Internet routing

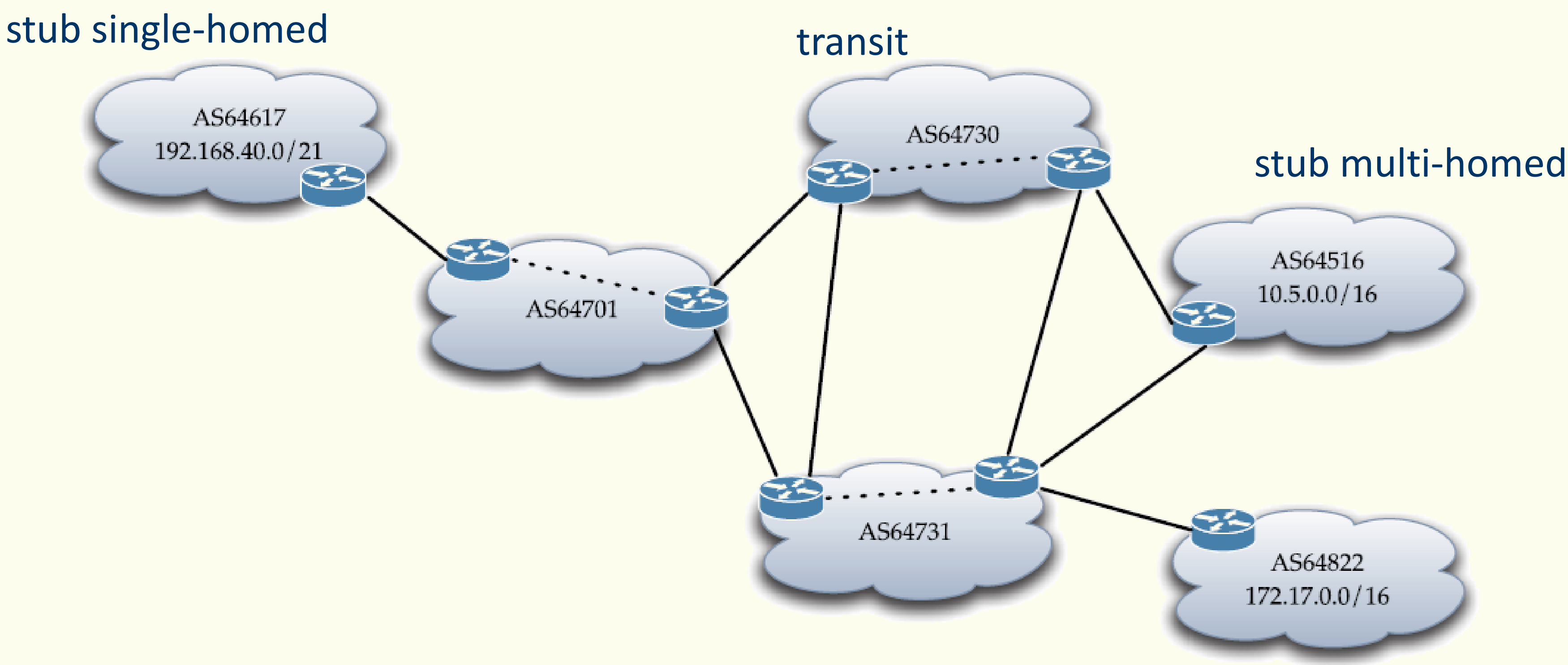
Enzo Mingozzi
Professor @ University of Pisa
enzo.mingozzi@unipi.it

Origins

- Exterior Gateway Protocol (EGP)
  - Adopted in the first global IP-based network (NSFNET, 1984)
  - **Strictly hierarchical**: two levels, one global core
  - **No direct connections** allowed between non-backbone ASes
- EGP's strict requirements demonstrated not to be applicable in practice
  - Other factors drive connectivity and routing decisions of an autonomously administrated network
- BGP development followed
  - No constraints on the hierarchy
  - Policy-based routing
  - Version 4 finally released to account for CIDR

Global routing

- Full **default-free** routing tables at **transit** ASes
- Not needed at **stub** ASes, though implemented

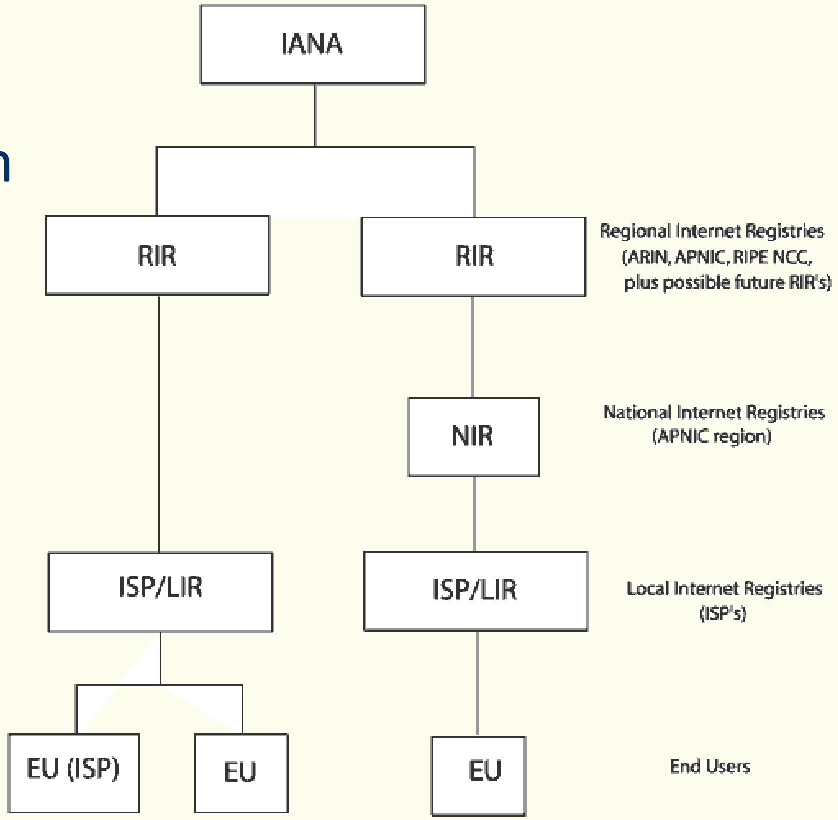


IP address allocation

- Internet Corporation for Assigned Names and Numbers (**ICANN**)

An **Internet Registry** is an organisation that is responsible for distributing IP address space to its members or customers and for registering those distributions

Each registry has its own rules and pricing for IP address allocation and AS number assignment



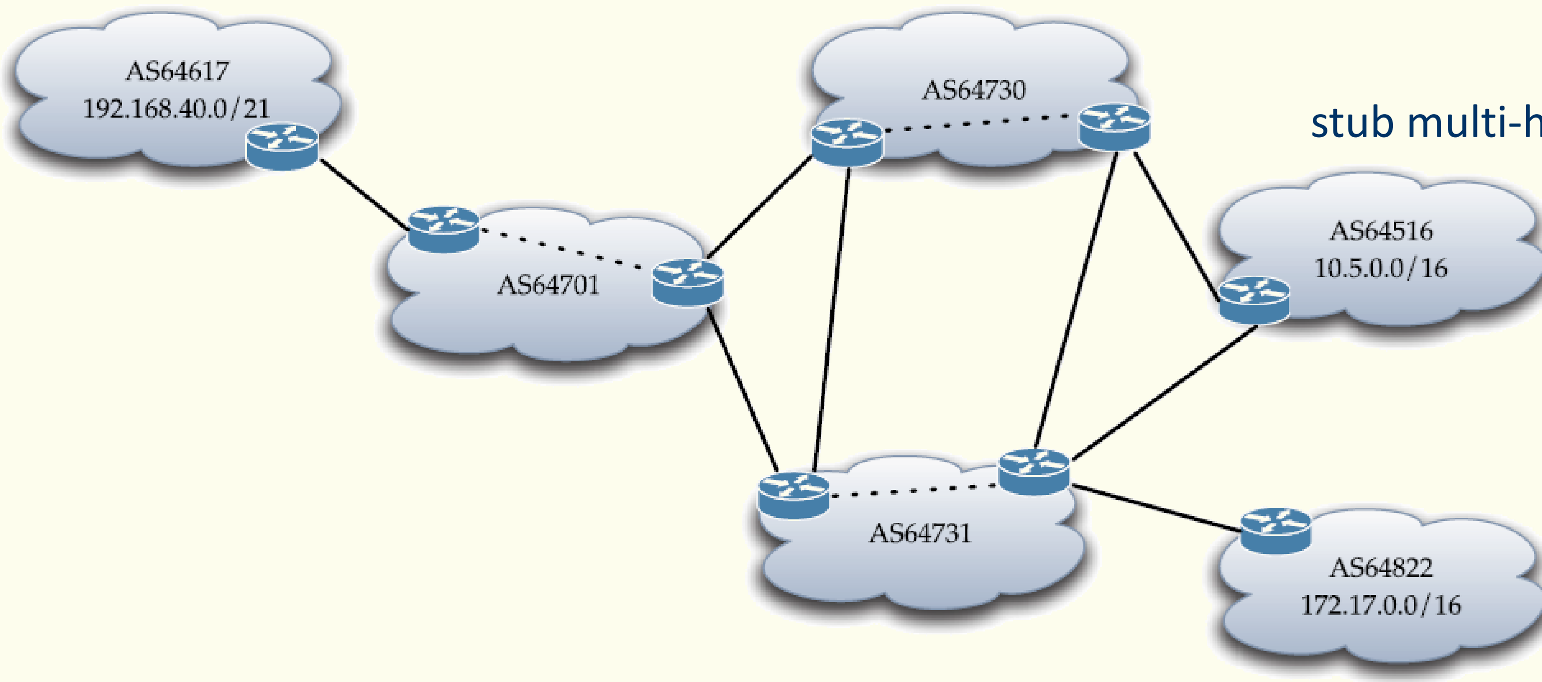
https://www.ripe.net/publications/docs/ripe-policies

IP address allocation

Case 1. An organization gets both an IP address block and AS number
- Establish multi-home connectivity
- Establish eBGP sessions with upstream ISPs

