Communication Networks

Spring 2019

Q&A Session

Alexander Dietmüller

Tobias Bühler

https://comm-net.ethz.ch/

ETH Zürich

August 6 2019

How do you guide IP packets
from a source to the destination?

Essentially,
there are **three ways to compute** valid routing state

	Intuition	Example
#1	Use tree-like topologies	Spanning-tree
#2	Rely on a global network view	Link-State SDN
#3	Rely on distributed computation	Distance-Vector BGP

Essentially,
there are **three ways to compute** valid routing state

	Intuition	Example
#1	Use tree-like topologies	Spanning-tree
#2	Rely on a global network view	Link-State SDN
#3	Rely on distributed computation	Distance-Vector BGP

Distance-vector protocols are based on Bellman-Ford algorithm

Every router maintains a table of all received distance vectors

X
routing table

dest.	via *Y*	via *Z*
Y	3	11
Z	13	1
U	30	2

Every router maintains a table of all received distance vectors

Let $d_x(y)$ be the cost of the least-cost path known by x to reach y

X
routing table

dest.	via *Y*	via *Z*
Y	3	11
Z	13	1
U	30	2

Every router maintains a table of all
received distance vectors

Let $d_x(y)$ be the cost of the least-cost path
known by x to reach y

Each node bundles these distances
into one message (called a vector)
that it repeatedly sends to all its neighbors

until convergence

X
routing table

dest.	via *Y*	via *Z*
Y	3	11
Z	13	1
U	30	2

X
vector

dest.	cost
Y	3
Z	1
U	2

Every router maintains a table of all received distance vectors

Let $d_x(y)$ be the cost of the least-cost path known by x to reach y

Each node bundles these distances into one message (called a vector) that it repeatedly sends to all its neighbors

until convergence

Each node updates its distances based on neighbors' vectors:

