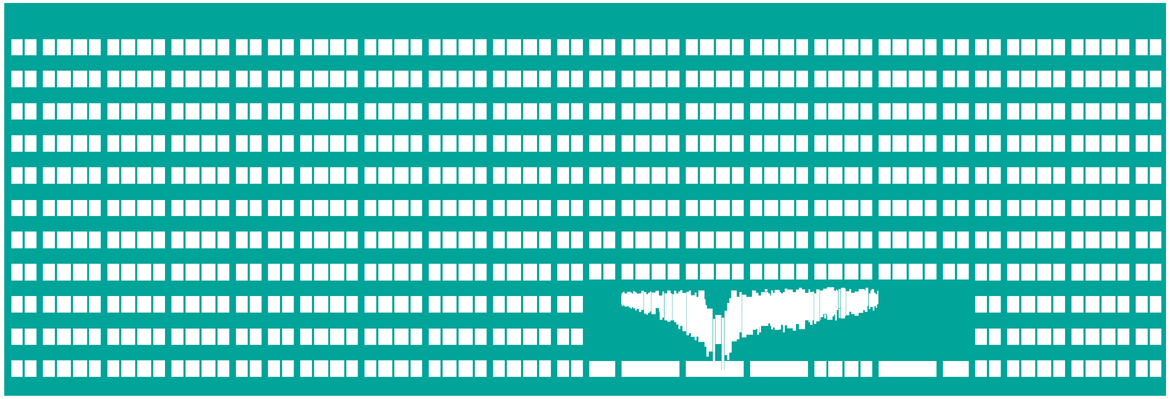
VSB TECHNICAL UNIVERSITY OF OSTRAVA | FACULTY OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE | DEPARTMENT OF COMPUTER SCIENCE

Routing, Routing Algorithms & Protocols



Computer Networks
Lecture 6

Circuit-Switched and Packet-Switched WANs

Circuit-Switched Networks

- Older (evolved from telephone networks), a node has to explicitly ask the network to establish circuit to the partner node and tear it down
  - A network returns a circuit ID if a node may establish multiple circuits at the same time
- Capacity is reserved for the whole existence of circuit
  - as all data travel through the single circuit, there is no risk of ordering data out-of-order
- Advantageous for users – guaranteed QoS
- Not economic for net. operator in case of bursty traffic
  - The channel is not utilized between data bursts
- In case of the circuit outage the node has to ask creation of the new circuit
  - A network has to find an alternate path that avoids usage of the failed paths/intermediate nodes
  - A circuit-establishment procedure can be time consuming (e.g. handshake of analog modems)

Packet-Switched Networks

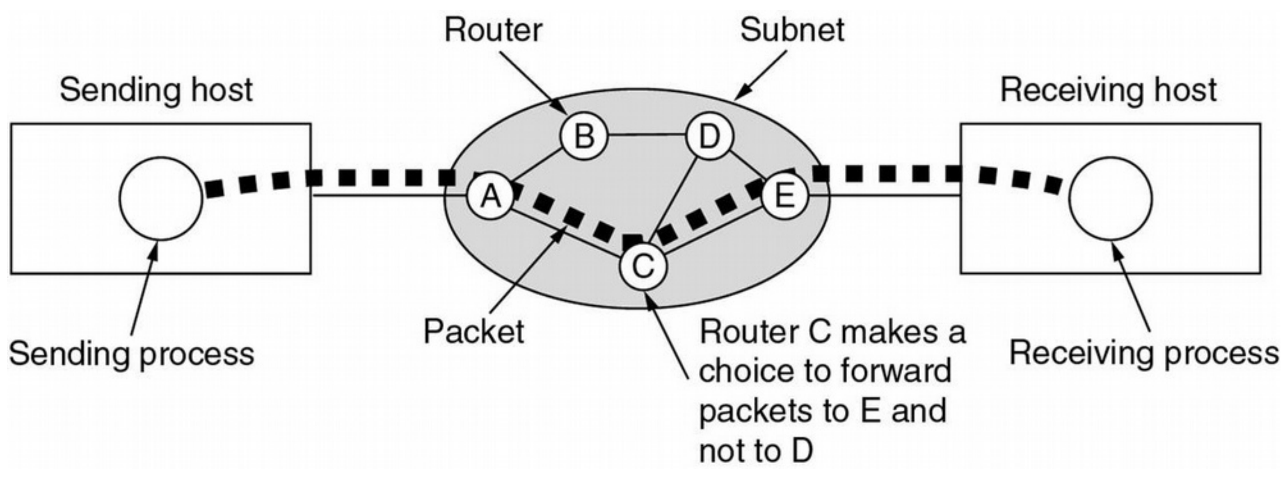
- Developed in scope of the military ARPA project
  - The goal was to reach a resiliency against link and switching element failures and a fast error recovery
  - Forms a basis of the today's Internet
- Hop-by-hop data passing over a polygonal structure
  - Routers interconnected with (redundant) links
- Data units (packets) are passed between routers independently of other packets
  - Each packets carries the receiver identification
  - Every packet may follow a different path
    - Protects against outages of links and routers
    - Packets may be received out-of-order and/or duplicated
  - A packet may wait in a router's queue for arbitrary time

Virtual Channel

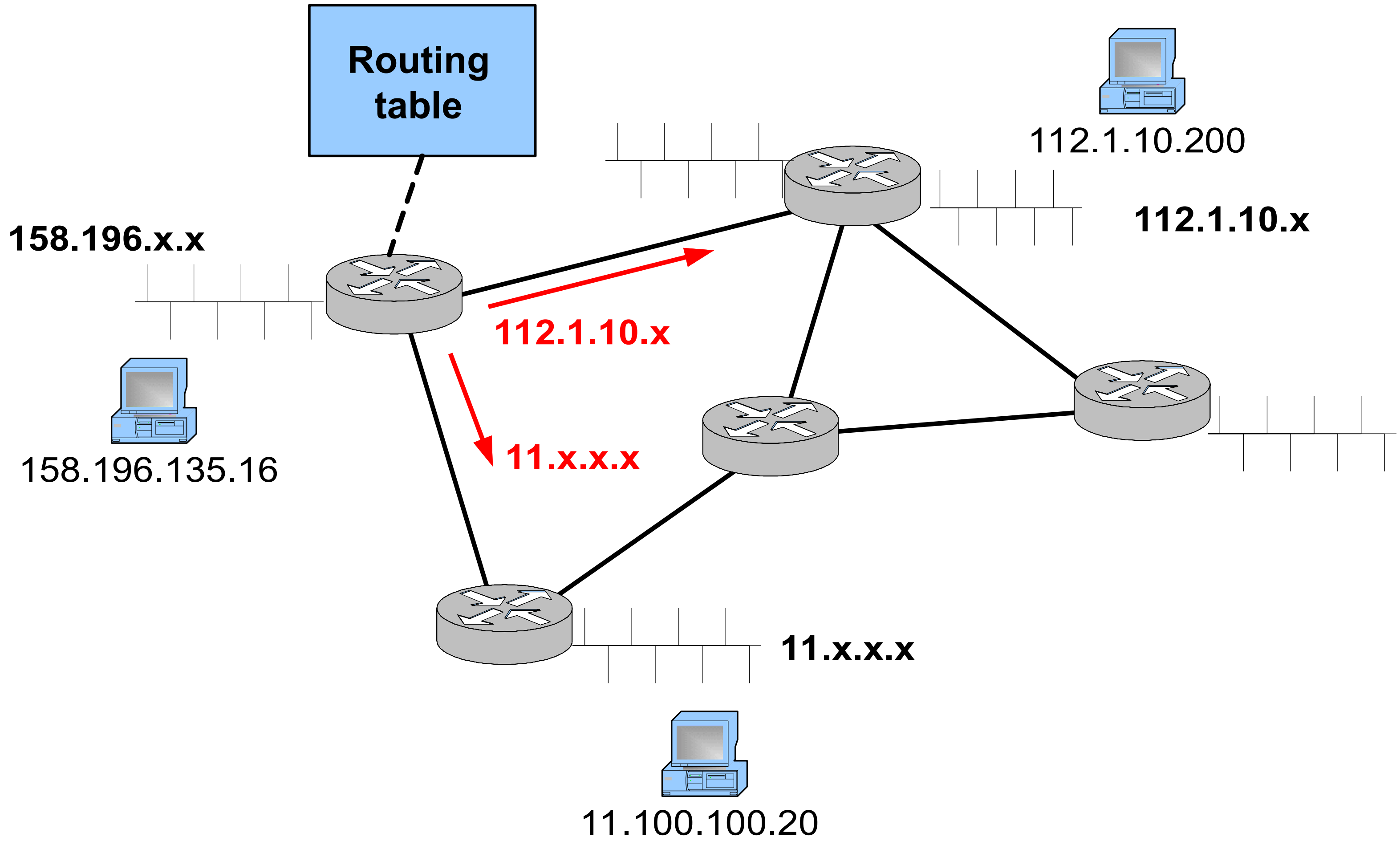
- A logical channel created on a packet-switched network infrastructure
  - A compromise between circuit switching and packet switching
- Created and abolished on the user node request
- Finding of path across the network takes place just during the creation of the virtual channel
  - A channel ID is allocated in each switching element and stored in a switching table
  - The ingress router provides a channel ID to the user node
  - User node places the channel ID into each packet
- All packets travel along the same path, so they cannot arrive out-of order