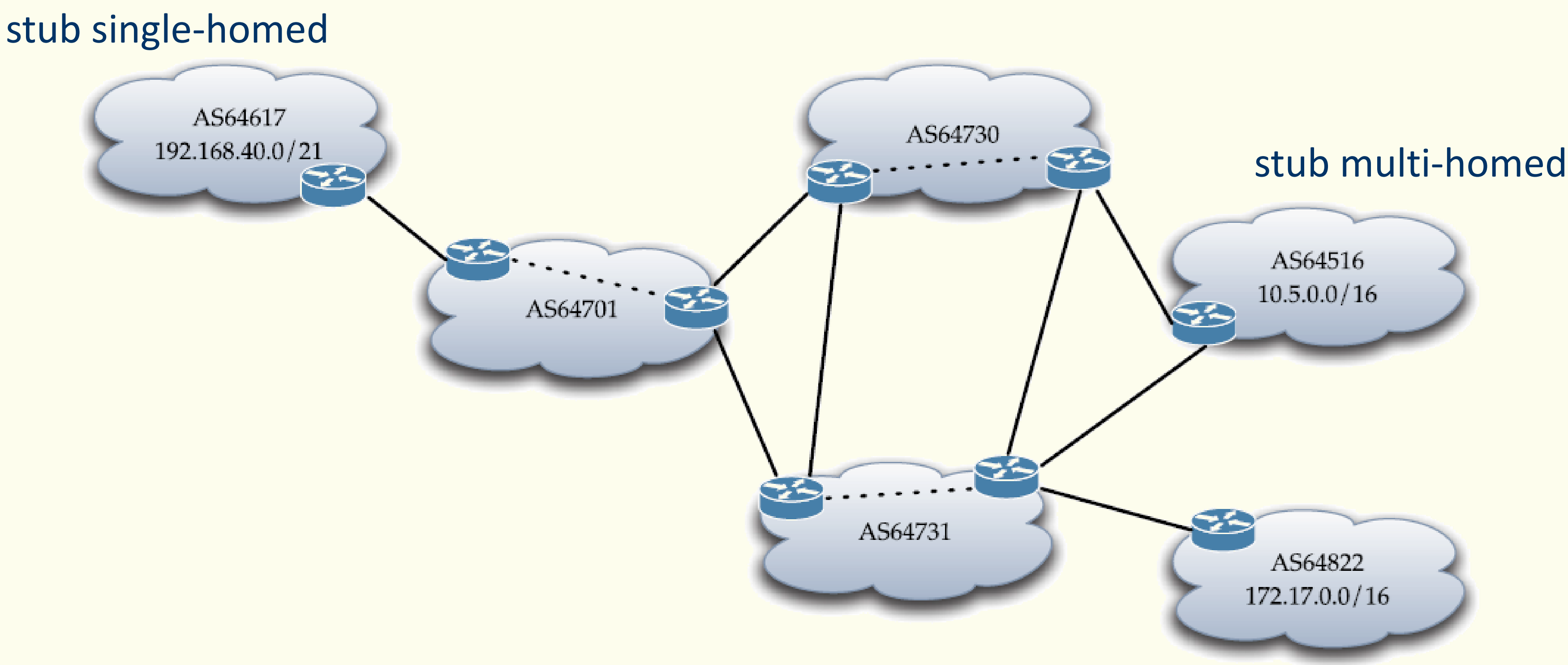
stub single-homed

stub multi-homed



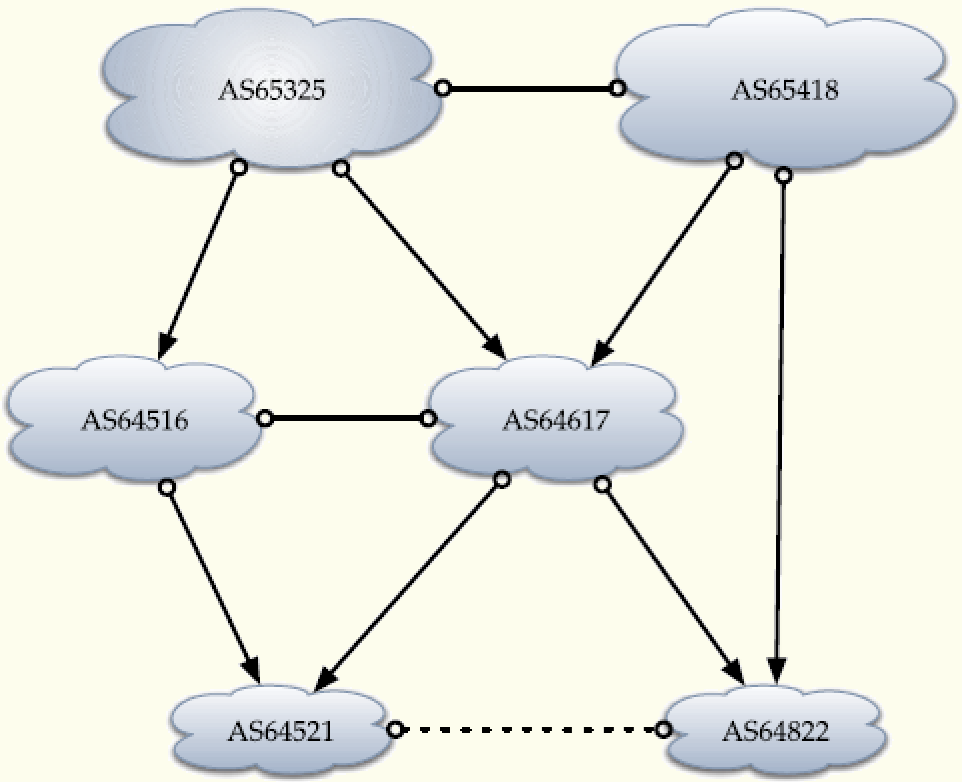
IP address allocation

Case 2. An organization gets an IP address block but not an AS number
- Establish an agreement with an ISP to serve as 'home' AS, then either
  a) setup private AS numbering + BGP peering
  b) exploit OSPF in the upstream ISP + route redistribution



ISP commercial agreements

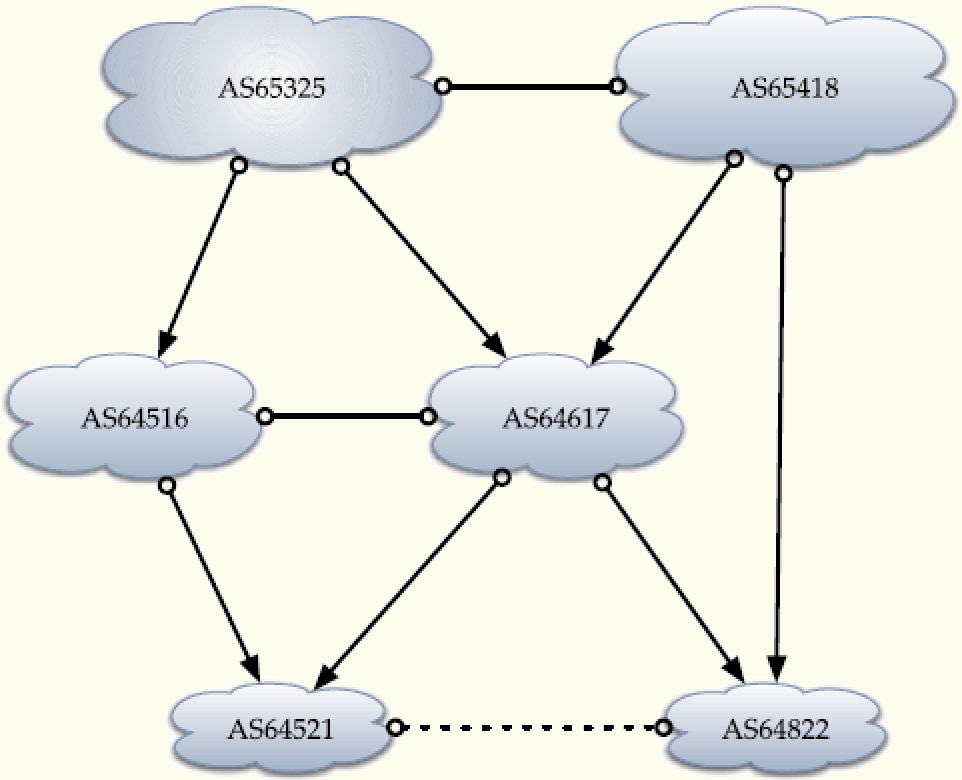
Business relationships between ASes arise from **contracts** that define the **pricing model** and the **exchange of traffic between ASes**



- **Multilateral agreement**: Several ISPs build/use shared facilities and share cost (e.g., public or private exchange points)
- **Bilateral agreement**: Two providers agree to exchange traffic if traffic is almost symmetric, or agree on a price, taking into account the imbalance in traffic swapped (e.g., in a private peering)
- **Unilateral agreement for transit**: A customer pays its provider an “access” charge for carrying traffic (e.g., a tier 4 ISP would pay a charge to the tier 3 ISP)
- **Sender Keeps All (SKA)**: ISPs do not track or charge for traffic exchange. This is possible in private peering and in some public peering. Usually true for tier 1 ISPs

ISP commercial agreements

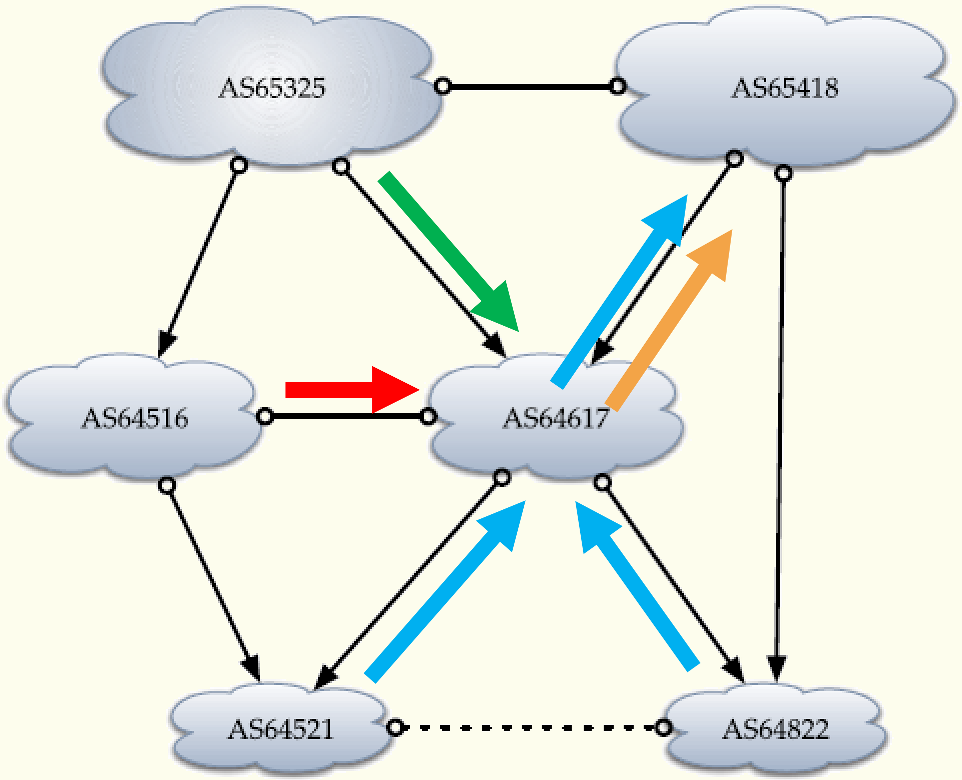
Business relationships between ASes are often schematically categorized in the literature into the following basic types