$$d_x(y) = \min\{ c(x,v) + d_v(y) \} \qquad \text{over all neighbors } v$$

Whenever a router uses another one,

it will announce it an infinite cost

The technique is known as poisoned reverse

Internet routing comes into two flavors:
intra- and *inter-domain* routing

inter-domain
routing

Find paths between networks

intra-domain
routing

Find paths within a network

inter-domain
routing

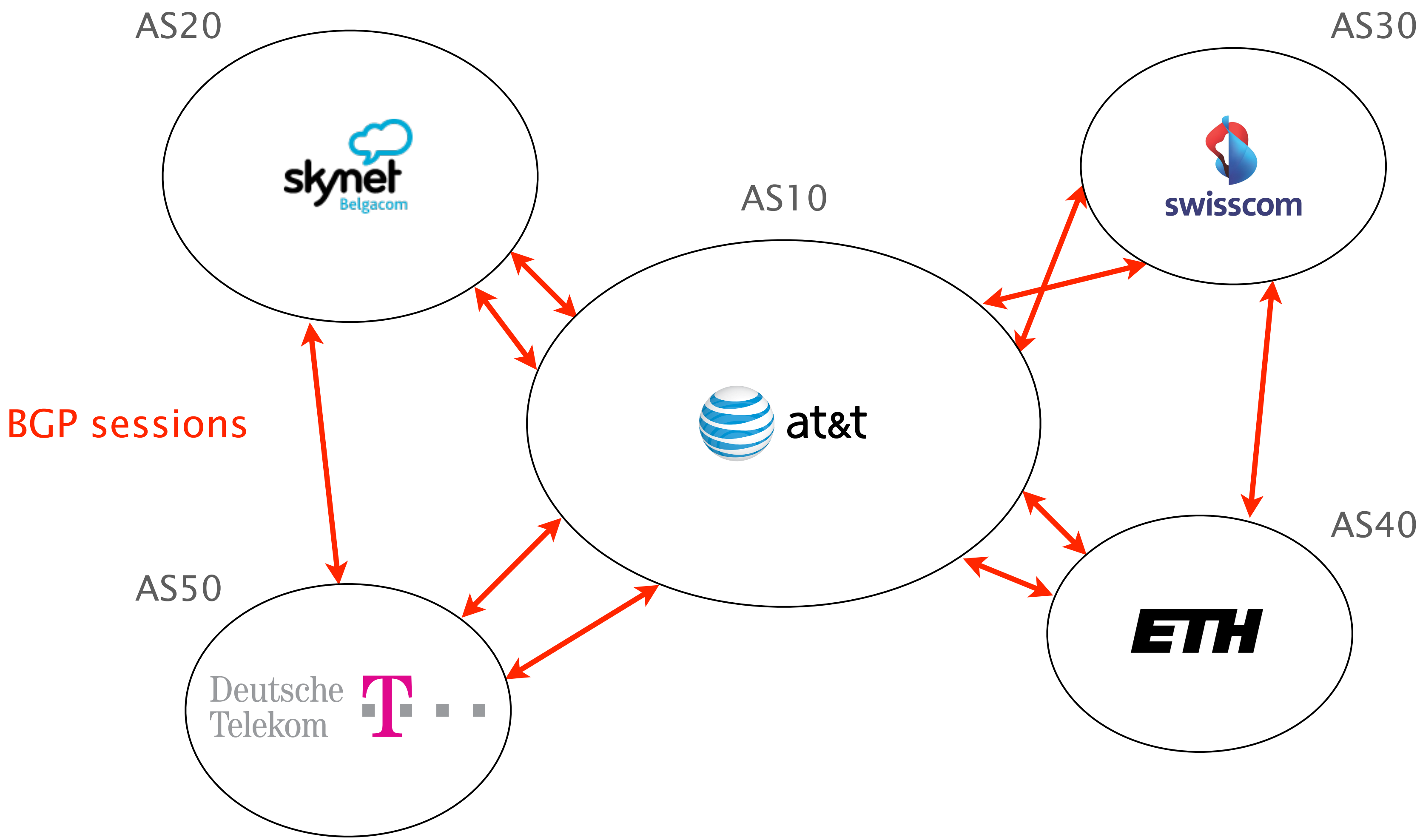
intra-domain
routing

Find paths between networks

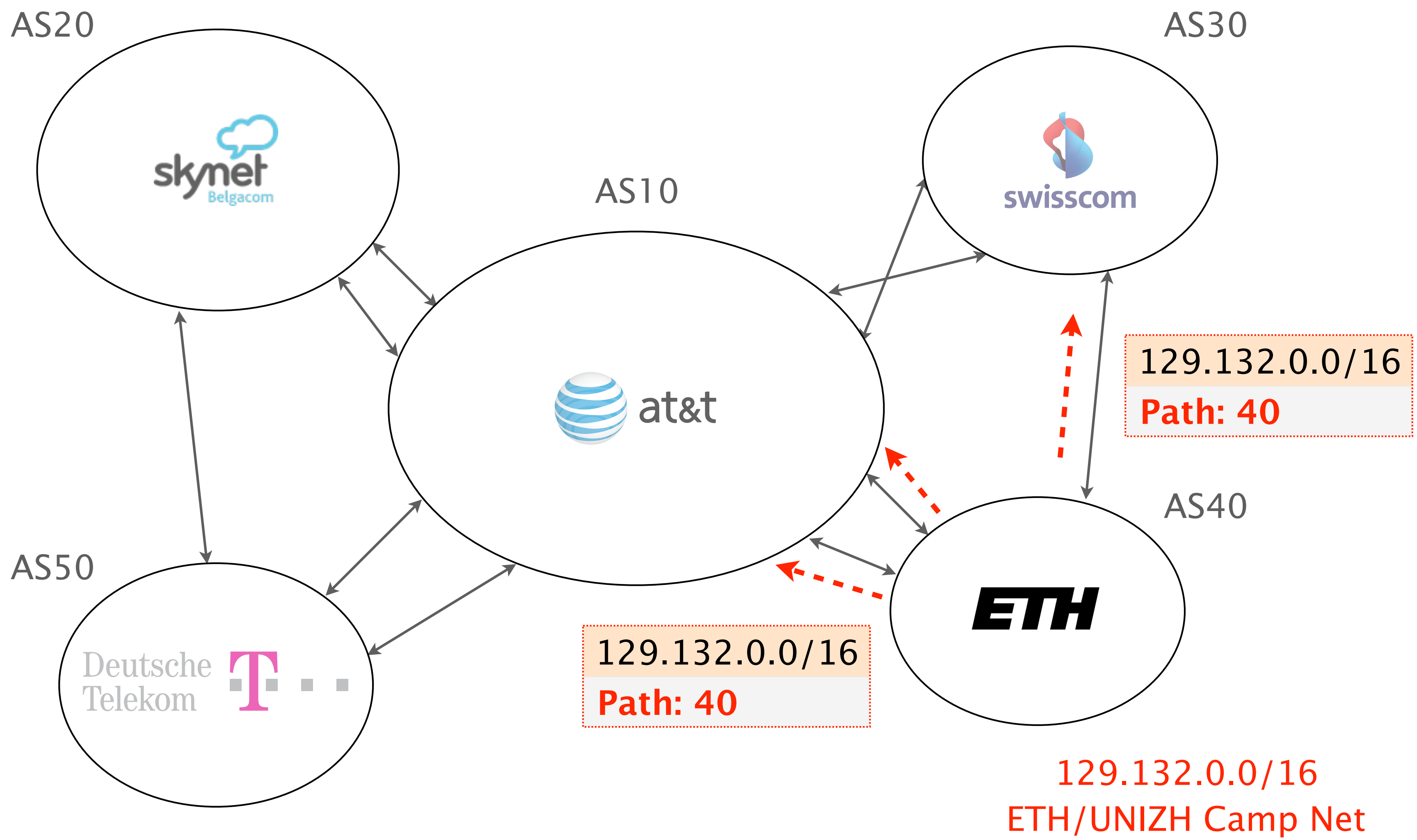
Internet

Border Gateway Protocol (BGP)

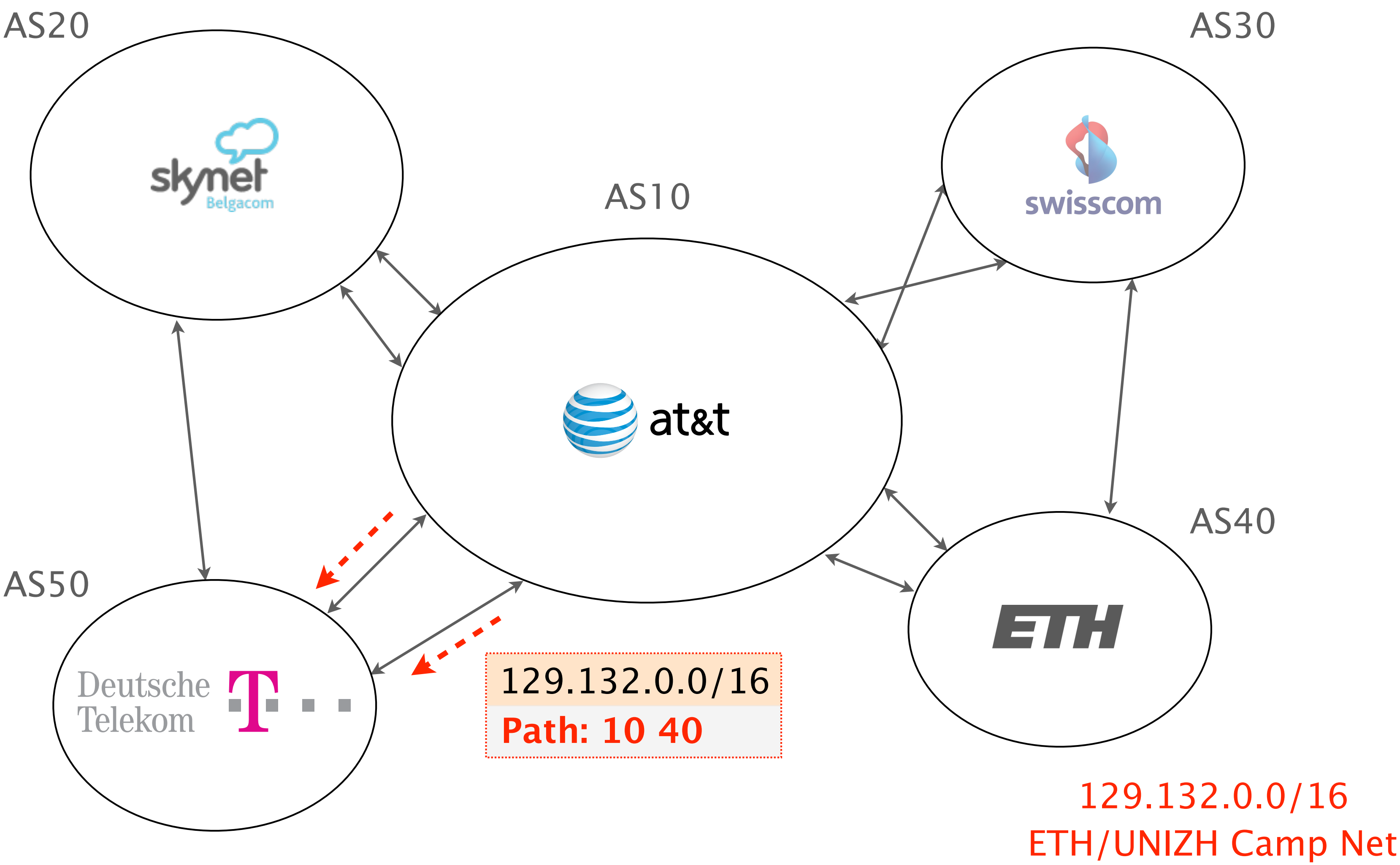
BGP is the routing protocol "glueing" the entire Internet together



BGP announcements carry complete path information instead of distances



Each AS appends itself to the path when it propagates announcements



There are 2 main business relationships today:

- customer/provider
- peer/peer

many less important ones (siblings, backups,…)

These policies are defined by constraining which BGP routes are *selected* and *exported*

Selection

which path to use?

Export

which path to advertise?

