Computer Networks (part 4)

Rémi Emonet – 2021
Université Jean Monnet – Laboratoire Hubert Curien

Computer Networks: global overview

1. Introduction to computer networks
2. Networking application layer (HTTP, FTP, DNS, …)
3. Data transfer layer (UDP, TCP, …)
4. Network layer (routing, IP, ICMP, NAT, …)
5. Lower layers, wireless and mobile (Ethernet, ARP, …)
6. Security (SSL, …)

Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast

application
transport
network
link
physical

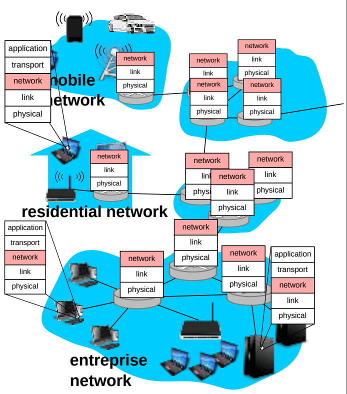
Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast

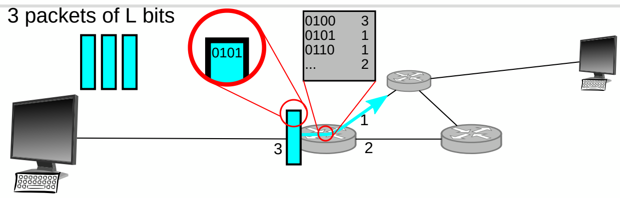
application
transport
network
link
physical

The Network Layer (IP)

- Goal: transport segments (TCP, UDP) between hosts
  - un-guaranteed transport
- Sending a segment
  - encapsulation in a datagram
- Receiving a segment
  - de-encapsulation of the segment
  - delivery of the segment to the transport layer
- At the core of the network
  - in all routers
  - routers manipulate IP headers

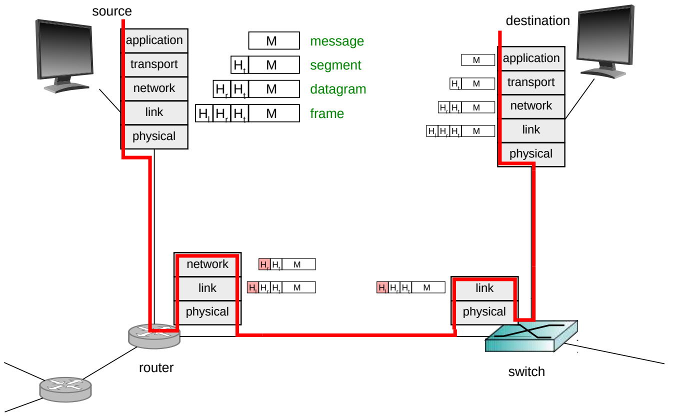


Router: reminders



- Router: routes/transfers packets
  - forwarding table
- Router: stores and forward
  - unit of transfer: a packet (datagram)
  - receives a whole packet
  - then, forwards it
- Queuing (buffering) if the output link is busy

Routing and encapsulation: reminder

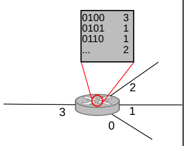


Datagram Network

- No concept of connection in the network
- Routers ignore (transport layer) connections
- Send only datagram
- Packet transfer based on the destination IP address
- Transfer using forwarding tables
  - association IP → interface (link)
  - created by routing algorithms

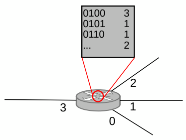
Forwarding Tables: address ranges

IP Address Range	Interface
from 11001000 00010111 00010000 00000000 to 11001000 00010111 00010111 11111111	0
from 11001000 00010111 00011000 00000000 to 11001000 00010111 00011000 11111111	1
from 11001000 00010111 00011001 00000000 to 11001000 00010111 00011111 11111111	2
default	3



- 4 millions possible addresses ⇒ using ranges
- Default destination

Forwarding Tables: longest prefix



IP Address Range	Interface
11001000 00010111 00010*** ********	0
11001000 00010111 00011000 ********	1
11001000 00010111 00011*** ********	2
******** ******** ******** ********	3

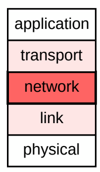
- Representing ranges as prefixes
  - routing using the longest-common-prefix rule
- Often, storing as a prefix-tree (or *trie*)
  - faster lookup and retrieval
  - more compact representation