- <u>**Customer-to-Provider**</u>: A ***customer*** AS pays its *upstream* AS ***provider*** to be connected to the rest of the Internet
- <u>**Peer-to-Peer**</u>*:* ***peer*** ASes exchange traffic, typically without any payment involved
- <u>**Sibling-to-Sibling**</u>: sibling ASes act like a provider to each other

Customer-to-Provider

A **customer** AS buys **transit access** to the **provider** AS to reach all networks it cannot reach otherwise



Export policies

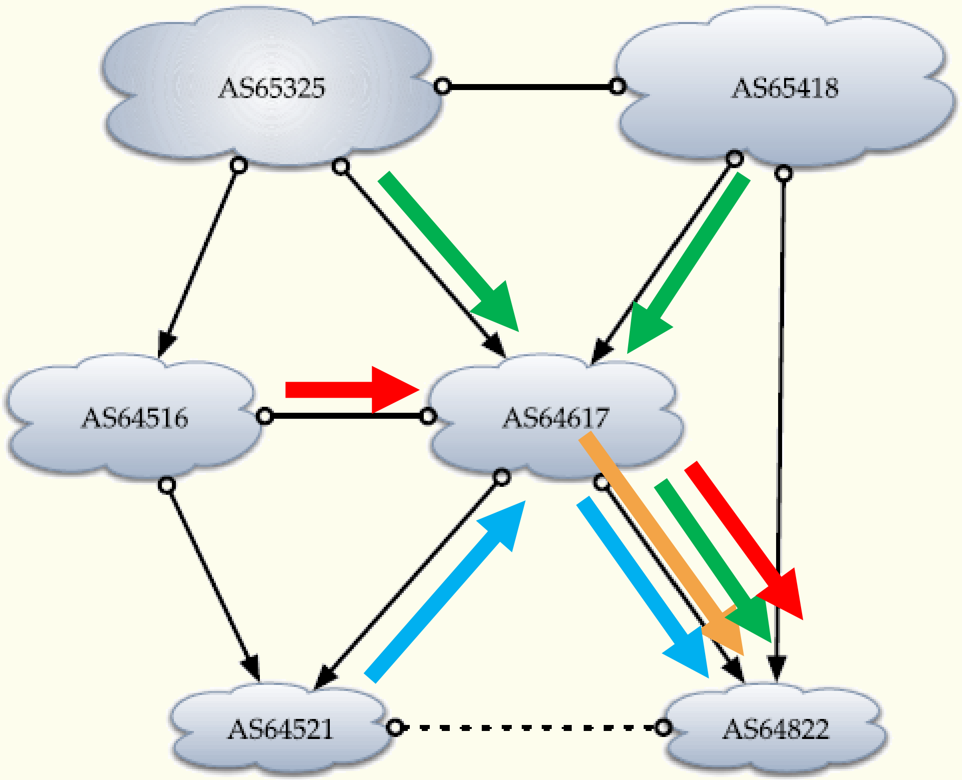
A **customer AS** exports to the **provider AS**
- its **routes**
- its **customer routes**

usually it does not export
- its (other) **provider routes**
- its **peer routes**

to **avoid to transit** their traffic

Customer-to-Provider

A **customer** AS buys **transit access** to the **provider** AS to reach all networks it cannot reach otherwise



Export policies

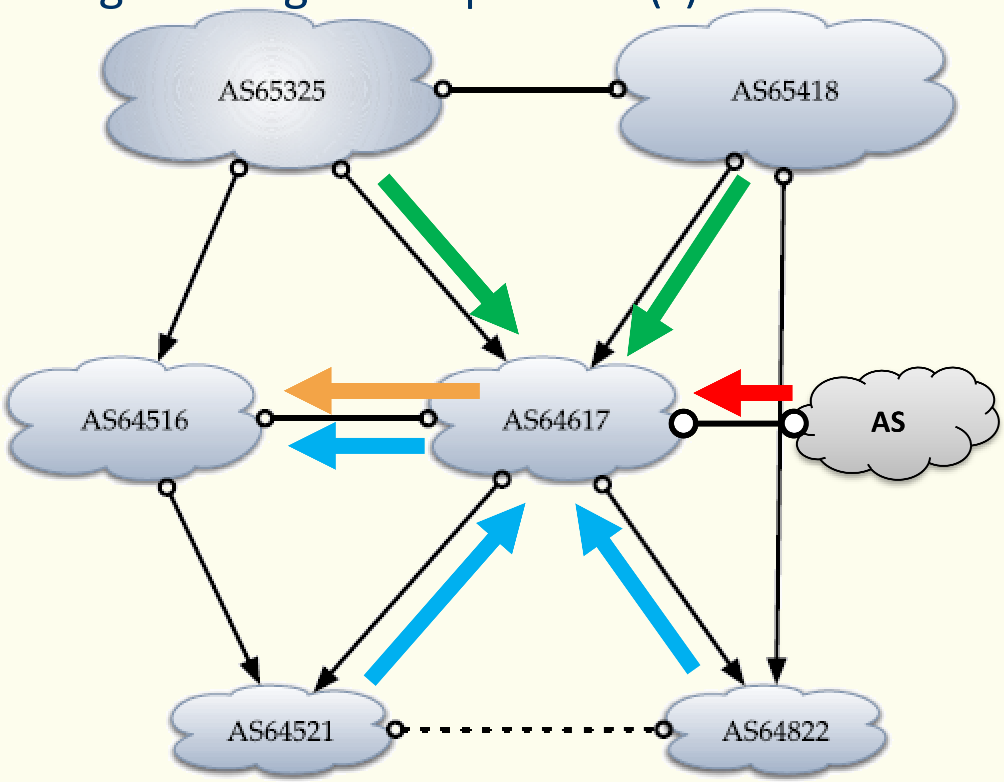
A **provider AS** exports to the **customer AS**
- its routes
- its customer routes
- its provider routes
- its peer routes

to allow the customer AS reach all the Internet destinations

In essence, the provider advertises to the customer either its **full routing table** or a **default route**

Peer-to-peer

Peer ASes agree on **exchanging** traffic (typically **without any payment** involved) in order to **reach each other** and **their respective customer routes**, without the need to go through their provider(s)



Export policies

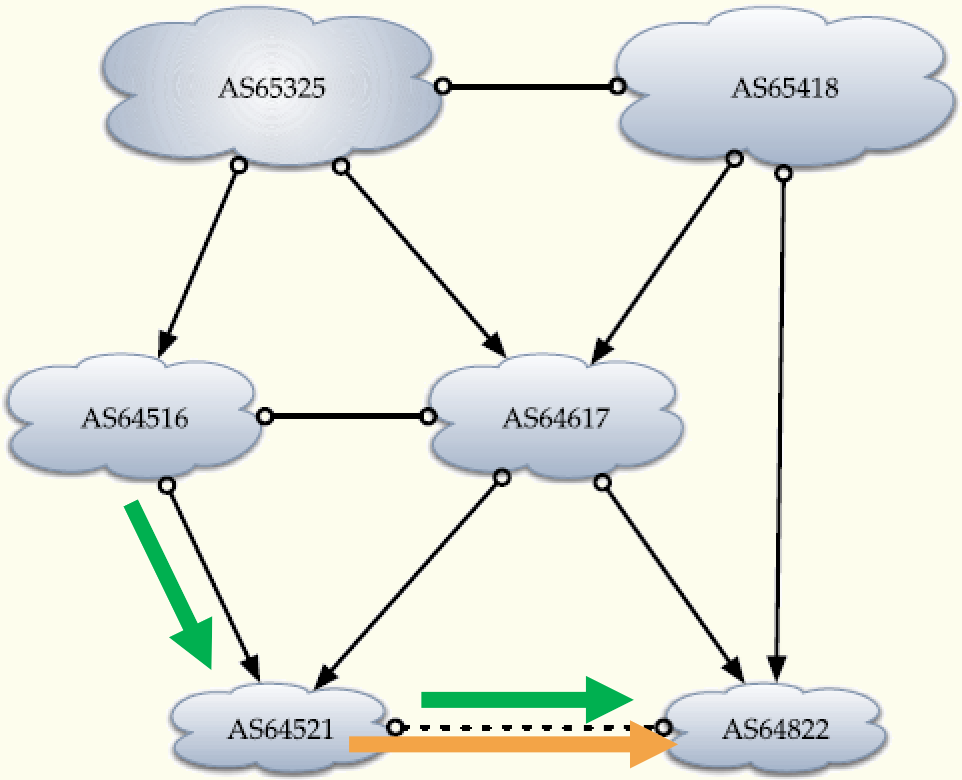
An **AS** exports to the peer AS
- its routes
- its customer routes

usually it does not export
- its provider routes
- its (other) peer routes

to **avoid to transit** their traffic

Sibling-to-sibling

Each AS acts like **a provider to the other**, e.g., the two ASes belong to the same ISP

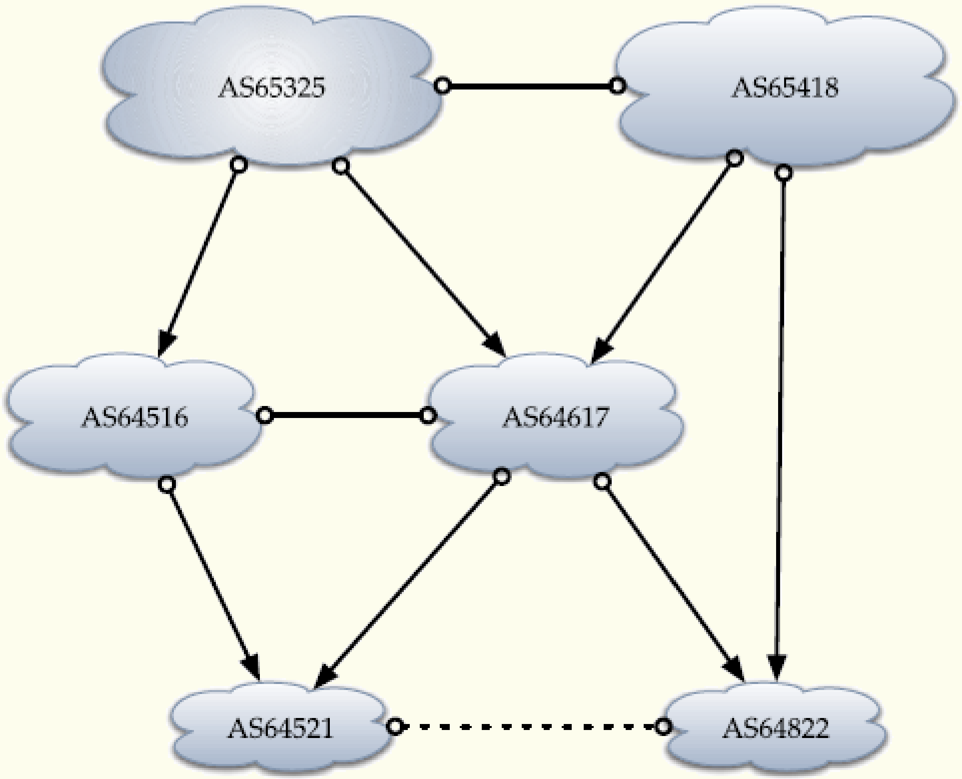


Export policies

An **AS** exports to the sibling AS

- its **routes**
- its **customer routes**
- its **provider routes**
- its **peer routes**