Selection

Export

which path to use?

which path to advertise?

control outbound traffic

Business relationships conditions

route selection

For a destination p, prefer routes coming from

- customers over
- peers over
- providers

route type

Selection

Export

which path to use?

which path to advertise?

control inbound traffic

Routes coming from peers and providers are only propagated to customers

		send to		
		customer	peer	provider
	customer	✓	✓	✓
from	peer	✓	-	-
	provider	✓	-	-

On the wire, BGP is a rather simple protocol composed of four basic messages

type	used to…
OPEN	establish TCP-based BGP sessions
NOTIFICATION	report unusual conditions
UPDATE	inform neighbor of a new best route a change in the best route
KEEPALIVE	inform neighbor that the connection is alive

UPDATE inform neighbor of a new best route

a change in the best route

BGP UPDATEs carry an IP prefix
together with a set of attributes

IP prefix

Attributes

Describe route properties

used in route selection/exportation decisions

are either local (*only* seen on iBGP)

or global (seen on iBGP *and* eBGP)

Attributes	Usage
NEXT-HOP	egress point identification
AS-PATH	loop avoidance outbound traffic control inbound traffic control
LOCAL-PREF	outbound traffic control
MED	inbound traffic control

Prefer routes…

with higher LOCAL-PREF

with shorter AS-PATH length

with lower MED

learned via eBGP instead of iBGP

with lower IGP metric to the next-hop

with smaller egress IP address (tie-break)

Each BGP router processes UPDATEs according to a precise pipeline

Life of a BGP router is made of three consecutive steps

while true:

- receives routes from my neighbors
- select one best route for each prefix
- export the best route to my neighbors

In practice,
multiple URLs can be mapped to the same IP

name	DNS ⟷	IP address
www.vanbever.eu		188.165.240.60
www.routeur.be		188.165.240.60

How does a web server receiving an HTTP request know, which website you want to access?

The host field tells the server which website it should serve

HTTP request:

```
GET / HTTP/1.1
Host: www.google.com
```

Anycast, Unicast, Multicast

„one-to-one-of-many“

Important, discussed in lecture

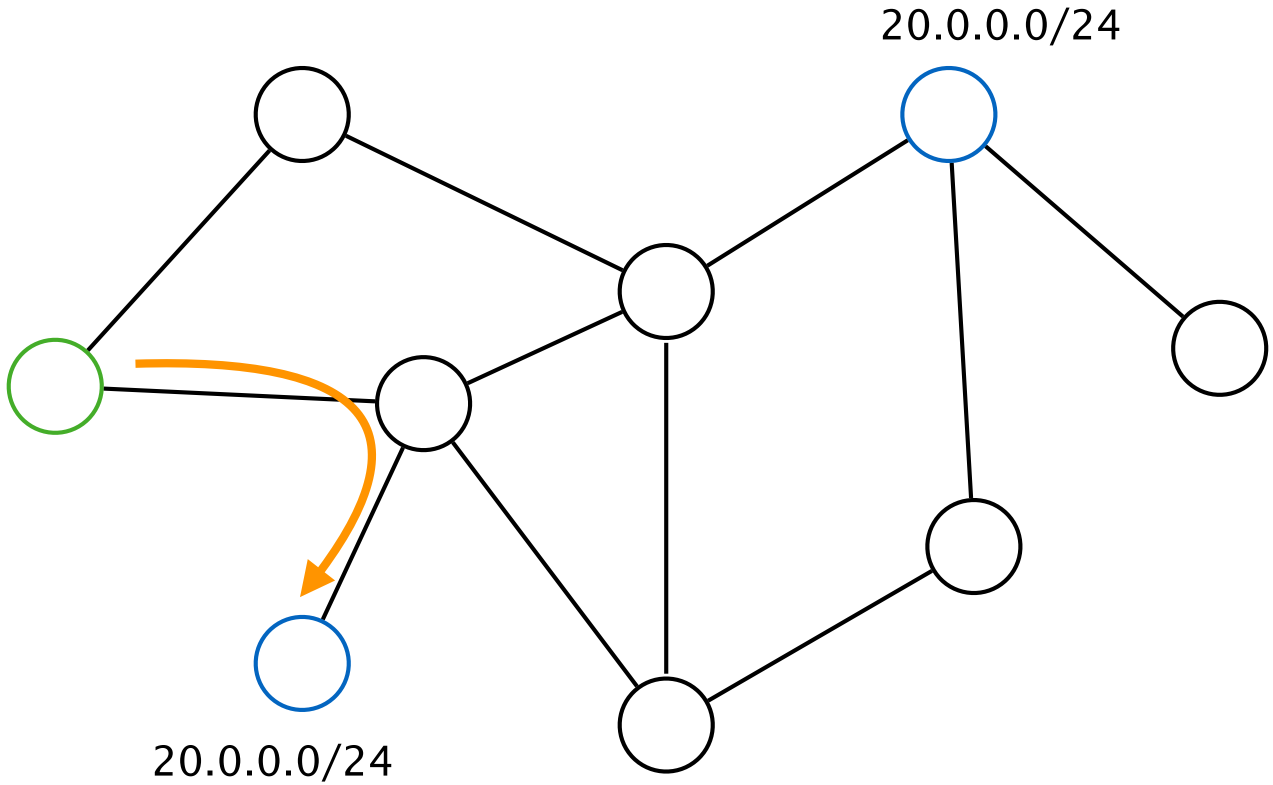
Used for scalability, load-balancing (e.g. DNS root server)