Routing

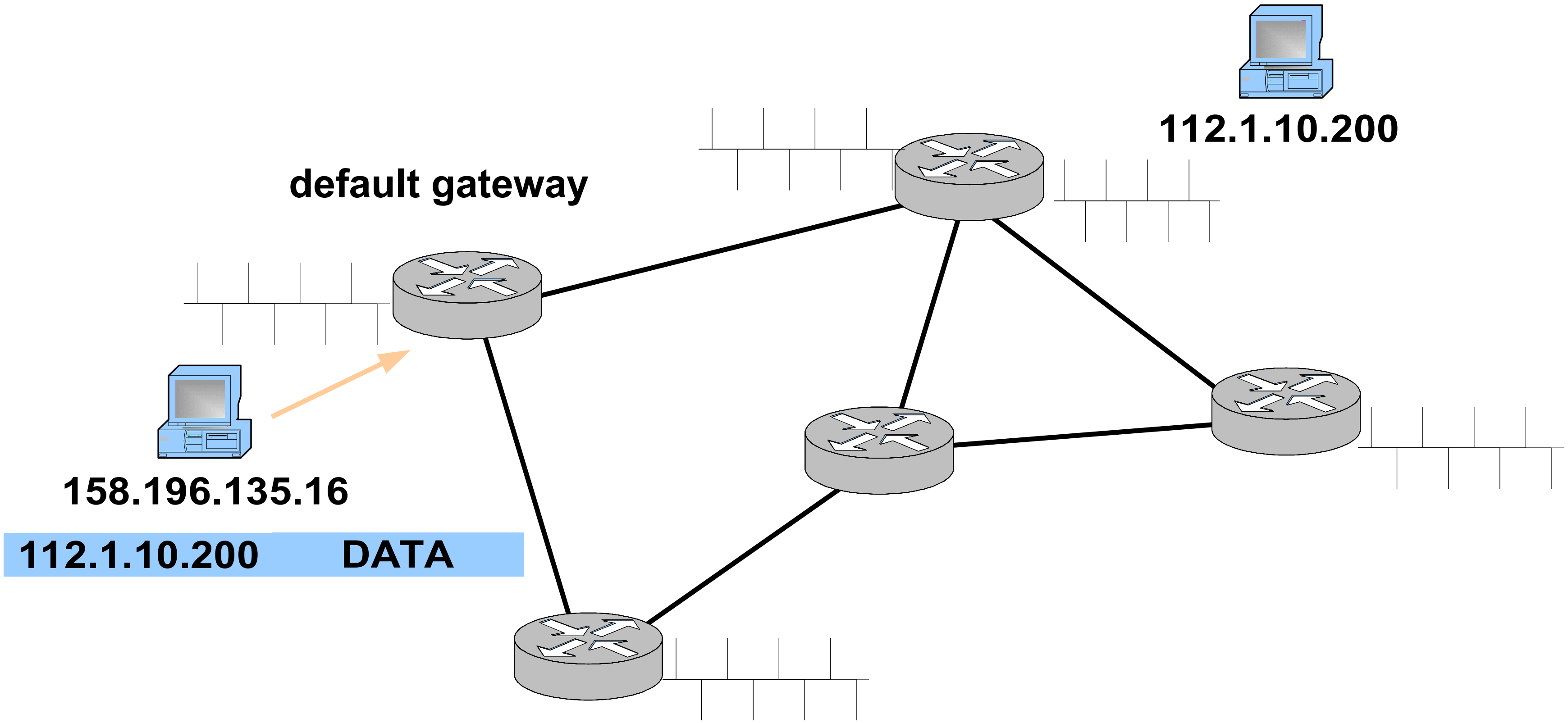
- Finding a path across the network
- Takes place
  - During the circuit setup in connection-oriented networks
    - Setup of switching circuitry of COs along the path of the physical circuit
    - Creation of entries of switching tables of switching elements along the path of the virtual circuit
  - During processing of individual packets in packet switched networks