Valley-free property



Valley-free property: "After traversing a provider-to-customer or peer-to-peer edge, the AS path cannot traverse a customer-to-provider or peer-to-peer edge"

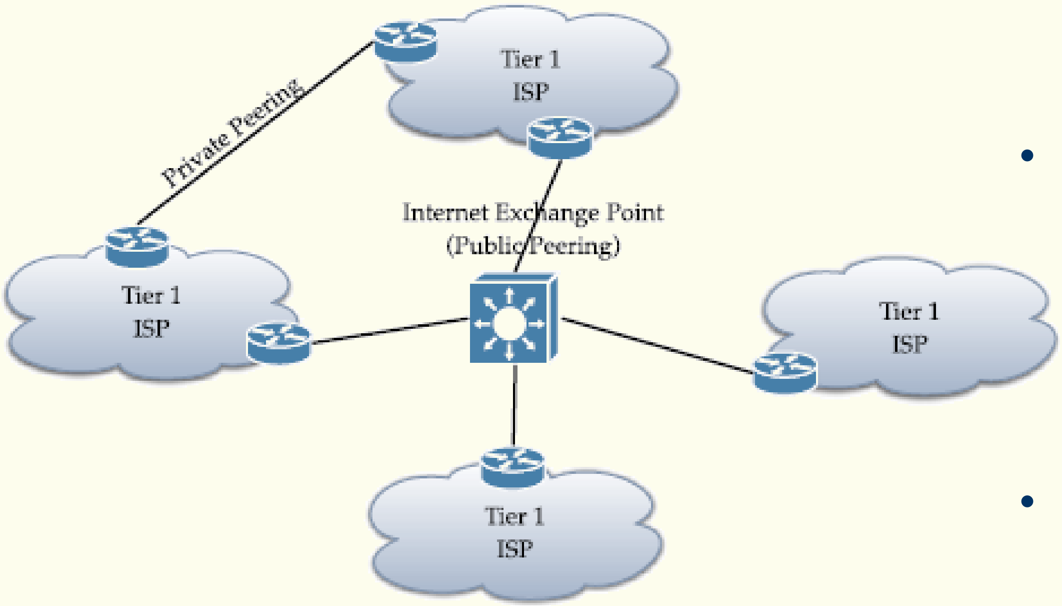
Examples

- Path (AS64516, AS65325, AS64617) is valley-free
- Path (AS64516, AS64521, AS64822, AS64617) is not valley-free

The valley-free property is demonstrated to be true for each advertised BGP path, if the corresponding export policies are obeyed by each AS

Lixin Gao, "On inferring autonomous system relationships in the Internet," in *IEEE/ACM Trans. on Networking*, vol. 9, no. 6, pp. 733-745, Dec. 2001

Internet exchange points (IXPs)



- ISPs peer **privately** or meet at an Internet exchange point (**IXP**)
- IXPs are usually operated by neutral entities that play the role of **providers for traffic exchange services** to ISP customers. Such arrangement is known as **public peering**
- Private network exchange points are also possible

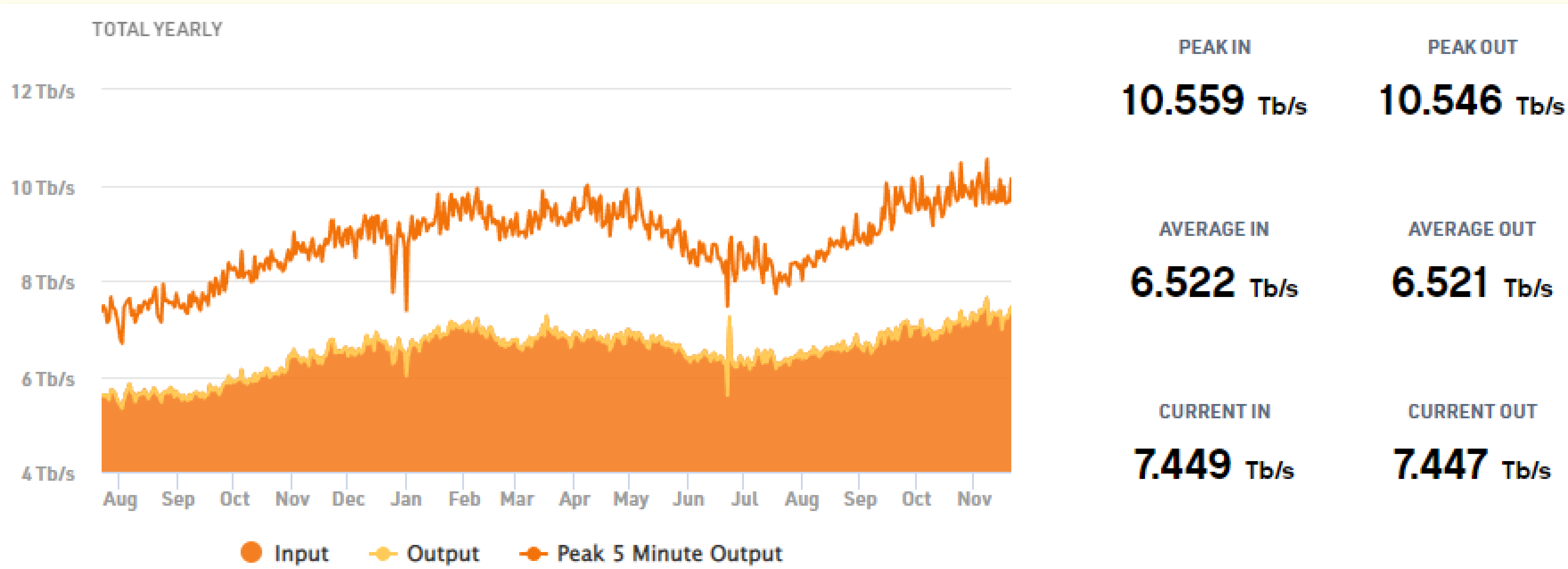
- Customers' routers are collocated in the **same physical (distributed) facility**
- In essence, an IXP is a **giant traffic switching point**. Connectivity is provided at very high speeds (1GE to 100GE), at **L2** or **L3**.

- IXPs have **well-defined policies** for joining the IXP
- Many **Content Providers** now also peer at IXPs to interconnect with major ISPs

Internet exchange points (IXPs)

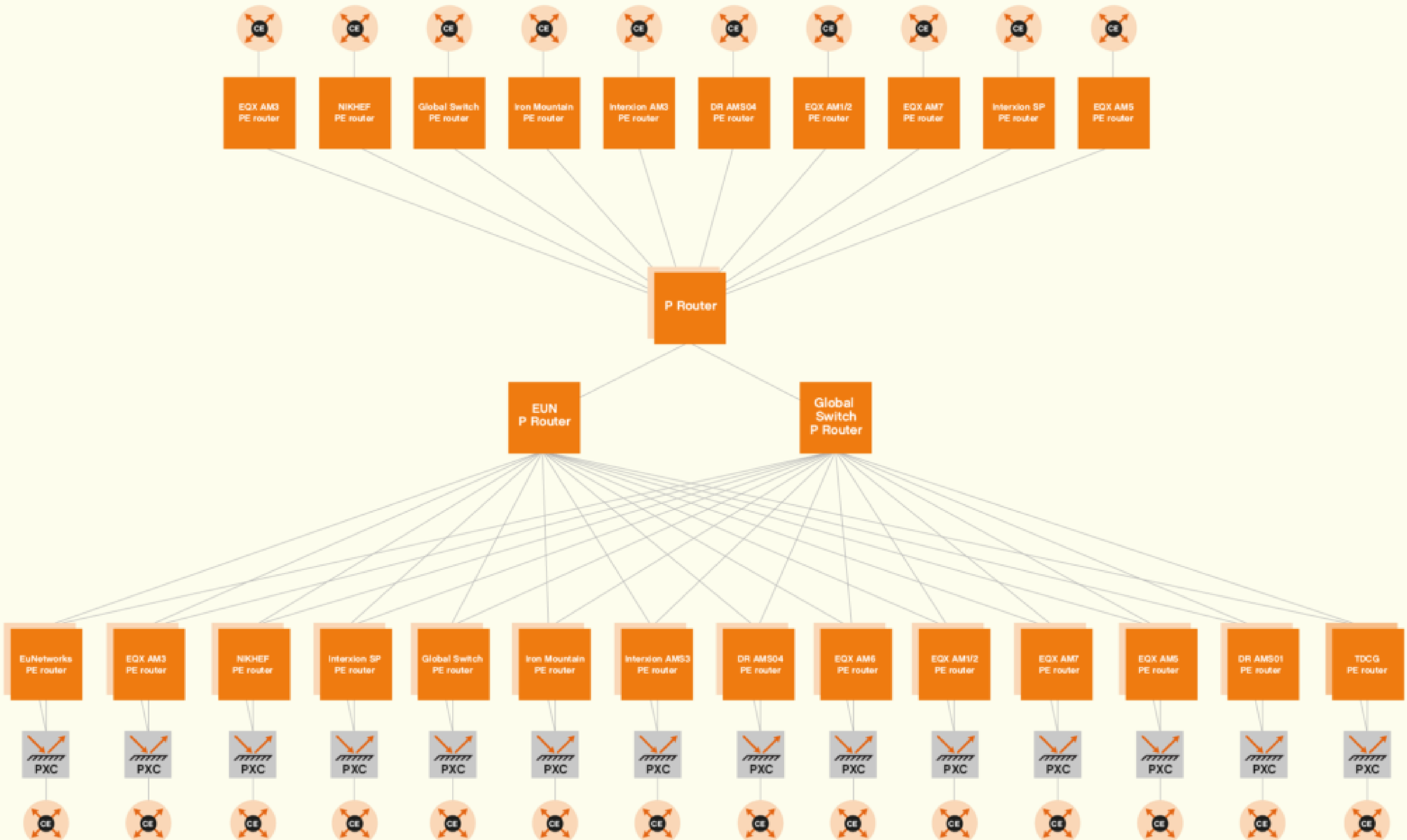
Amsterdam Internet Exchange (AMS-IX) is considered the largest IXP worldwide

