E7311 SDN Fundamentals and Techniques

Spring 2022, PIII, PIV

Professor Raimo Kantola

MyCourses

Course staff and contact information

Lectures 1…3: Prof. Raimo Kantola

raimo.kantola@aalto.fi tel: 040-750 1636

Office: CS: B258, Wed 14.15 -16.00

Lecture 4 : Juha Järvinen

Lectures 5… Aiman Nait Abbou and Hireche Otmane

ONOS and SDN: Aiman Nait Abbou

Hireche Otmane

(dr Jose Costa will visit on one lecture)

Lectures on E7311, Spring 2022: in Zoom

18 Jan 22	Tue	10:15-12:00	Internet, IP, Addressing
25 Jan 22	Tue	10:15-12:00	Network Address Translation
01 Feb 22	Tue	10:15-12:00	Interior Routing in IP
08 Feb 22	Tue	10:15-12:00	BGP
15 Feb 22	Tue	10:15-12:00	SDN, ONOS etc…
	:		:
	:		:

Material

- Course handouts
  – On course page in MyCourses
  – Slide copies and Lecture Notes
- Papers on SDN
- Possibly specifications, RFCs, and Internet-drafts
  – Links on course page or on slides or in MyCourses (or use google search)

Course requirements

- Grading will have 3 components
  - Returned Quiz answers 30%
  - Practical Exercises 40%
  - Exam. 30%

- A Quiz on a lecture is typically published on the week of the lecture and will be returned by the next lecture.

Course requirements

- Prerequisites:
  - Some 5cr course on Internet basics, understanding of transmission, and *packet transport*
  - Virtualization in Linux - Labwork on E7….
  - This lecture and its lecture notes try to fill in the biggest possible gaps in background

Addressing in Networks
Elements of Network stucture
Some Routing concepts --> motivation for SDN

Invariants in networks: something every network has

- Addressing
- Routing

Two use cases for SDN: 1) routing and 2) admission

Definition: Address is an identifier of a place in a network.

Types of addressing systems in Networks

Type	Examples	Pros	Cons or cost
Flat	Ethernet	Simple to move devices, entities with address Min configuration/OPEX	Does not scale to large/global network
Hierarchical	Internet Old telephony networks	Reduction in routing table size and frequency of changes	Address tied to topology → changes impact remote communication parties
Location independent (identifies Subscriber)	Mobile network phone Nr	Mobility	Need authentication, translate every nr to ***Routing Nr***
Fixed	Classical IP Router and server addresses	simple	Rigid, changes are hard Scalability limited
Dynamic	Dynamic IP addresses in Internet	Only active devices need an address	Address is not an ID: receiver does not know sender→ lack of trust

Three (main) ways to forward packet data

1. **Routing:** Proactively created ***routing table*** + no flow state = RT changes with topology changes (link or node failures, additions, removals)
2. **Switching**: ***flow state*** is used to mangle the packet that is sent to outgoing interface
   - E.g. SDN controller used to create the flow state entries
3. No routing table: **path carried in packet header**: SCION
   - Security keys proactively distributed;
   - 1 AES/packet/router authenticates packet (faster than longest match prefix search)
   - IP also has source routing but since the addresses are not trustworthy, it is very rarely used

What is the shift in technology to SDN?

Starting point is a dynamically routed IP Network or the "Internet" characterized by

- addressing tied to topology
- dynamic routing protocols give little control to Admin over the network
- Best effort service
- Lack of security

Target is a network where

- network can be automatically reconfigured at will as the IT systems are reconfigured based on-demand: scaling up and down etc.
- E.g. migrating a container (that has an IP address) from one host to another or even from one data center to another should be possible without changing the container IP address
- Network resource allocation is well managed in a uniform manner across the protocol stack

SDN generalizes state in networks into entries that have

- **Addresses, address-masks, ports, timeout**
  - Addresses can include number of interface, wavelength, MAC-addresses, MPLS labels, VLAN tags, IP addresses
  - Masks can handle address aggregation
  - Ports= TCP/UDP ports
- This model allows representing flow state and routing table entries in the same generic format and takes the heat out of the old debate: "*routing is good, flow state and switcing is bad*"
- SDN is about creating and managing state entries of the above form in SDN switches.

