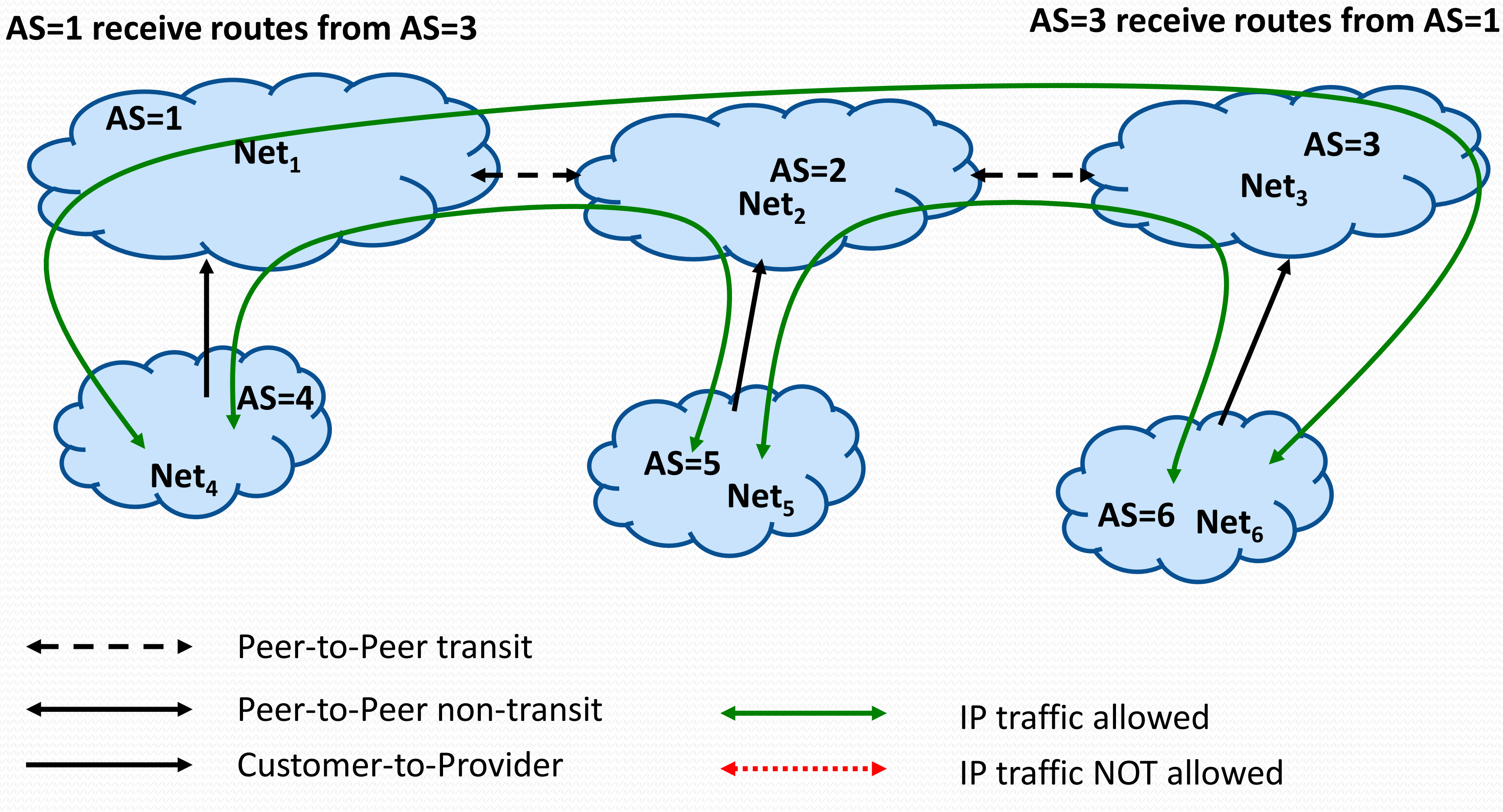
- **Peer-to-Peer non-transit:** two AS's agree in exchange their routes and their customer routes but do not transit others



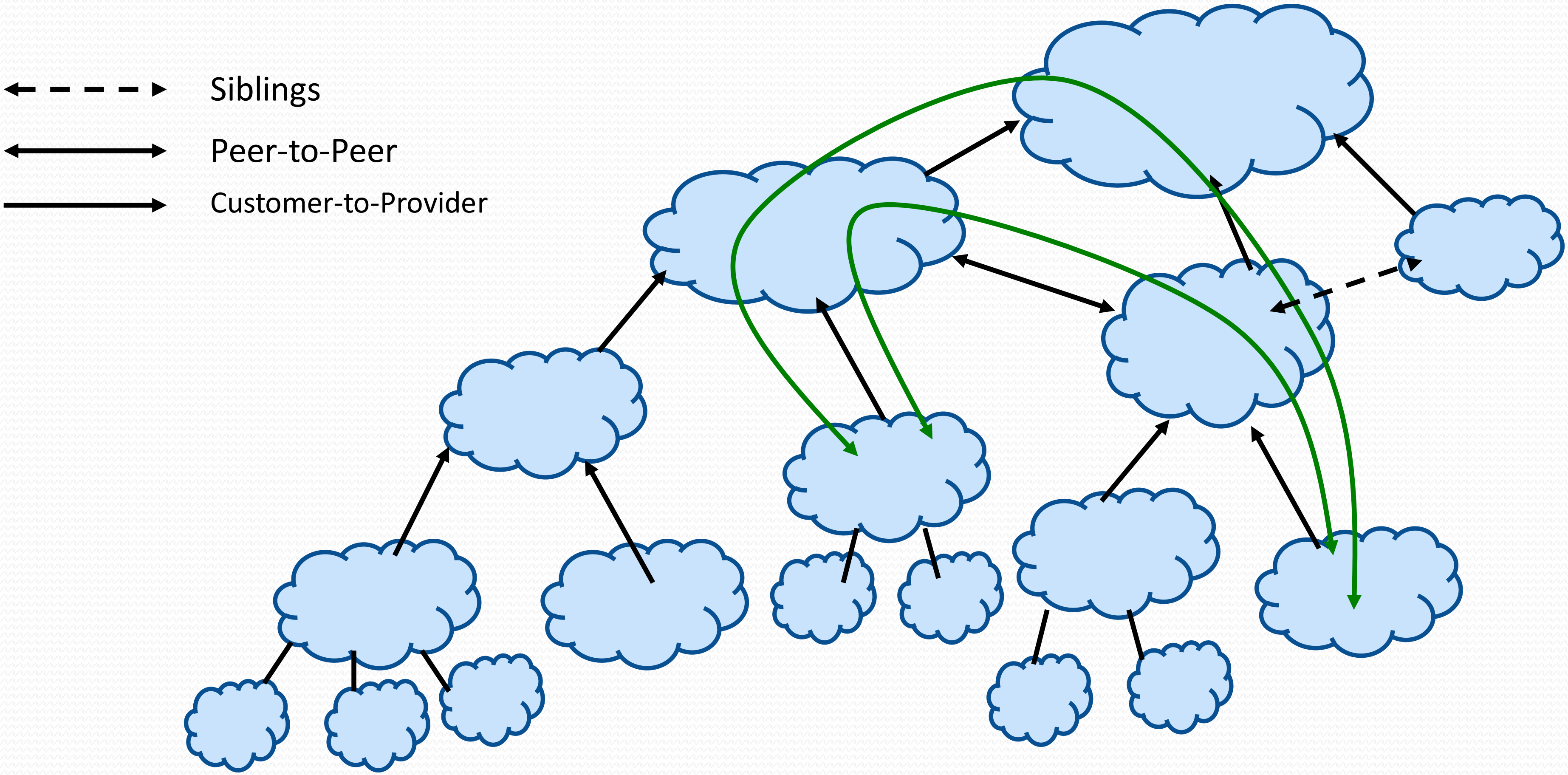
- **Peer-to-Peer transit (siblings):** two AS's agree in exchange their routes and transit any other route



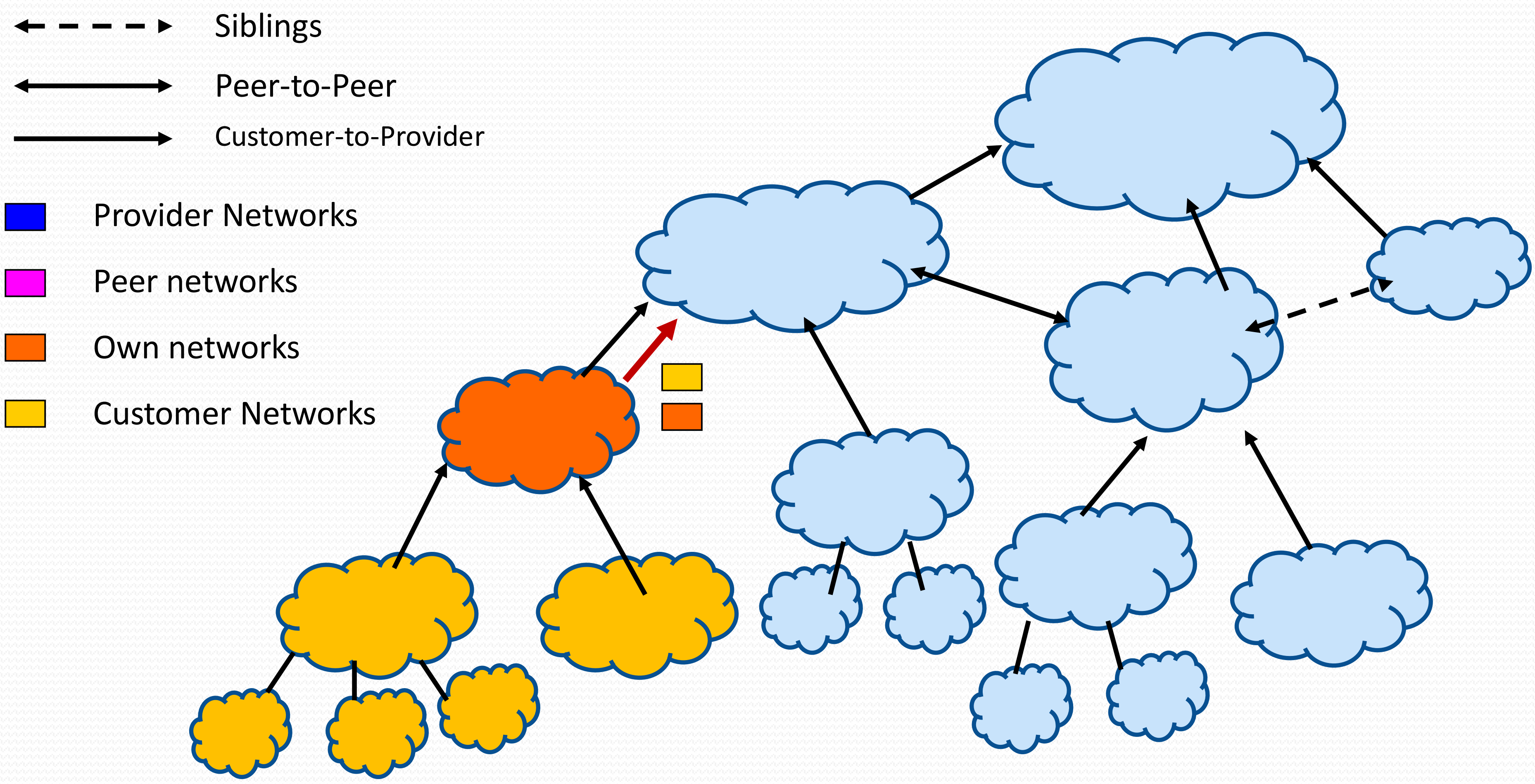
- **Peer-to-Peer transit (siblings):** If bi-directional traffic, there should be bi-directional sibling relationships



- Peer-to-peer and siblings offer short cuts in the Internet architecture

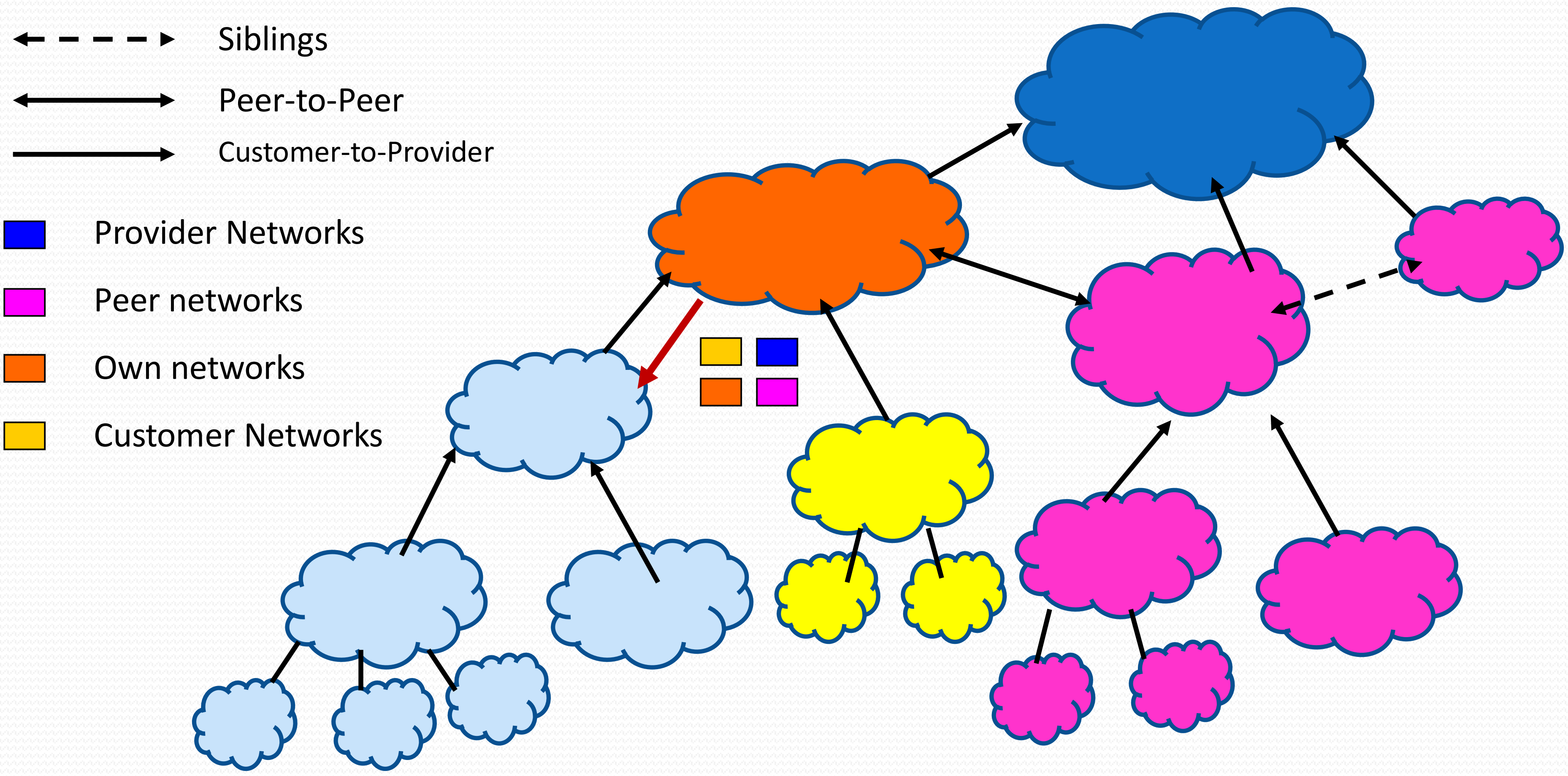


- **Export Routing Policies:**
  - **A customer exports to his provider** all his routes and his customer networks but not his peers and providers

