Routing

(/udd/bcousin/ITI-Caire/Cours/1.Routing-eng.fm- 17 April 2002 14:38)

PLAN

- Introduction
- Addressing
- Routing
- Conclusion

Bibliography:

- Cnristian Huitema, Routing in the Internet, Prentice Hall, 2000.

1. Introduction

Routing is the most characteristic function of Network layer (ISO Layer 3):

- enabling the forwarding of data packets from sender to destination(s)
- through a network:
  - . hosts
  - . routers
  - . links

• Data forwarding is the network function

• Routing is a network management function, but essential

• Addressing is one among numerous network management functions which is generally required to assure routing and forwarding

2. Address

2.1. Introduction

Network layer uses address to identify network systems (hosts and routers).

- IP forwarding process uses address to determine the next hop
- Address criteria
  - address uniqueness
    - . unique association between an address and a network system
  - address flexibility
    - . address can identify many kinds of network entities (hosts, services, etc.)
  - address extensibility
    - . addressing system must manage large(r) network

Address characteristics:

- address structure
- address field semantic

2.2. Address structure

Many different address structures exist:

- structured/non-structured
  - . structured: IP address = address class+netid+hostid
  - . non-structured: Ethernet address (6 bytes number)
- fixed/variable field: field length, field position in address
  - . fixed: IP address length = 32 bits
  - . variable: IP hostid = variable in length and position
- hierarchical structure: address, fields, sub-fields, etc.

2.3. Address field semantic

Many different semantics can be given to address fields:

- none
  - . simple identification number
- functional
  - . service identification
  - . e.g.: multicast IP address = 224.0.0.9 (RIPv2 routers), TCP port number
- geographic
  - . system location
  - . e.g.: national telephone prefix - 202 = Egypt

2.4. Address characteristics

Address characteristics versus network functions:

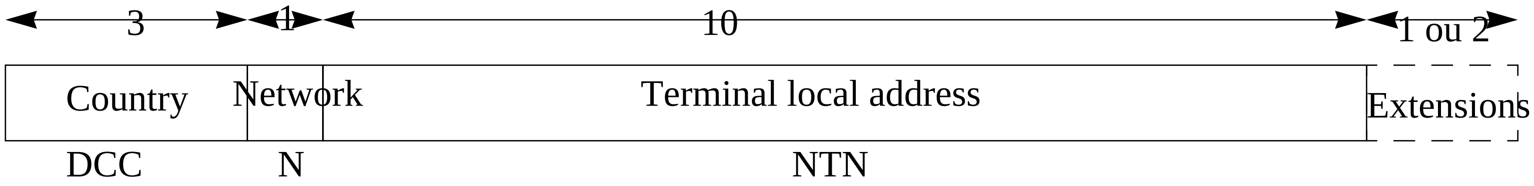
- Routing
  - ☞ hierarchical addresses with geographic semantic
- Mobility
  - ☞ non geographic semantic
- Broadcasting (multicasting)
  - ☞ non geographic semantic
- Address density
  - + unstructured address

2.5. One ISO address format

■ X.121 address format: public network for digital data transmission (80's)

Hierarchical addressing with geographic and functional semantics

14/16 half bytes (DCB representation: one digit = 4 bits)



Reserved prefixes: 0=E.163 (telephone), 1=reserved, 8=F.69 (telex), 9=E.164 (ISDN)

- DCC field ("data country code")
  - e.g. US = [310-329], France = [208-212]
- N field ("network code"): no more than 10 networks
  - e.g.: Transpac = (208)0
- NTN field ("network terminal number"): 10 digits (optionally 12).
  - internal structure of NTN field is locally determined by network administrator
- E.g.: 20 80 11 42 07 19 68

2.6. Internet address

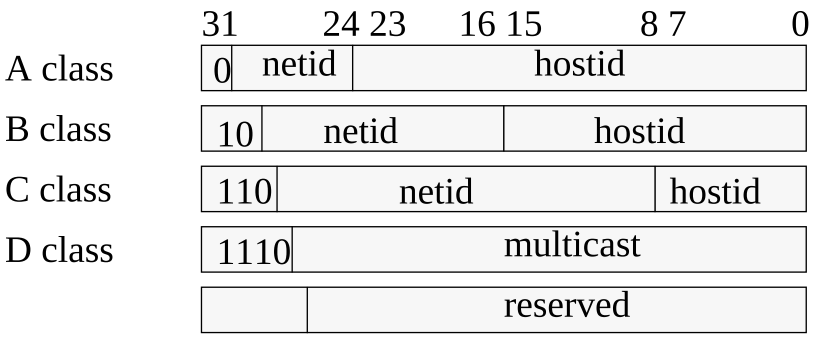
IPv4 address: fixed length of 32 bits.