Datagram Network and Alternatives

- Internet (datagram)
  - exchanges between computers (hosts)
  - best effort: no guarantees, no hard constraints
  - heterogeneity (a lot of different types of links)
  - heterogeneity → difficulty to offer a uniform service
  - intelligent hosts (computers)
    - can adapt, do verification, correct errors
    - complex hosts ⇒ simpler network
- ATM (Asynchronous Transfer Mode, network of virtual circuits)
  - inspired by phone infrastructures
  - hard constraints
    - strict delays
    - guaranteed transfer
  - simplistic terminals (phones)
    - complexity moved inside the network
- NB: we won't cover virtual circuit networks

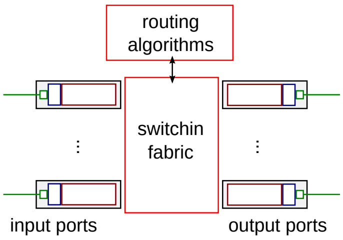
Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast



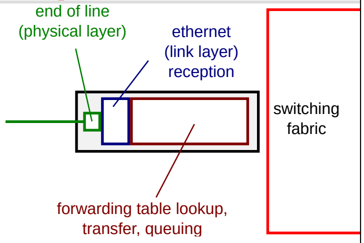
Router: role and architecture

- 2 roles
  - forward packets
    - using tables
    - very fast and often using dedicated hardware circuits
  - execution of routing algorithms
    - updating the forwarding tables
    - often implemented in software
- Structure
  - forwarding tables in input ports
  - switching fabric (transfer circuit)
  - buffering in output ports



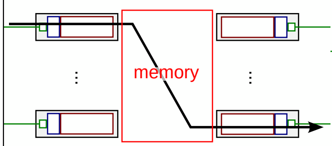
Router: input ports

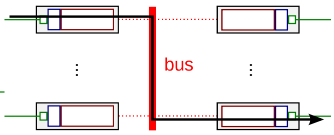
- Bit reception
- Ethernet frame reception
- Packet transfer
  - based on the destination address
  - using the forwarding table (cached in the port)

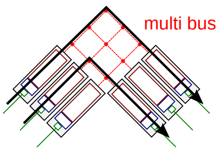


- Goal: process packets at least at the input link speed
- Buffering
  - if the switching fabric is busy
  - if the switching fabric is too slow

Router: switching fabric



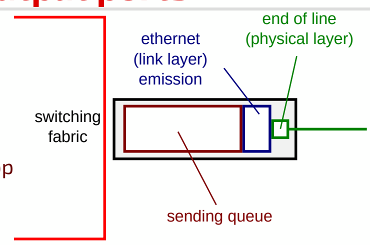




- Role: transfer a packet from an input to an output queue
- Transfer rate?
  - ex: 5 input ports at 1 Gb/s, packet size of 1.5 kB
- 3 main types

Router: output ports

- Queuing, if the line is busy
- Sequencing
  - choice of the next packet to send
  - choice of the packets to drop
  - tune-able policy (not necessarily FIFO)



- Ethernet frame emission
- Bit emission

Router: queuing situations

- Output capacity overload
  - delays while the previous packet is sent
  - creation of a queue in the output port
  - packet loss
- Switching fabric overload
  - too many simultaneous inputs
  - creation of queues in the input ports
- Head-of-the-Line (HOL) blocking
  - one of the output queue O, is full
  - a packet from an input port I should go to O
  - the packet blocks the input port I
  - all other packets in this port I are blocked
  - buffering occurs in input port I

Router: choosing buffer sizes

- Rule of the thumb
  - RFC 3439
    - buffer size =
    - example: 200ms 1Gb/s
    - pessimistic
  - More recent recommendation, with input ports
    - buffer size =

Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast

application
transport
network
link
physical

Internet Protocol: headers

Bit	0 1 2 3	4 5 6 7	8 9 10 11 12 13	14 15	16 17 18	19 20 21 22 23 24 25 26 27 28 29 30 31
0	Version	IHL	DSCP	ECN	Total Length	
32	Identification				Flags	Fragment Offset
64	Time To Live		Protocol		Header Checksum	
96	Source IP Address					
128	Destination IP Address					
160	Options (if IHL > 5)					
	Data (UDP Segment, TCP Segment, etc) … … …					