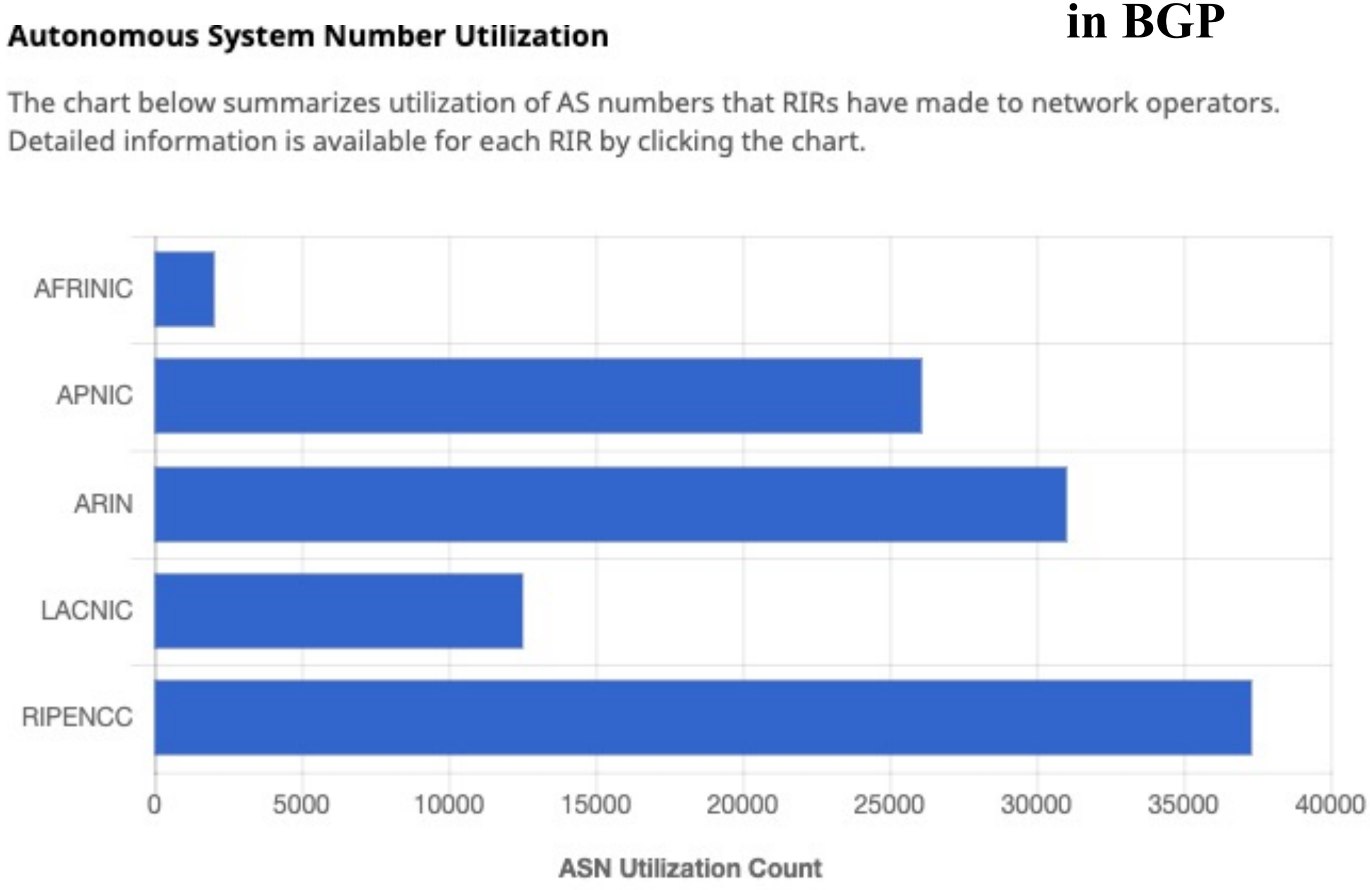
Internet consists of Autonomous Systems

/E7310/Internet_AS.svg

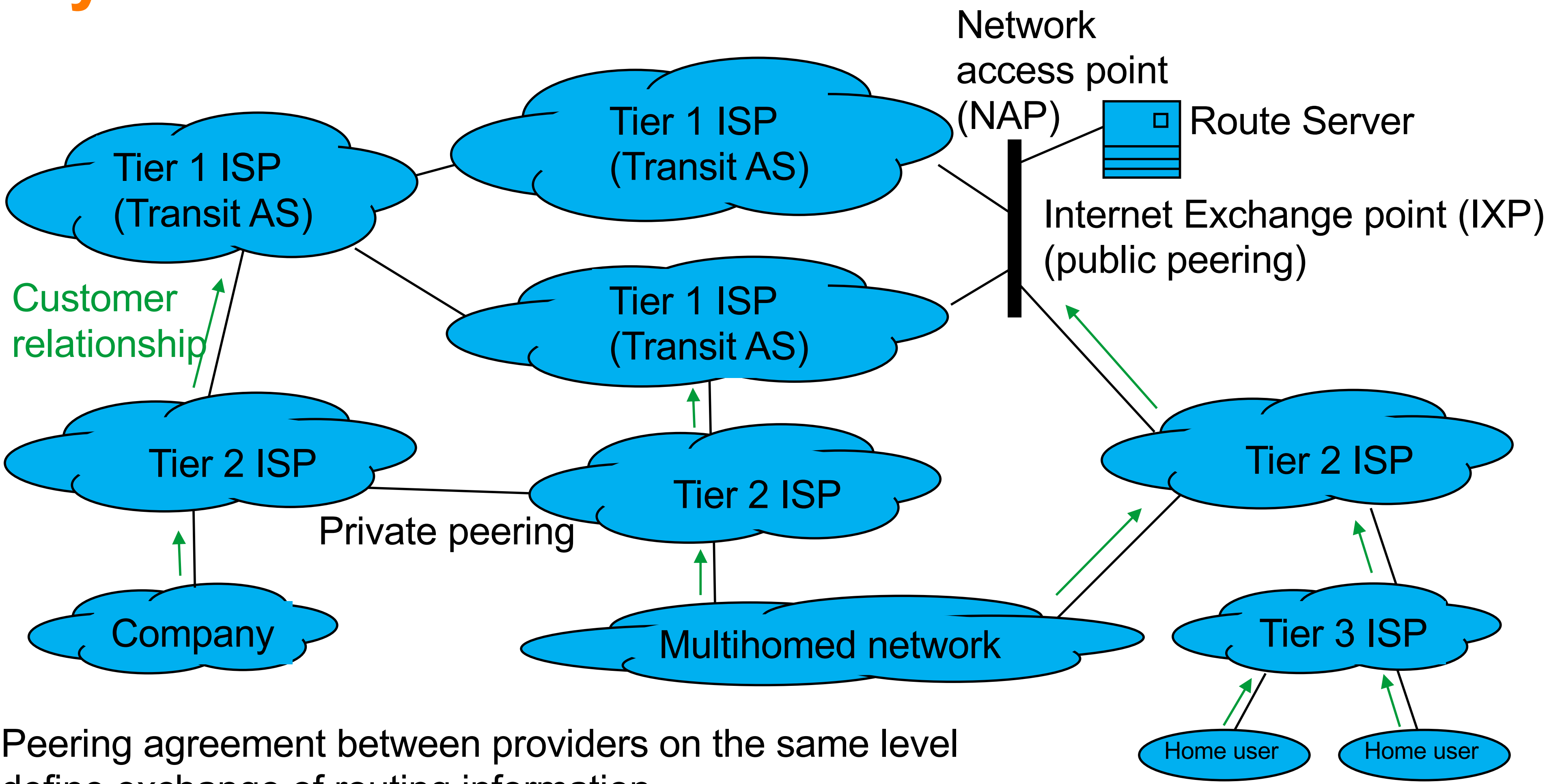
- An *Autonomous System (AS)* is a part of the Internet owned by a single organization.
  - Stub AS, Multihomed AS, Transit AS
  - Identified by Autonomous System Number (*ASN*)
  - 32 bit number since 2007
  - In an AS, usually one *interior routing protocol* is used
    - e.g. OSPF or IS-IS
- An *exterior routing protocol* is used between ASs
  - Currently *Border Gateway Protocol version 4 (BGPv4)*