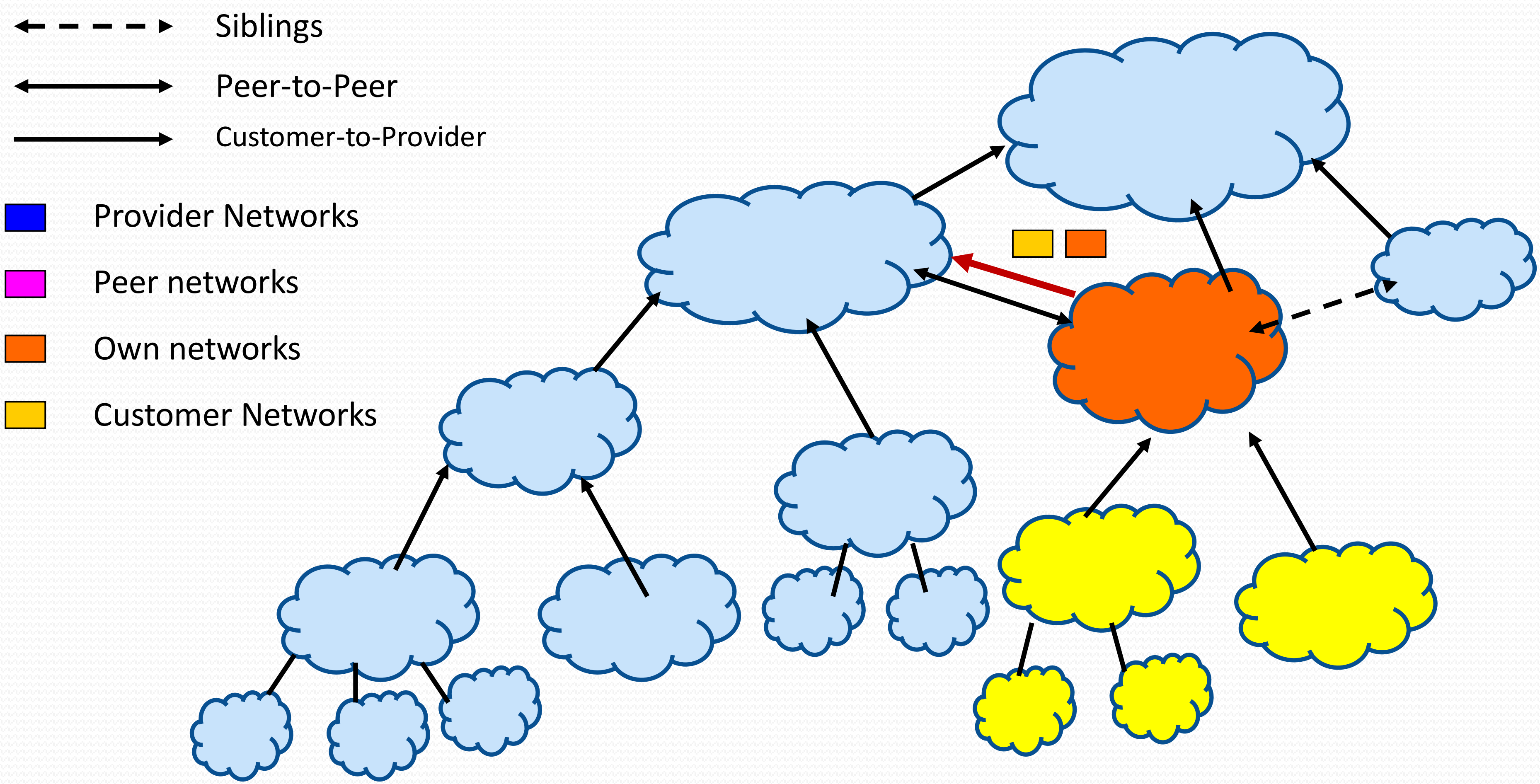


Export Routing Policies:

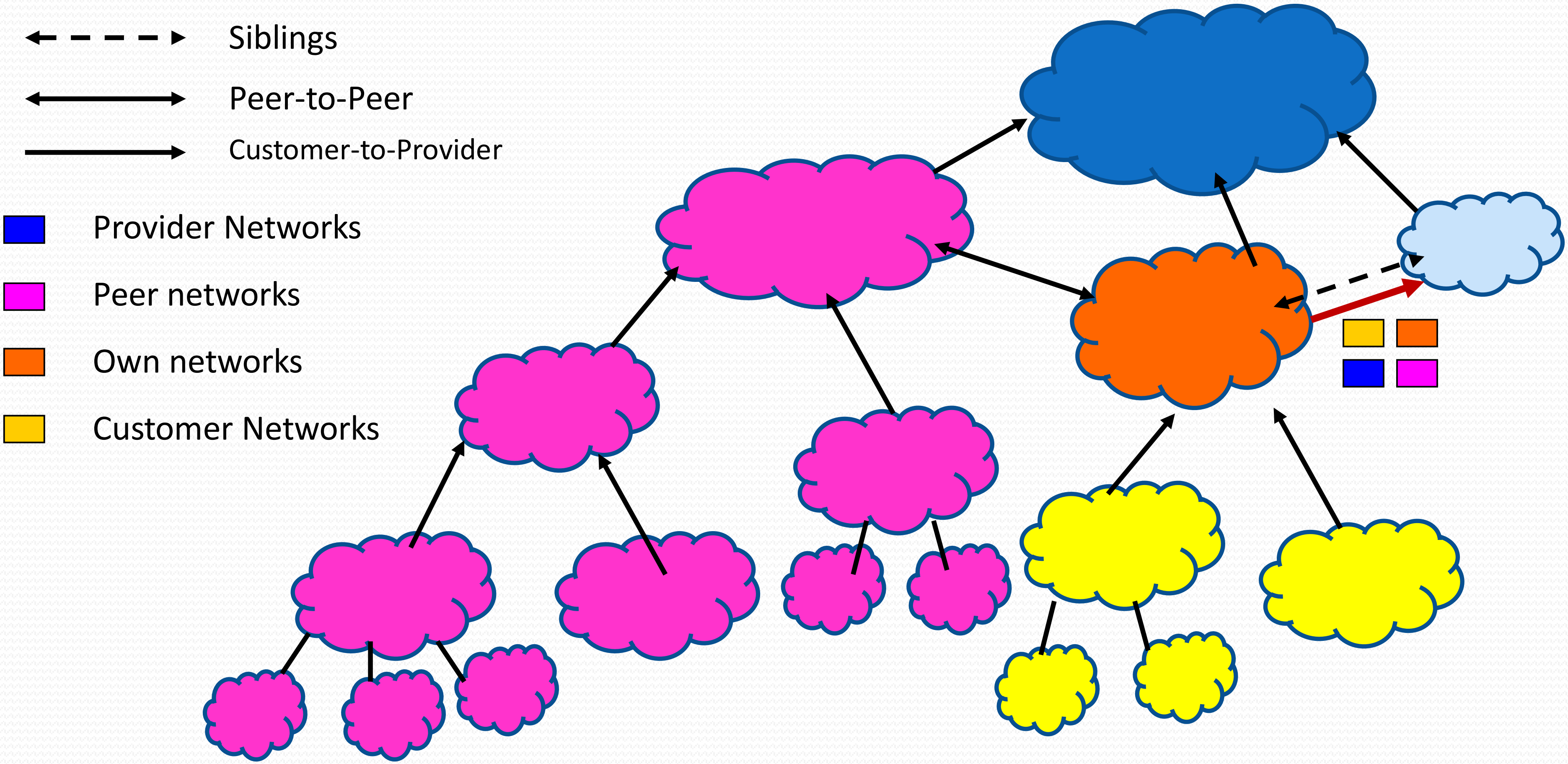
- A **provider export to his customers** his whole routing table (other customers, peers, providers and own networks) except customers of the customer



- **Export Routing Policies:**
  - **A peer to peer non-transit exports to other peers** all his routes and his customer networks but not his peers and providers

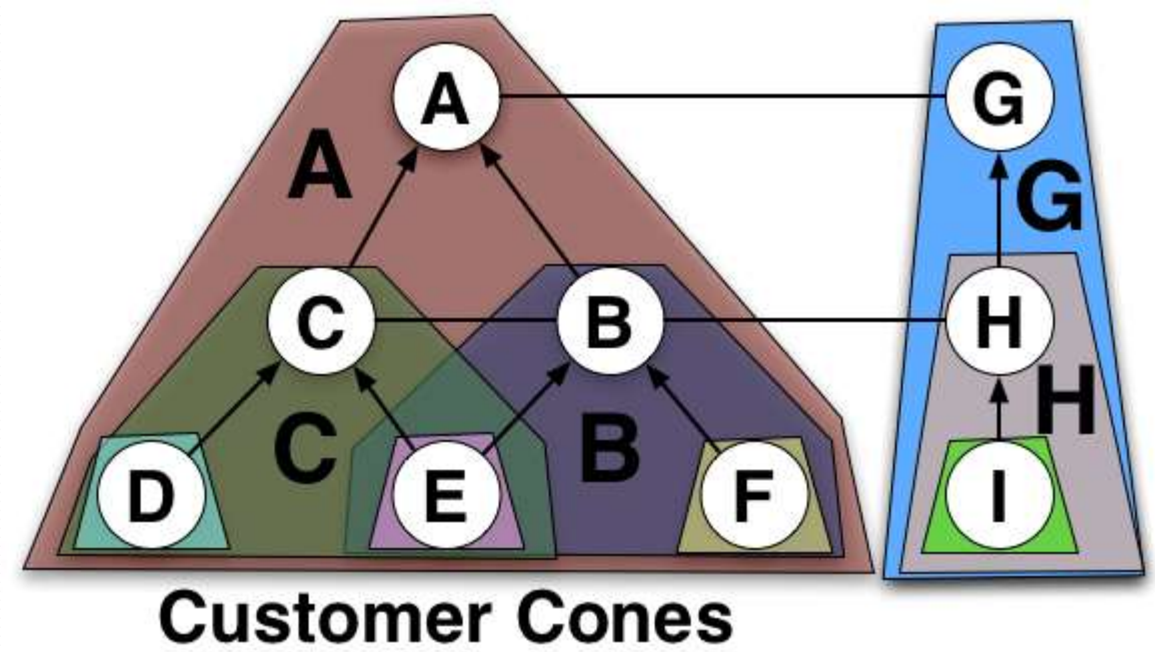


- **Export Routing Policies:**
  - **A peer to peer transit exports to other peers** his whole routing table (other customers, peers, providers and own networks) except customers of the peer



- **The Cooperative Association for Internet Data Analysis (CAIDA)**: **http://www.caida.org/home/**
  - CAIDA investigates practical and theoretical aspects of the Internet in order to:
    - provide macroscopic insights into Internet infrastructure, behavior, usage, and evolution,
    - foster a collaborative environment in which data can be acquired, analyzed, and (as appropriate) shared,
    - improve the integrity of the field of Internet science,
    - inform science, technology, and communications public policies.
  - E.g.
    - AS core IPv4 & IPv6 maps
    - AS ranking
    - Visualization of the IPv4 address space
    - AS rank by **customer cone**
    - DNS workload maps
    - …

- **Customer Cone:**
  - **Definition**: The set of ASes, IPv4 prefixes, or IPv4 addresses that can be reached from a given AS following only customer links (recursevely)
  - E.g., AS rank by **customer cone** (https://www.caida.org/data/as-relationships/)



- **Customer Cone size:**

  http://as-rank.caida.org/?n=50&ranksort=1&mode0=as-intro#customer-cone

  - **Peering cone size ratio** as the ratio in customer cone sizes of a pair of ASes if they (hypothetically) peered.
  - Similar customer cone sizes will have this ratio closer to 100, also an indication the ASes have more incentive to peer.
  - The closer this ratio is to zero, the larger the difference in customer cone sizes, and the less incentive the larger provider will have to peer with the smaller.

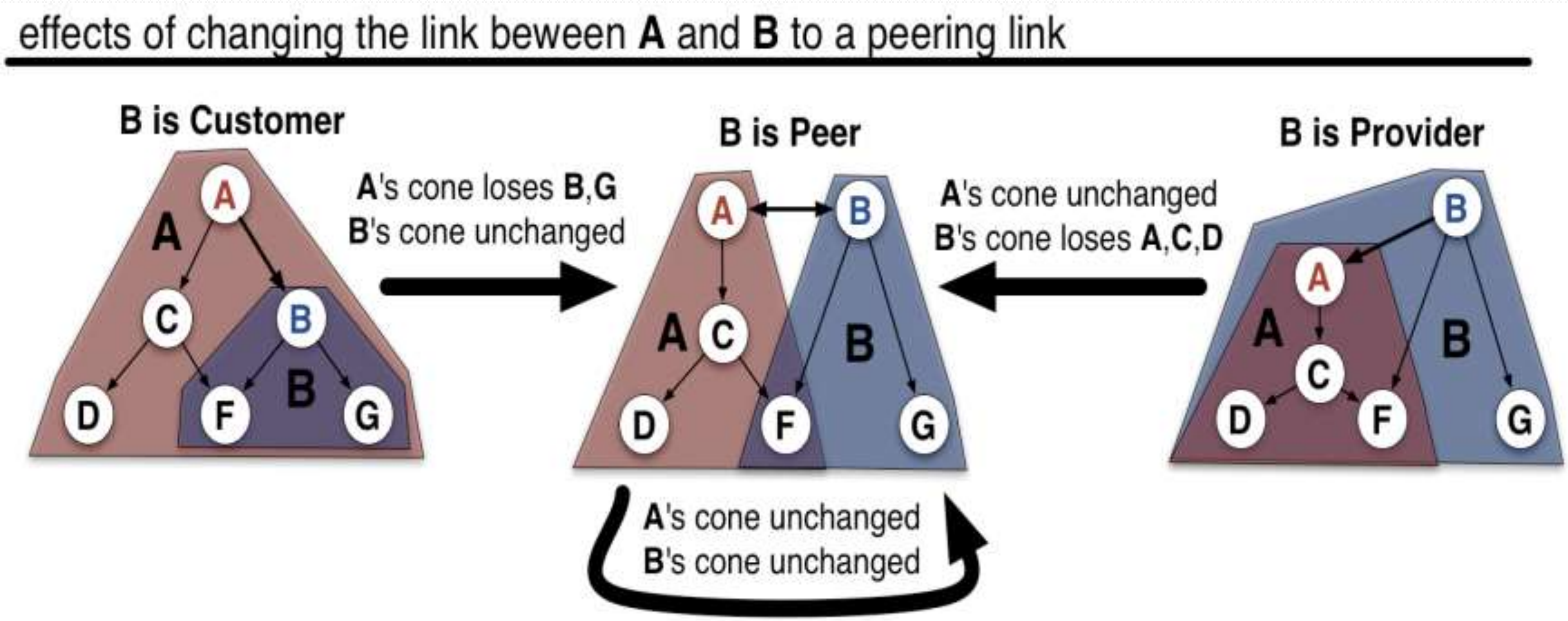
  - To compare magnitude of differences, the **peering cone size ratio** always uses the larger customer cone as the denominator.

    - For example, for AS pair S and N, with customer cone sizes C(S) and C(N), respectively , if C'(S) and C'(N) were their respective customer cone sizes if S and N became p2p peers (with other links unchanged), then the **peering cone size ratio** is C'(N)/C'(S) if C'(S) > C'(N), otherwise C'(S)/C'(N).

Customer Cone size:

http://as-rank.caida.org/?n=50&ranksort=1&mode0=as-intro#customer-cone

If the original graph had **B** as a customer of **A** then **A**'s cone contains 7 ASes: **A,B,C,D,E,F,G**. **B**'s cone contains three ASes:**B,F,G**. If the link between **A** and **B** is changed to a peering link, **A** loses customers **B** and **G**, which it had access to exclusively through its customer relationship with **B**. **A**'s cone does not lose **F**, since it can still reach it through its customer relationship with **C**. **A**'s cone size thus shrinks to 4 ASes: **A,C,D,F**. Since AS **B** did not previously reach any customers through **A**, its customer cone is unaffected by this change.



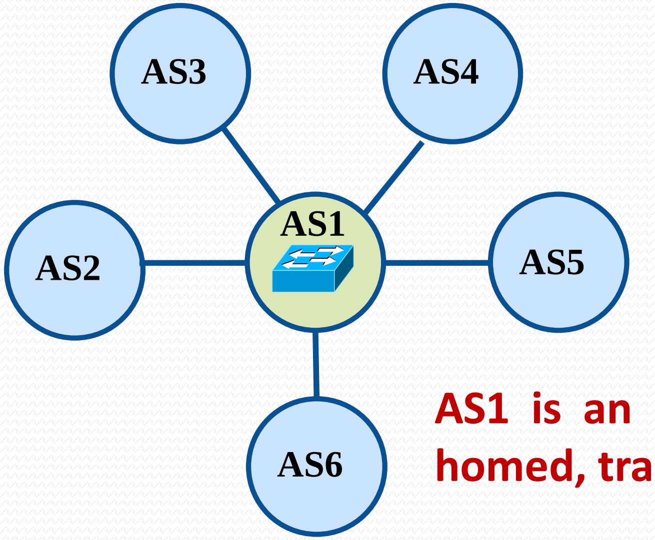
Topic 1: Internet Architecture & Addressing.