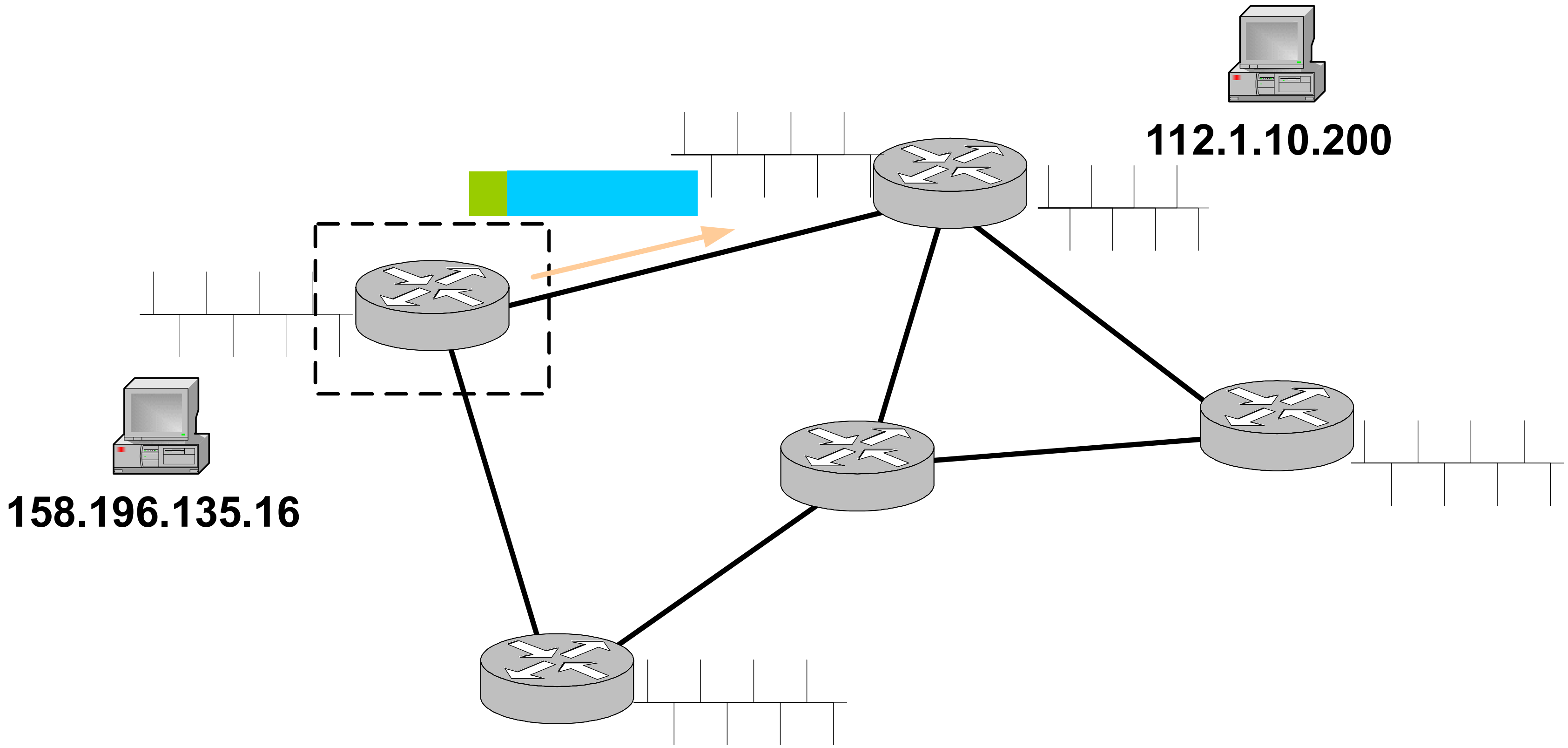
Routing in Packet-Switched Network (1)



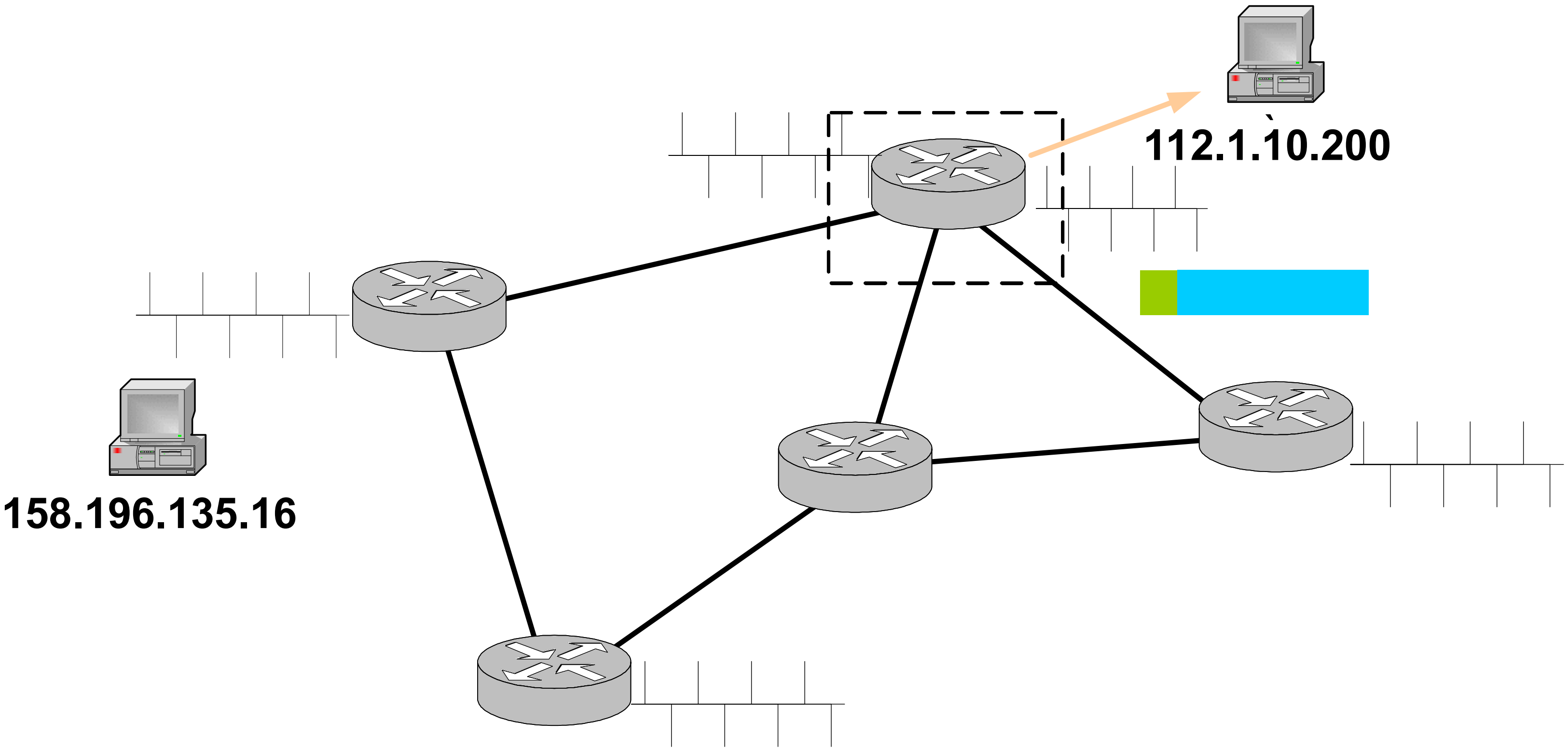
Routing in Packet Switched Network (2)