Number of Autonomous Systems

Unique ASes **>100 000 in 3/2021 in BGP**

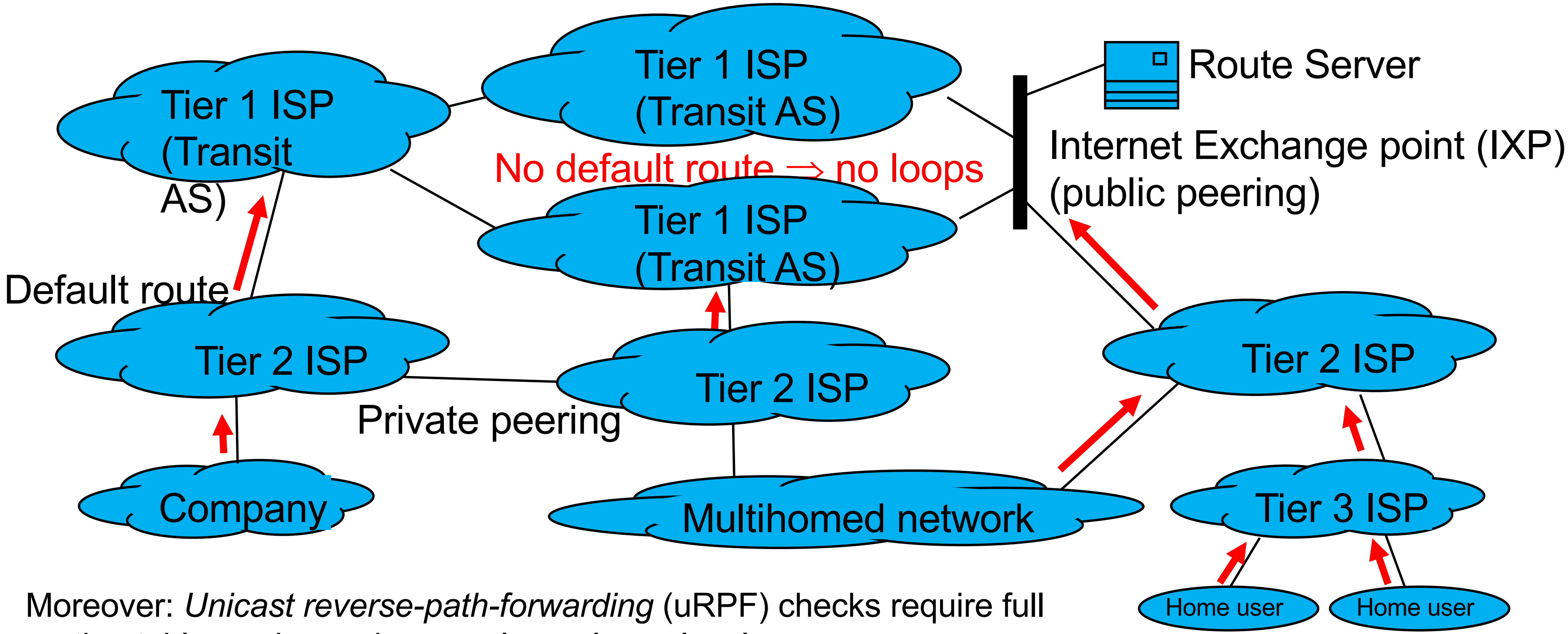


Organization of the Internet as Autonomous Systems



Peering agreement between providers on the same level define exchange of routing information

Stub ASs use default routes to reduce routing table size, the transit ASs are default-free