Routing finds shortest-paths

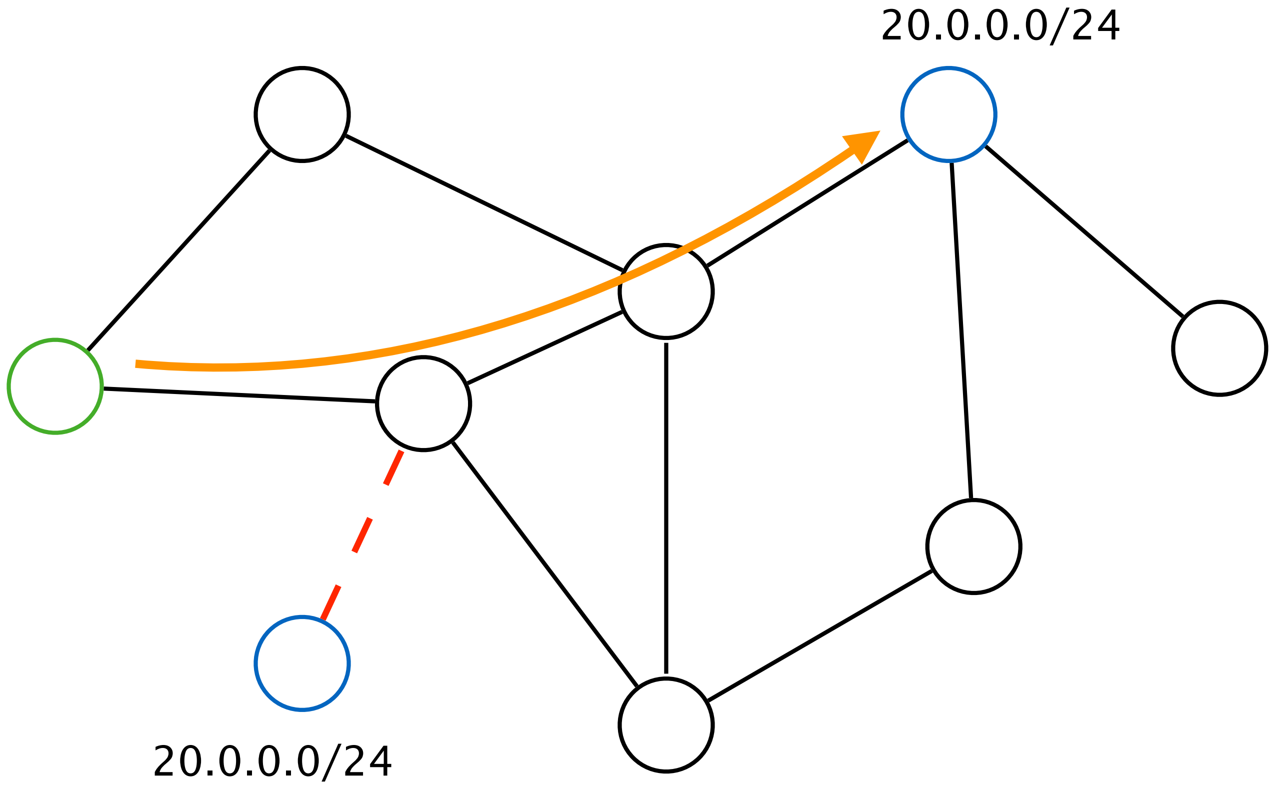
Seamless replication

But, potential problems for stateful applications

Anycast, Unicast, Multicast



Anycast, Unicast, Multicast



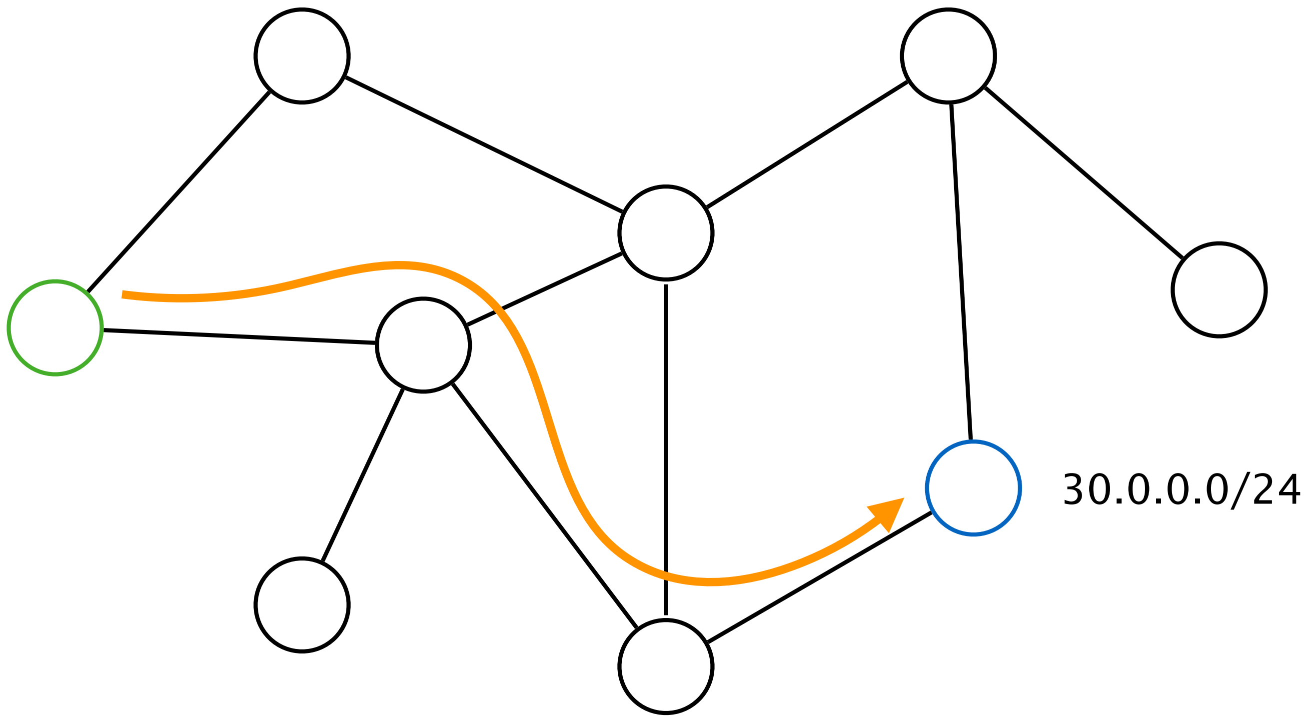
Anycast, Unicast, Multicast

„one-to-one“

Destination address uniquely identifies a single receiver

No replication

Anycast, Unicast, Multicast

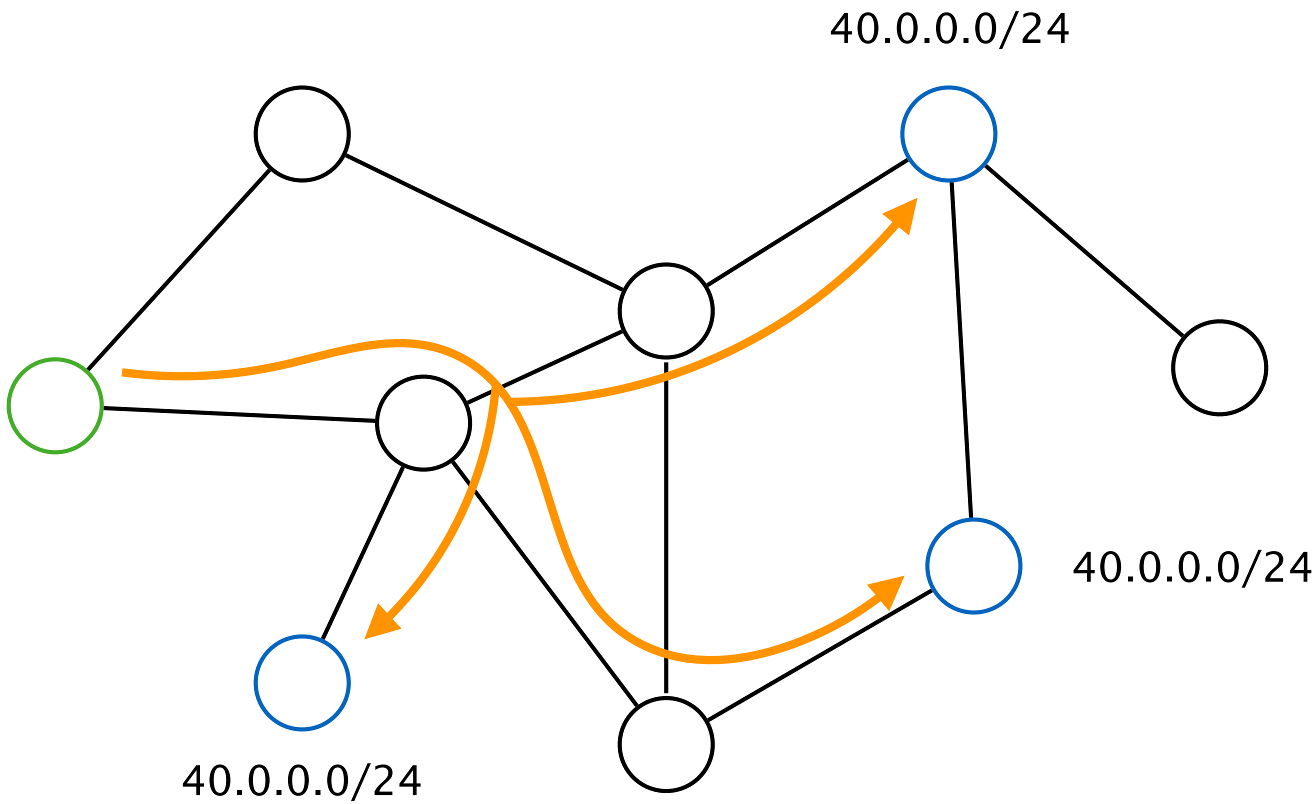


Anycast, Unicast, Multicast

„one-to-many-of-many“ („many-to-many-of-many“)

E.g. useful to stream the same video to multiple receivers

Anycast, Unicast, Multicast



Exercise 4 - Detective Work

SRC MAC Address	DST MAC Address	SRC IP Address	DST IP Address
6a:00:02:49:a1:a0	11:05:ab:59:bb:02	65.222.11.1	65.222.8.2
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.16.1
da:15:00:00:01:11	11:05:ab:59:bb:02	129.132.103.40	65.222.8.2
11:05:ab:59:bb:02	40:34:00:7a:00:01	65.222.8.2	65.222.15.254
11:05:ab:59:bb:02	ac:00:0a:aa:10:05	65.222.8.2	65.222.9.99
ac:00:0a:aa:10:05	01:05:3c:34:00:02	65.222.9.99	65.222.13.255
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.8.1

Exercise 4 - Detective Work

SRC MAC Address	DST MAC Address	SRC IP Address	DST IP Address
6a:00:02:49:a1:a0	11:05:ab:59:bb:02	65.222.11.1	65.222.8.2
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.16.1
da:15:00:00:01:11	11:05:ab:59:bb:02	129.132.103.40	65.222.8.2