767 route server peers, 883 total connected ASes



https://www.ams-ix.net/ams/documentation/total-stats

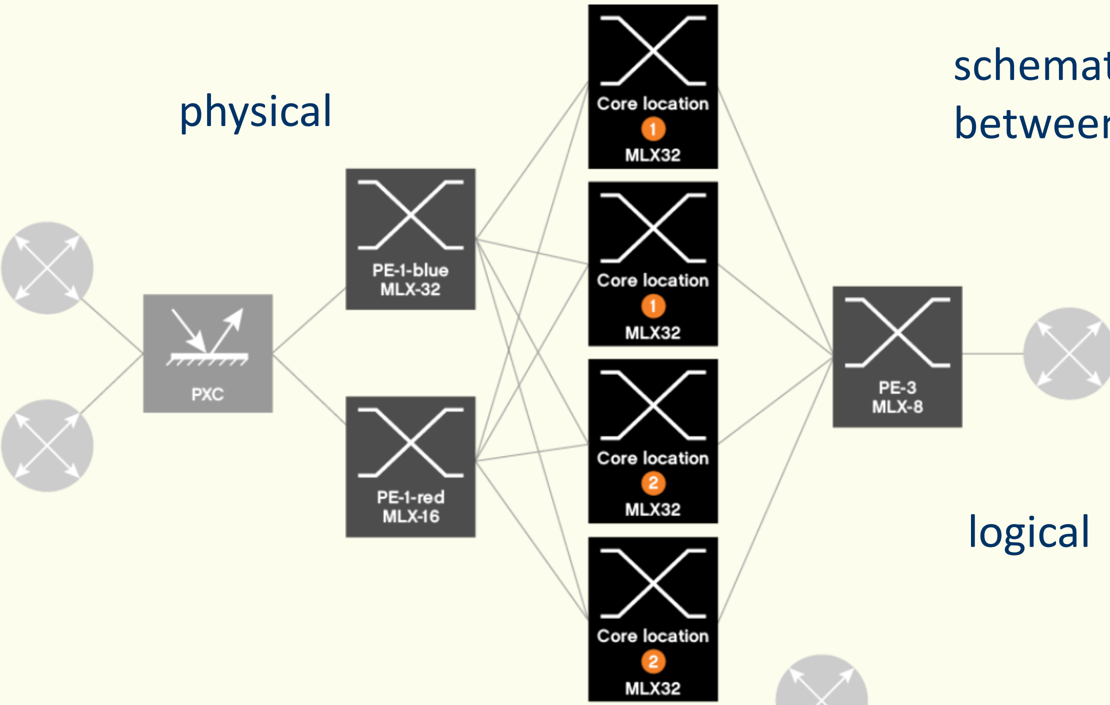
Internet exchange points (IXPs)



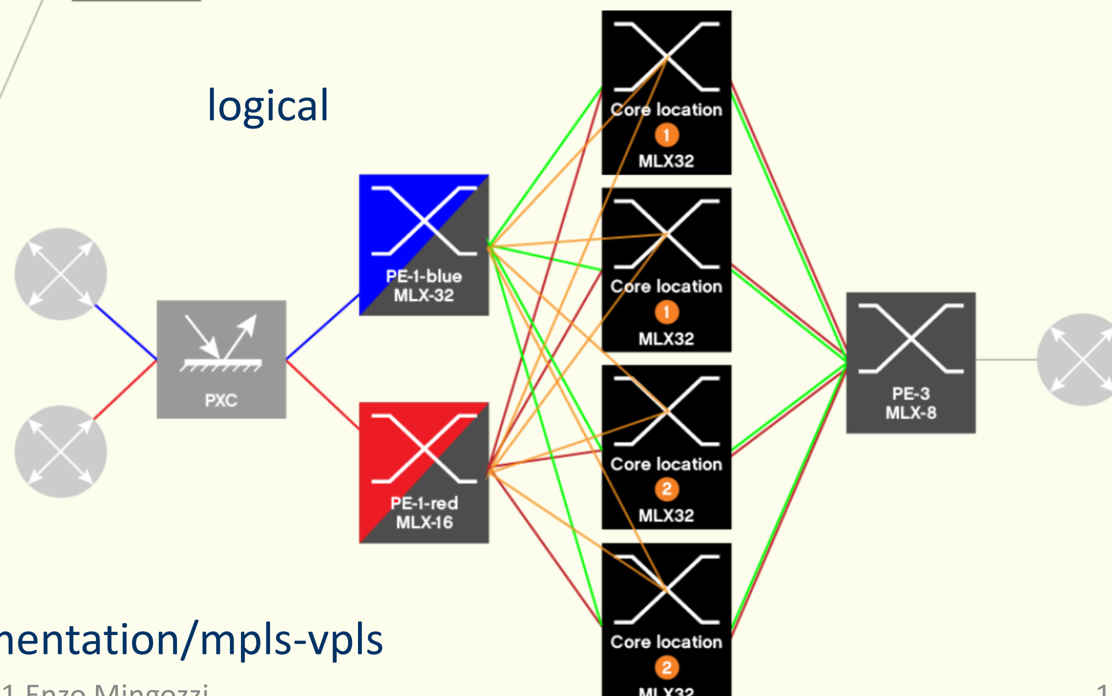
Internet exchange points (IXPs)