- **The Cooperative Association for Internet Data Analysis (CAIDA)**: **http://www.caida.org/home/**
  - E.g., **Ranking of AS's**: https://asrank.caida.org
  - E.g. AS rank examples (check them):
    - **Telefonica backbone (AS 12956)**
    - **Telefonica (AS 3352)**
    - **Jazztel (AS 12715)**
    - **Ono (AS 6739)**
    - **Orange (AS 12479)**
    - **Vodafone-ES (AS 12430)**
    - **Others ….**

- **How AS's established peering relationships ?**
  - Using an inter-domain routing protocol called BGPv4.
  - BGPv4 establishes sessions between routers using **private peering connections** or **public peering connections** (**Exchange Points**).
- **Private peering connections**
  - **Point-to-point connections** among two AS in order to transport routes and traffic
    - Increase link reliability and offer high bandwidth but at cost increases
  - **Alleviate the traffic that crosses public exchange points,**
  - **Connections are bi-laterally negotiated among ISP's** (peer-to-peer basis),
  - Routing and business treats are confidential and only known among partners (peers).

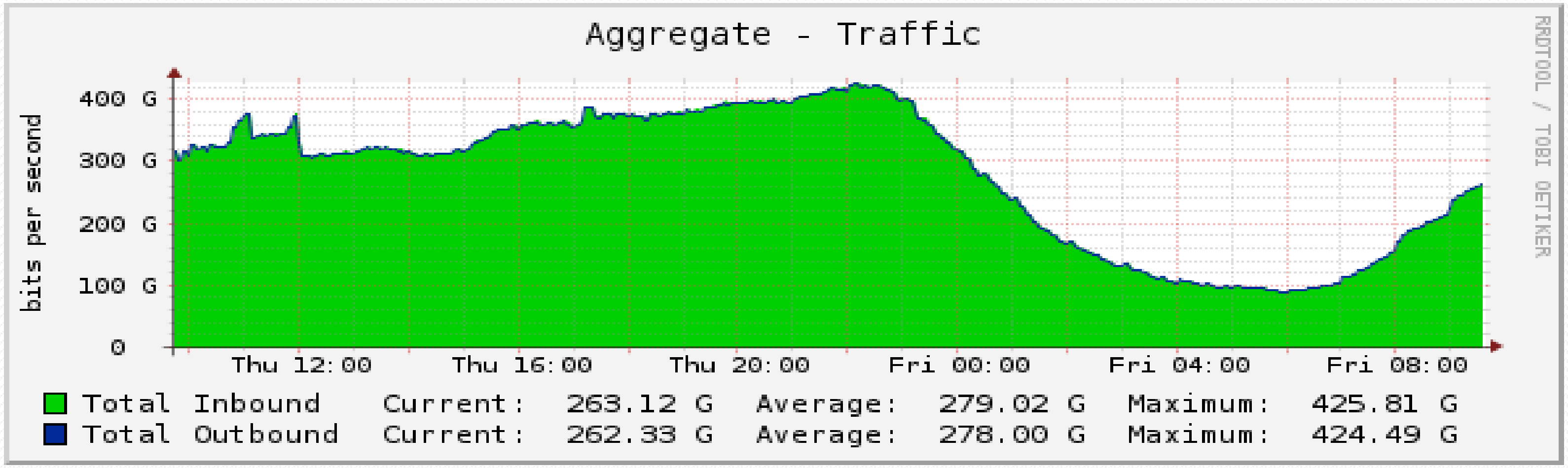
Public Peering: Exchange Points

- **National**: inter-connect ISP belonging to the same country
  - Espanix (www.espanix.net), Catnix (www.catnix.net), EuskoNix
- **International**: inter-connect National IXP; EuroIX (www.euro-ix.net) with 105 European IXP
  - The inter-connectivity matrix between IXP will give us information about ISP peering
- Not all ISP are connected to IXP

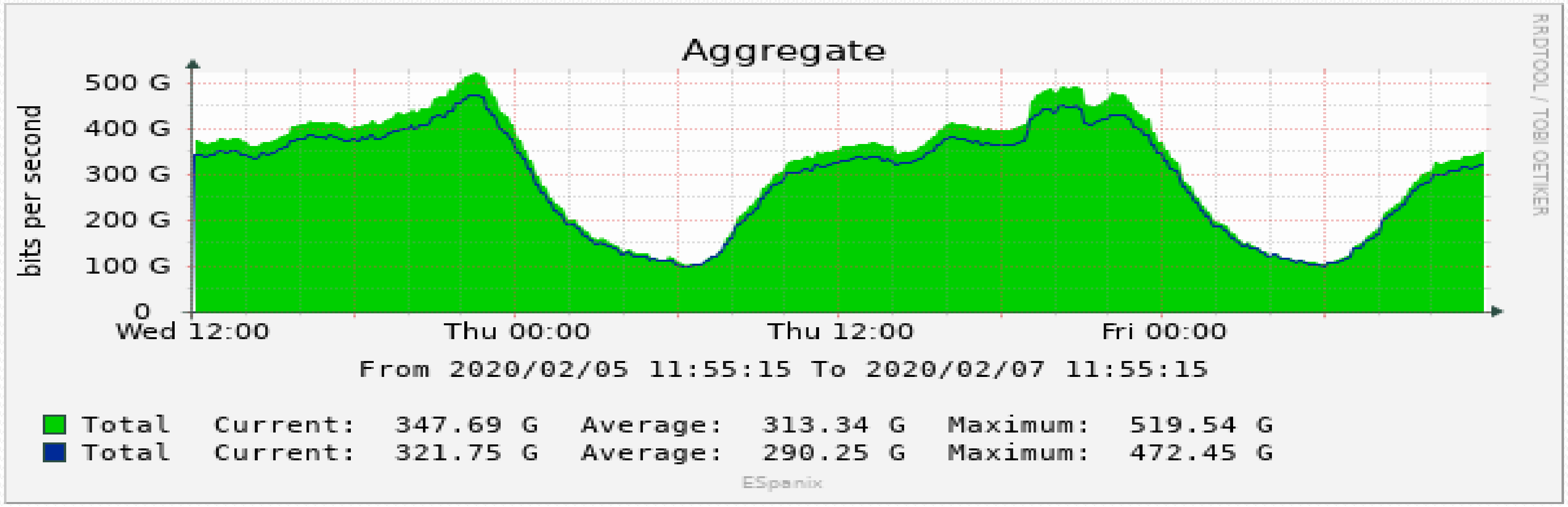


AS1 is an IPX (multi-homed, transit)

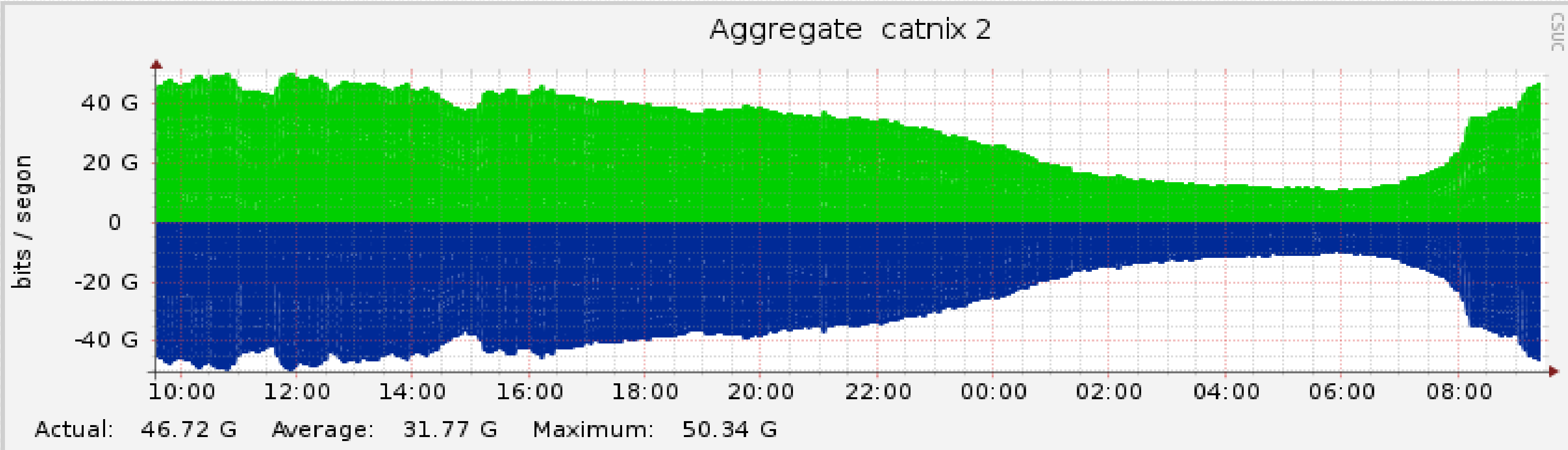
- **Espanix Traffic (07/Feb/2019)**



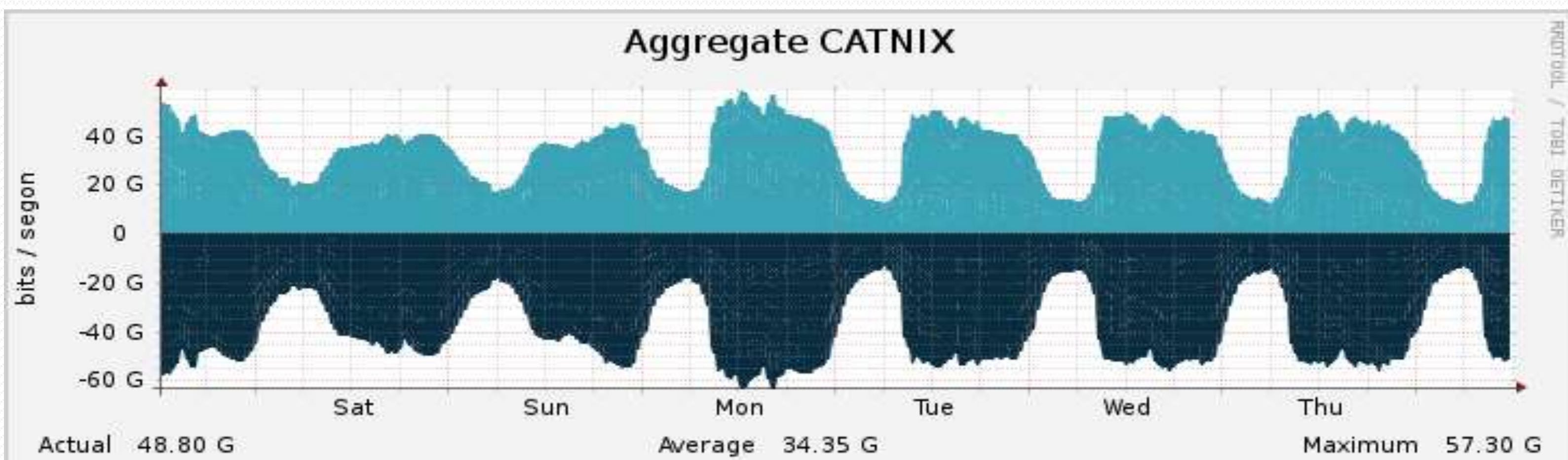
- **Espanix Traffic (07/Feb/2020)**



- **Catnix Traffic (07/Feb/2019)**



- **Catnix Traffic (07/Feb/2020)**



Internet Scalability

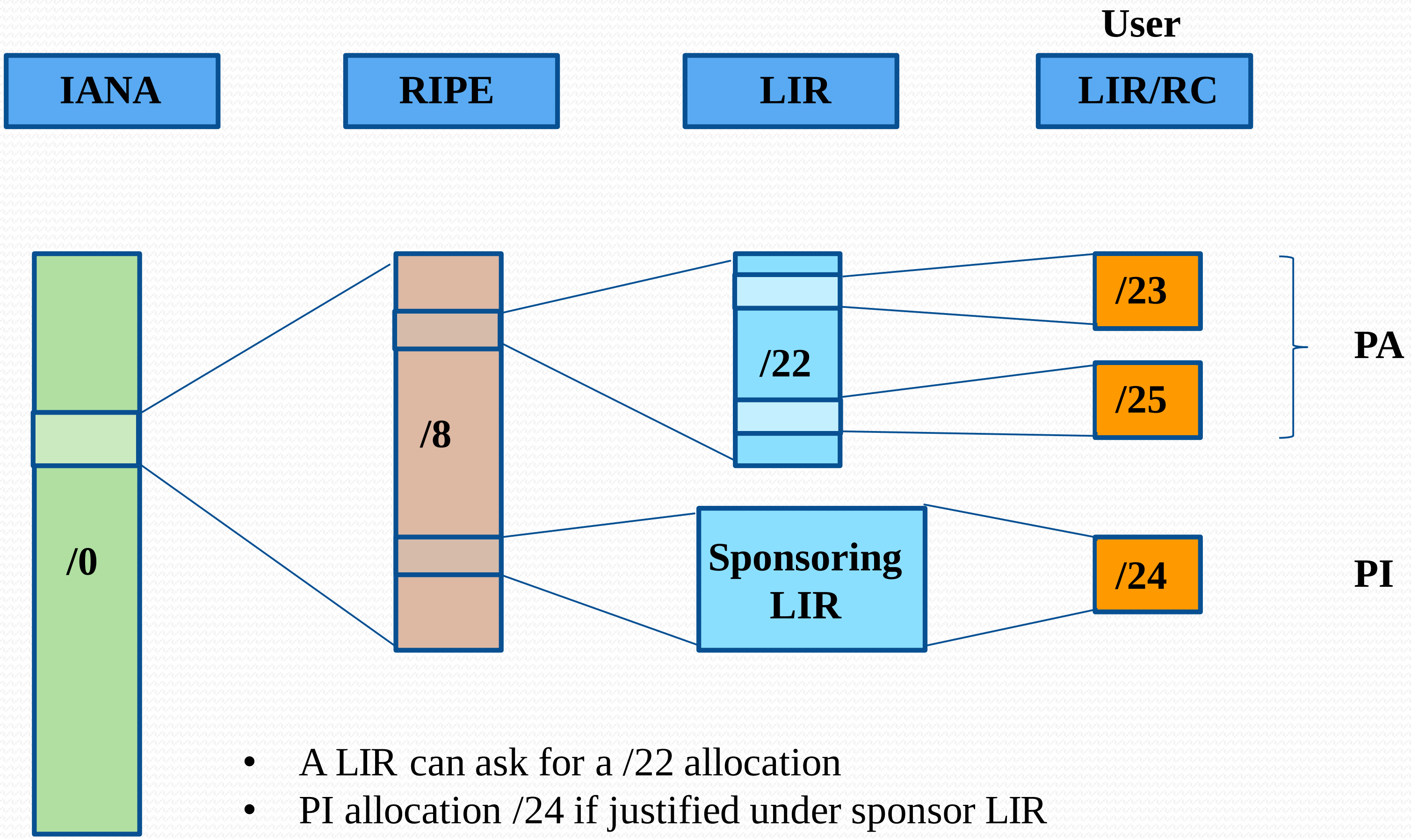
- **Scalability** (how many IPv4@ are available) impacts Internet architecture,
- **IP addresses**: define two issues in Internet
  - End-point identity (host identity)
  - Forwarding or location identity (network identity)
- **Address organization**:
  - IANA had blocks of /8 that leases to RIR (Regional Internet Registers),
  - RIRs leased subneting blocks of /22 from /8 to LIRs,
  - ISPs use the addresses or lease parts of blocks to dial-up end users and IP-Net clients.
  - **NAT**: intermediate solution to leverage exhaustion in the use of IP addresses
  - **IPv6** should be the final solution

- **IPv4 address space management:**
  - **IANA IPv4 blocks: 221 blocks of /8,** since from the 255 x /8 blocks, we have to remove multicast (224-241), reserved (0,242-255), private (10, 172.16/12, 192.168/16), etc.
  - IP address blocks **<u>are not</u>** bought, they are a public resource shared by LIR that obtain from RIR and that they can use always that (i) fulfil RIR rules, (ii) are a LIR member (annual maintenance payment).
    - If you are not a member (LIR) you can not get an IP block from RIR, and you will have to contact with an ISP.
  - RIR do not assign IP classes (A,B,C), since they use **CIDR** (*Classless Inter-Domain Routing).*
  - Last 5 /8 blocks were delived by IANA at Feb 2011, one block to each RIR. Thus, RIPE currently has the remaining of a single /8 block.
    - One /8 is equal to 16.8 million IPv4 addresses.