Hierarchical address of 2/3 fields:

- (address class,) netid, hostid

3 unicast address classes:

- variable length and position of netid and hostid fields.



Eg : 146.84.32.07
B class, netid = 146.84, hostid = 32.07

1 multicast address class:

- station group identifier

CIDR (Classless InterDomain Routing)

- address classes A, B, C have been dropped
- IP addresses are structured on a bit basis by:

- prefix address = netid + subnetid
- suffix address = hostid

• E.g.: 131.254.8.76 --> “Armor” team subnet of “IRISA” net: 131.254.8/22 + “pondichery” host: 76

• free IP address exhaustion

• IP address aggregation in routing tables

IPv6 “new generation”:

• new packet format

• new address format

- address length = 128 bits!
- anycast type address (+ unicast type+ multicast type)

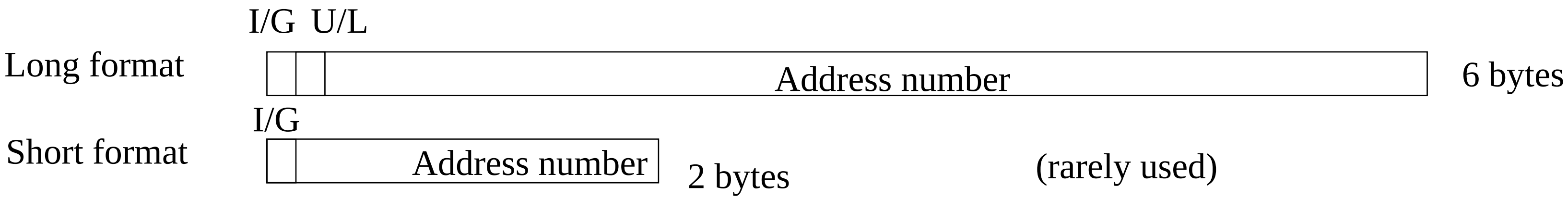
2.7. IEEE 802 address

Local Network address (Medium Access Control layer): IEEE 802 standard

- Ethernet, Token Ring, Token Bus, FDDI.

Unstructured address

2 formats:



- I/G bit: individual or group address
- U/L bit: universal or local address

Reserved address:

- Broadcast = 0xFFFFFFFFFFFF

Address allocation (by the network card manufacturers):

- 3 first bytes: manufacturer identification number (Sun= 080020, 3COM= 02608C, etc.)
- 3 last bytes: serial number of the network card

2.8. Addressing

Address processing

- address allocation: uniqueness
  - . network management
- address decoding: efficiency
  - . datagram forwarding

Address allocation

• Each entity (company, university, etc.) would like to manage its addressing without interference of outside constraints.

- internal choices: topology, organizational, functional, political, etc.

But

• Addressing must be globally coordinated

- to assure uniqueness and efficiency during datagram forwarding

☞ one (autonomous) address domain is attributed to each entity

☞ hierarchical structure of address + delegation

2.9. Address domain

Each address domain is autonomously managed by an entity which acts as authority of the address domain.

A hierarchical tree of address domain can be built.

Each address domain authority knows:

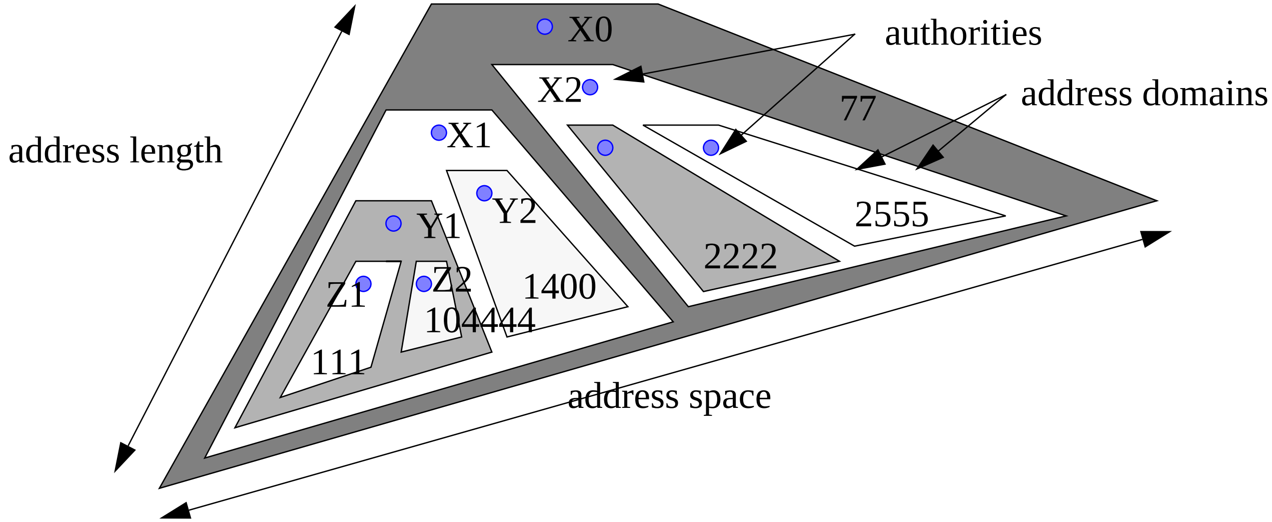
- the address set inside its address domain
- the authority name of its upper address domain
- the authority names of its address sub-domains (if they exist)
- the address partitions which are relevant to these sub-domains

Address domains are hierarchically mapped using prefixes and suffixes:

- prefixes identify the address domain (its authority)
- suffixes identify objects managed by the address domain
  - either a concrete object
  - or the authority of a sub-domain from which the object management have been delegated

Generic example:

```
x∈{1*} --> X1 authority
    x∈{1[0-2]*}--> Y1 authority
         x∈{10*}--> Z2 authority
         x∈{1[1-3]*}--> Z1 authority
    x∈{1[4-9]*}--> Y2 authority
x∈{2*} --> X2 authority
x∈{[3-9]*]} --> X0 objects
```



Concrete examples:

- Telephone numbering
- IP addressing
- Internet name space (Domain Name System)

2.10. Conclusion

Hierarchical address domain

- unique address allocation: strict partition of sub-domain
- efficient decoding: local processing using prefix knowledge

Delegation

- each network entity (national country, ISP, company, etc.) can choose its own addressing plan, with few coordination
- length, semantic and location of (sub-)fields are very flexible and variable

3. Routing

3.1. Introduction

Main function of Network layer is to transmit data packets from sender to destination(s). Packet travels through a network based on routers.

At each packet reception, router selects the next router where the packet has to be forwarded. The selected router is the next step on the (shortest) path to the destination.

In each router, the selection is done based on

- packet header information
  - . address but also virtual channel identification, label, etc.
- routing table
  - . routing information: next step on the (shortest) path to destination

3.2. Routing functions

Routing is built up of 2 functions:

- Packet forwarding
- Routing table update

Packet forwarding processing phases:

- Datagram reception and header processing
- Routing table lookup using IP destination address
- Datagram forwarding

Routing table update

- Network topology changes
  - . numerous station, router and link insertions or failures, anytime
  - . automatic management
- routing table update protocol(s)
  - . routing messages

Routing table update protocol --> Routing protocol

3.3. Packet forwardings

Packet forwarding could be based on:

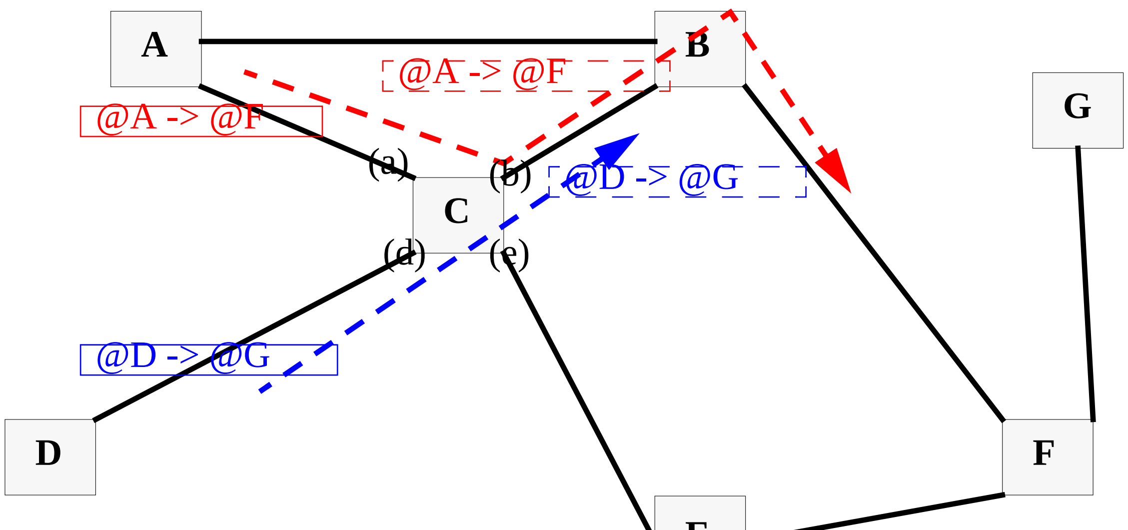
- datagram forwarding
  - . forwarding is determined on a datagram by datagram basis.
- virtual channel forwarding
  - . forwarding is determined by the Virtual Channel information.

3.4. Datagram forwarding

Each router has an updated routing table. For each destination, the routing table entry points to the next router.

On packet reception, router performs:

- destination address extraction from the packet header,
- routing table lookup, and extraction of the next router address from the selected entry,
- packet forwarding to the next router



datagram frowarding

C routing table

destination address	next routeur	additional information
@A	@A	(a)
@G	@B	(b)
...		
@F	@B	(c)
@X	@E	(e)
...		

3.5. Datagram forwarding characteristics

Datagram forwarding:

- No connection phase, no disconnection phase
- The number of entries in routing table is function of the number of destinations
- Routing table is modified only on topological change
- Next router is determined on a datagram basis:
  - Successive datagrams coming from the same sender going to the same destination could follow different paths
  - Packet ordering are not assured

Network example

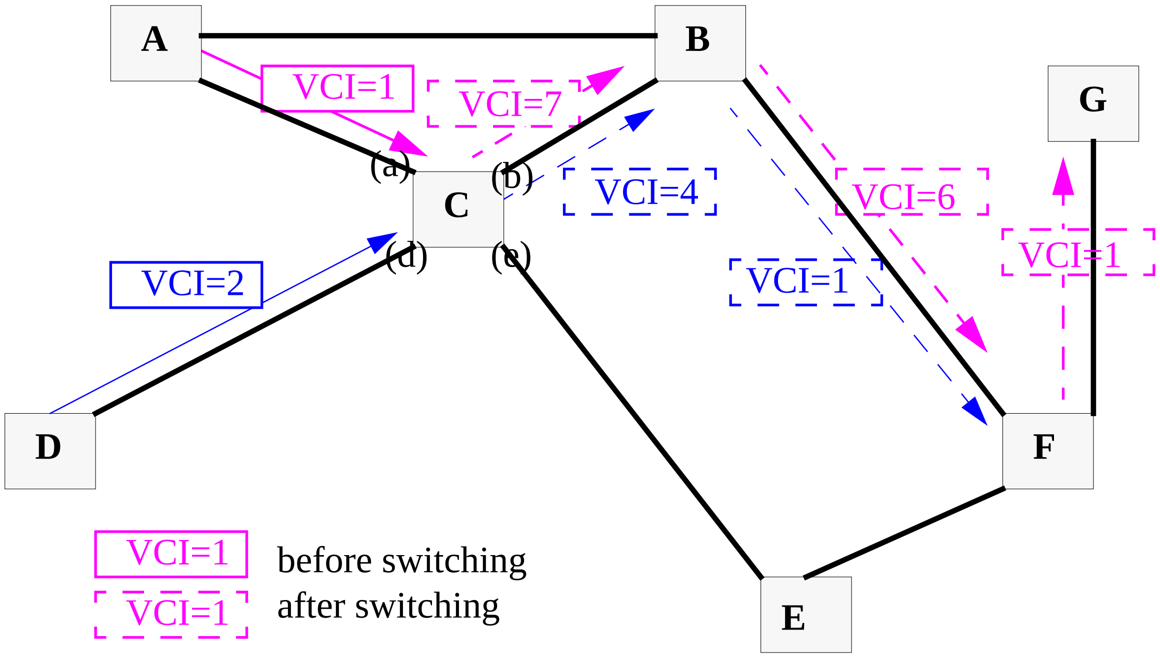
- IP: Internet protocol

3.6. Virtual channel

Virtual channel (VC) links sender to destinations. Each VC segment (between 2 adjacent switches) is identified by a number (VCI : VC Identifier).