schematic drawing of the typical connections between customers

physical

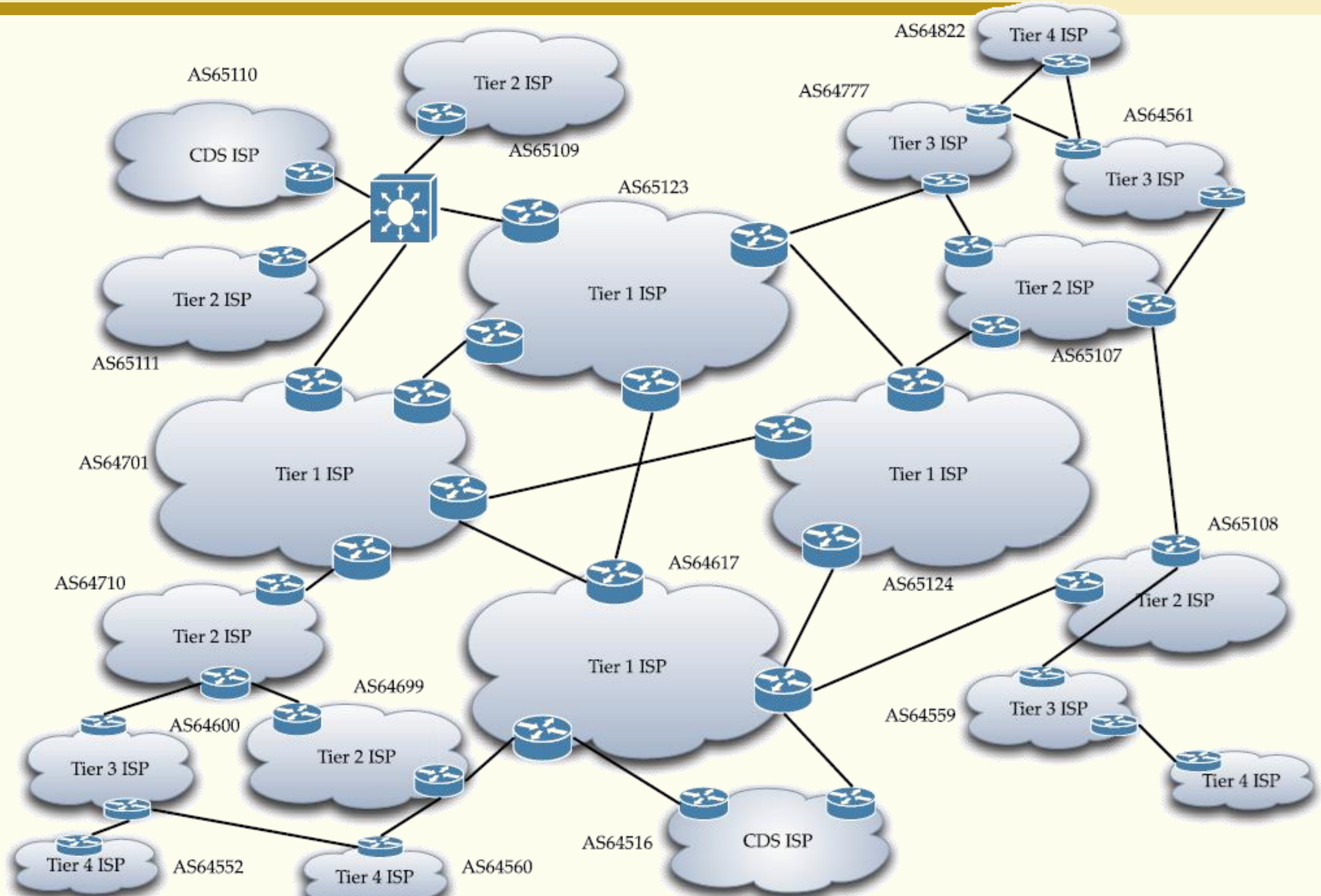


logical



https://www.ams-ix.net/ams/documentation/mpls-vpls

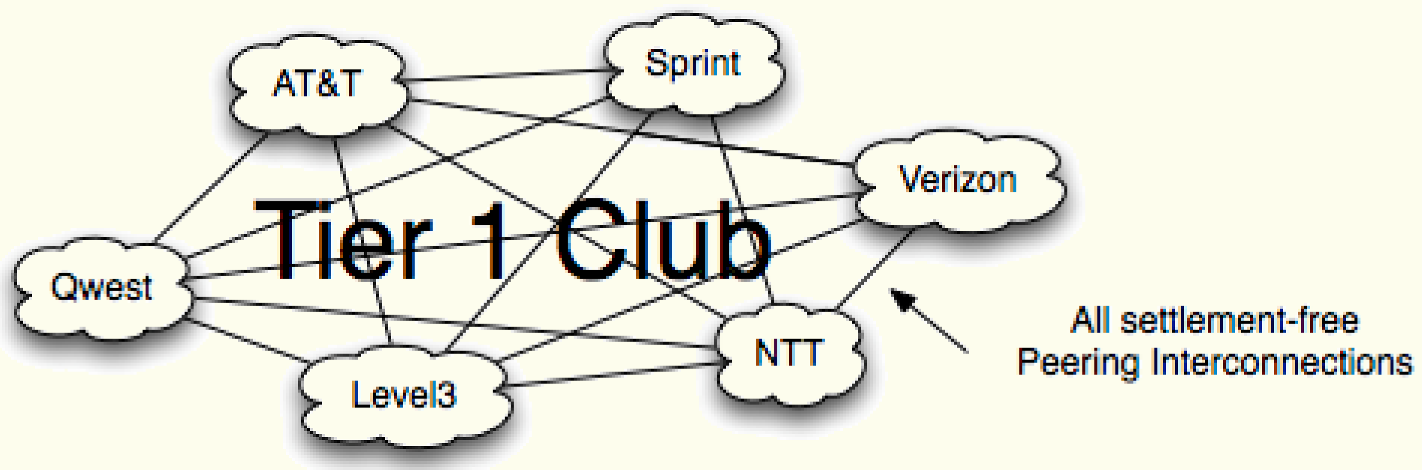
Internet connectivity



Tier 1 ISPs

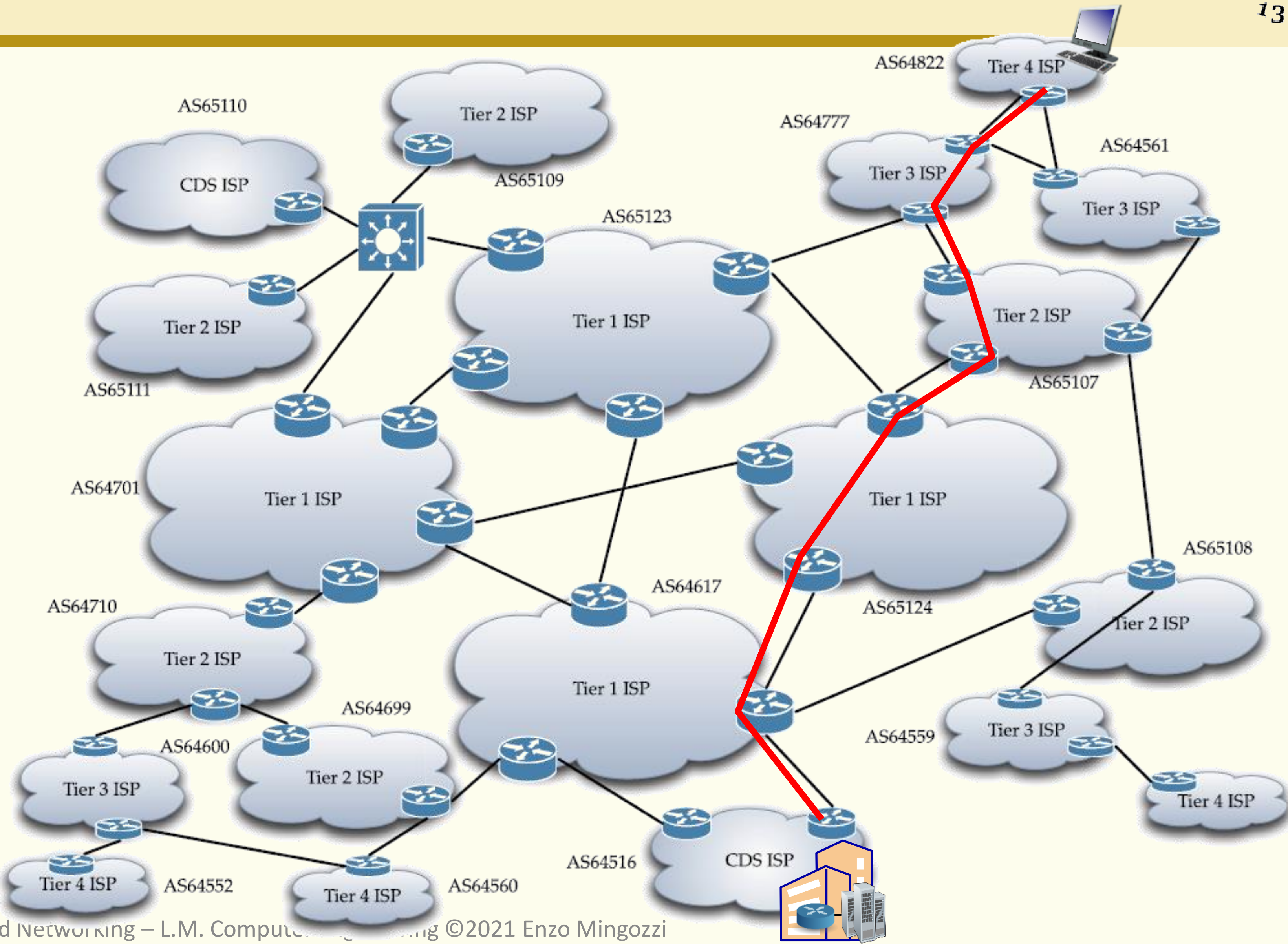
A ***Tier* 1 *ISP*** can reach every other network on the Internet **without purchasing IP transit** or **paying settlement**

A *Tier* 1 *ISP* has **no providers**

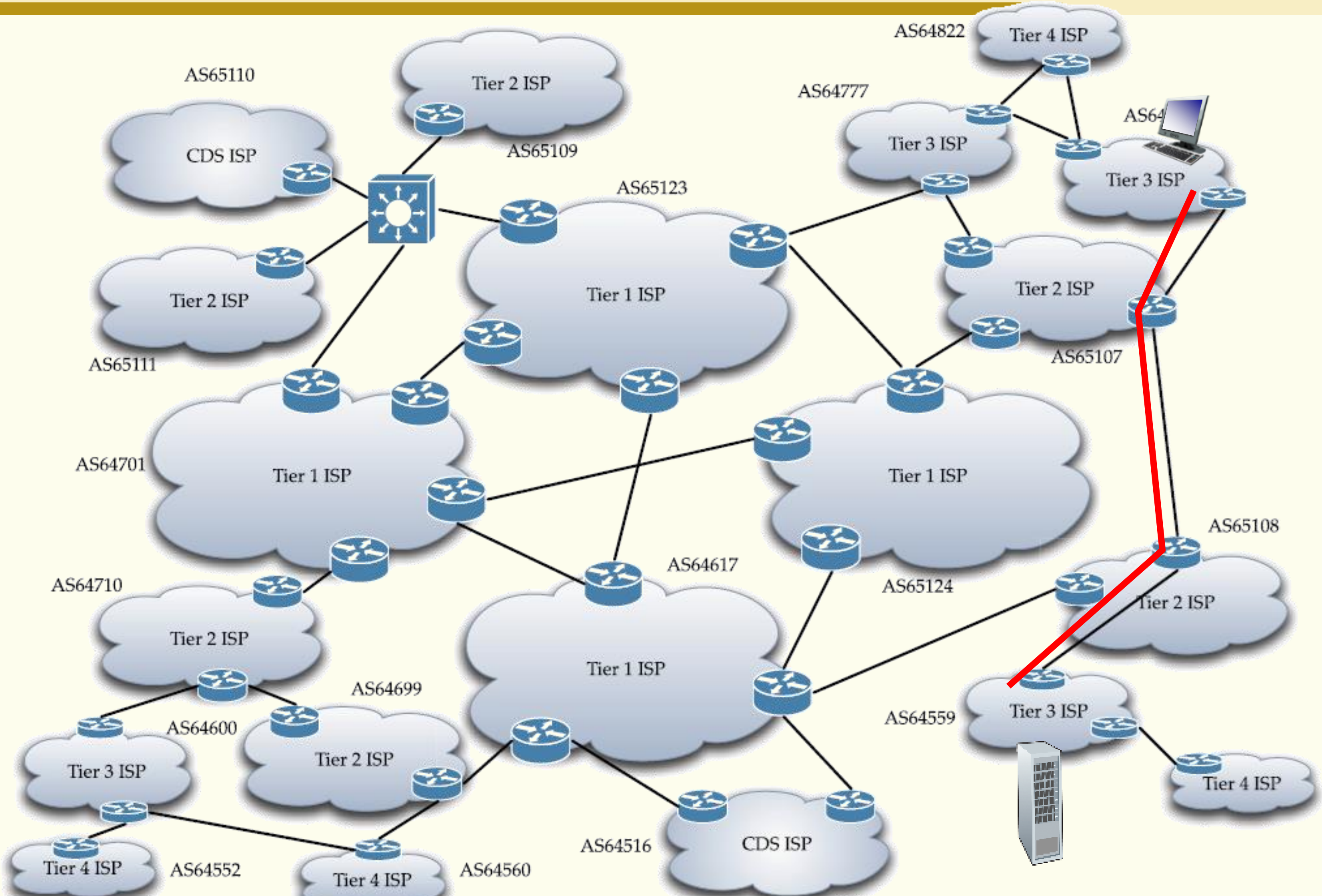


A list of universally recognized Tier 1 networks is available here:
https://en.wikipedia.org/wiki/Tier_1_network