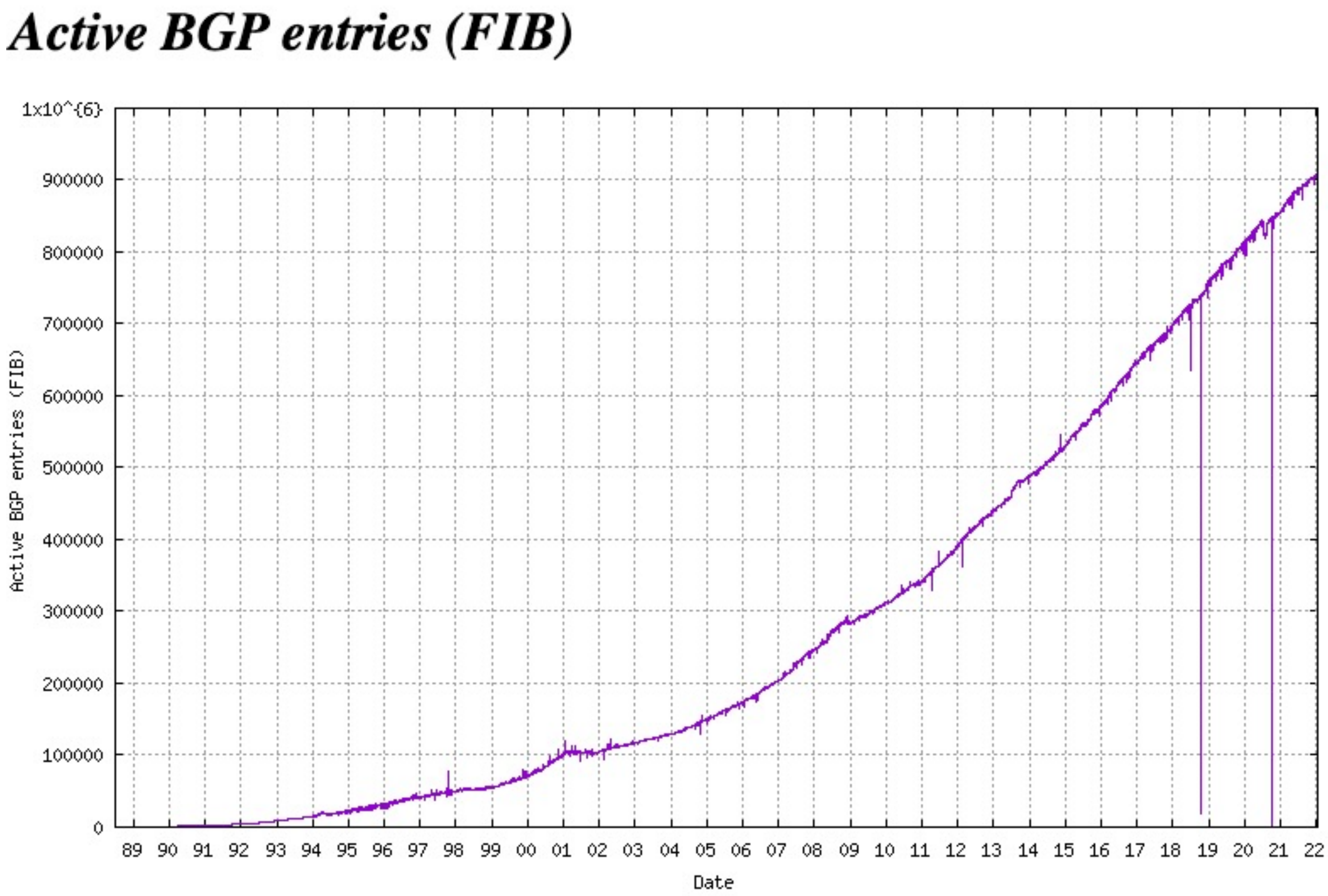


Moreover: *Unicast reverse-path-forwarding* (uRPF) checks require full routing tables ⇒ becoming popular on lower levels

Tier 1: No default route. Instead they must have an entry for each valid address. Otherwise the packet to an unknown hop would loop forever with the default routes.

Aalto University

Entries in BGP routing tables



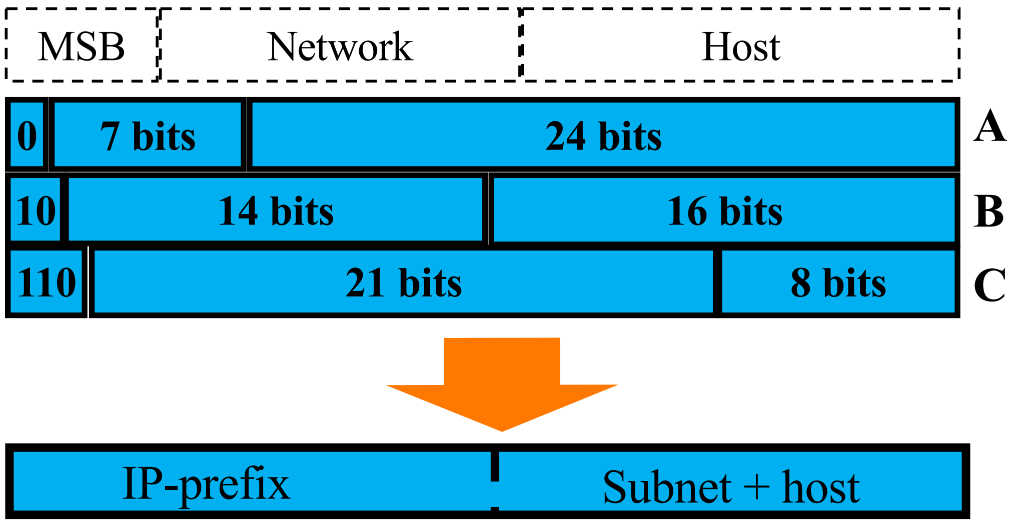
https://bgp.potaroo.net/as2.0/bgp-active.html

A Use case for BGP: Data centers

- BGP routers can handle > 1M route entries in Routing table
  - Use them for host routes = /32 prefixes
  - Can handle a set of data centers with 1M containers
  - Can move any container anywhere and update BGP routing tables everywhere to complete the migration
- This competes with the idea to use SDN in those same data centers to obtain the same migration capability.

CIDR – Classless Inter-Domain Routing

CIDR allows splitting 32-bit IP-addresses freely into prefix and tail



- A sequence of C-class networks can be represented:

  194.51.120.0 - 194.51.12**7.255** = (8 C-class networks)

  network = 194.51.120.0

  mask = 255.255.248.0 or /21

  (hexa FF FF F8 00

Repetition: address arithmetic

- Example

	192.24.134.23	address
AND	255.255.248.0	mask
	192.24.128.0	network

	192.24.134.23	address
AND	0.0.7.255	NOT (mask)
	0.0.6.23	host (or subnet+host)

$$\underbrace{\texttt{11000000.00011000.10000}}_{\text{network}}\,|\,\underbrace{\texttt{110.00010111}}_{\text{host (subnet+host)}} \quad \text{address}$$

$$\texttt{11111111.11111111.11111}\,|\,\texttt{000.00000000} \quad \text{mask}$$

Aalto University

Repetition: address arithmetic

- Example

	192.24.134.23	address
AND	255.255.248.0	mask
	192.24.128.0	network

	192.24.134.23	address
AND	0.0.7.255	NOT (mask)
	0.0.6.23	host (or subnet+host)