On packet reception, switch performs:

- VCI extraction from the packet header,
- switching table lookup, and extraction of the output link and next VCI from the selected entry,
- packet forwarding to the output link with the new VCI



C switching table

input		output	
link	VCI	link	VCI
(a)	1	(b)	7
(b)	1	(a)	1
(b)	2	(d)	1
(d)	1	(b)	2
(d)	2	(b)	4
(e)	4	(d)	2

3.7. Virtual Channel characteristics

Virtual channel switching:

- Routing equipment is called “switch”
- VCI are allocated during connection phase. They are freed during disconnection phase.
- Number of entries in switching table is function of the number of virtual channels using the switch.
  - VCI field could have short length
- Switching table is modified, every time a new virtual channel is set
- If the network topology is changed (link or router failure), many virtual channels could be incorrect.
- During connection phase, a switch use routing table to determine the next link
- Switching table is called “hard state” information
- Packet ordering is naturally kept by virtual channel
- All packets inside a virtual channel follow exactly the same route

Network examples:

- X25, frame relay, ATM, etc.

3.8. Routing types

Routing table update!

- Two types of routing:
  - Static routing
    - no routing modification without human intervention
  - Adaptive routing (dynamic, automatic):
    - automatic routing adaptations based on network topology modifications, or more rarely, network traffic variations

3.9. Routing algorithms

- centralized:
  - routing data are processed in one central node
  - the central node is a single point of failure
  - all routing traffic goes to and comes from the central node: point of congestion
- distributed:
  - routing data are processed in several nodes (generally all the routers).
  - routing data synchronization between routers is required
    - . current solution: the same algorithm processes the same data on every router
    - . theoretically true but practically impossible:
      - ☞ Internet routing instability (TTL, tolerant applications to instability, etc.)

3.10. Routing data

- Routing data Location
  - local to routing node
    - . immediate access to data but rough information
  - from distant node, transmitted to routing node
    - . message transmission delay: data incoherence.
- Routing data Type
  - topology information
    - . link or node existence, link or path nominal bandwidth, economic cost, etc.
    - ☞ the most frequently used type of routing data
  - traffic information
    - . link or path available bandwidth, link or path delay, link or path error rate, buffer occupancy, etc.
    - ☞ Traffic Engineering
  - application requirements
    - . data flow QoS (bandwidth, delay, jitter, error rate, economic cost, etc.)
    - . ☞ Connection Access Control

3.11. Main properties of routing algorithms

- Accuracy
- Fault tolerance
- Stability
- Fairness
- Efficiency
- Complexity

☞ Good routing algorithm is a trade-off between all these properties.

Routes must fulfill:

- application requirements (individual criterion)
- optimum network resources management (global criterion).

3.12. Multiple paths routing

❑ The shortest path notion is based on evaluation criterion (generically called metric, or distance, or cost):

- geographic distance, hop count, throughput, delay, error rate, financial cost, etc.
- separate criteria could lead to opposite routes
- several routing tables could be used to take into account the different routes from different criteria.

❑ Several routes to the same destination:

- each routing table entry could have a list of next routers
- route with exactly the same cost
  . load balancing
- second choice route
  . alternate route in case of failure