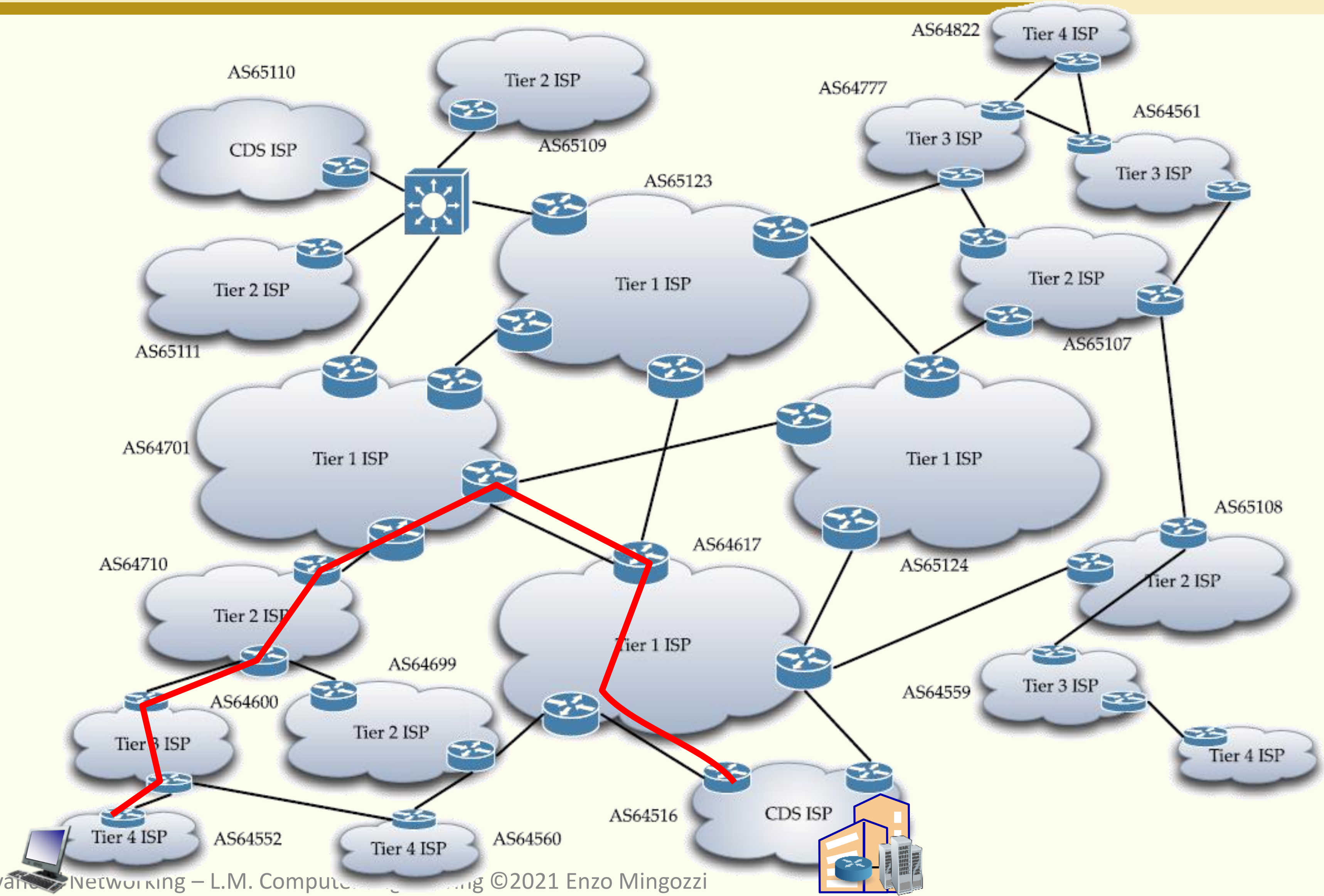
Example sessions



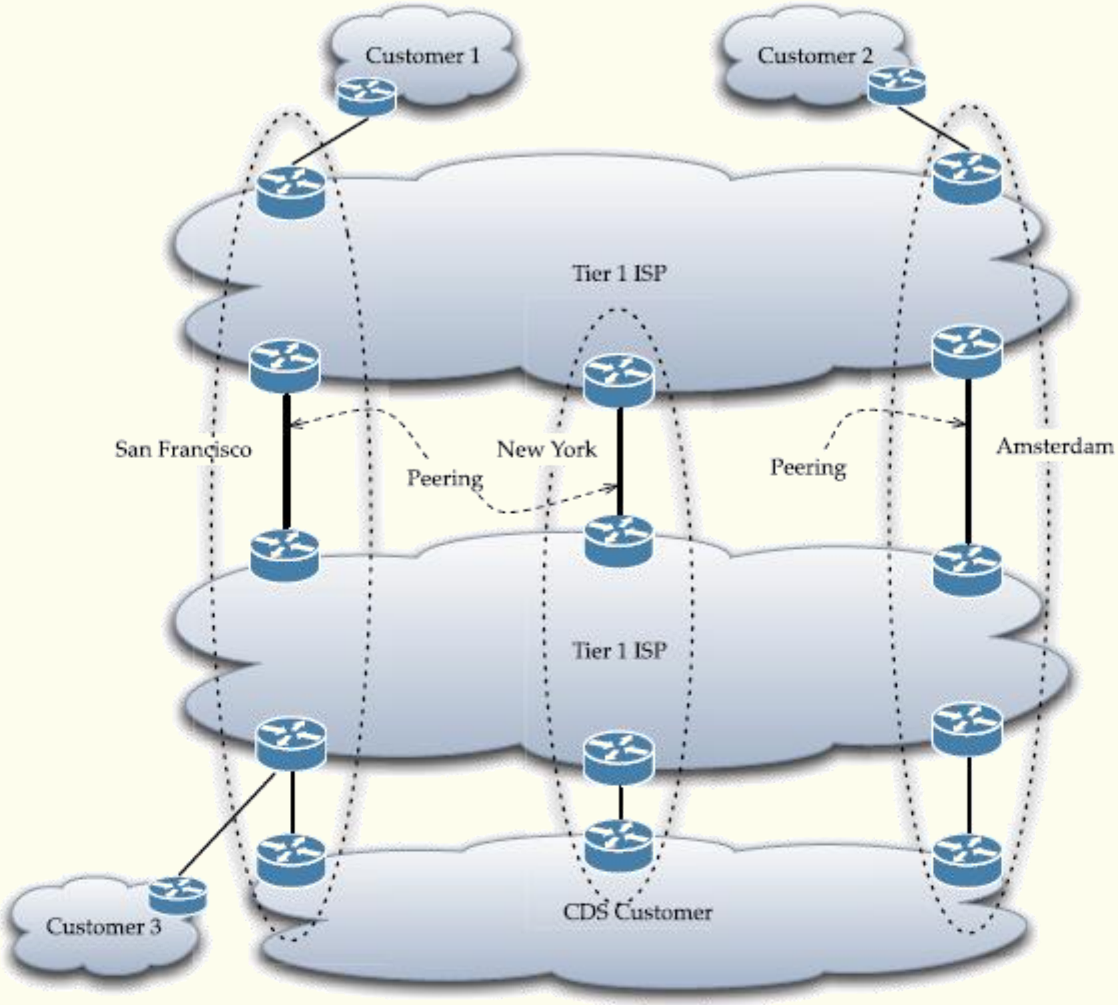
Example sessions



Example sessions



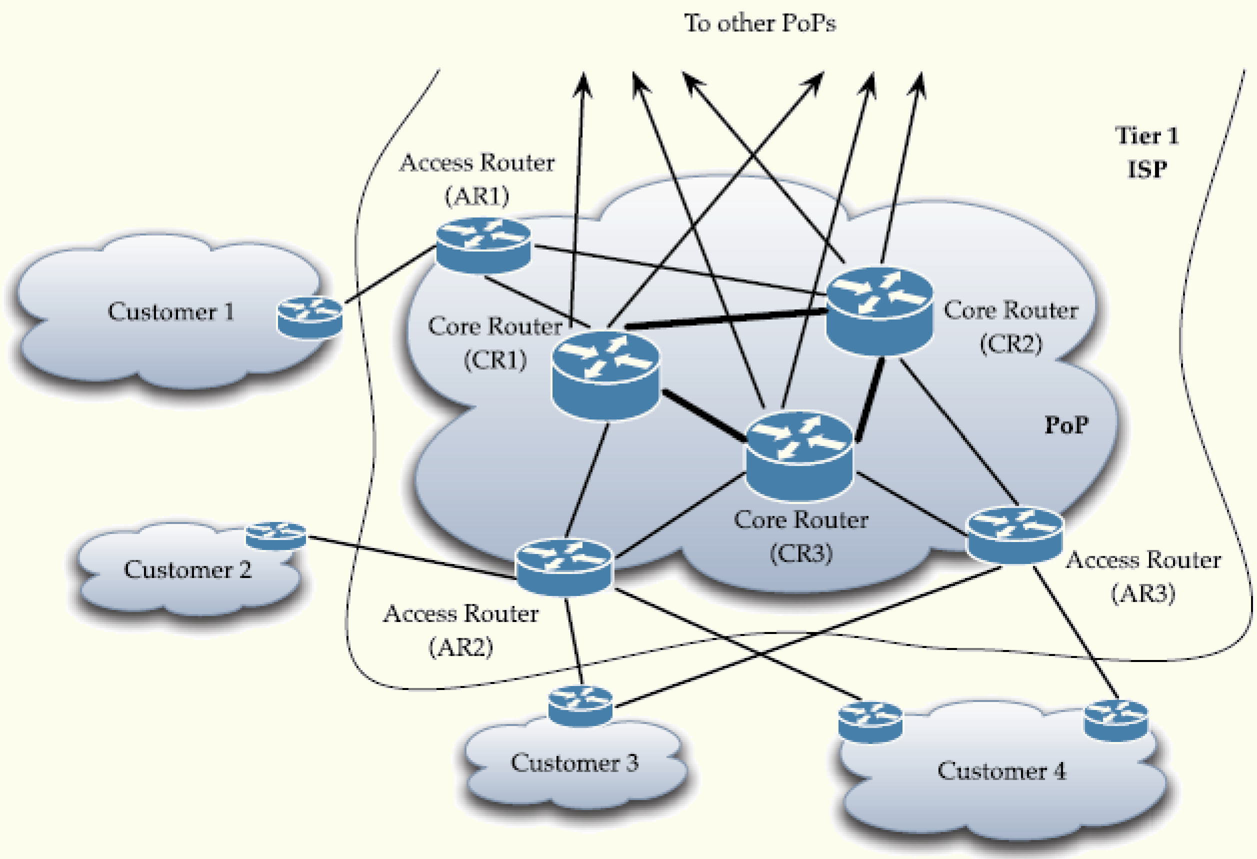
Geographic perspective



1. Customer 1 to Customer 3
2. Customer 2 to Customer 3
3. Customer 1/2 to CDS

https://peering.google.com/#/

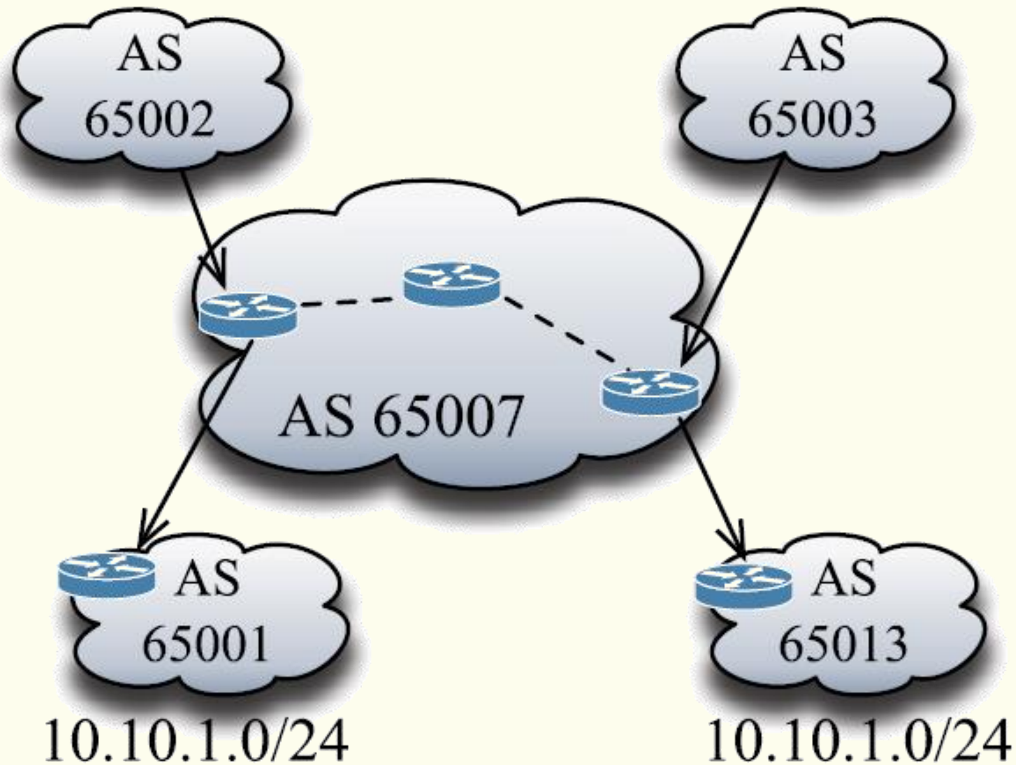
Point of Presence (PoP)