11:05:ab:59:bb:02	40:34:00:7a:00:01	65.222.8.2	65.222.15.254
11:05:ab:59:bb:02	ac:00:0a:aa:10:05	65.222.8.2	65.222.9.99
ac:00:0a:aa:10:05	01:05:3c:34:00:02	65.222.9.99	65.222.13.255
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.8.1

Router interface MAC address

Exercise 4 - Detective Work

SRC MAC Address	DST MAC Address	SRC IP Address	DST IP Address
6a:00:02:49:a1:a0	11:05:ab:59:bb:02	65.222.11.1	65.222.8.2
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.16.1
da:15:00:00:01:11	11:05:ab:59:bb:02	129.132.103.40	65.222.8.2
11:05:ab:59:bb:02	40:34:00:7a:00:01	65.222.8.2	65.222.15.254
11:05:ab:59:bb:02	ac:00:0a:aa:10:05	65.222.8.2	65.222.9.99
ac:00:0a:aa:10:05	01:05:3c:34:00:02	65.222.9.99	65.222.13.255
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.8.1

Router interface MAC address

Dst 65.222.8.2 does not go over router => internal

Exercise 4 - Detective Work

SRC MAC Address	DST MAC Address	SRC IP Address	DST IP Address
6a:00:02:49:a1:a0	11:05:ab:59:bb:02	65.222.11.1	65.222.8.2
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.16.1
da:15:00:00:01:11	11:05:ab:59:bb:02	129.132.103.40	65.222.8.2
11:05:ab:59:bb:02	40:34:00:7a:00:01	65.222.8.2	65.222.15.254
11:05:ab:59:bb:02	ac:00:0a:aa:10:05	65.222.8.2	65.222.9.99
ac:00:0a:aa:10:05	01:05:3c:34:00:02	65.222.9.99	65.222.13.255
6a:00:02:49:a1:a0	da:15:00:00:01:11	65.222.11.1	65.222.8.1