Types of IPv4 addresses in Internet

- **PA (Provider Aggregatable):**
  - Blocks of addresses that may be sub-assigned to other ISPs or to other companies that also may leased the addresses to their customers
  - May be aggregated by routing protocols
  - If an entity changes its ISP and its IP block is PA, he has to **return** the IP block to the original provider and obtain a new block from the new ISP, except if the block was directly obtained from RIPE
- **PI (Provider Independent):**
  - Blocks of addresses that can not be sub-assigned to other ISP (RIPE does not already assign this kind of blocks) and only may be assigned to end users
    - If a Corporative network wants this kind of blocks , he may ask to RIPE (via its ISP)
  - Can not be aggregated by routing protocols
  - These blocks are **portable** (changing ISP implies that you may keep the block)

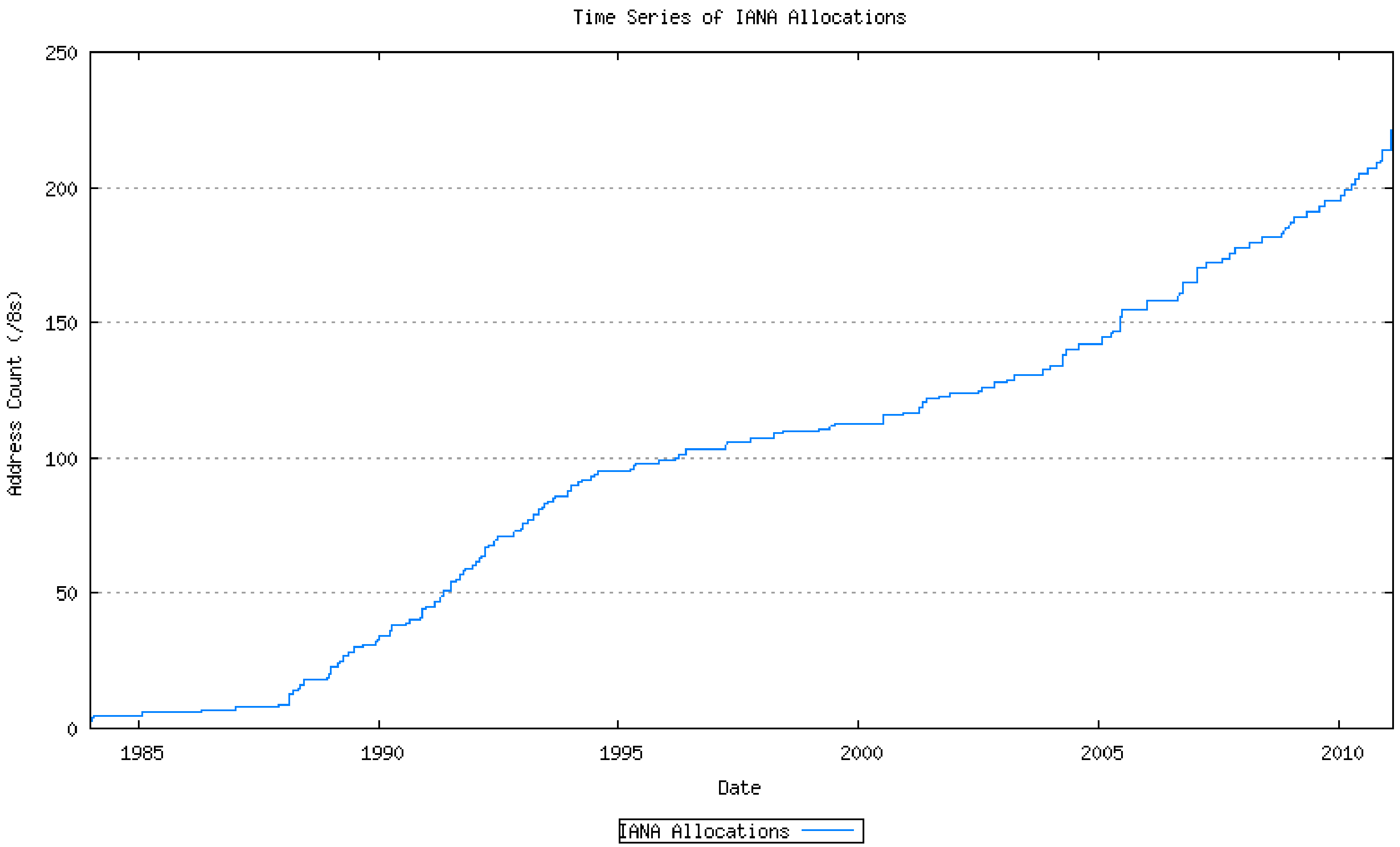
- **IPv4 addresses:** RIPE assigment



- A LIR can ask for a /22 allocation
- PI allocation /24 if justified under sponsor LIR

Topic 1: Internet Architecture & Addressing.

- **IANA IP blocks (221 blocks of /8) (YEAR 2010)**

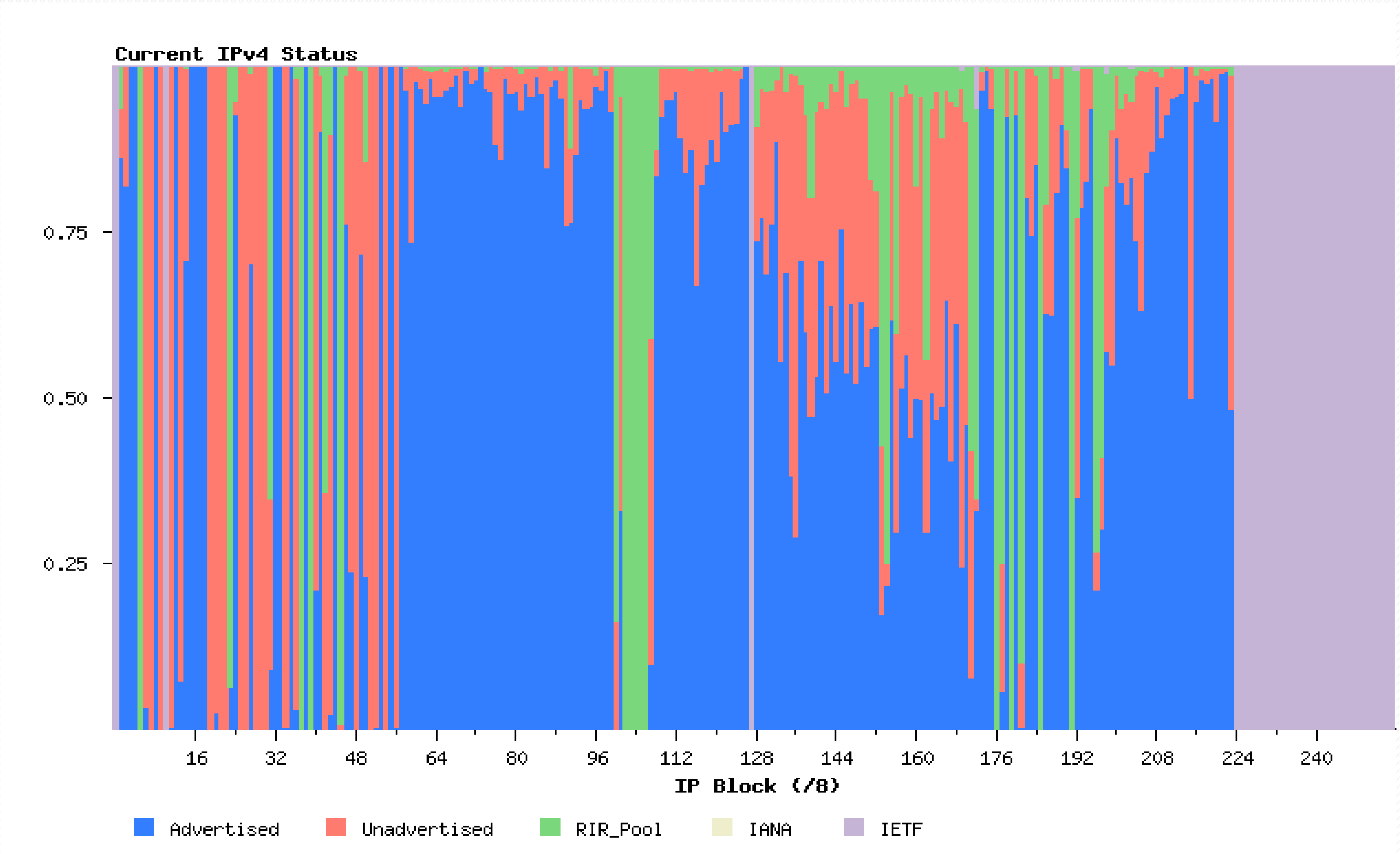


- **IPv4 address space management**
  - **IANA IP blocks (221 blocks of /8)**

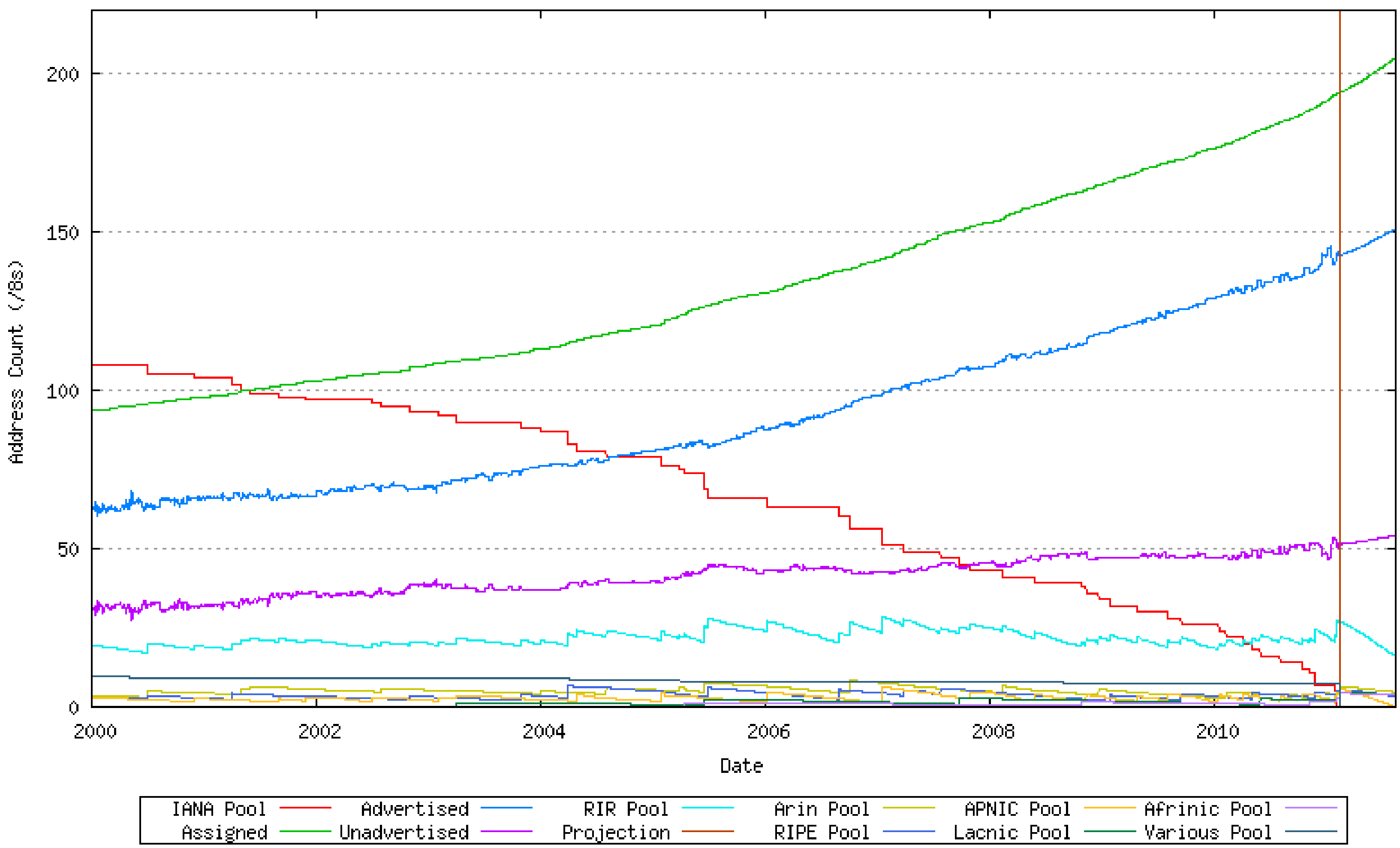
  - Any individual IPv4 address could be in any one of five states:
    - reserved for special use, or
    - part of the IANA unallocated address pool,
    - part of the unassigned pool held by an RIR,
    - assigned to an end user entity but not advertised in the routing system, or
    - assigned and advertised in BGP.

Topic 1: Internet Architecture & Addressing.

- **IPv4 blocks** advertised and unadvertised in 2010



- **IPv4 exhaustion → IANA (Feb/2011), RIR (08/2011)**



- Look at RIPE statistics and RIPE IPv4 exhaustion webpage

RIPE web page (run-out, pool and waiting list):

- https://www.ripe.net/manage-ips-and-asns/ipv4/ipv4-run-out?pk_vid=b419d928bf32ffaa15984529274e152d

  (run-out/depletion process)
- https://www.ripe.net/manage-ips-and-asns/ipv4/ipv4-pool (pool of IPv4@)
- https://www.ripe.net/manage-ips-and-asns/ipv4/how-waiting-list-works (waiting list for obtaining a /24 net for any recovered IPv4 net)

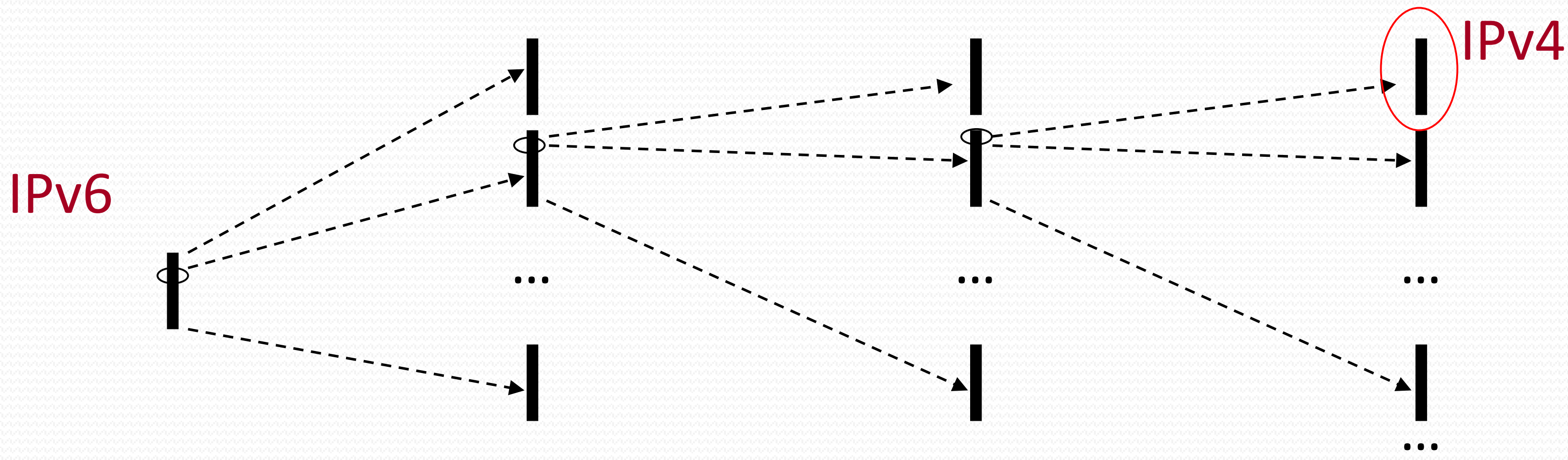
IPv4 address used

- https://www-public.imtbs-tsp.eu/~maigron/RIR_Stats/RIR_Delegations/World/ASN-ByNb.html