Prefix hijacking

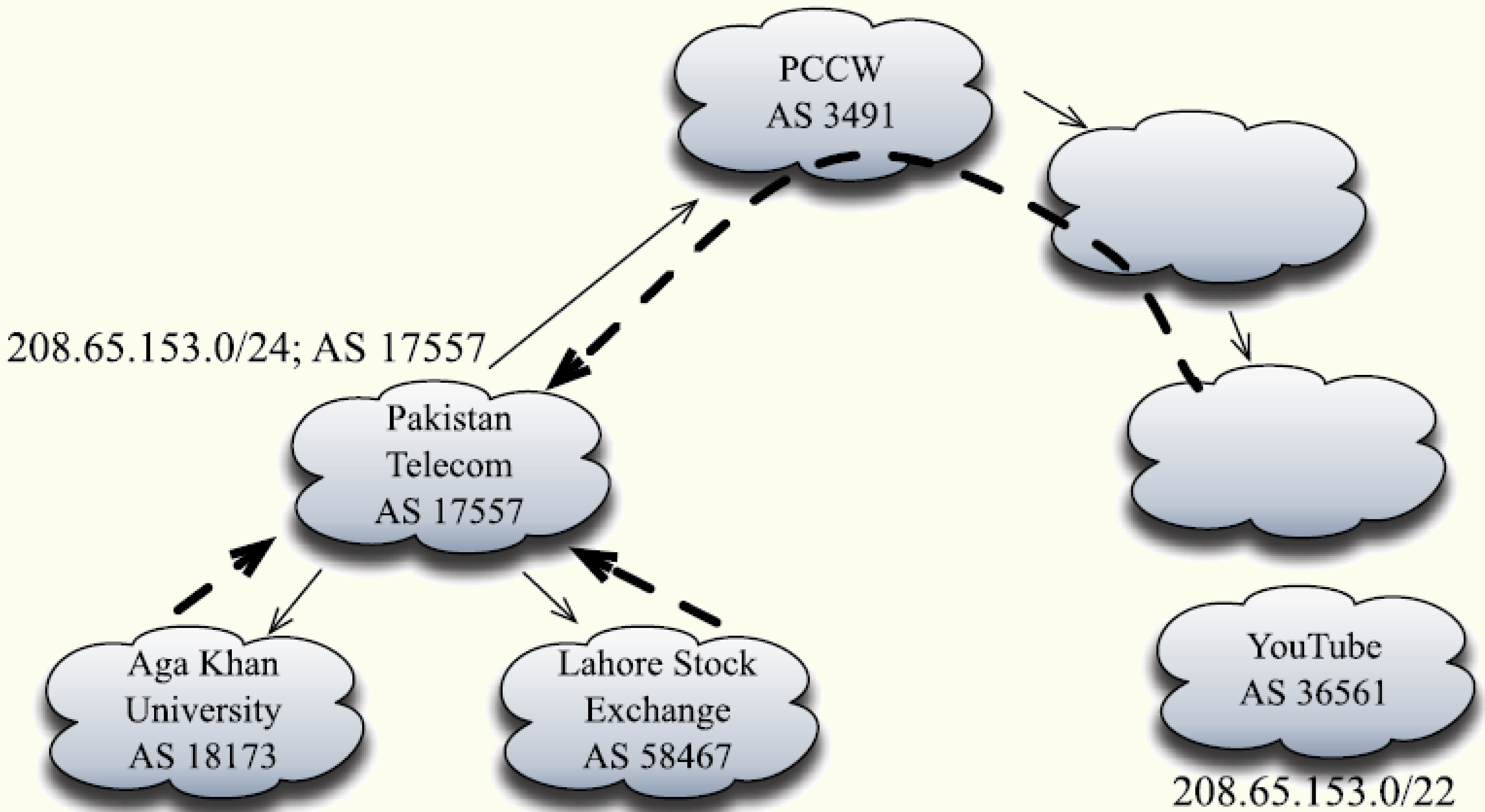
The entire BGP announcement process works on trust: the AS that announces an IP prefix **is assumed to be the owner** of the address

IP prefix hijacking refers to the situation when an AS announces an IP prefix without any authorization



Prefix hijacking

YouTube hijacking by Pakistan Telecom (February 2008)



Hiran R., Carlsson N., Gill P. (2013) Characterizing Large-Scale Routing Anomalies: A Case Study of the China Telecom Incident. In: Roughan M., Chang R. (eds) Passive and Active Measurement. PAM 2013. Lecture Notes in Computer Science, vol 7799. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-36516-4_23

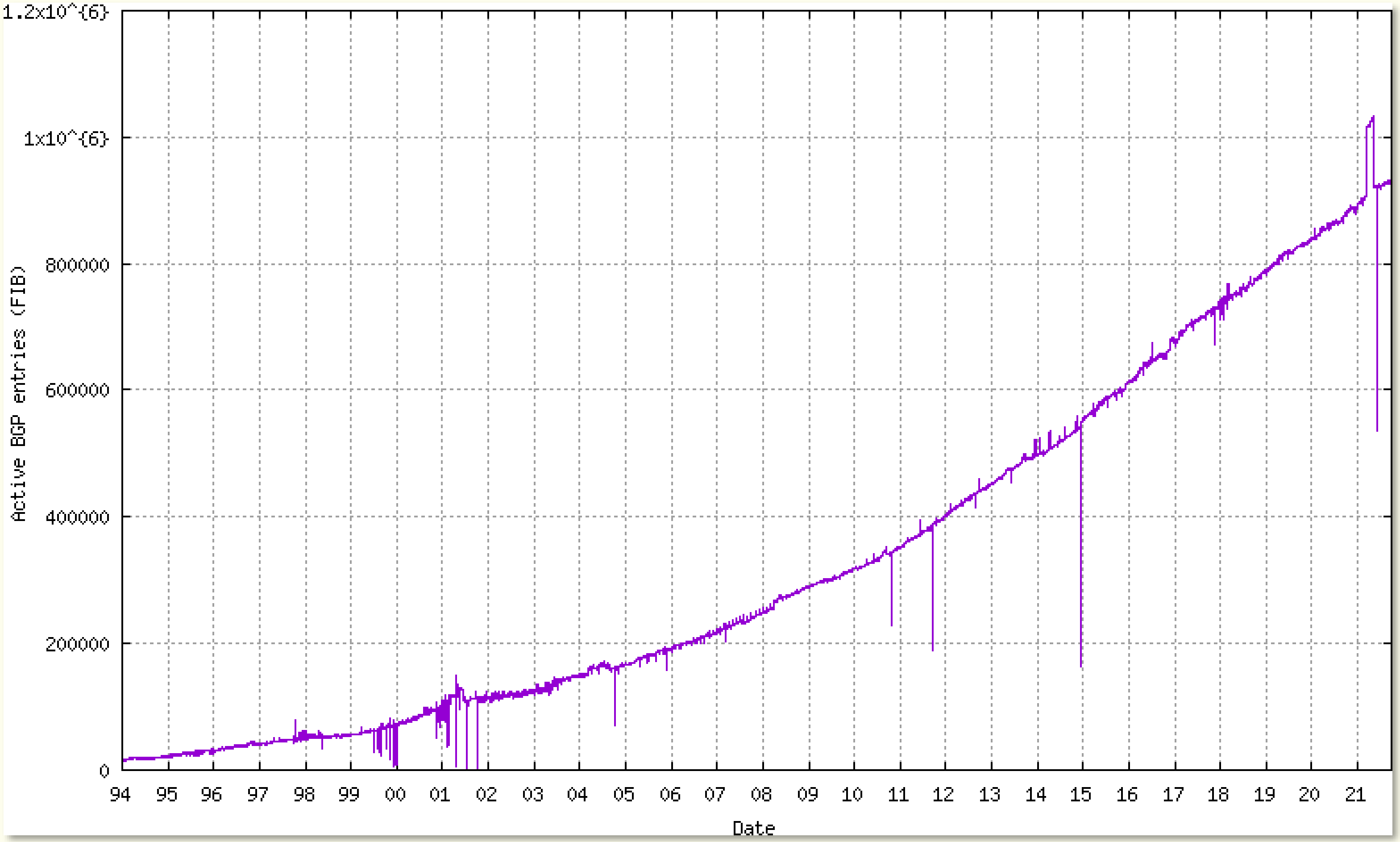
Size and growth of the Internet

The time trend of the number of unique AS numbers advertised within the BGP table. The associated metrics characterise AS's into origin only, transit only, or mixed (origin and transit), and also look at the trends of advertised address space as related to the number of ASes. [BGP data obtained from AS6447, https://bgp.potaroo.net/as6447/]

AS Reports (plots)	Data Sets(txt)
Unique ASes	70786
Origin only ASes	59493 (84,04%)
Transit only ASes	401 (0,57%)
Mixed ASes	10892 (15,39%)
Multi-Origin Prefixes	6894
ASes originating a single prefix	24687
Average entries per origin AS	12.6264
Average address range span for an origin AS	40755.0398

Size and growth of the Internet

Active BGP entries. [BGP data obtained from AS6447, https://bgp.potaroo.net/as6447/]



References

- D. Medhi, K. Ramasamy, **Network Routing: Algorithms, Protocols, and Architectures**, 2nd/ed. Morgan Kaufmann, ©2018
- RFC
  - **RFC4271**, A Border Gateway Protocol 4 (BGP-4), Jan. 2006
  - **RFC4360**, BGP Extended Communities Attribute, Feb. 2006