Router interface MAC address

Dst 65.222.8.2 does not go over router => internal

Dst 65.222.8.1 reaches router and has to be in the same subnet as 192.168.8.2

=> 65.222.8.1 is the IP of the router

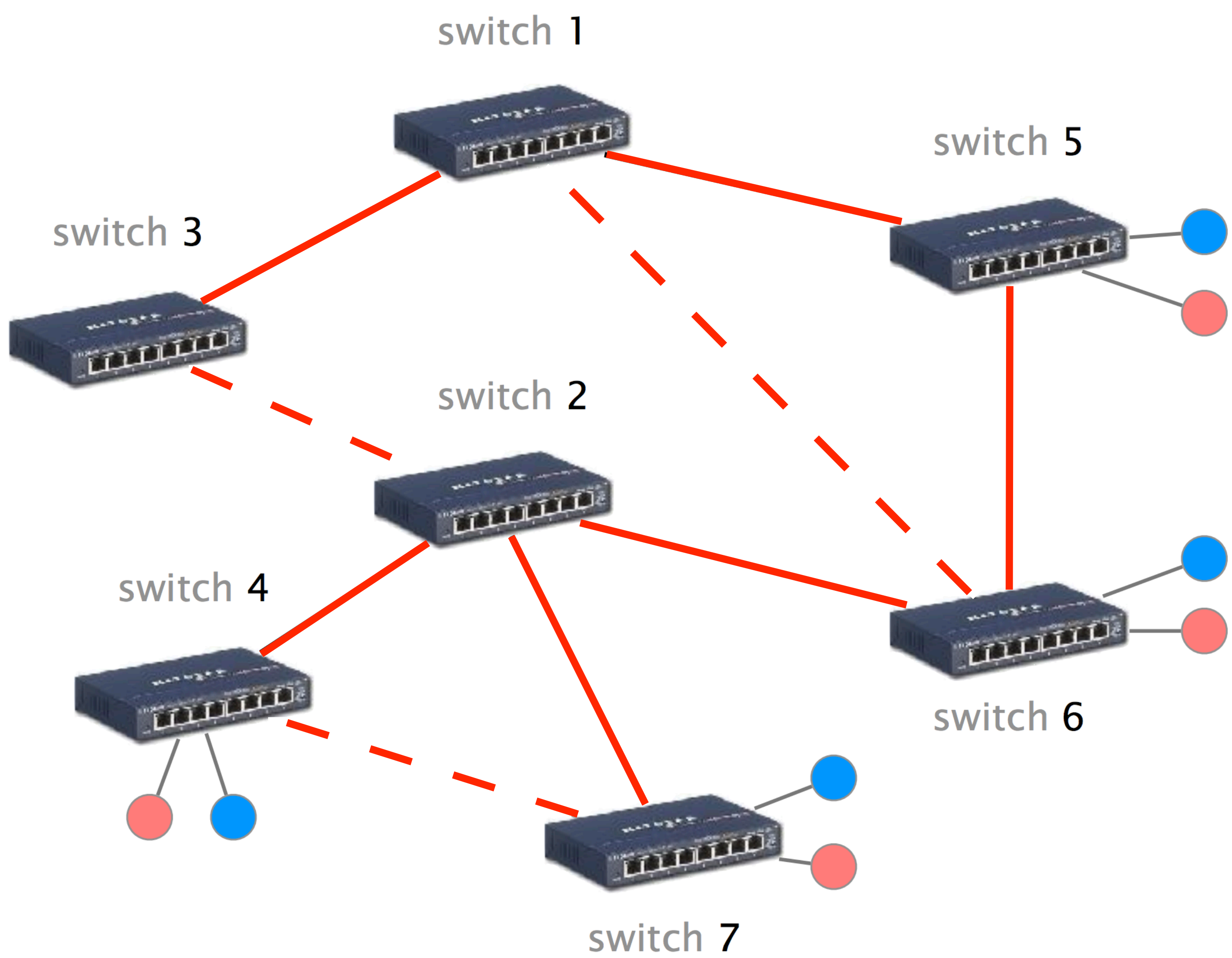
VLAN

Access link: part of only one VLAN
normally connects hosts with switches (to get „access“)

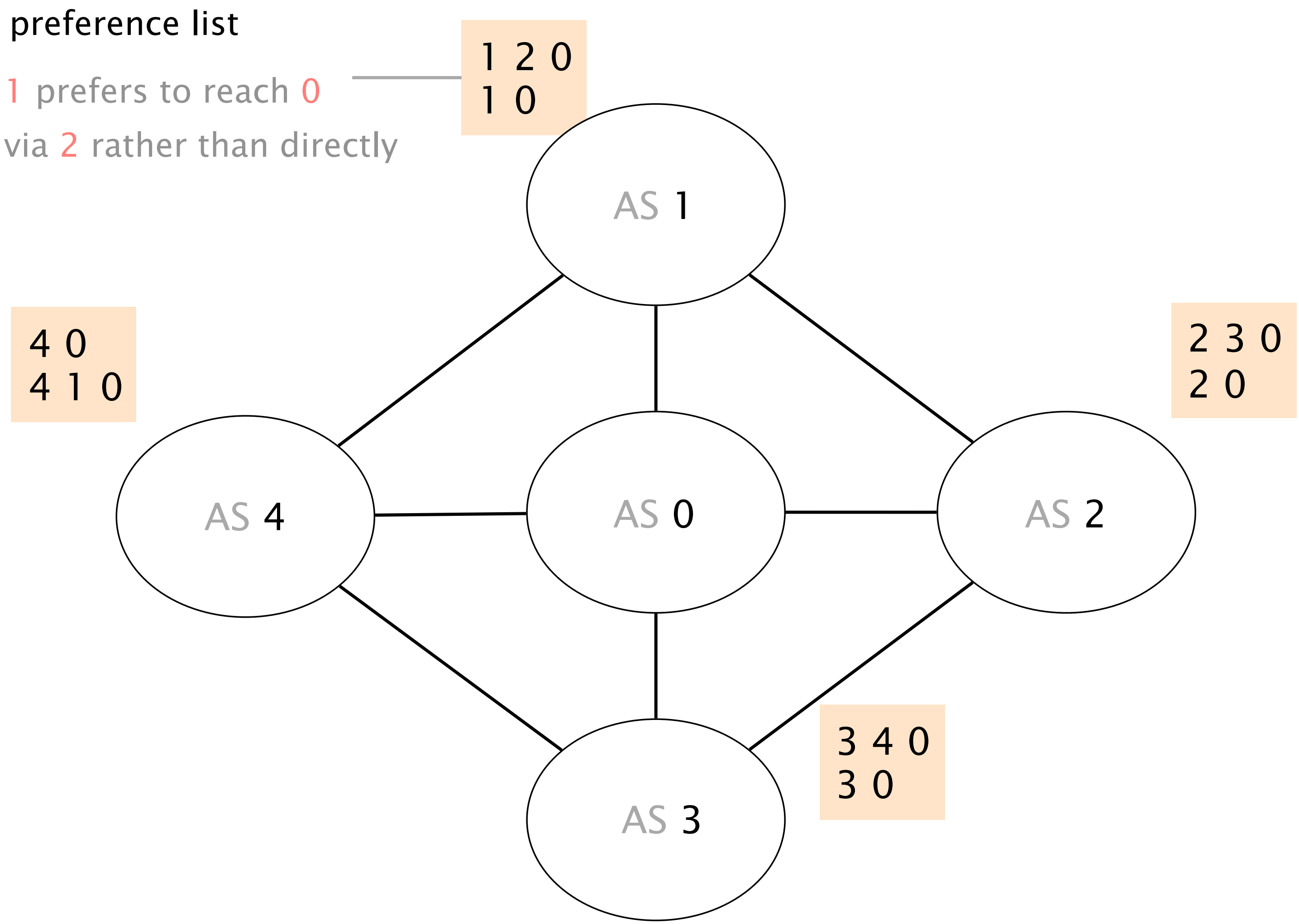
Trunk link: **can** carry traffic for multiple VLANs
normally connects switches to other switches or routers

The per-VLAN spanning tree still spans the **entire** network
even if some of the switches do not have hosts in all VLANs
=> better optimized paths for hosts in one VLAN
=> ready for new hosts in the future