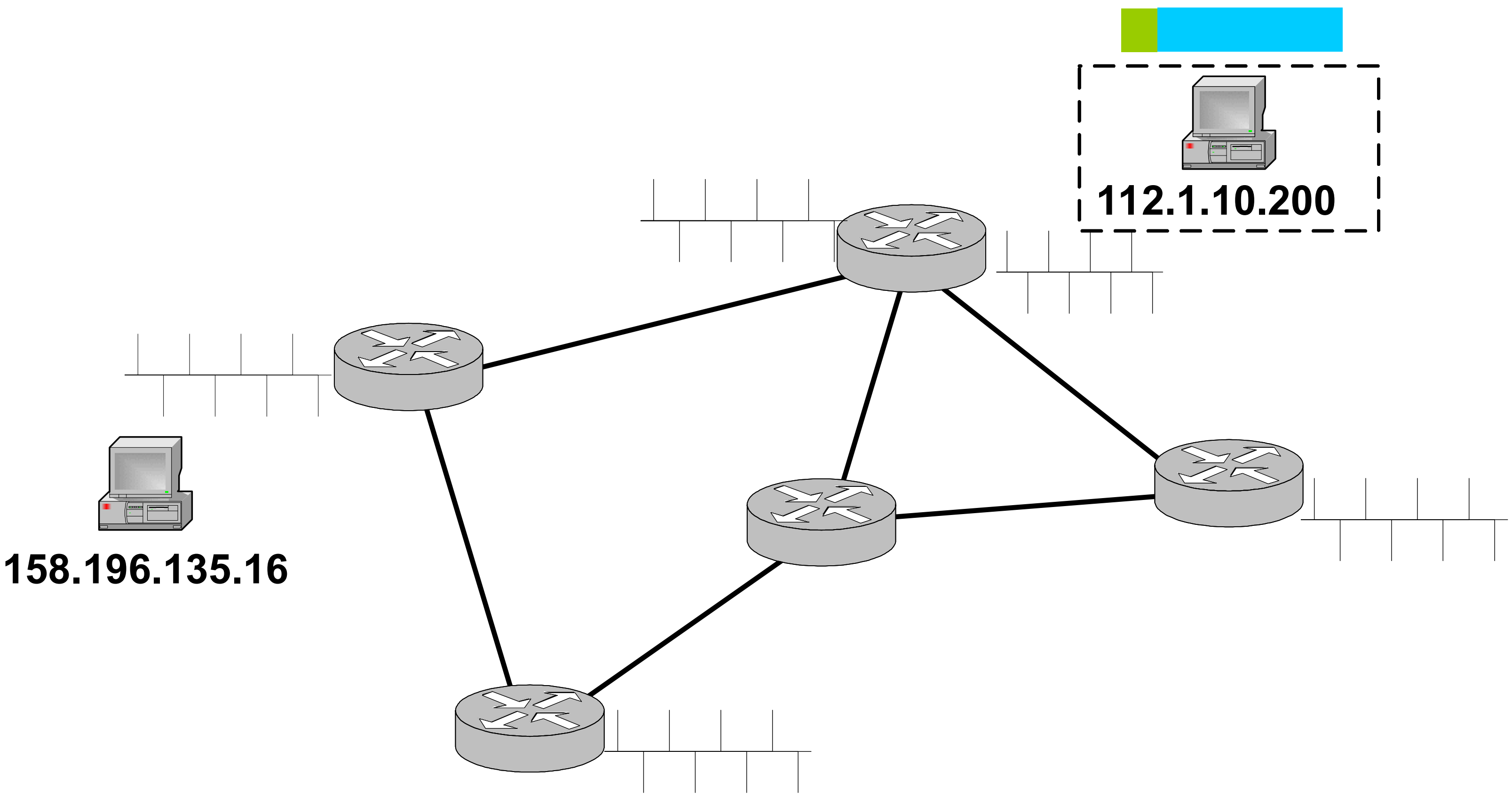
Routing in Packet Switched Network (3)



Routing in Packet Switched Network (4)



Routing in Packet Switched Network (5)



Routing Algorithms

Routing Algorithm

- A part of the OSI RM Layer 3 software that decide which interface to send an incoming packet from or to what neighboring switching device should (a virtual) circuit be routed
- Makes decisions based on the information pre-configured manually or obtained from a **routing protocol**
- Most commonly implemented using a **routing table**
  - Creation and maintenance of the routing table using a routing protocol may take place at the same time as actual data forwarding

Required Properties of the Routing Algorithm

- simplicity
- robustness
- stability
- fairness
- optimality

Some of the properties are in contradiction, so we need to decide what ones will be preferred

Routing Table

- Maintains records with the following format:
  - *<Destination(network+mask), outgoing-interface/next_hop, metric>*
- A particular node or a network prefix network may be used as a destination
  - In IP environment we may use address prefixes of various lengths
- In IP networks, a record that matches as much bits of the destination address as possible is used to route a packet
    - → **all records of the routing table have to be checked for each packet** → time consuming process

Default Route

- Typically used for networks connected by a single interface to a network of a higher hierarchy level
- All packets that do not match any routing table record are routed via the default route
- The purpose is to limit a number of records in a routing table
- Denoted as 0.0.0.0/0 in IP networks
  - The number of bits of the destination address that have to match is 0

Approaches to the Routing Problem

- Centralized
- Distributed
- (Isolated)

- *Nonadaptive* – static
- *Adaptive (dynamic)* – reacts on
  - The current network topology
  - The current load of the network
    - Implemented rarely as there is a danger of the routing instability

Centralized and Distributed Routing

Centralized Routing

- A central Routing Control Center (RCC) collects information about all routers' neighborhoods, combines together the network topology, calculates and distributes routing tables for all routers
- All routers periodically report their neighborhood states to the RCC
  - List of alive neighbors, interface queue lengths, total processed load, ...
- A distribution of routing tables to individual routers is problematic
  - inconsistent routing information during the distribution of the new routing tables version

Not used in practice today

Distributed Routing

- Every router knows its neighborhood
  - List of active neighbor routers
  - State of all connected links
    - Dynamic: operational status, current load
    - Static: Bandwidth, configured cost,MTU

- Every router exchanges its routing information with other routers
  - Using directly connected neighbors

- Every router independently constructs its routing table according to the obtained routing information

Used today in both in the Internet and intranets

Static and Dynamic Routing

Static Routing

- Routing tables are configured manually
  - requires a lot of administrator's work
- Incurs No routing protocol overhead
  - Consumed bandwidth, CPU utilization
- Safer
  - Avoids forging of the fake routing information and network topology eavesdropping
- Non-adaptive
  - Administrator's action is necessary in case of a router or link outage
- May be applied if the topology does not change too much
  - because of outages or the network modification

Used rather often in small intranets

Dynamic Routing

- Automatically adapts to the current network topology (and sometimes to the current traffic)
- Routing protocols must be applied
- Inevitable in networks with frequent changes and/or with the unknown topology
    - typically in the Internet

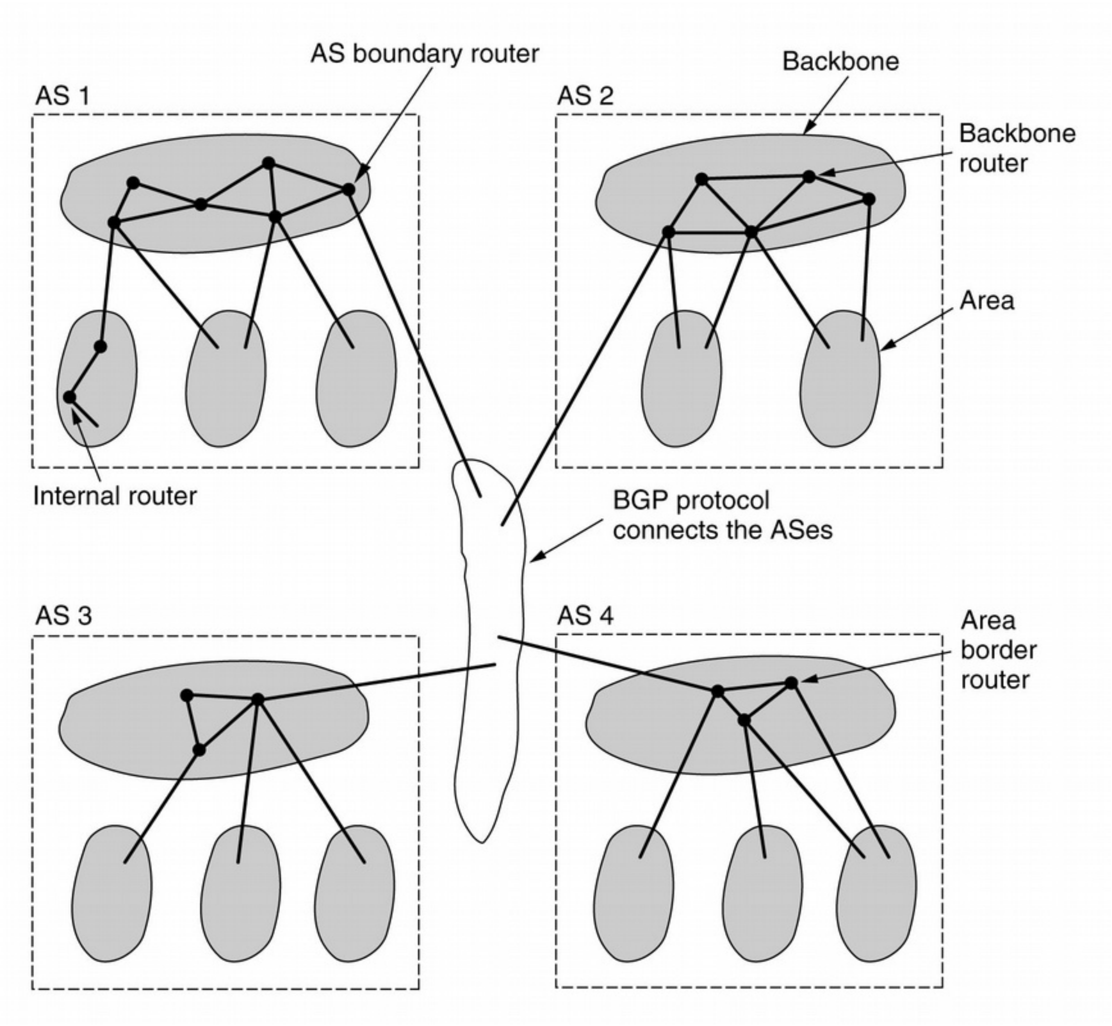
In practice, a combination of static and dynamic routing is applied