	192.24.134.23	address	(alternative way)
–	192.24.128.0	network	
	0.0.6.23	host (or subnet+host)	

CIDR – Classless Inter Domain Routing

- Problems caused by the growth of the Internet
  - Not enough B-class addresses
    - A few thousand addresses required for an average organization
    - Class A too big (16M addresses), class C too small (256 addresses)
    - Only 16384 class B networks
  - Addresses in class B are used inefficiently
    - Class B is usually too big too (65534 addresses)
  - Solution: use several class C networks
  - But: Growth of routing table size
- Internet growth has forced the adoption of CIDR address arithmetic to improve the efficiency of using IP address space. CIDR was adopted 1992.
  - General Longest match prefix search/packet was needed → ASIC implementation.
  - IP routing data plane on general purpose CPU tends to be slow…

Depletion of IP addresses is a problem but somehow we are surviving…

- IP provides 2^{32} (about 4.3 billion) addresses
- May be 70% of this address space is actually advertised
- Earlier consumer addresses were typically shared between several users (with dial-up connections), now users are always connected
- Organizations have been assigned too large address blocks during the classfull addressing era, still many of these addresses are unused or not publicly accessible
- Host-density ratio low due to subnetting
- Virtualization increases the need of addresses
- CIDR and NAT are officially short-term fixes(?)
- Official story: IPv6 is a long-term solution

On 3 February 2011, IANA allocated the last two /8 address blocks to Regional Internet Registries!

IPv6 adoption has been slow

- Deployment started >20 years ago;
  - one of the slowest adoption rates for new technology in the history of networking technology
  - DNS stats from March 2017 tell that
    - IPv4 DNS queries over UDP per day were about 4Billion
    - IPv6 DNS queries over UDP per day were about 500M.
- Adoption led by countries that are really short on IPv4 addresses
  - 37 countries have at least 5% of their traffic in IPv6.
- In terms of architecture IPv6 is very close to IPv4, added value is quite limited…

On the role of routing

- Every network must have **routing:** a use case for SDN
- How does the SDN controller know what to write into SDN switch entries
  - One application that tells the controller what to write is Routing:
  - Calculate a path from source to destination, manipute entries in all switches on the path
- Since the controller has limited capacity (nrof new entries it is able to add to switches/s), it is best to maximize the entries that are created *proactively*, before any host tries to do anything.

Two contexts of using routing

- Routing individual user's traffic (flows and packets)
  - In Internet usually based on **proactive** approach: a before made routing table is used for packets
  - Hop count most common metric
- Provisioning:
  - Admin sets paths into the underlying infra for delivering the traffic of many users/customers
  - E.g. provisioning the mobile backhaul network eNBs to Internet
  - ***Residual capacity*** is a good metric when using predicted traffic volume numbers – using this metric requires a centralized approach in practice (Path Computation Element, Connection Manager, SDN controller etc.)
  - Allocations are long term – may be adjusted from time to time based on measured user traffic

Source routing vs. Hop-by-Hop routing/distributed vs. centralized

w
h
o

- **Hop-by-hop:** each node is autonomous, uses its own forwarding information base (FIB) to choose the next hop
- **Source routing** = source node (host or edge) calculates a route, puts it into the packet and intermediate nodes just follow the instructions in the packet
  - If host calculated – security risk
  - Can be useful if properly used: SCION uses 1 AES per packet/node to ensure that path is genuine

h
o
w

- **Distributed:** Routing Tables (RT) are created by each node independently (like in the Internet)
  - In circuit networks we have had also hop-by-hop but centralized systems
- **Centralized:** a centralized node calculates RTs for each node – e.g. SDN