- Total length: size in bytes
- Time to live (TTL): number of routers before dropping a packet
- Protocol: protocol used in the transport layer
- Source IP and destination IP address: 32-bit addresses of hosts
- Fragment…: split of a single datagram in several datagrams

IPv4 Addressing

- IPv4 address
  - 32 bits
  - identifies a network interface (network adapter on a host, port on a router, …)
  - unique in the network (unless there is some NAT)
- IPv4 addressing: example
  - 223.1.1.1 (4 bytes)
  - `11011111 00000001 00000001 00000001` (binary)
- Interface
  - connection between a host (or router) and a physical link
  - hosts often have one or more interfaces at a given instant (wifi, cable, loopback, …)
  - routers most of the time have multiple interfaces
- Important: each interface has an IP address

Network, sub-network and connections in a sub-network

- Sub-network
  - group of addresses
  - physical access without a router
  - common most significant bits
- CIDR: Classless InterDomain Routing
  - subnet mask
  - example: `192.168.10.0/24`
    - 24-bit subnet
    - 8-bit for hosts in the subnet
  - example: `223.1.16.0/23`

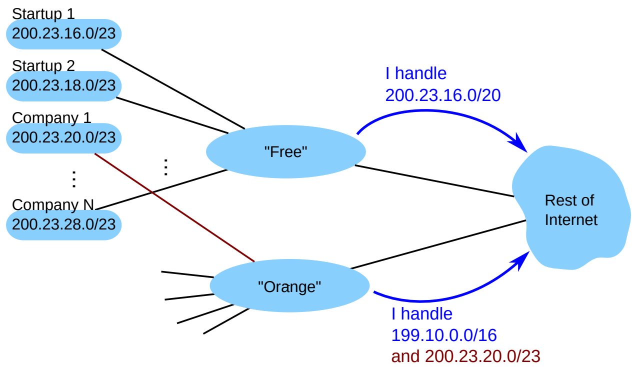
How to obtain an IP address? What entities and/or machines are involved in giving you your IP?

Dynamic Host Configuration Protocol

- Goal: dynamic allocation of IP addresses (by a server)
  - IP address allocation
  - bail renewal
  - reuse of unused addresses
- Principles
  - the client broadcasts "DHCP discover"
  - the server answers with an address in a "DHCP offer"
  - the client asks the address with "DHCP request"
  - the server acknowledges with "DHCP ack"
- Other information sent by the server
  - gateway address
  - name and IP of DNS servers
  - subnet mask
- DHCP uses UDP

How does the client communicates with the DHCP server? How to use UDP without IP address?

Hierarchical Addressing and Aggregation



- Efficient and compact routing handling
- Based on the longest common prefix principle

Allocation of IP Ranges

- How do an Internet Service Provider (ISP) get its IP addresses?
- ICANN: Internet Corporation for Assigned Names and Numbers
  - allocation of IP addresses
  - management of DNS
  - allocation of domain names
  - conflict resolution
- Typical situation: quick summary
  - the ICANN allocates IP ranges to ISPs
  - ISPs allocates sub-ranges to their clients
  - a DHCP server is configured for a given address range
  - the DHCP server allocates addresses dynamically to the hosts

NAT: Network Address Translation

- Motivation: many local networks with multiple devices
  - a single address given by the ISP (not a big range)
  - freedom to reorganize a local network
  - freedom to change ISP without changing the local network
  - make local hosts invisible from the outside (security)
- Generic principle, having an external IP **ext-IP**
  - on connection from **int-IP1:port1**, select a free external port **port2**
  - change source in outgoing messages **int-IP1:port1** → **ext-IP:port2**
  - change target in inbound messages **ext-IP:port2** → **int-IP1:port1**
  - and more
    - static port forwarding to allow connections from outside
    - various "optimization"
    - traversal: ICE, STUN, TURN, …
- Private subnets
  - `10.0.0.0/8` ; `172.16.0.0/12` (255.240.0.0) ; `192.168.0.0/16`
- Problem?