- routers typically prefer static routes over the dynamic ones

Hierarchical Routing – the Principle

- Organizes (a large) network into hierarchical parts
  - Also advantageous for division of competences of the network administration
- Routers in individual parts know just the topology of the part, a route into the higher hierarchy level and a list of networks in parts of a lower hierarchy level
  - But not the topology of the parts at the lower hierarchy levels
  - The aim is to limit number of routes in routing tables
  - Route aggregation and default route

Hierarchical Routing – Example



Routing Algorithms (Routing Protocols)

What are the Differences between Routing Protocols ?

- Utilized metric
- Level of routers' knowledge of the network topology
- Scheme of the propagation of the routing information
  - exchange between neighbors, flooding to all routers, diffusion algorithm, ...
- Technical implementation of sending of the routing information
  - broadcasting/multicasting, update period
- The aim is to reach a fast convergence

  = a time between topology change and stabilization of routing tables of all routers

Classes of the Routing Protocols

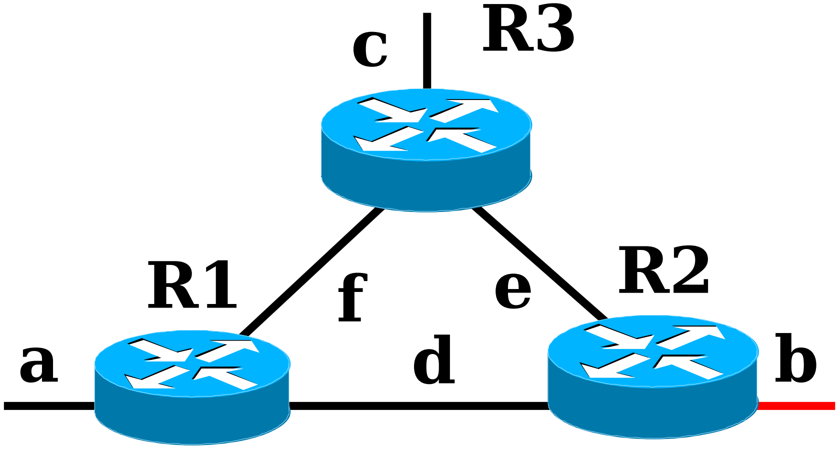
- **Distance Vector**
  - Older
  - Easier implementation
- **Link State**
  - More complex implementation
  - Faster convergence
  - More deterministic behavior in „special cases“

Distance Vector Algorithms (DVA)

DVA Basic Principle

- Routers do not know the network topology, just the outgoing interfaces for individual destination networks + distances to those networks (distance vectors)
  - in practice, we need a next-hop address
- The routing table originally contains just the directly connected networks
  - Manually pre-configured by administrator
- The routing table is periodically broadcasted to all neighbors
- A router combines/adapts is routing table from distance vectors propagated by its neighbors
  - If a route fails to be advertised by the neighbor, it is removed from the routing table

An Example: Propagation of network b (1)

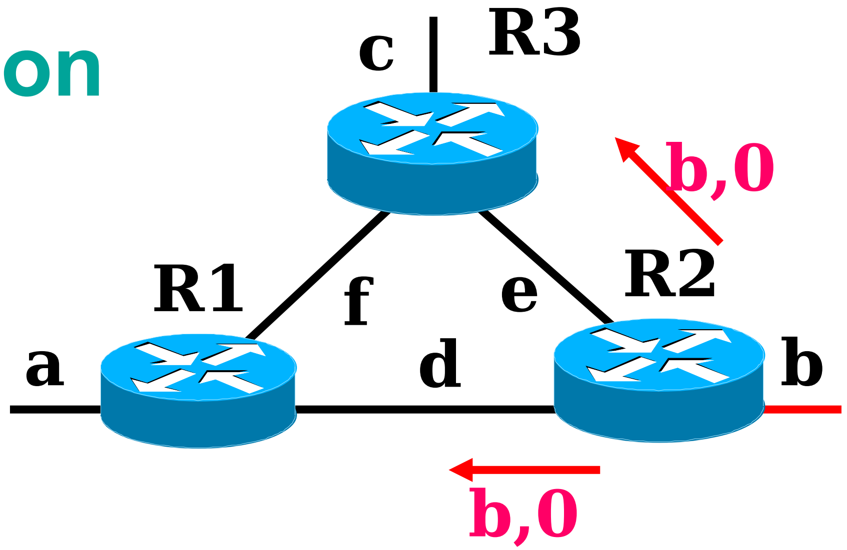


Pre-configured information:

```
R1:        R2:        R3:
-----      -----      -----
a 0 -      b 0 -      c 0 -
f 0 -      d 0 -      e 0 -
d 0 -      e 0 -      f 0 -
```

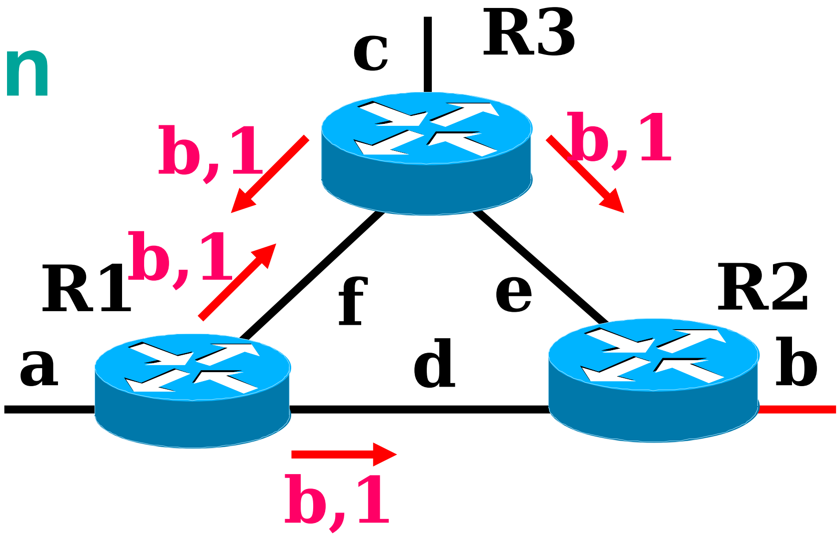
An Example: Propagation of network b (2)