IPv6:

- Address scheme that increases the number of IPv4 addresses from 32 bits (2^{32} space) to 128 bits (2^{128} space)
- Improves IPv4 Address space and other issues such as security (IPSEC)
- Check IPv6 address space use in:

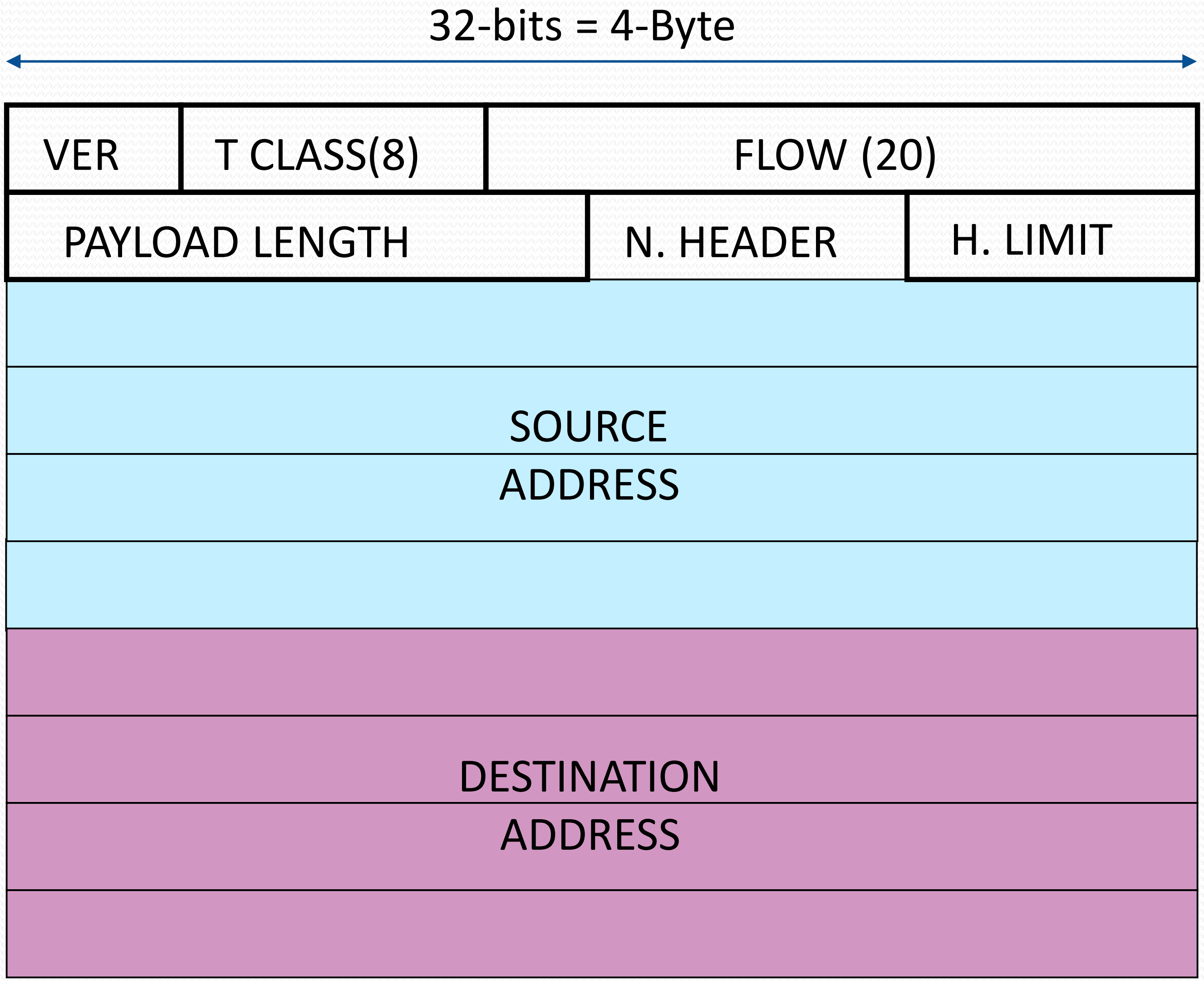
https://www-public.imtbs-tsp.eu/~maigron/RIR_Stats/RIR_Delegations/World/ASN-ByNb.html



IPv6 deployment:

- ISPs and telecom operators use IPv4
- IPv6 islands
  - IPv6 in IPv4 encapsulation, **dual stack IPv4-IPv6** and **IPv6 translation to IPv4**
- From IPv4 to IPv6
  - Technologies that begin with IPv6. E.g UMTS, sensors, vehicular, etc !!! in order to force the massive IPv6 deployment
  - Migrate IPv4 to IPv6 is very costly and should be progressive

- **IPv6 Header:**


32-bits = 4-Byte
VER
T CLASS(8)
FLOW (20)
PAYLOAD LENGTH
N. HEADER
H. LIMIT
SOURCE
ADDRESS
DESTINATION
ADDRESS

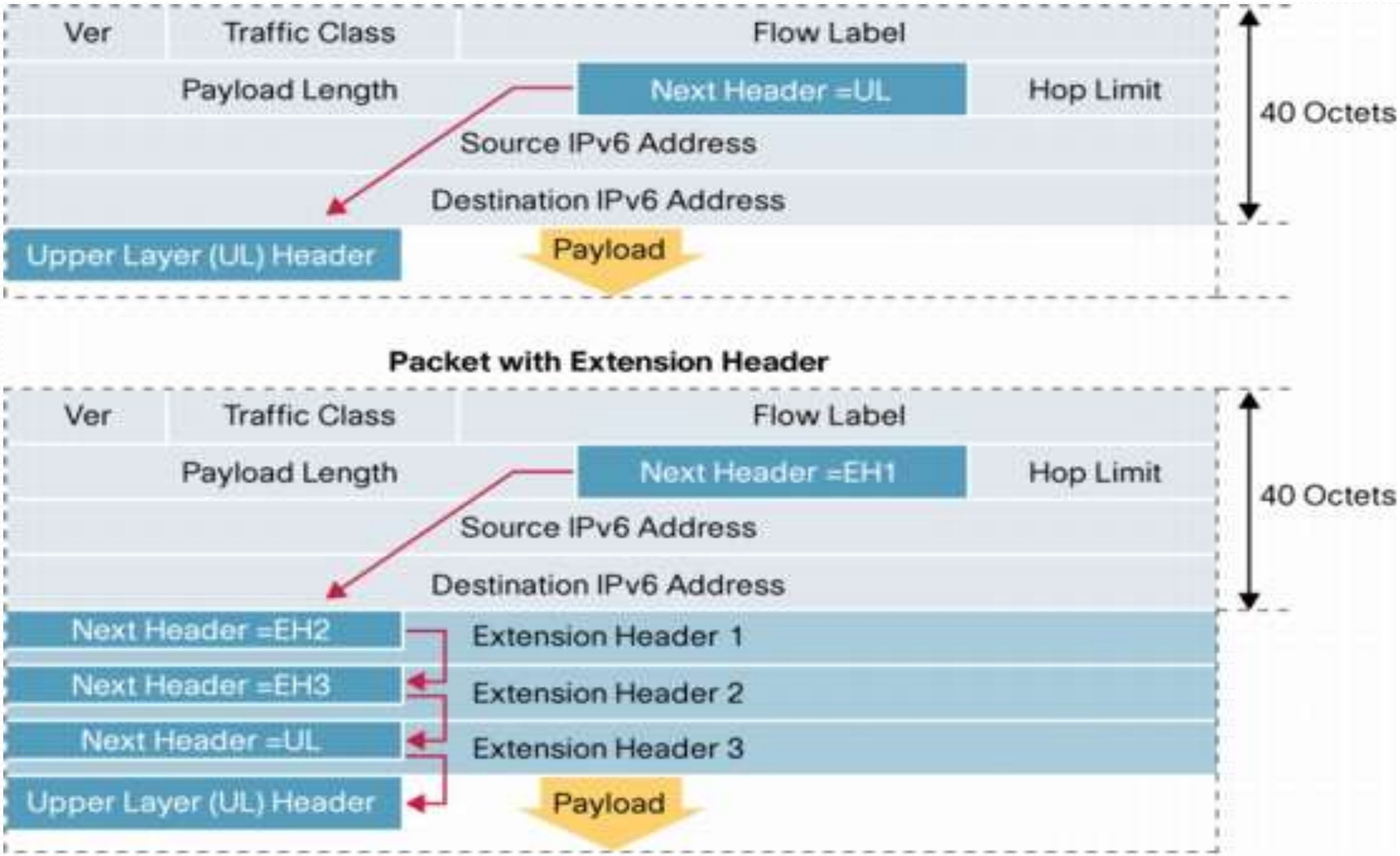
IPv6 Header:

- **Version (4-bit) :** value of 6,
- **Type of class (8-bit):** type of class/service, equivalent to TOS in IPv4,
- **Flow Label (20-bit):** (@source, FL) Identifies data flows. Assigned by the origin: should be a random number between 00000 and FFFFF (00000: means that FlowLabel is not used),
- **Payload Length (16-bit):** size of the data (Extension Headers + Payload),
- **Next Header (NH field, 8-bit):** next header (embbeded headers) in the IPv6 header,
- **Hop Limit (8-bit):** decremented each time a packet is forwarded,

- **IPv6 header options:**

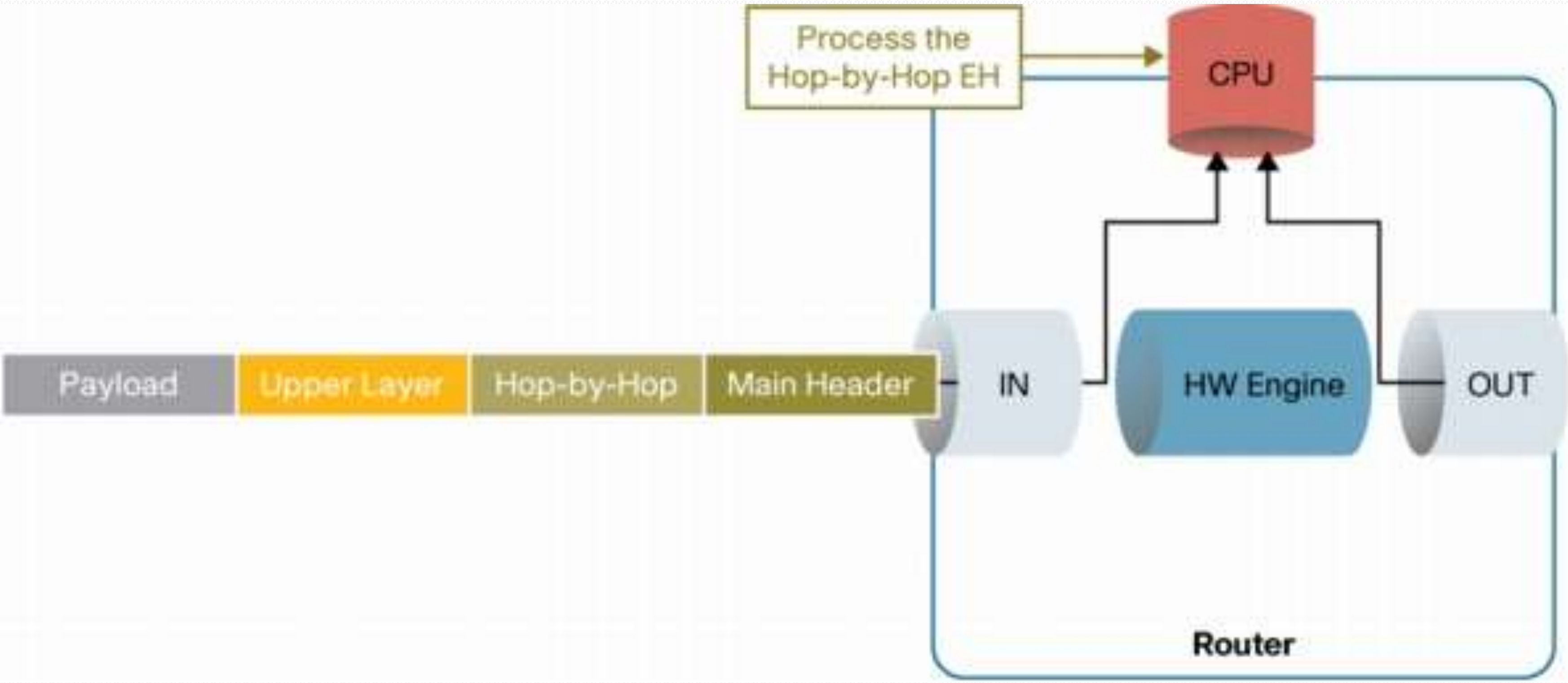
Extension Header	NH	Description
Hop-by-Hop Options	0	Options that need to be examined by all devices on the path (Router Alert, RSVP).
Destination Options (before routing header)	60	Options that need to be examined only by the destination of the packet at L3 (used with MIPv6).
Routing	43	Methods to specify the route for a datagram (used with MIPv6).
Fragment	44	Contains parameters for fragmentation of datagrams.
Authentication Header (AH)	51	Contains information used to verify the authenticity of most parts of the packet.
Encapsulating Security Payload (ESP)	50	Carries encrypted data for secure communication.
Destination Options (before upper-layer header)	60	Options that need to be examined only by the destination of the packet (at L5)
Mobility (currently without upper-layer header)	135	Parameters used with MIPv6.

- **IPv6 header options:**



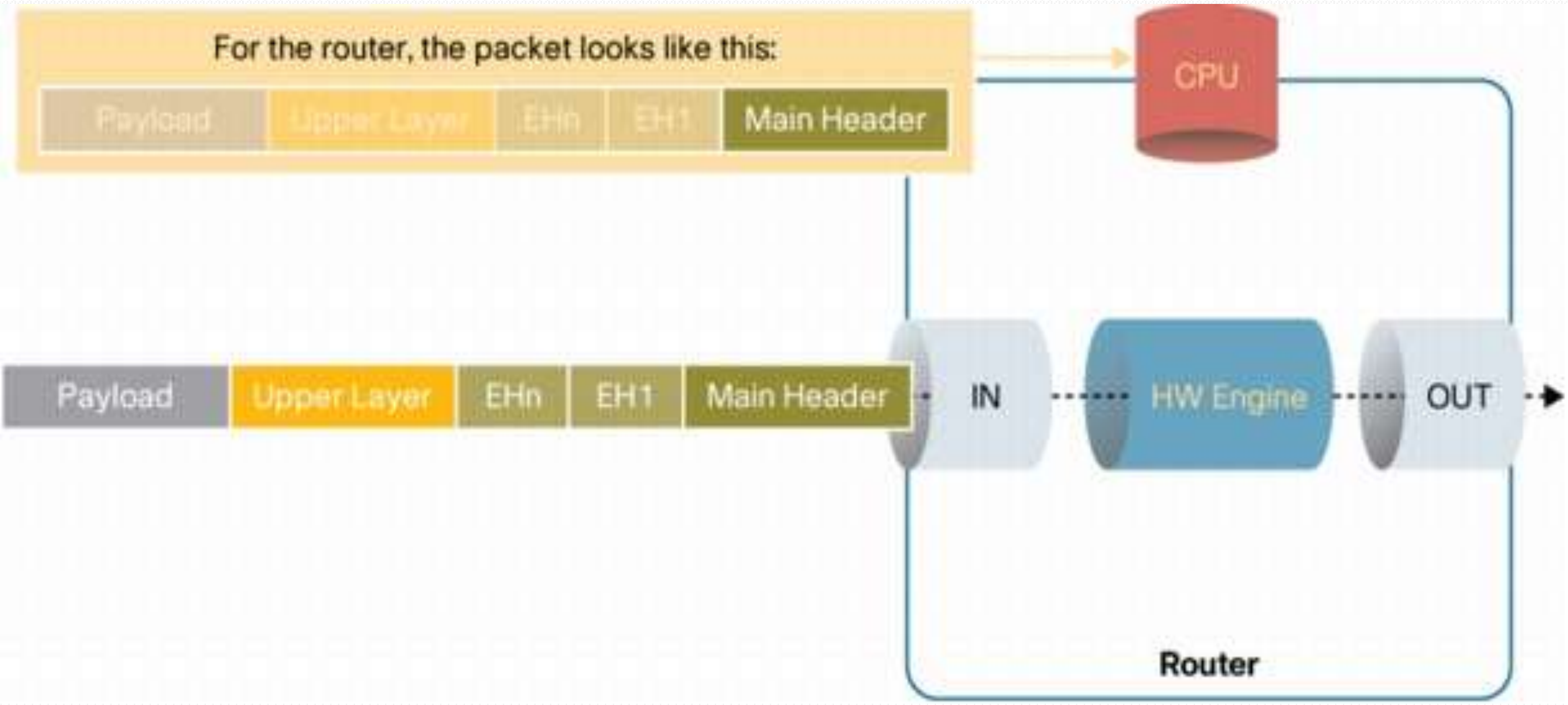
- **IPv6 header options:**

Hop-by-hop headers SHOULD be examined by each router → packets can not be forwarded by HW