Energy Efficient Routing

- Technologies with sleep modes are emerging: Energy Efficient Ethernet;
  - Energy consumption is a function of carried traffic
  - When there are no packets for time *t*, link goes to sleep. When a packet is offered, link quickly wakes up
  - Variants for copper and fiber
  - SDN can be used to control the placement of flows onto paths
- Optimizing for energy is quite different from the traditional additive (eg hop count)/concave (eg BW) optimizition
  - Must maximize the number of links that sleep (from the most power hungry links) under constraints of capacity, delay etc.
  - Instead of optimizing *Capital Expenditure* like in traditional methods, Energy Efficient Routing tries to optimize *Operational Expenditure*.

When is routing optimal?

From the user point of view: Minimum probability of blocking, delay, jitter (stability), loss or maximum bandwidth …

From the network point of view: Maximum network throughput. Requires short routes, while excess traffic needs to be directed to least loaded parts of the network. At the same time user service requirements need to be met.

It follows that routing is a complex optimization problem. Most times the optimum cannot be found in a closed form. Therefore, we are interested in near-optimal, heuristic approximations.
SDN advantages: centralized residual capacity routing AND Energy efficient routing are possible

Traffic Engineering

- IP ROUTING (Internet) uses Shortest paths only OR *equal cost multipath* (ECMP)routes.
- However, between points A and B in the network there may be many parallel almost equidistant paths while the shortest paths are all heavily loaded.
  - This happens when the current traffic is different from predicted and the network is highly meshed
  - This state may persist for a long period of time (>10 min)
- Under this situation, it would be beneficial
  - to use not only the shortest paths – this is called traffic engineering.
  - In packet networks this requires a switching layer such as MPLS – can be managed eg by SDN
  - Be able to use massive multipath like in SCION

Packet network routing classification (given a dimensioning)

- Proactive vs. reactive
  - Proactive: Routing table formation is a background process, individual user's flow is forwarded using the existing forwarding table
  - Reactive (or on-demand): (MANET, NATs) route (binding state) is formed as a reaction to user action; Ethernet MAC learning.
- Interior vs. Exterior
  - Interior: RIP, OSPF, IS-IS – for one admin domain
  - Exterior: BGP – for the whole Internet (between admin domains)
- Source vs. hop-by-hop
- Distributed vs. centralized

Advantages of distributed routing for IP

- Self configuration
  - If nodes are given the IP addresses for their interfaces, they can tell each other of their existence and parameters → the whole network becomes capable of routing packets automatically
  - The network recovers automatically of link and node failures: if a path exists from node X to Y, it will be found automatically
  - IP networks were designed to survive a nuclear strike
- Best effort model/distributed routing for IP have been a huge success and has scaled much beyond the original design that was meant for connecting university networks.
- Now moving away to SDN….

Why are we moving to SDN?

- Distributed routing supports Best Effort service→ low profit
- Operators want to make the best use of all network resources they have invested in under all traffic fluctuations and instead of BE offer Carrier Grade (CG) services
  - CG: From point A in network to point B carry traffic X, else drop. == for all traffic that is carried such a rule must be present in the network, if not, drop the packet → hope for higher profits.
- SDN allows
  - centralized routing for provisioning using ***residual capacity*** as optimization criteria
  - Traffic Engineering: use all network resources to maximize network throughput and thus revenues
  - Implementing any centralized algorithms (for any purpose) on general purpose computing platforms (Datacenters: linux+KVM + virtualization)
  - Managing migration of VMs and containers – fits together with orchestration in cloud computing.

Please register on the course before Jan. 24th, 2022!

Thanks for your participation and enjoying the course!