ICMP: Internet Control Message Protocol

- Used between hosts and routers
  - "echo", used by the `ping` command
  - used also by `traceroute`
  - designed for error reporting
- ICMP Message
  - on top of IP
  - content
    - type
    - code
    - and the first 8 bytes of the datagram that caused the error

```
Type  Code  description
0       0       echo reply (ping)
3       0       dest. network unreachable
3       1       dest host unreachable
3       2       dest protocol unreachable
3       3       dest port unreachable
3       6       dest network unknown
3       7       dest host unknown
4       0       source quench (congestion
                  control - not used)
8       0       echo request (ping)
9       0       route advertisement
10      0       router discovery
11      0       TTL expired
12      0       bad IP header
```


Traceroute and ICMP

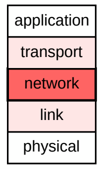
- Successive emission of UDP segments
  - with a random port
  - first with TTL = 1
  - second with TTL = 2
  - third with TTL = 3
  - …
- When datagram n arrives at router n
  - the router drops the datagram
  - the router sends an ICMP(type=11, code=0) (TTL expired)
  - the message contains the router address
- RTT = time between UDP emission and ICMP message reception
- Stopping criterion
  - we reached the destination
  - materialized by ICMP(type=3, code=3) (dest. port unreachable)

IPv6: quick overview

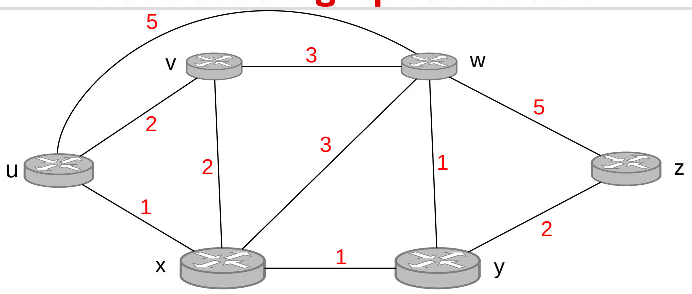
- Main motivations
  - **no more IP addresses…**
  - IPv4 is sub-optimal at the header level (options, …)
- Differences with IPv4
  - 128-bit addresses (instead of 32 bits)
  - fixed-size headers (40 bytes)
  - no fragmentation allowed
  - ICMPv6
    - "packet too big"
    - multicast groups
  - no more checksum
- Transition
  - progressive transition
  - cohabitation of the two protocols
    - IPv4 tunnels to connect IPv6 islands
    - IPv6 datagram inside a IPv4 datagram
  - slow adoption

Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast

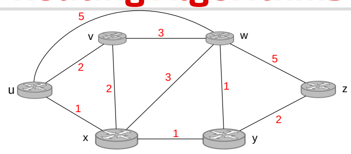


Abstraction: graph of routers



- Cost: different cases
  - constant to 1
  - inverse of the bandwidth
  - depending on the congestion
  - combinations of these
- Goal: find the shortest path

Routing Algorithms

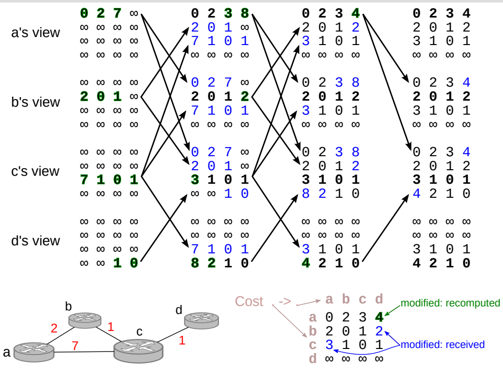


- Dijkstra: link-state (wikipedia: Dijkstra's algorithm)
  - global algorithm
  - requires knowledge about the complete network topology
- Distance Vector
  - distributed/decentralized algorithm
  - knowledge of the neighbors only
  - iterative process
- Periodic updates
  - necessary when the topology changes
  - necessary when costs change

Distance Vector (DV)?

- Each node estimates the optimal distance
  - **from**: itself and each of its neighbors
  - **to**: all nodes
- Each node communicates
  - to each of its neighbors
  - a vector of distances
  - **from**: itself
  - **to**: all (known) nodes
- Message-passing algorithm
  - distributed optimization, iterative process
  - message exchange between neighbors
  - on reception by x of a DV from a neighbor
    - $\forall y, D_x(y) = \min_v(c(x,v) + D_v(y))$
    - + in case of change, notification to all neighbors
- Example with 4 nodes: a, b, c, d

Distance Vector (DV)