Routing tables after reception of an update from R2



R1:	R2:	R3:
a 0 -	b 0 -	c 0 -
f 0 -	d 0 -	e 0 -
d 0 -	e 0 -	f 0 -
b 1 R2		b 1 R2

An Example: Propagation of network b (3)



Routing tables after reception of an update from R1 and R3

```
R1:       R2:       R3:
-----     -----     -----
a 0 -     b 0 -     c 0 -
f 0 -     d 0 -     e 0 -
d 0 -     e 0 -     f 0 -
b 1 R2              b 1 R2
b 2 R3              b 2 R1
```

I already have a better route

I already have a better route

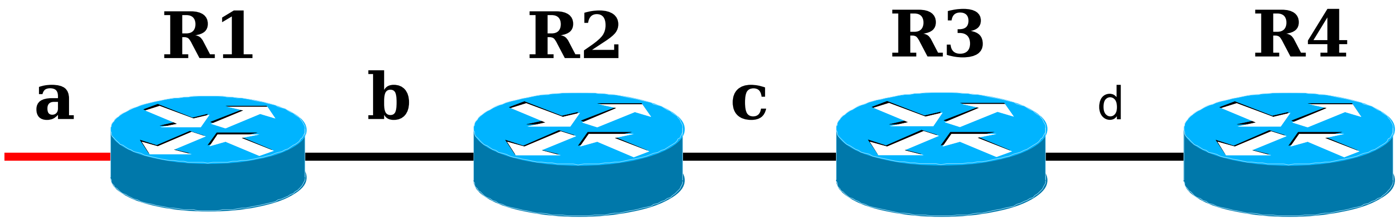
Building of Routing Tables from the Received Routing Information

- If a router receives a route to yet unknown network, it install it into the routing table
- If (an alternative) route with a better metric to already known network is received, it replaces the current route
  - A route with the same metric will be either ignored or used for load balancing
- If a route's metric becomes worse but there is no better route, the router has to change route's metric in the routing table

Properties of DVA

- The hop count (number of routers on the path) is used as a metric
  - Does not take link parameters into account
    - e.g. link bandwidth, delay, load, …
    - suitable in the early days of Internet with equal links
- Slow convergence after a topology change
  - The change advertisement has to wait until the next routing table broadcast
- Periodic routing table broadcasts incur a substantial network load
- Too "optimistic"
  - router learns new routes quickly, but forgets routes that became unusable slowly
  - after the expiration of the invalidation timer when the route is not advertised anymore

DVA Convergence

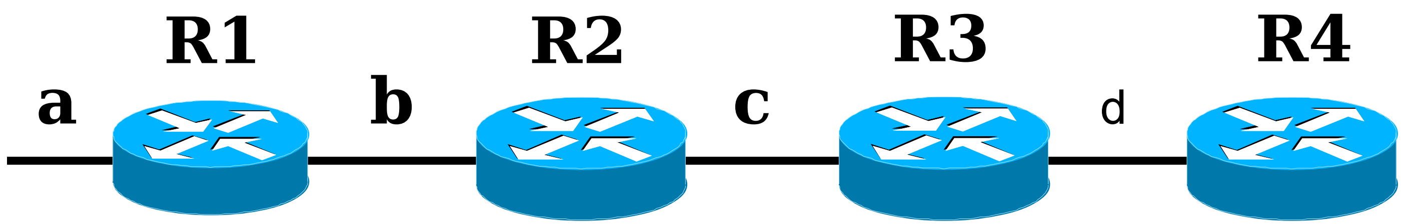


Network a goes up.
Consider the update period of 30s (RIP):

- R2 gets to know about a after 30s (max)
- R3 gets to know about a after 60s (max)
- R4 gets to know about a after 90s = 1.5 min (max)

Count-to-Infinity Problem(1)

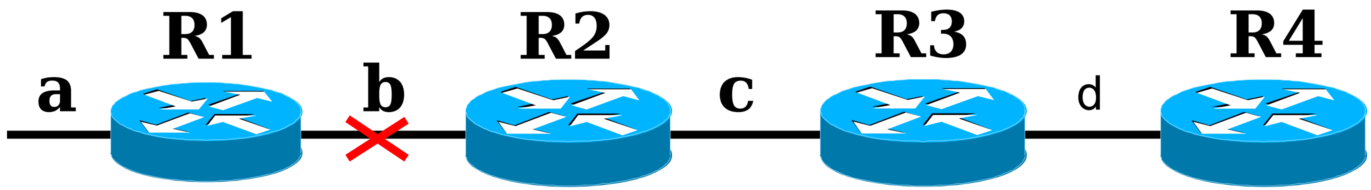
If a route fails to be advertised and a router hears the alternative route, it does not know that the alternative route back to it:



Consider the routing table entries for network **a**:

R2: 1,R1 **R3: 2,R2** **R4: 3, R3**

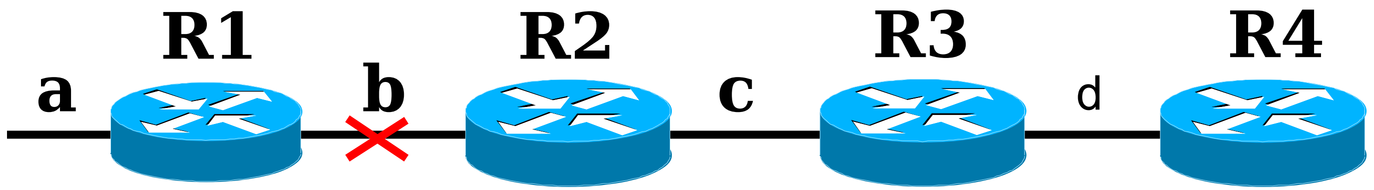
Count-to-Infinity Problem (2a)



Now consider the link between R1 and R2 fails

R2: ∞,- **R3: 2,R2**

Count-to-Infinity Problem (2b)



Now consider the link between R1 and R2 fails

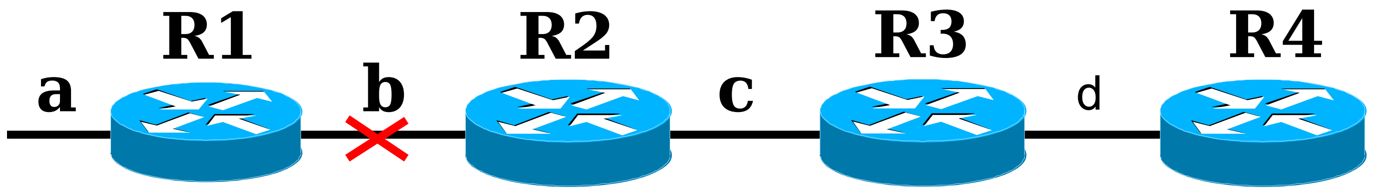
R2: ∞,-   R3: 2,R2

R3 advertises to R2 that it can reach a with 2 hops

(but does not tell that the route goes via R2 ;-))

R2: 3,R3   R3: 2,R2

Count-to-Infinity Problem (2c)