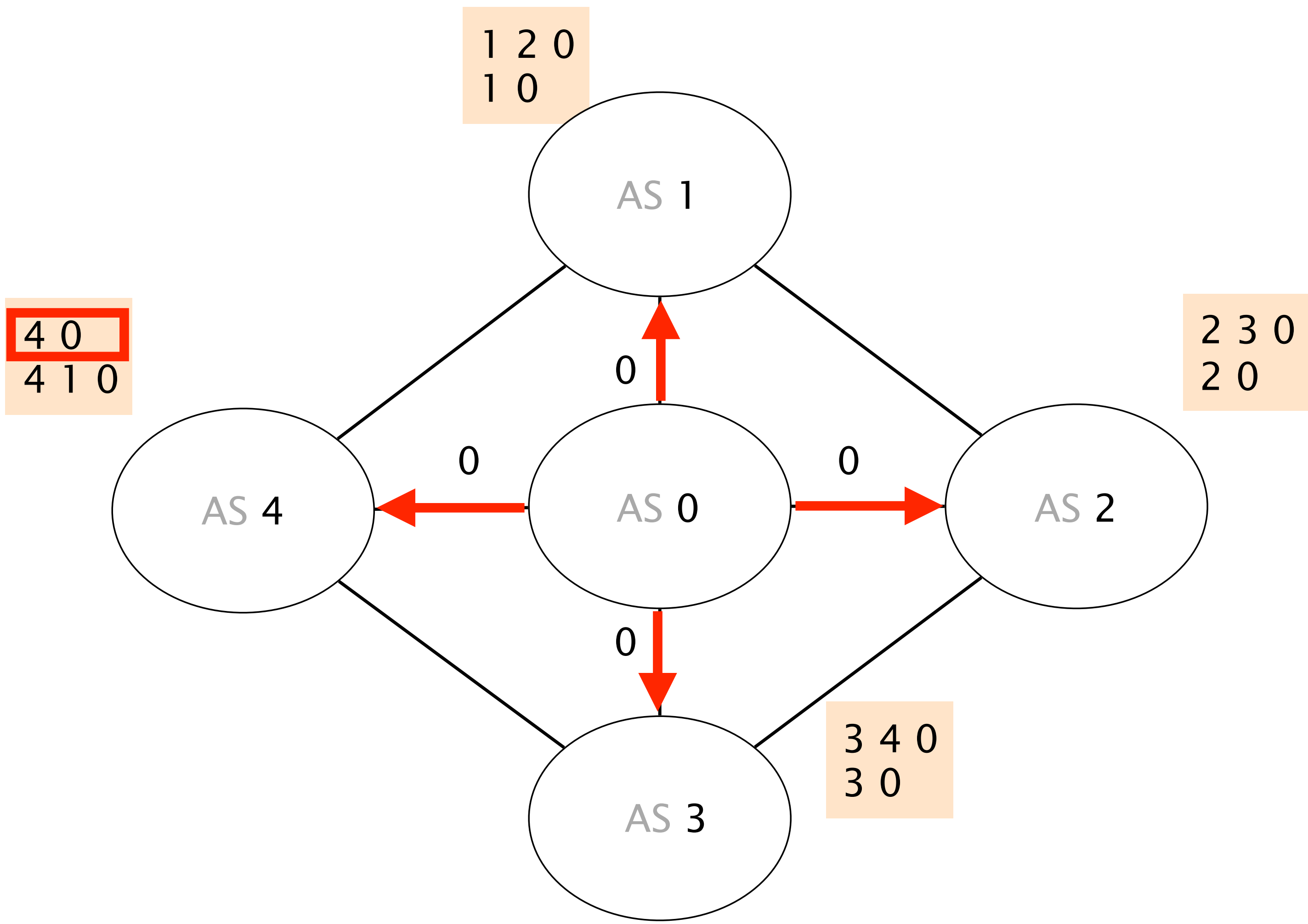
VLAN (spanning tree from the slides)