- on reception by x of a DV from a neighbor
  - $\forall y, D_x(y) = \min_v(c(x,v) + D_v(y)) \Rightarrow$ notification on change

Computer Networks 4: Plan

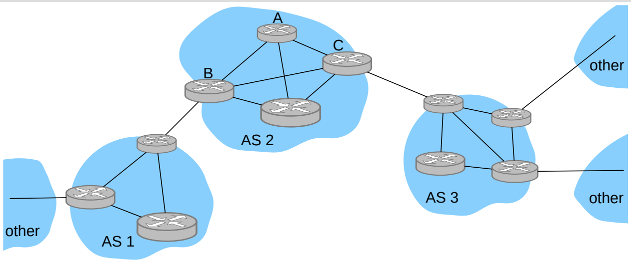
- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast

application
transport
network
link
physical

Hierarchical Routing and Autonomous Systems (AS)

- Routing algorithms
  - do not work with 2^{32} addresses
  - too many entries, too much data exchanged (size of DVs)
- Using address ranges
- Internet = network of networks
  - autonomy of each network
  - control of the internal routing
- Aggregation of routers in regions or Autonomous Systems (AS)
  - internal routing algorithms: intra-AS protocols
  - gateway router at the border of an AS
  - global routing algorithms: inter-AS protocols
- Forwarding tables = mixed of info from different algorithms
  - intra-AS ⇒ table for internal destinations
  - inter-AS ⇒ table for remote destinations

Routing to Another AS: inter-AS protocols



- Router A from AS2 receives a packet for a remote destination
  - to which router should it be forwarded? B or C?
- Inter-AS protocol: AS2 must
  - learn what addresses can be reached via AS1 and AS3
  - propagate this information, internally and externally
- When multiple possibilities, select the faster AS-exit

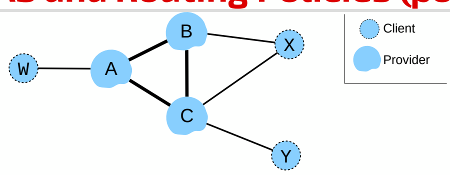
Routing Protocols for Internet

- Inside an AS (intra-AS), Interior Gateway Protocols
  - RIP: Routing Information Protocol
    - distance vector, with a cost of 1 for each link
  - IGRP: Interior Gateway Routing Protocol (kind of DV, by Cisco)
  - OSPF: Open Shortest Path First, link-state (Dijkstra)
  - IS-IS: Intermediate System - Intermediate System (Dijkstra)
- Between AS (inter-AS), Border Gateway Protocol (BGP)
  - cement between multiple AS
    - interconnection
    - routing policies (network management)
      Do we accept to transfer traffic for another given entity?
  - eBGP
    - exposes to another AS the accessible prefixes
      ⇒ promise that we will forward these packets
    - receives (from another AS) what prefixes are accessible
    - exchanging "routes": prefixes + list of AS to traverse
  - iBGP
    - propagates this information inside an AS

Intra-AS Differences and Details

- RIP: Routing Information Protocol
  - distance vector, constant cost of 1
  - sends DV every 30 seconds
  - send at most 25 sub-networks (prefixes)
  - uses UDP, executable: `routed`
- IGRP: Interior Gateway Routing Protocol
  - proprietary
  - distance vector
  - advertises every 90 seconds
  - hop count of 100
  - multiple metrics
- OSPF: Open Shortest Path First
  - link-state (Dijkstra)
  - all routers of the AS learn the topology
  - directly over IP
- IS-IS: Intermediate System to Intermediate System
  - very close to OSPF

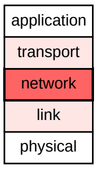
Inter-AS and Routing Policies (political)



- X has two providers (redundancy, failure safety)
  - X does not want to be a "pipe" between B and C
  - ⇒ X do not advertise itself as a route between B and C
- Advertisement propagation
  - A advertises route AW to B
  - B advertises route BAW to X
- B advertises route BAW to C?
  - no! it paid to provide a service to its clients, not to others
  - forces the C → W traffic to pass through A directly

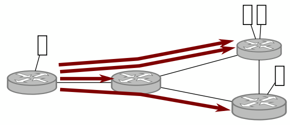
Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast

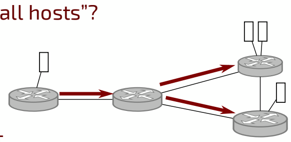


Broadcast and Multicast

- Motivation
  - send a message to all hosts in a network
  - or, to a set of hosts (multicast)
- Duplication by the source
  - explicitly sending to all hosts
  - inefficient…
  - who is "all hosts"?