Now consider the link between R1 and R2 fails

R2: ∞,- R3: 2,R2

R3 advertises to R2 that it can reach **a** with 2 hops

(but does not tell that the route goes via R2 ;-))

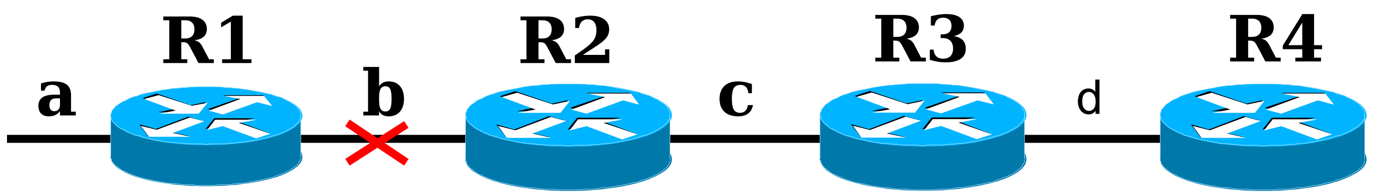
R2: 3,R3 R3: 2,R2

R2 informs R3 that it can reach a with 3 hops

(R3 had a route through R2 with metric 2, but R2 is announcing different value → original value is overwritten)

R2: 3,R3 R3: 4, R2

Count-to-Infinity Problem (2d)



Now consider the link between R1 and R2 fails

R2: ∞,- R3: 2,R2

R3 advertises to R2 that it can reach **a** with 2 hops

(but does not tell that the route goes via R2 ;-))

R2: 3,R3 R3: 2,R2

R2 informs R3 that it can reach **a** with 3 hops

(R3 had a route through R2 with metric 2, but R2 is announcing different value → original value is overwritten)

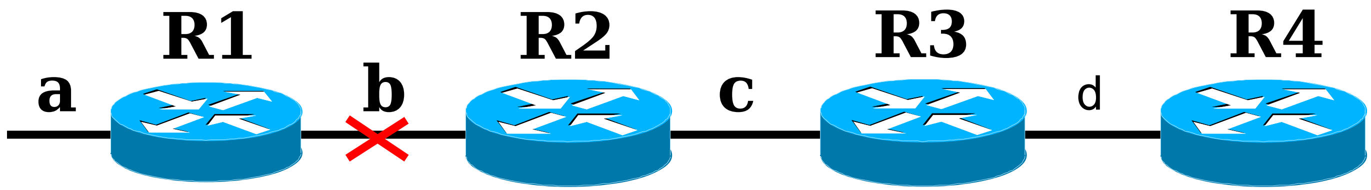
R2: 3,R3 R3: 4, R2

R3 informs R2 that it can reach a with 4 hops

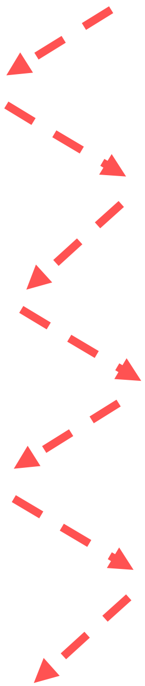
R2 had a route via R3 with 3 hops, but R3 now advertises 4 hops, so R2 has to update it to 5 hops

R2: 5,R3 R3: 4, R2

Count-to-Infinity Problem (3)



→ The metric of network a still increases on R2 and R3 (up to the infinity)



Solution of the Count-to-Infinity Problem (1)

- We limit the maximum value of a route's metric
  - a finite value that represents infinity
  - diameter of a network topology graph + 1
- Routes with metric equal to „infinity“ are not used

Solution of the Count-to-Infinity Problem (2)

- **Split horizon rule**
  - routes are not advertised out of the interface from that they have been learned

The first solution also solves loops over multiple routers that cannot be broken with a Split horizon rule

→ In practice, both solutions are implemented together

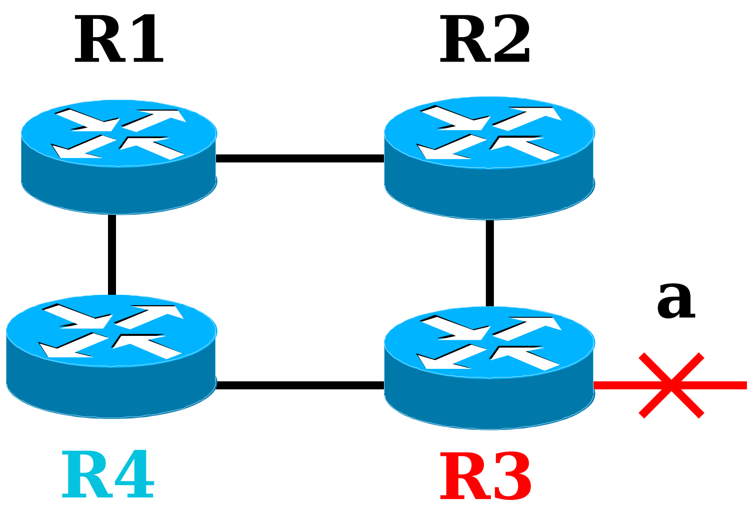
Improvements of DVA Routing Protocols (1)

- **Triggered update** – the router broadcasts (a new) routing table just after it detects a change, i.e. Without waiting to the update timer
    - Interface goes up or down
    - A new route is learned from the neighbor
- **Poisson reverse**
    - Improvements of the split horizon
    - advertises routes out of the interface they came from with an infinite metric

Improvements of DVA Routing Protocols (2)

- **Hold down** – after a route is lost, the router does not accept advertisements of the same route from other routers for a duration of a hold down interval
  - avoids learning of routes that really do not exist, but the other routers did not mentioned their unavailability yet

Hold down – An Example



Hold-down technique applied on R4

- R3 informs about network **a** unavailability (triggered update)
- R1 offers a route to **a** with 3 hops
  - R4 will ignore that route for a hold down period
  - as R1 probably did not noted that the route it advertises is unavailable

Examples of DVA Protocols

- Routing information protocol (RIP)
  - Old, but still used in small networks
    - Uses 16 to represent infinite metric
  - Easy to implement
  - No special requirements on administrator's knowledge
- Interior Gateway Routing Protocol (IGRP)
  - Cisco proprietary
  - Composite metric
    - bandwidth, delay, (reliability, load)
  - Not limited to 16 hops (increased to 255)
- …

Link State Algorithms (LSA)

Properties of LSA

- Calculates shortest paths with the global knowledge of states of all links
  - operational state, cost
- Routers (graph vertices) know the topology of the whole network (a graph) and costs of individual links (edge labels). They maintain that information in a topology database
  - all routers have the same content of the topology database
- Every router calculates shortest paths tree to all other routers (and networks connected to them) using Dijkstra algorithm
  - in contrast to DVA, all routers calculate routing tables from the same information

Operation of LSA Protocols

- Every router continually checks status of all directly connected links
  - Checks the availability of neighboring routers using Hello protocol
- If a change is detected, the updated information about router's neighborhood is immediately flooded to other routers that updates it in their topological database
  - No periodic routing table broadcasts, only the information about topology changes is sent

Instant reaction on link state changes + recalculation of SPF → fast convergence

Topological Database

- Consists of records containing:
  - Router ID
  - List of alive neighboring routers (IDs)
    + costs and network addresses related to corresponding links
  - List and network addresses of stub networks connected to the router
    The network address is always appended to the link