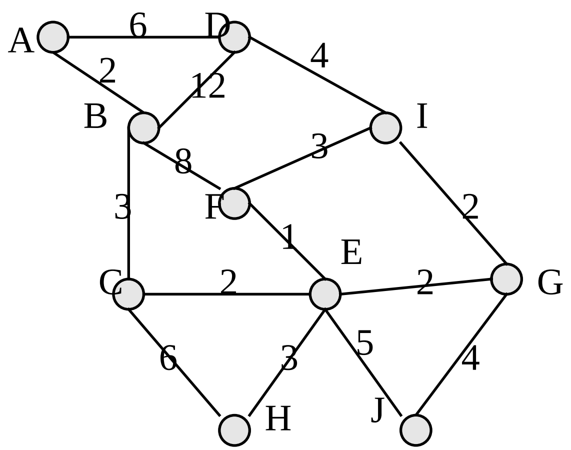
3.13. Shortest path first algorithm

Centralized algorithms

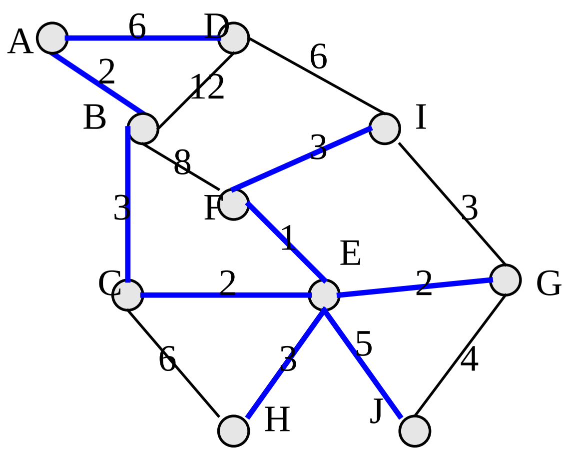
- positive valuated graphs [Dijkstra (1959), Moore (1957)],
- matrix representation [Floyd],

• Determination of shortest paths from a node to every node

- to build a tree whose root is the initial node
- to append the link which connects a new node and which has the shortest path



initial graph



shortest path tree from root A

3.14. Some simple routings

- No routing table!
- Use few or none local information!

3.14.1 Broadcasting by flooding

- received packet are sent to all output links but its input link
  - simple, robust, fast (find all routes hence the shortest),
  - heavy overload
  - multiple occurrence of the same packets in cyclic graph
  - example: path explorer of Source Routing

3.14.2 Random routing

- random selection of the next links
  - selective flooding
  - simple, same inconveniences as flooding, less of overload than flooding
  - but random performance.

3.14.3 Hot potatoes [Baran, 1964]

- next link selection based on minimum delay:
  - . e.g. buffer occupancy.
  - - local optimum /= route (global) optimum.

3.14.4 Backward learning

- Route learning uses information extracted from data packets.
  - - sender address in the data packet header

    ☞ sender could be reached using the input link (backward)
- no overhead traffic: no routing message
- but inference information is not stable
- information lifetime is limited
- example: Transparent bridging

3.15. Distance vector algorithm

- Adaptive and distributed routing

Synthetic routing data transmission between neighbor routers.

Principle:

- every node knows all its neighbors,
- every node sends periodically routing message with a list of all shortest path destinations
    . routing message is a distance vector where every entry is a distance to a destination
- On routing message reception, router learns new destinations or shortest routes.

Example:

- Internet RIP (“Routing information protocol”)

Simple routing, limited scalability, mono-criterion (“hop-count”).

3.16. Link state algorithm

❑ Adaptive and distributed routing

Two phases protocol:

- topological data broadcasting between all routers
- local processing of routes using topological knowledge.

Principle:

- Broadcasting of all the local links and neighbors:
  . all routers get the same view of the current network topology
- local processing of routes from the router to all destinations:
  . e.g.: Shortest path first algorithm
  . several routes to the same destination, using several criteria, could be processed

Examples:

- OSPF: Open Shortest Path First (Internet), IS-IS (ISO)

3.17. Scalability

Routers have to process thousands packets per second

Datagram forwarding optimization:

- fast hardware
- algorithm complexity depends of the routing table size
  - search algorithm complexity: n log(n)
  - n: number of destinations inside the network --> some billions!

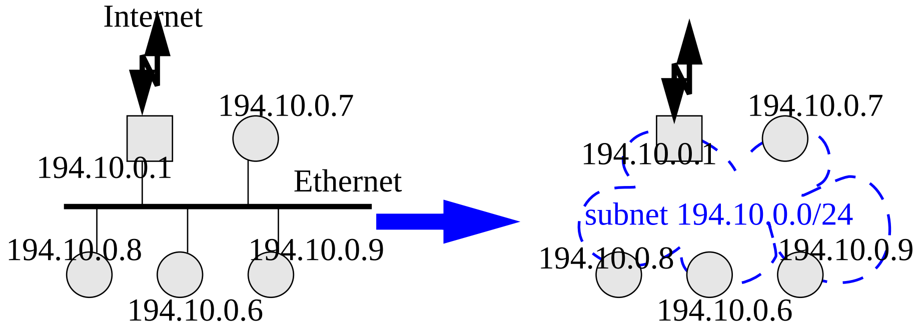
Solutions:

- Aggregation
- Hierarchy

3.18. Aggregation

• Subnetworking

- every host with same address prefix belongs to the same subnet ("netid")
- e.g.: any host on a Ethernet LAN could belong to the same subnet



• CIDR ("Classless Inter Domain Routing")

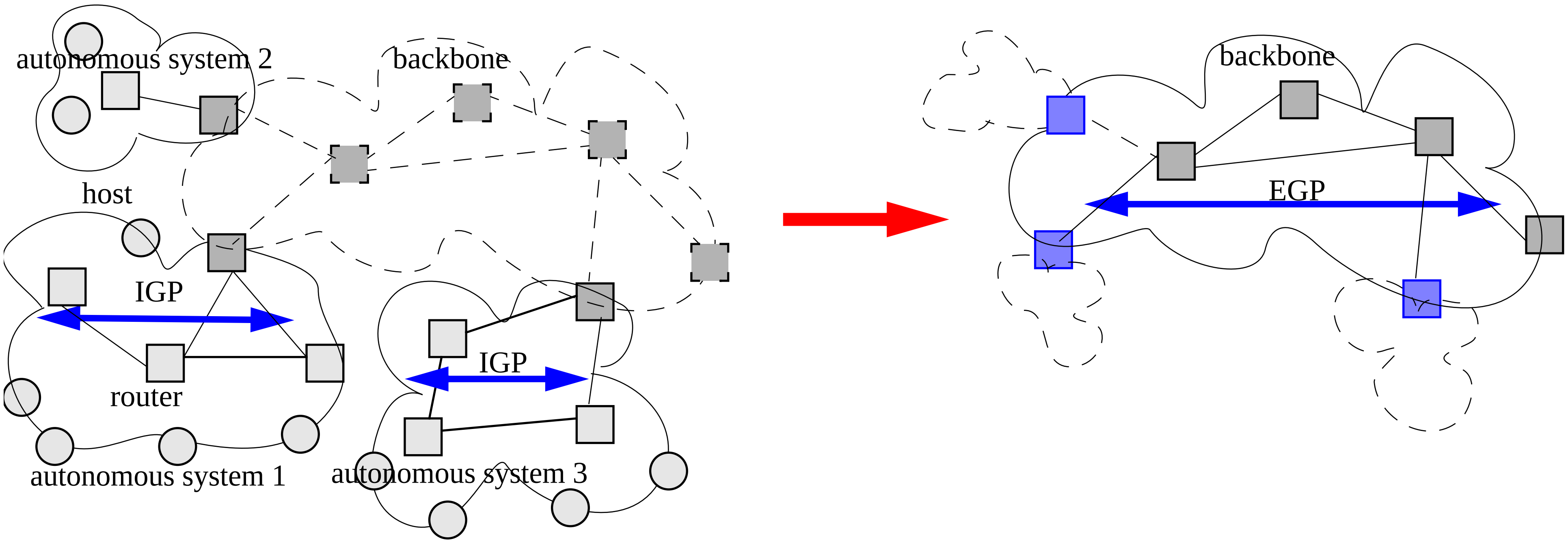
- extension of the previous concept
- one routing table entry for all host destinations with consecutive IP addresses and same next router

  . european prefix: 194.0.0.0/7 = [194.0.0.0, 195.255.255.255]

3.19. Hierarchical routing

2 (main) hierarchical routing levels:

- Interior Gateway Protocol
  - routing data are transmitted between routers inside a routing domain
  - every router inside a routing domain knows the route to reach every destinations inside the same domain
  - Interior routers process IP addresses
  - e.g.: RIP, OSPF, etc.
  - routing domain == “autonomous (routing) system”
- Exterior Gateway Protocol
  - routing data are transmitted between backbone routers
  - backbone routers use generic informations: #A.S.
  - e.g.: BGP, EIGP, etc.



Note: some protocols have a supplementary routing level

- “routing area”: routing sub-domain
  - . e.g. OSPF

4. Conclusion

- Routing
  - Routing: Network layer essential
  - Addressing: complementary function
- Packet forwarding
  - Datagram forwarding
  - Virtual Channel
- Routing table
  - routing table update
  - routing table update protocols
  - routing messages
- Routing protocols
  - flooding
  - Distance vector algorithm
  - Link state algorithm.