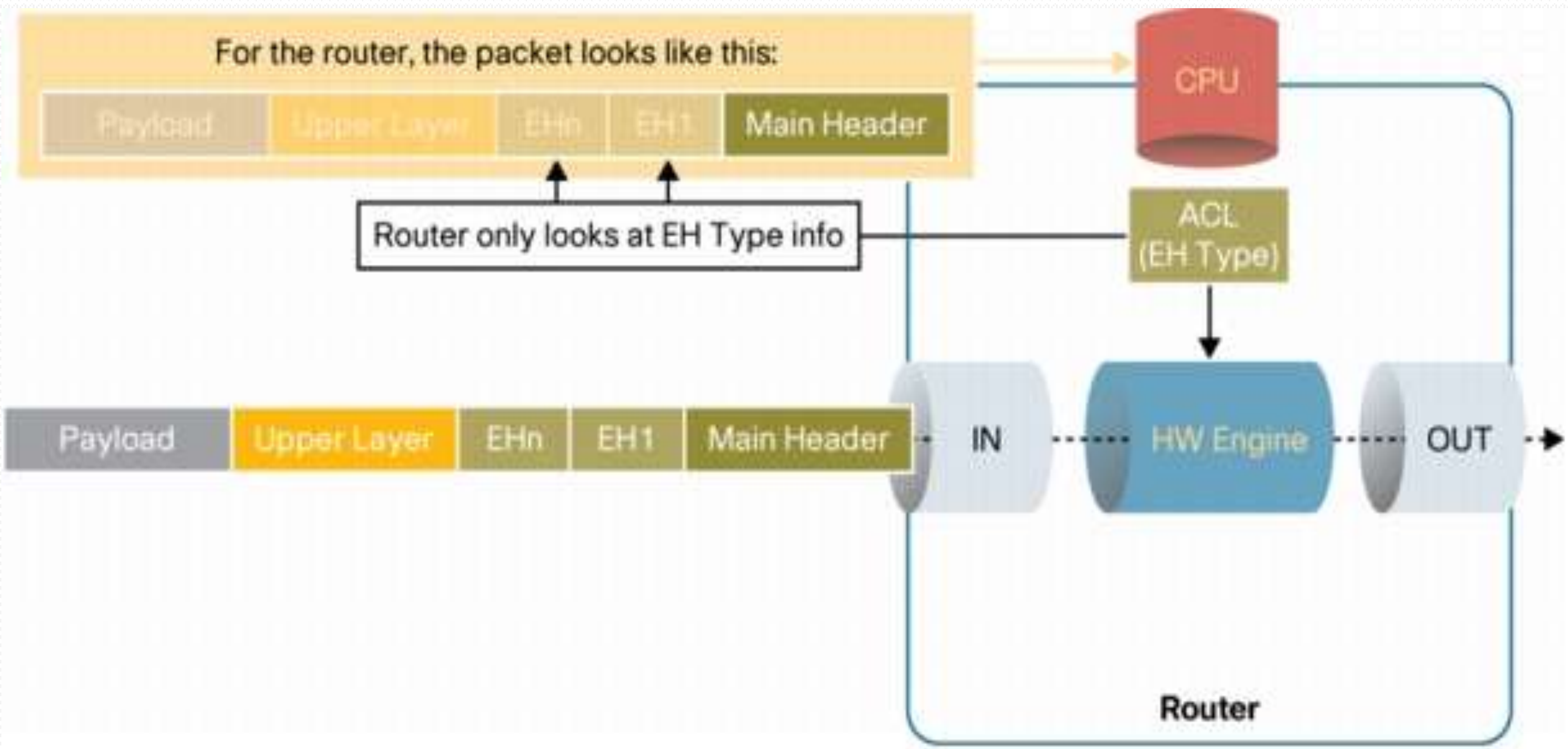
- ## IPv6 header options:

The other option headers ARE NOT REQUIRED to be examined by each router
→ packets can be forwarded by HW



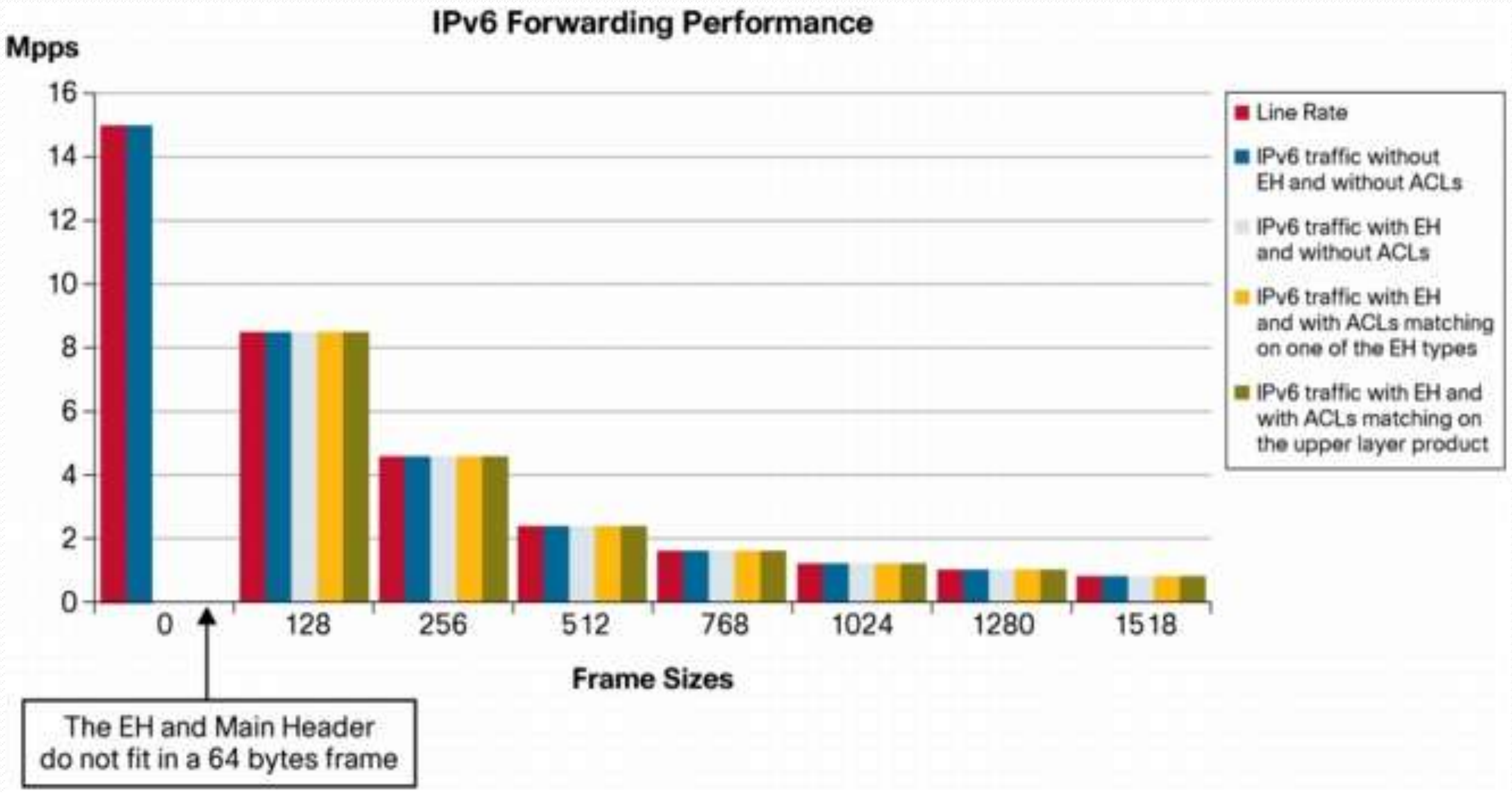
IPv6 header options:

EH headers can be filtered via ACL's → useful if you want to avoid for example source routing, in this case, since the router only takes forwarding decisions, the packet can be handle via SF or HW.



- **IPv6 header options:**

Performance in a CISCO E5 10 Gigabit Ethernet Line Cards (**mpps** is millions of packets per second).



- **IPv6 Address Notation (128 bit addresses):**

Host: xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx/64

Routing-prefix
(máximum of 64-bit)

Host-Id
(always 64-bit)

Routing-prefix size

Host: 2001:708:310:52:202:2dff:fe4b:a120/64

Routing-prefix

Host-Id

Routing-prefix size

Net or prefix: 2001:708:310:52/64

Net or prefix: 2001::/48

- **IPv6 Address Notation (128 bit addresses):**

Host: **2001:708:310:52**:**202:2dff:fe4b:a120**/64

Routing-prefix Host-Id Routing-prefix size

Compressing the IPv6 address:

fedc:ba78:0000:0000:0001:0000:1212:1111 →

fedc:ba78::1:0:1212:1111

Examples:

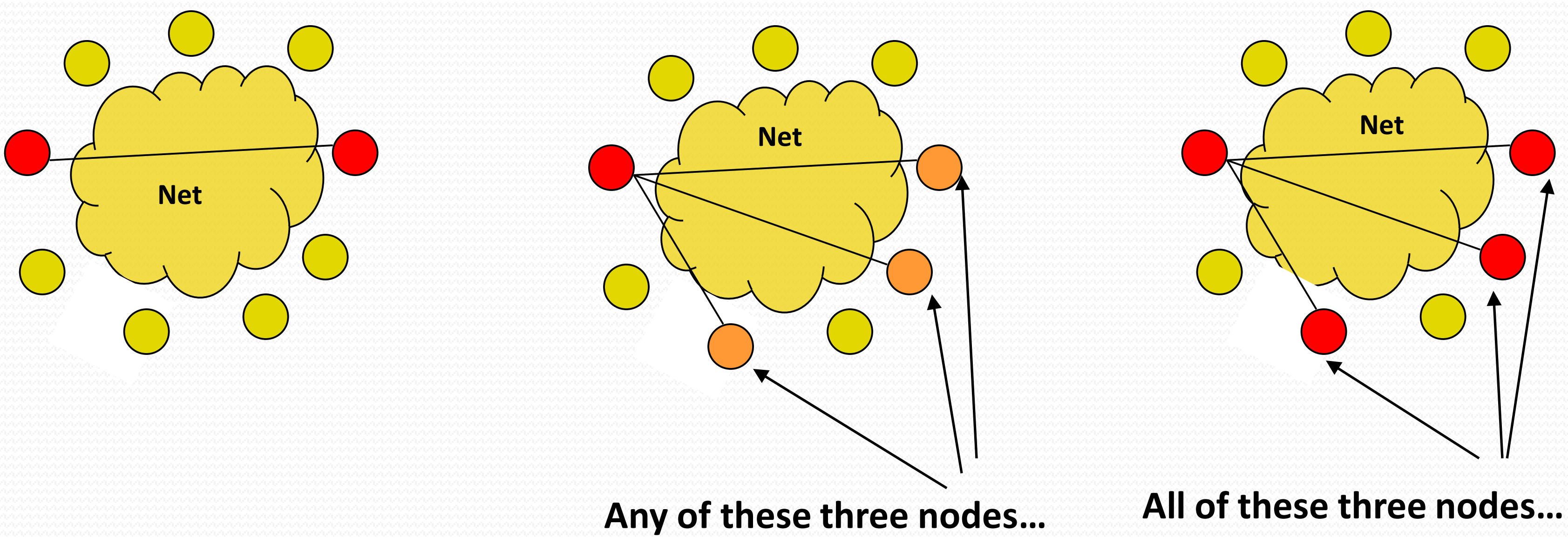
0000:0000:0000:0000:0000:0000:0000:0001 → **::1** **(loopback)**

ff02:0000.0000.0000.0000:0000:0000.0002 → **ff02::2**

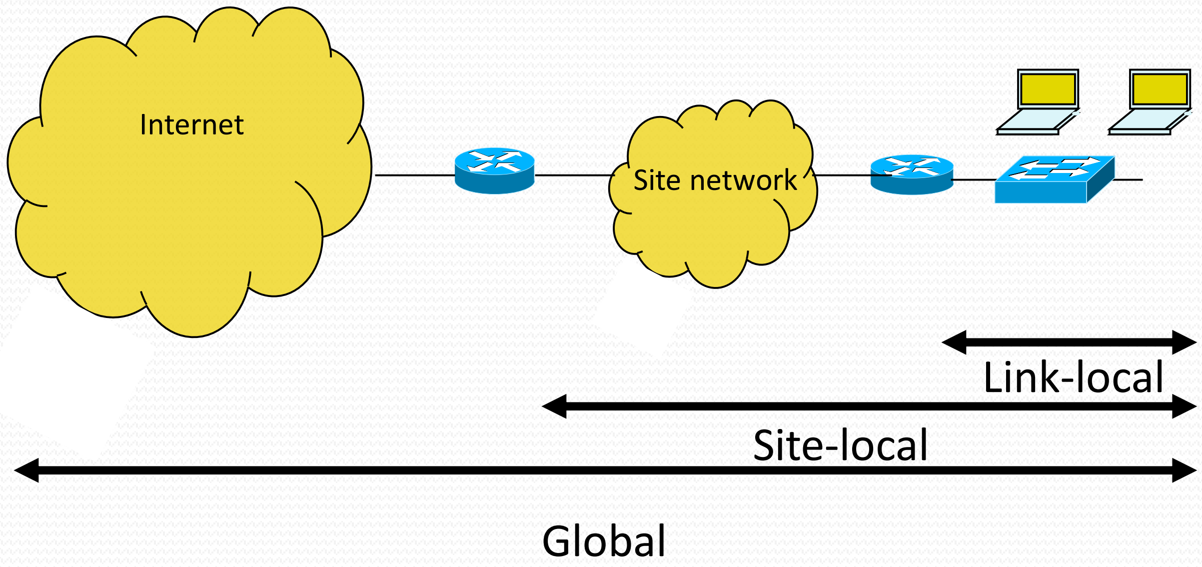
(multicast, all routers, link-local scope)

Types of IPv6 addresses:

- **Unicast:** only one destination
- **Anycast:** any destination among a group of destinations
- **Multicast:** several destinations (**includes broadcast**)

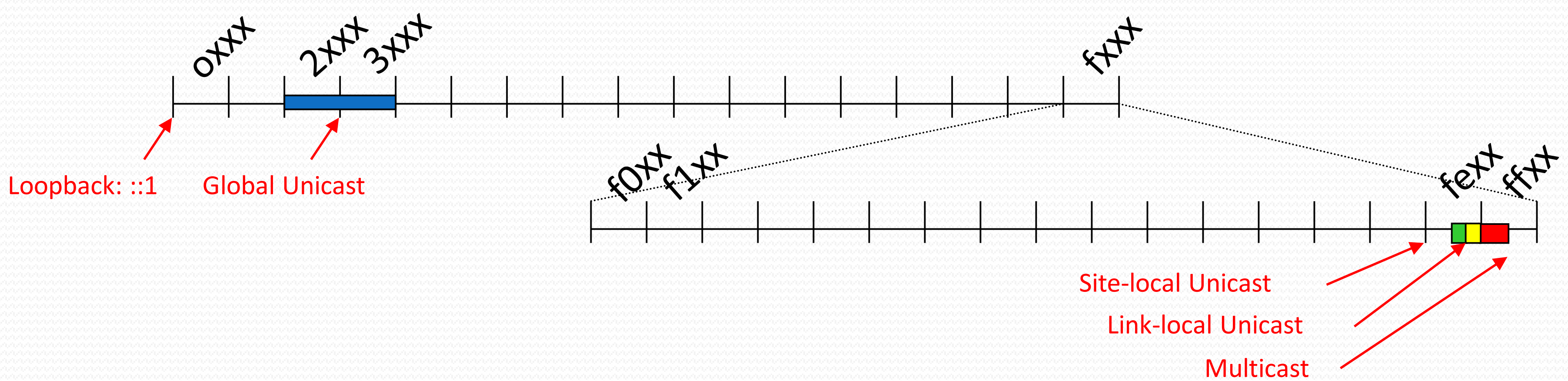


- **IPv6 addresses:**
  - Addresses have a scope:
    - Global (routed in the whole Internet)
    - Site Local (No routed outside the local network)
    - Link Local (No routed by routers, allows a Plug&Play with only communication inside a link)



Topic 1: Internet Architecture & Addressing.