- Records are generated by individual routers after any topology change and flooded to the whole network
- The topology graph may be constructed using topology database records

Examples of LSA Protocols

- Open Shortest Path First (OSPF)
  - open standard
  - Router first calculates the shortest paths tree and transforms it to the routing table
  - Supports the hierarchical routing (areas)
  - Very popular in today's networks
- IS-IS
  - ISO standard
  - similar principle as OSPF

Algorithms of Graph Theory used by Routing Protocols

Basic Algorithms used in Routing Protocols

- Dijkstra
  - Calculates a shortest paths tree from a given vertex to all other vertices of a edge-labeled graph
  - Used in OSPF (a little bit modified)
- Floyd
  - Calculates routing tables for individual routers from a cost matrix of the network topology graph
- Ford-Fulkerson
  - Calculates distances from all nodes to a given node
  - In distributed version used as a basis of RIP

Routing in TCP/IP Networks

TCP/IP Routing Protocols

Interior routing protocols (IGP)

- Used to route inside autonomous systems
- Open Standards
  - RIP (Routing Information Protocol)
  - OSPF (Open Shortest Path First)
- Proprietary
  - IGRP, EIGRP - Cisco
  - NLSP – Novell

Exterior routing protocols (EGP):

- Used to route between autonomous systems
  - BGP (Border Gateway Protocol) - path vector protocol

Classful Routing Protocols

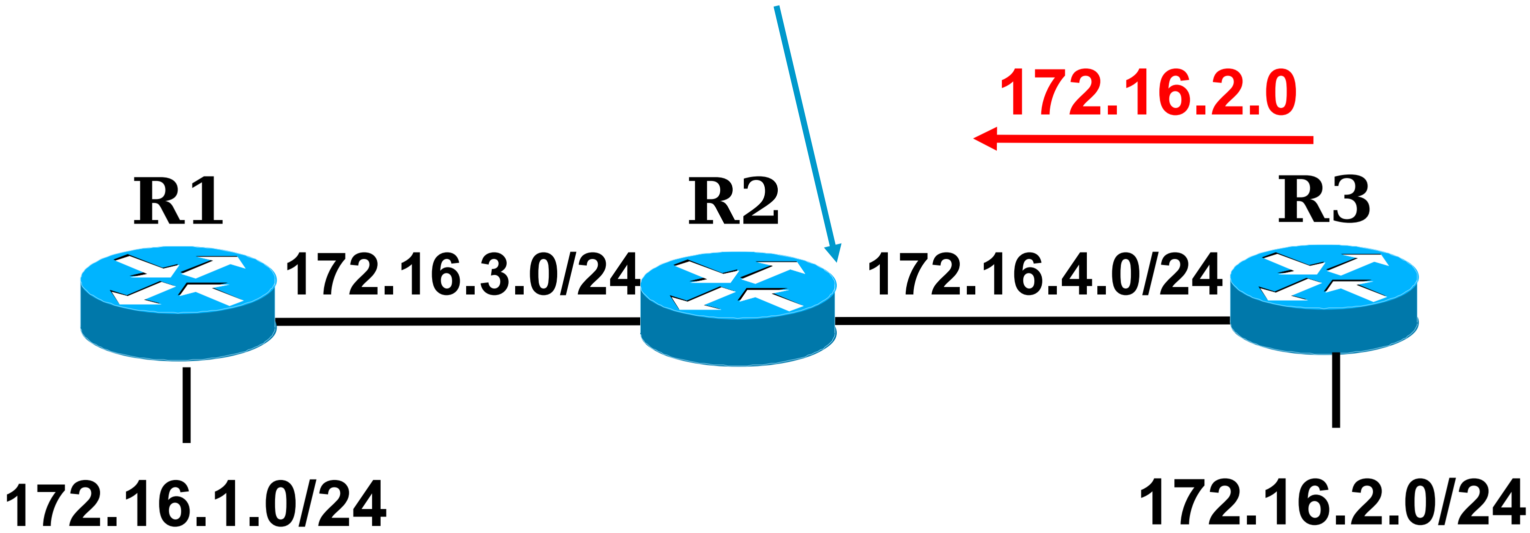
- Do not propagate subnet masks but rely on IP address classes
- Require a common (constant) subnet mask for all subnets of the single major network
- Require subnet continuity
  - Subnets of the same major network may not be separated by other networks
  - As the subnets are aggregated to a classful address on the major network boundary

- Old DV protocols (RIP v.1, IGRP)
  - As a result not of much use nowadays

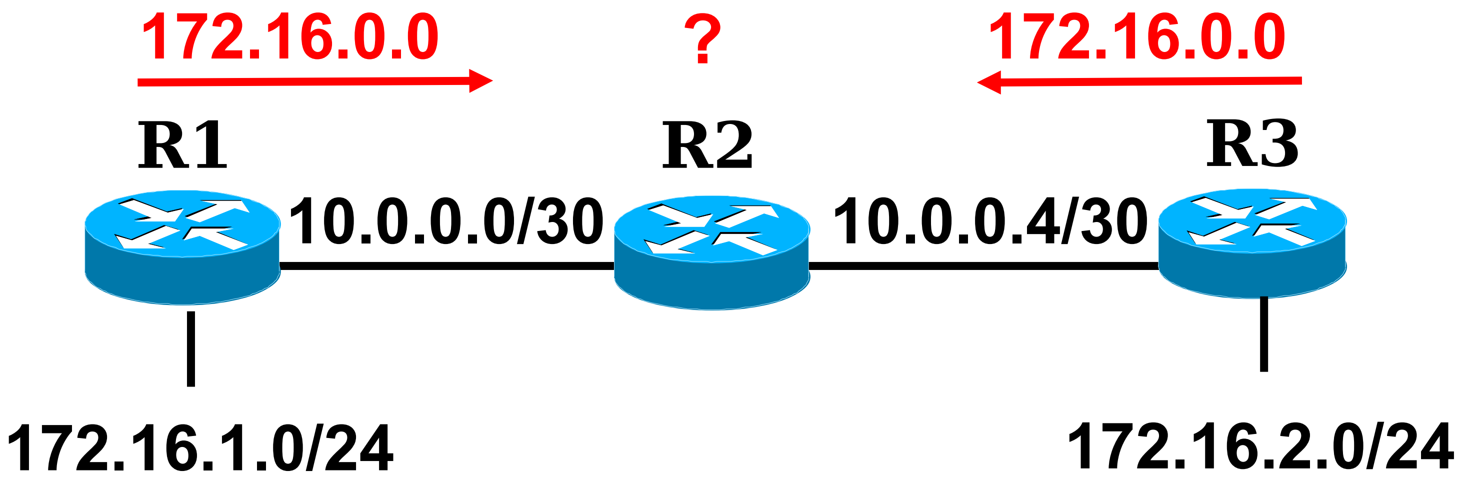
Classful Addressing and Constant Length Subnet Mask

172.16.2.0, but with what ?

I will take the mask of the interface from which the routing information arrived !



Problem of Discontinuous Networks with Classful Addressing



VLSM: Variable-Length Subnet Mask

- The constant subnet mask is inefficient if we have very different number of nodes on individual segments
  - e.g. 2 stations on P2P links and tens of stations on the Ethernet LAN segment
- VLSM (RFC 1009) allows to have different subnet masks on subnets of the same network
  - addresses still have to be unique
- May be used only with routing protocols that propagate subnet masks together with network addresses (OSPF, ISIS, RIP v.2).
  - Network addresses are stored in routing tables together with subnet masks