- Duplication by the network
  - efficient and operational
  - problems, issues?



Approaches for Broadcast

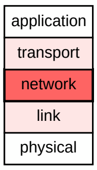
- Flooding
  - each router sends a copy of the packet on each link
  - problems with cycles and flooding (too many copies)
- Controlled flooding
  - a node only broadcasts a message the first time
  - need to keep the message identifier for each broadcasted message
  - or, *reverse shortest path*, does the message arrives from the shortest path?
- Use of a spanning tree
  - tree: no cycles
  - spanning: covers all nodes
  - to be created automatically
- IP Broadcasting (IPv4)
  - an address representing all hosts
  - "all ones" host
  - e.g. `192.168.255.255` when in a subnet `192.168.0.0/16`
  - special case: `255.255.255.255` means only this network

IP Multicasting

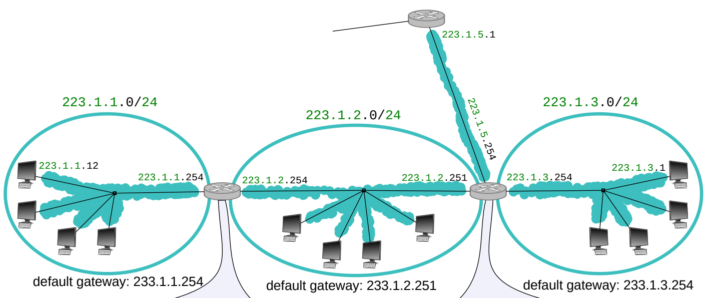
- `224.0.0.0/4`
  - set of multicast reserved addresses
- `224.0.0.0/24`
  - only for the local network
  - routers should not forward these
- `224.0.1.0/24`
  - some standard protocols (PTP, service discovery, …)
- `239.0.0.0/8`
  - private in an organization
  - blocked by routers at the frontier
- and more…

Computer Networks 4: Plan

- Goal: network layer
  - understand the principles of the network layer
  - understand their implementation inside Internet
- Overview
  - Network layer: role, transfer, routing
  - How a router works: forwarding, buffering, …
  - IP: Internet Protocol
  - Routing algorithms
  - Routing in Internet
  - Broadcast and multicast



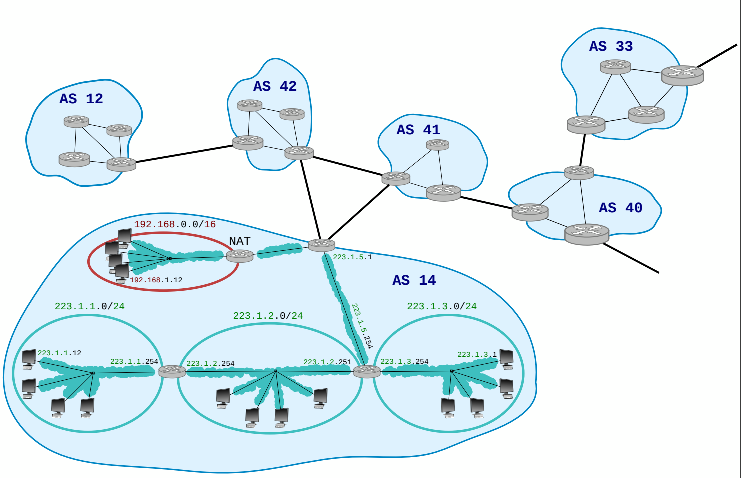
Network, Subnets, Routing, Gateway



destinations	interface	gateway
223.1.1.0/24	223.1.1.254	223.1.1.254
223.1.2.0/24	223.1.2.254	223.1.2.254
223.1.3.0/24	223.1.2.254	223.1.2.251
0.0.0.0/0	223.1.2.254	223.1.2.251

destinations	interface	gateway
223.1.3.0/24	223.1.3.254	223.1.3.254
223.1.2.0/24	223.1.2.251	223.1.2.251
223.1.1.0/24	223.1.2.251	223.1.2.254
0.0.0.0/0	223.1.5.254	223.1.5.1

Network, Subnets, Routing, Gateway



End Of Part