	Unicast	Anycast	Multicast
Global	[2000:…/64, 3fff:…/64]	[2000:…/64, 3fff:…/64]	ff0e::/16 ff1e::/16
Site-local	fec0::/64	fec0::/64	ff05::/16 ff15::/16
Link-local	fe80::/64	fe80::/64	ff02::/16 ff12::/16



IANA IPv6 address assigments:

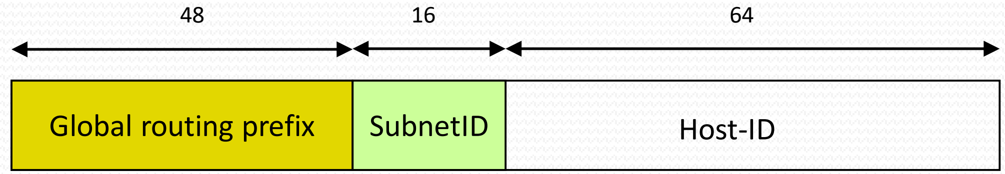
http://www.iana.org/assignments/ipv6-address-space/ipv6-address-space.xml

RIR IPv6 assignments:
http://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.txt

- **IPv6 addresses - Unicast addresses:**

Prefix + Interface ID

- Example: Global Unicast:



- Example : Link-local Unicast → fe80::/10
- Example : Site-local Unicast → fec0::/10
- Example : Loopback → ::1/128
- Example : No specified (no assignable) → ::0/128

- **IPv6 addresses:** Example of assignment rule for a **Global Unicast Aggregatable IPv6 address (RFC 2450)**
  - The **prefix** (64-bits) forms a tree of organizations

64-bit

001	TLA ((13+8)-bit)	NLA (24-bit)	SLA (16-bit)

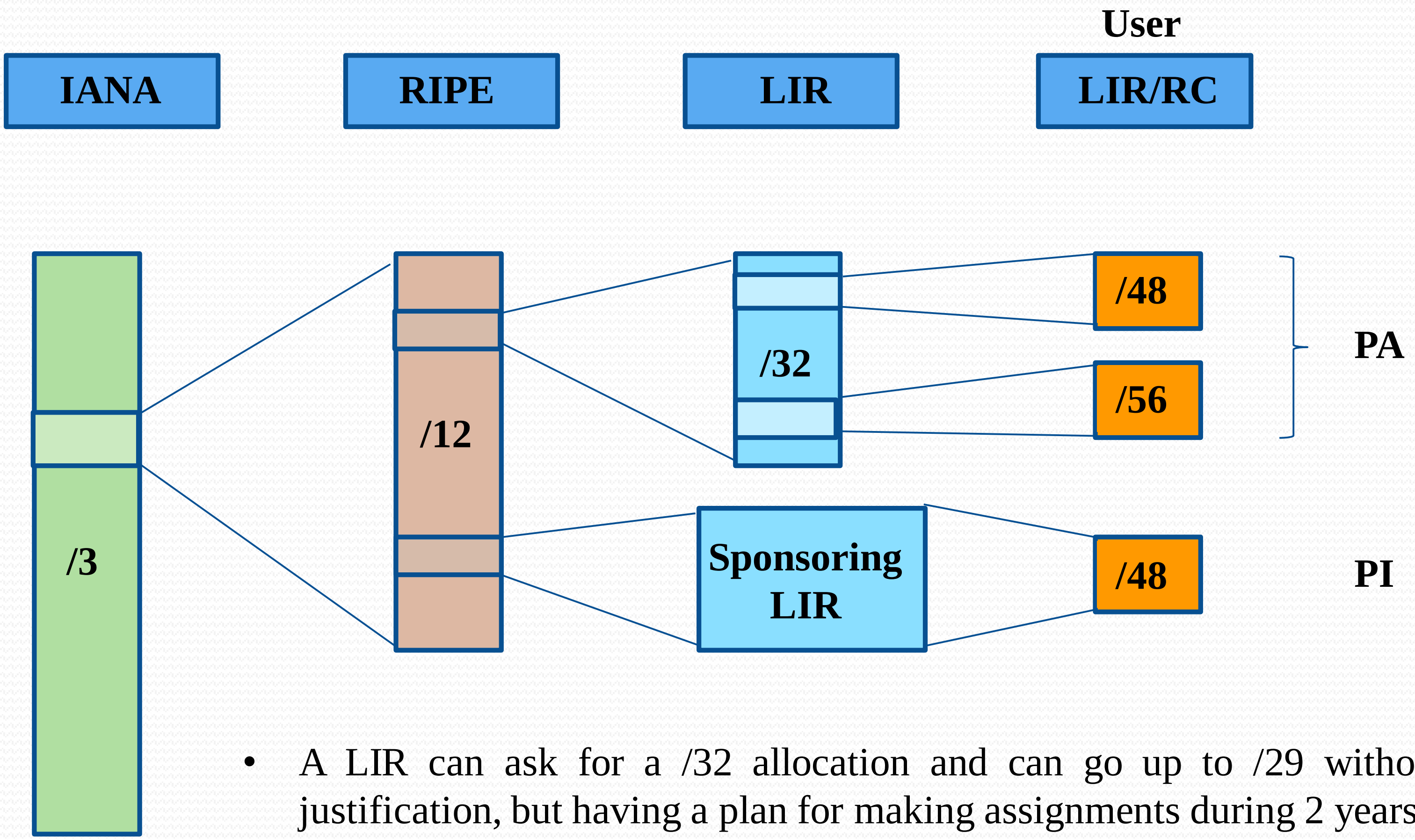
  - TLA: Top Level Aggregator. E.g., Tier-1 provider (Movistar),
  - NLA: Next-Level Aggregator. E.g., Tier-2 provider, (RedIRIS: UPC provider),
  - SLA: Site-Level Aggregator. E.g., Corporate Network (UPC).
- **IPv6 addresses:** Example of assignment rule for a **Global Unicast Aggregatable IPv6 address (RFC 3587)**

64-bit

Global routing prefix (n-bits)	Subnet ((64-n)-bits)

Topic 1: Internet Architecture & Addressing.

- **IPv6 addresses: RFC 3587 makes obsolete RFC 2450. Now the IPv6 address organization is left to RIR's:** RIPE assignment



- A LIR can ask for a /32 allocation and can go up to /29 without justification, but having a plan for making assignments during 2 years
- More tan /29 if justified (numbre of customers or size of infrastructure)

IPv6 addresses - Several ways of building an IPv6 address (InterfaceID)

Example: **Auto-configuration address in Ethernet**

Prefix/64 + EUI-64 (Extended Unique ID MAC address)

However MAC Address is 48 bits: **34:56:78:9a:bc:de** →

(**34** becomes **36**) → **3656:78ff:fe9a:bcde**

Bit Universal/Local (inverse of the 7th bit): 0011 0100 → 0011 0110

Constant: **fffe**

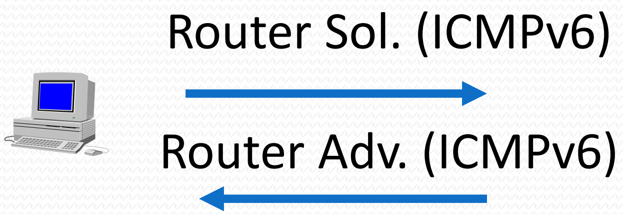
Imagine fe80::2 (all routers) and some NIC with MAC 00:00:00:00:00:02

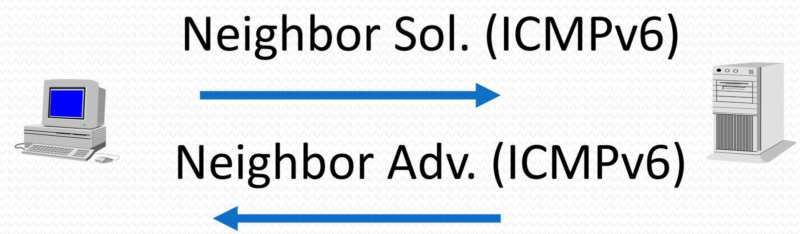
→ It would build a **fe80::2** instead of the **fe80::0200:0:0:2**

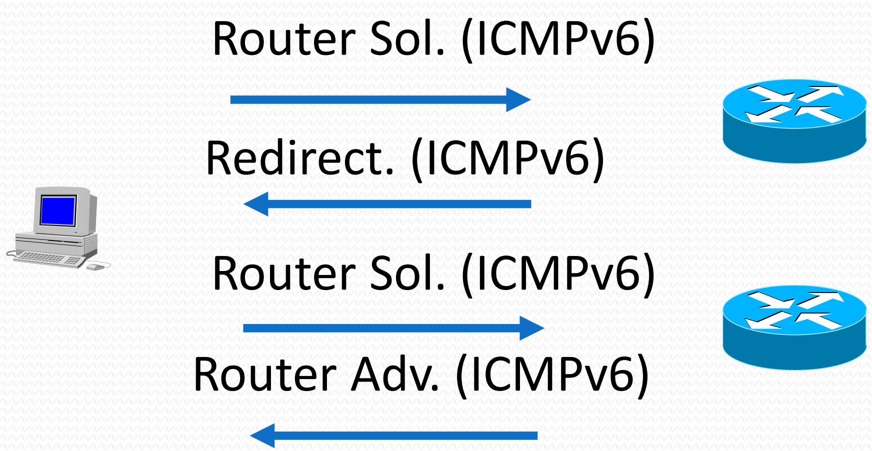
- Others
- Example: **IPv4 mapped IPv6 addresses**
- 0 (80 bits) : ffff (16 bits) : IPv4 (32 bits, in Hex)

- **Neighbor discovery protocol (NDP)** is a function – coded in ICMPv6 messages - that allows among other functionalities address resolution or address auto-configuration.
  - NDP defines 5 types of ICMPv6 packets:
    - **Router Solicitation (Type 133):** Hosts inquire with Router Solicitation messages to locate routers on an attached link
    - **Router Advertisement (Type 134):** Routers advertise their presence together with various link and Internet parameters either periodically, or in response to a Router Solicitation message.
    - **Neighbor Solicitation (Type 135):** Neighbor solicitations are used by nodes to determine the Link Layer address of a neighbor, or to verify that a neighbor is still reachable via a cached Link Layer address.
    - **Neighbor Advertisement (Type 136):** Neighbor advertisements are used by nodes to respond to a Neighbor Solicitation message.
    - **Redirect (Type 137):** Routers may inform hosts of a better first hop router for a destination.

Topic 1: Internet Architecture & Addressing.





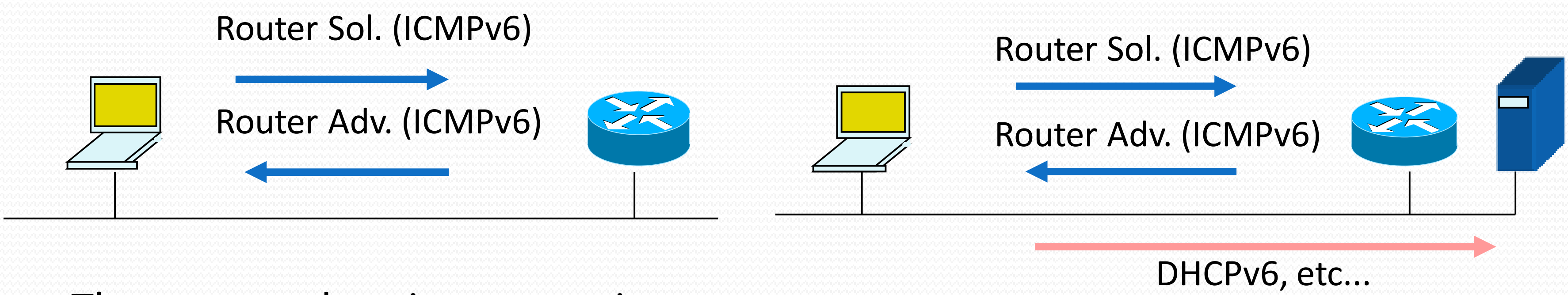


Topic 1: Internet Architecture & Addressing.

- These messages are used **to provide the following functionalities**:
  - **Router discovery:** hosts can locate routers residing on attached links.
  - **Prefix discovery:** hosts can discover address prefixes that are on-link for attached links.
  - **Parameter discovery:** hosts can find link parameters (e.g., MTU).
  - **Address auto-configuration:** stateless configuration of addresses of network interfaces.
  - **Address resolution:** mapping between IP addresses and link-layer addresses.
  - **Next-hop determination:** hosts can find next-hop routers for a destination.
  - **Neighbor unreachability detection (NUD):** determine that a neighbor is no longer reachable on the link.
  - **Duplicate address detection (DAD):** nodes can check whether an address is already in use.
  - **Recursive DNS Server (RDNSS) and DNS Search List (DNSSL) assignment** via a router advertisement (RA) options. This is a new feature and not widely supported by clients.
  - **Packet redirection** to provide a better next-hop route for certain destinations.

Topic 1: Internet Architecture & Addressing.

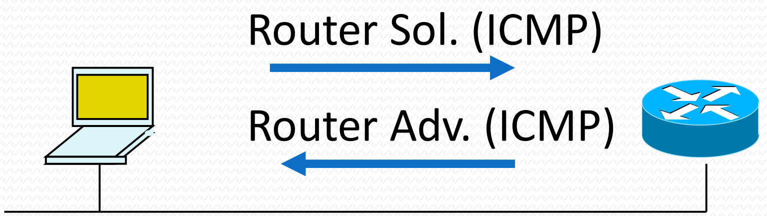
- # IPv6 address configuration:
  - ## RS (Router Solicitation)- RA (Router Advertisement)
    - **A router can assign a global address, indicating a prefix to which the host adds its Host-ID** (Prefix/64 + Host-ID) **or can indicate to use a DHCPv6 server.**



- The router advertisements gives:
  - Router's address (default Gateway),
  - Zero or more prefix addresses,
  - Slaac allowed (yes/not)
  - DHCPv6 availability → if yes, the host can use DHCPv6 to get addresses and DNS server addresses
  - DNS server addresses
  - MTU size (optional)

- **IPv6 address configuration: RS (Router Solicitation)- RA (Router Advertisement)**
  - A host can get several following possible ways for configuring an IPv6 address,
  - **Flag 'M'** (0 indicates stateless, 1 stateful), while the **flag 'O'** (1 indicates to obtain other configuration information such as DNS server addresses from a server). Default values: 'O'= 0, and 'M'= 0.
    - **RA + Only SLAAC** (StateLess Address Auto-Configuration): the router provides prefix (flags 'M' = 'O' = 0), and additional information is configured manually,
    - **RA + SLAAC + 'O' flag** (StateLess Address Auto-Configuration, DHCPv6 provides the additional information), (flag 'M'= 0 and flag 'O' = 1),
    - **RA + 'M' flag + 'O' flag** (DHCP managed, Statefull address configuration), (flags 'M' = 'O' = 1),
    - **RA + SLAAC + flag 'M'= 1 and flag 'O'= 0** (DHCPv6 provides the IPv6 prefix, nobody provides other info, e.g. DNS server), → NOT USED

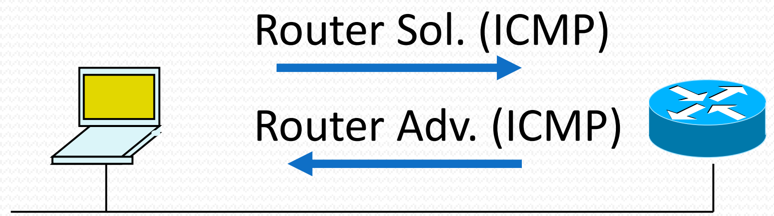
- **IPv6 address configuration: SLAAC (stateless address auto-configuration). ICMPv6 router solicitation**



NDP: Router Solicitation (ICMPv6)

IPv6	Hops = 255 @IPsource = **fe80::3656:78ff:fe9a:bcde** (link-local host address) @IPdst = **ff02::2** (multicast link-local = ALL ROUTERS)
ICMPv6	Type, Code = **133,0** (Router Solicitation ICMPv6 type/code)
ICMP options	SLAAC + ‘O’ option

- **IPv6 address configuration: SLAAC (stateless address auto-configuration). ICMPv6 router solicitation**



NDP: Router Advertisement (ICMPv6)

IPv6

- Hops = 255
- @IPsource = **fe80::3656:78ff:fe9a:aaaa** (link-local router address)
- @IPdst = **ff02::1** (multicast link-local = ALL HOSTS)

ICMPv6

- Type, Code = **134,0** (Router Advertisement ICMPv6 type/code)

ICMP options

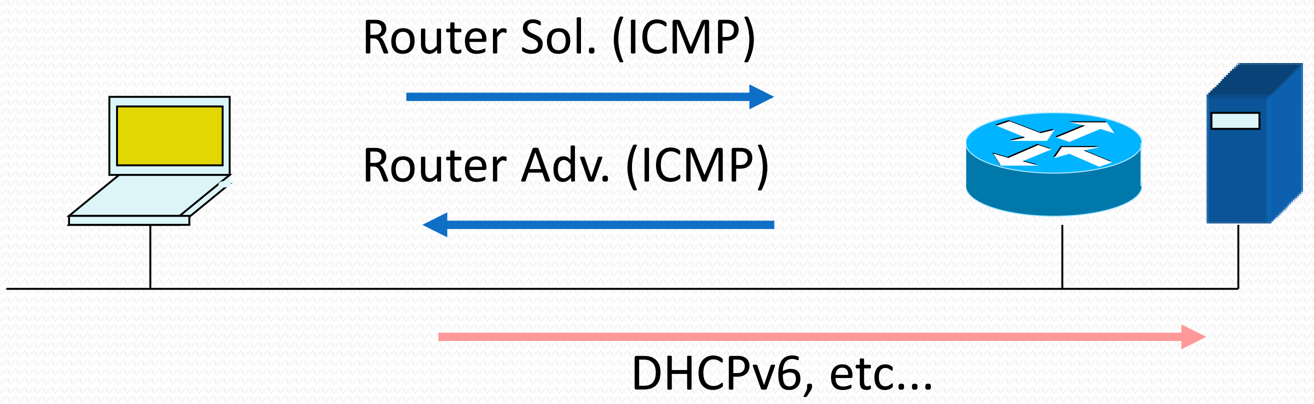
- SLAAC + 'O' option

The router gives back a **Router Advertisement** with a prefix and DHCPv6 gives the options (e.g., DNS address).

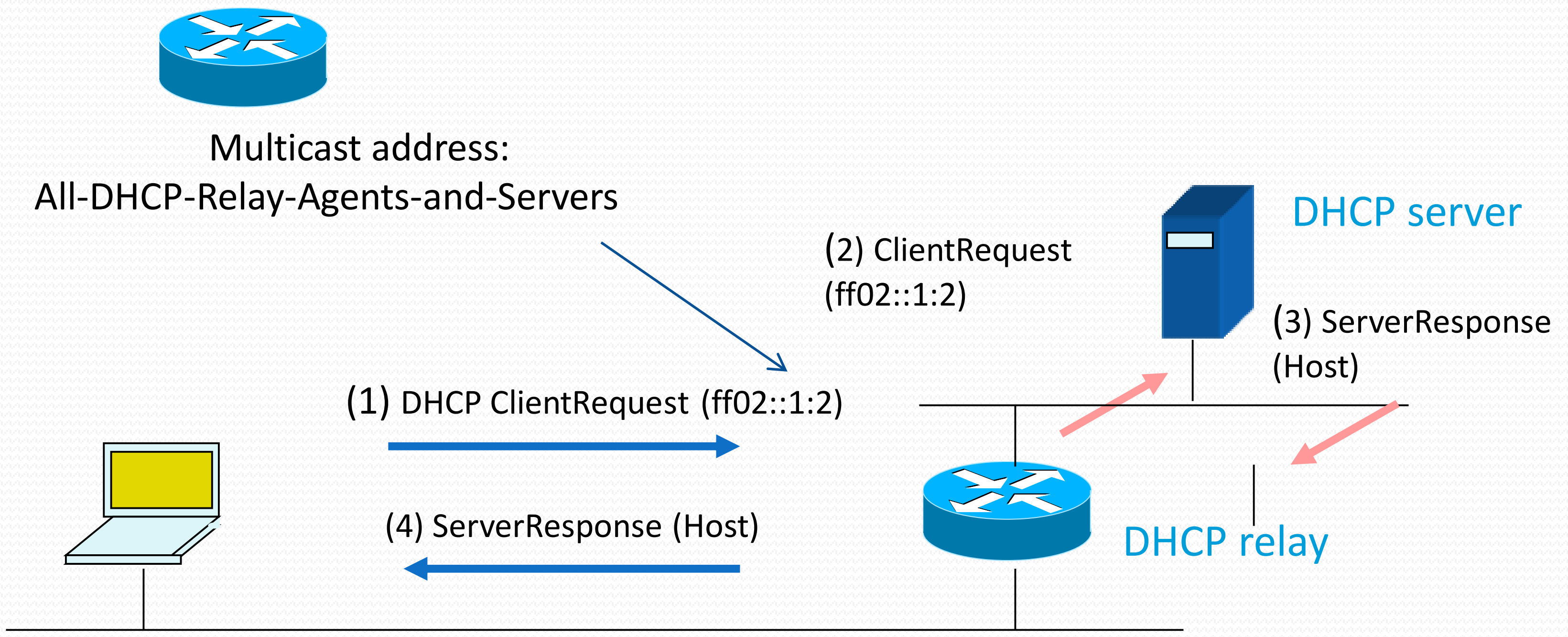
IPv6 address configuration:

IPv6 Stateful Address configuration

- **Stateless address configuration (RA + ’O‘= 1 + ’M’= 1)** requires an adequate configuration in all routers. Thus, the best strategy is to configure a **DHCPv6 server** (Stateful address configuration) that centralizes sending the required information at the cost of reliability (centralized solution), however reducing complexity and adding flexibility.
- RA (router advertisement) tells the host that a DHCPv6 is available for getting and IPv6 address. The RA can force to disable SLAAC and force DHCPv6,
- DHCPv6 servers can jointly work with DNS servers



- **IPv6 address configuration:**
  - **IPv6 Stateful Address configuration**



IPv6 Address configuration: DAD (Duplicate Address Detection/Discovery)

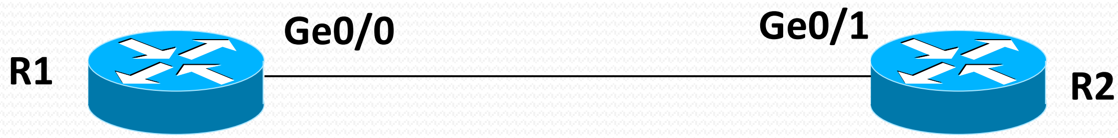
- Since an IPv6 host can have many IPv6 addresses

  (ex: link local + unicast + several multicast)

- Send a **Neighbor Solicitation** message with **source address ::** (null address) and **destination the tentative address (address to be checked)**.
- If somebody has the destination address will answer with a **Neighbor Advertisement** to address **ff02::1** (all nodes on the local network segment)

after auto-configuration processes the host perform **Duplicate Address Discovery** in order to check whether there are several duplicated IPv6 addresses → mechanism similar to Gratuitous ARP

Topic 1: Internet Architecture & Addressing.



R1 acts as a Router giving RA + only SLAAC service (RA + 'M'=0 + 'O'=0).

R2 acts as host receiving an IPv6@.

!!!!! enable IPv6 routing on a Cisco router using

R1(config)# ipv6 unicast-routing

!!!!! Select interface Ge0/0

R1(config)# int Ge0/0

!!!!! Give IPv6 address prefix to interface. If you omit omit the eui-64 parameter, you will need to configure the entire address manually. After you enter this command, the link local address will be automatically derived. It is a RA + only SLAAC configuration (no DNS).

R1(config-if)# ipv6 address 2001:0BB9:AABB:1234::/64 eui-64

R1(config-if)# no shutdown

!!!!! Configure R2. Select interface Ge0/0

R2(config)# int Ge0/1

R2(config-if)# ipv6 enable

R2(config-if)# ipv6 address autoconfig

R1# show ipv6 interface Ge0/0

GigabitEthernet0/0 is up, line protocol is up

IPv6 is enabled, **link-local address** is FE80::201:42FF:FE65:3E01

Global unicast address(es):

2001:BB9:AABB:1234:201:42FF:FE65:3E01,

subnet is 2001:BB9:AABB:1234::/64 [EUI]

R2# show ipv6 interface Ge0/1

GigabitEthernet0/0 is up, line protocol is up

IPv6 is enabled, **link-local address** is FE80::A8BB:CCFF:FE00:800

Global unicast address(es):

2001:BB9:AABB:1234:A8BB:CCFF:FE00:800

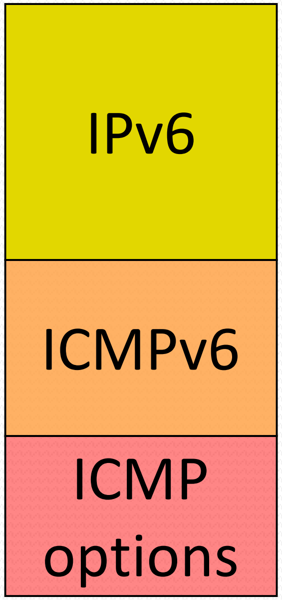
subnet is 2001:BB9:AABB:1234::/64 [EUI]

IPv6 address resolution:

- The equivalence of **ARP** is done via ICMP Neighbor solicitation/ Neighbor advertisement messages (remember that ARP is done at L2 while Neighbor solicitation is done at L3)

- The destination IP address of the ICMP message that carries the Neighbor Solicitation is the "**multicast solicited-node @**": it is formed using prefix **FF02::1:FF00:/104 + the last 24 bits of the IP address** that we are looking for, that means **FF02::1:FFxx:xxxx**
  - The use of this multicast address saves multicast addresses to the node, since multiple prefixes can be attached to a link, thus this multicast can be used for representing any of those prefixes,
- Periodically the host sends "Neighbor Solicitation" messages to check whether the hosts are reachable

IPv6 address resolution

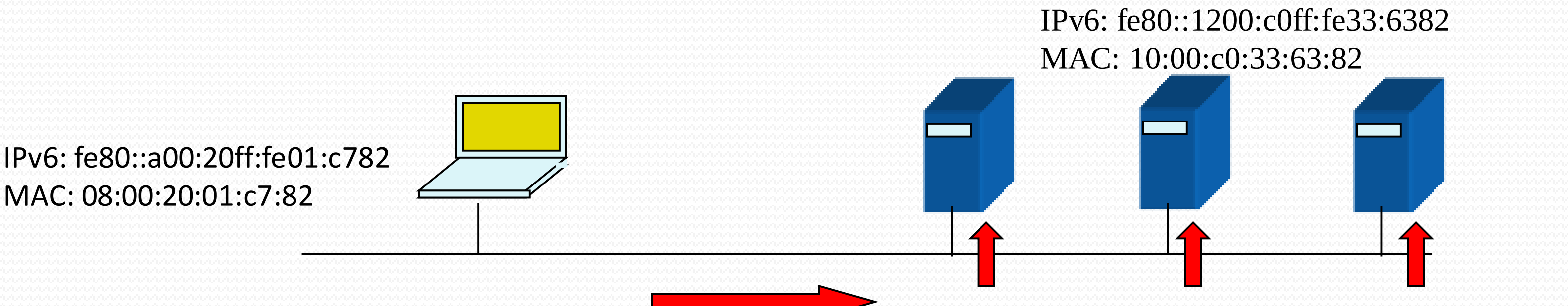


Multicast,
Link-local

Hops = 255
@IPsource= fe80::a00:20ff:fe01:c782 (link local)
@IPdst= **ff02::1:ff33:6382**

Type, Code= 135,0 (Neighbor Solicitation)
Target Address= fe80::1200:c0ff:fe33:6382

Opt Code= 1
Sender's Link Address= 08:00:20:01:c7:82



IPv6-dest: ff02::1:ff33:6382
Who has L2 address that mapps to IPv6 address fe80::1200:c0ff:fe**33:6382** ?

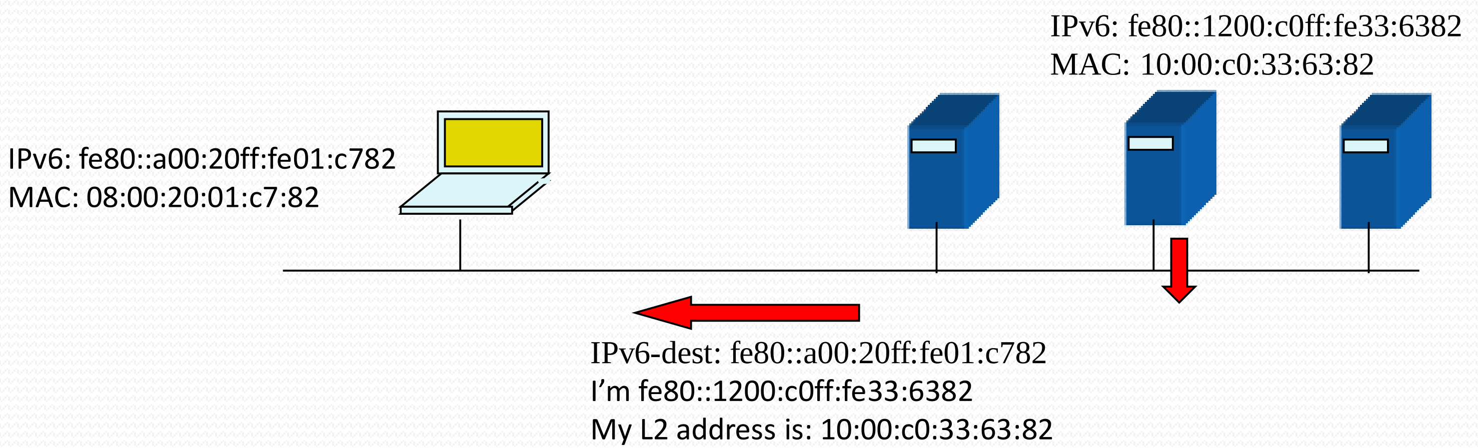
IPv6 address resolution

IPv6
ICMPv6
ICMP options

unicast

Hops = 255
@IPsource= fe80::1200:c0ff:fe33:6382 (link-local)
@IPdst= fe80::a00:20ff:fe01:c782

Type, Code= 136,0 (Neighbor Advertisement)
R= 0 (I'm not a router), S= 1 (I'm answering a question)
O= Override cached link layer @
Target Address= fe80::1200:c0ff:fe33:6382
Opt. Code= 2
Target's Link Address= 10:00:c0:33:63:82



- **Homework**
  - Look information on Content Delivery Networks (CDNs), P2P, farm servers architectures
  - Look information on SLA
  - Look at RIPE webpage (RIR governance, IPv4 exhaustion, IPv6 policies, AS maps, LIRs, etc)
  - Look at CAIDA web page (customer cone, peering, AS rank, AS-level Internet graph map. etc)
  - Look information at Exchange Points Espanix, Catnix and Euro-NIX webpages (members, policies, peering matrices, traffic)
  - Look information on Neighbour Discovery protocol and functionalities such as ARP in IPv6 and IPv6 address configuration