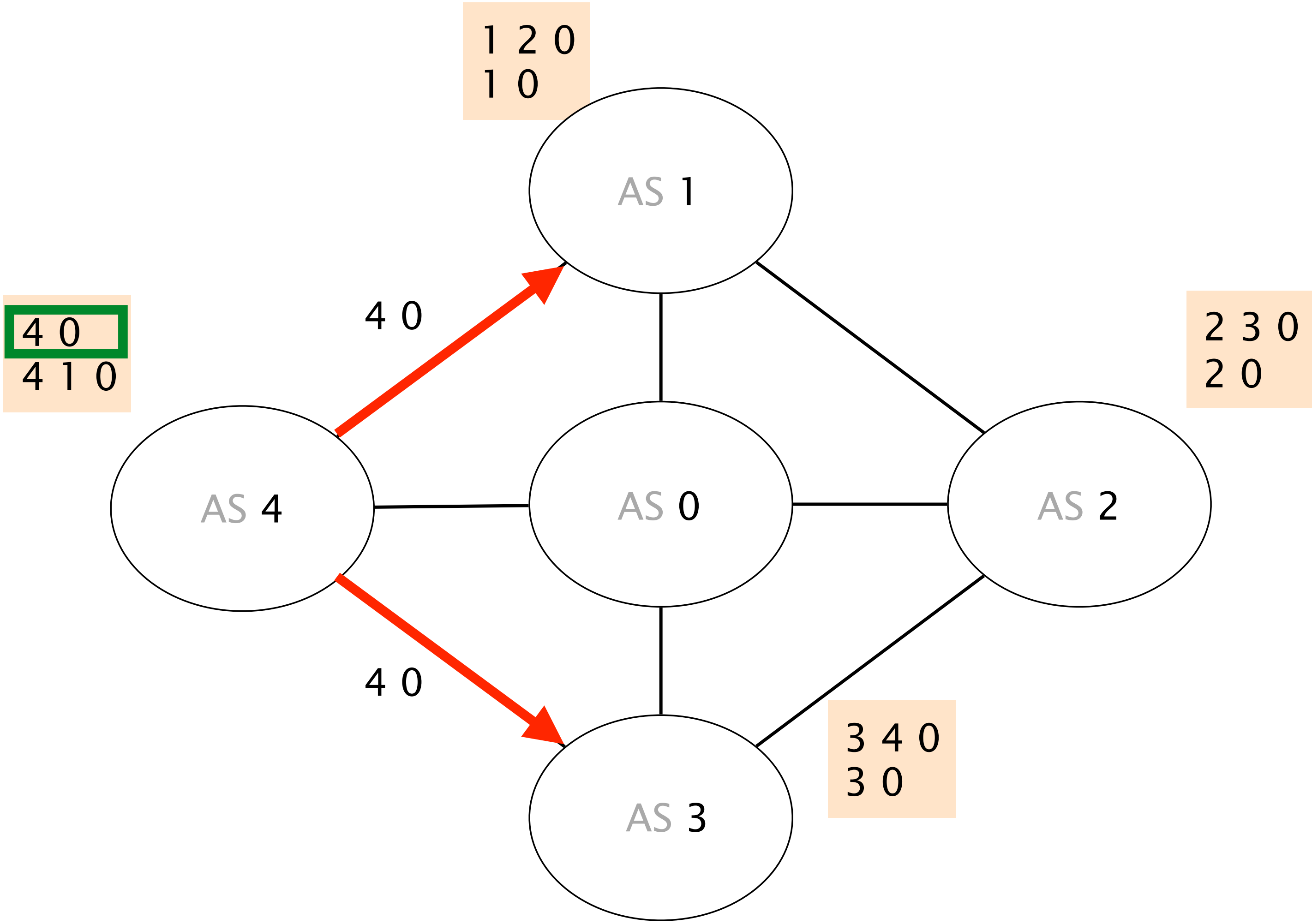
Exercise 6 - Convergence



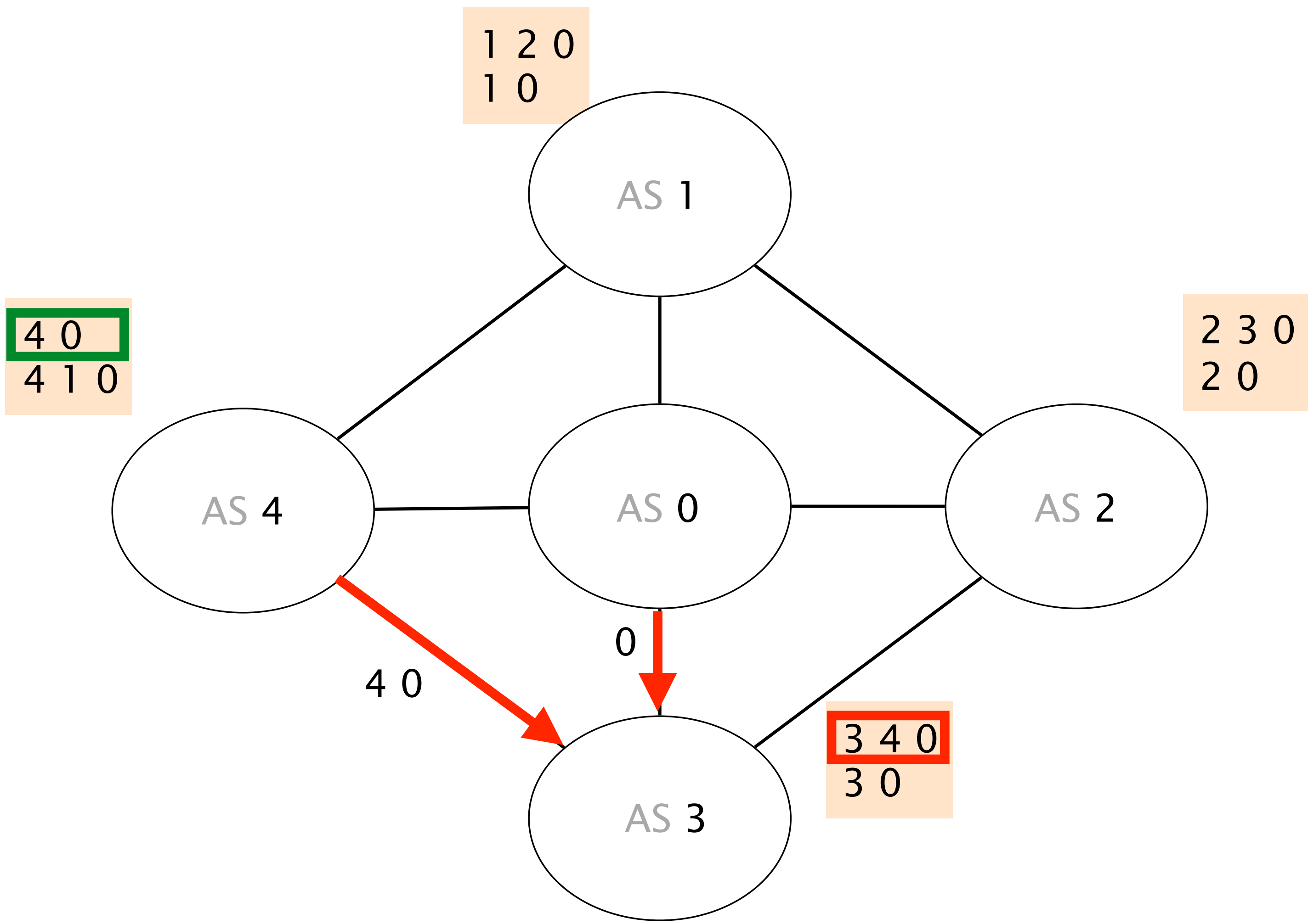
Exercise 6 - Convergence



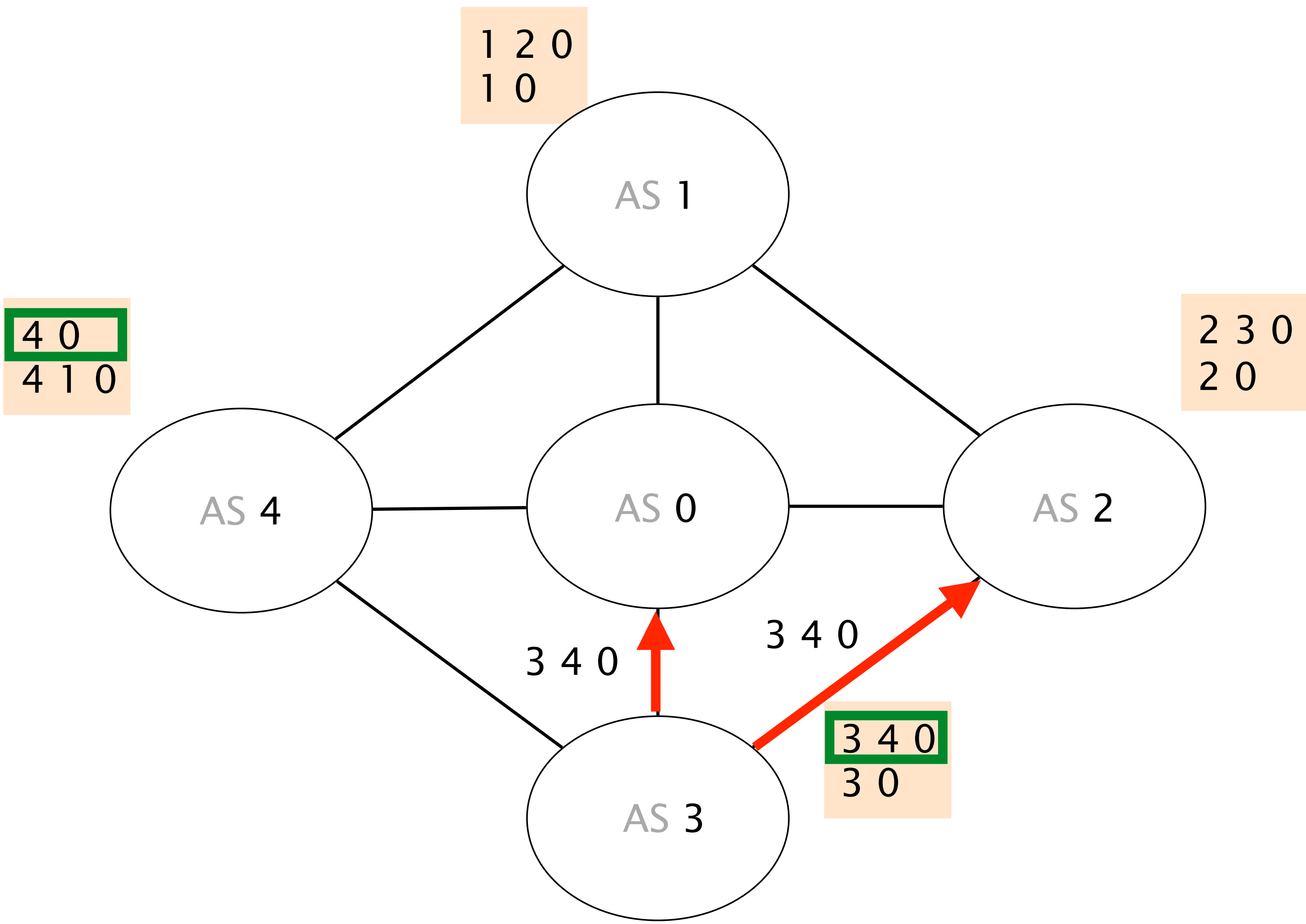
Exercise 6 - Convergence



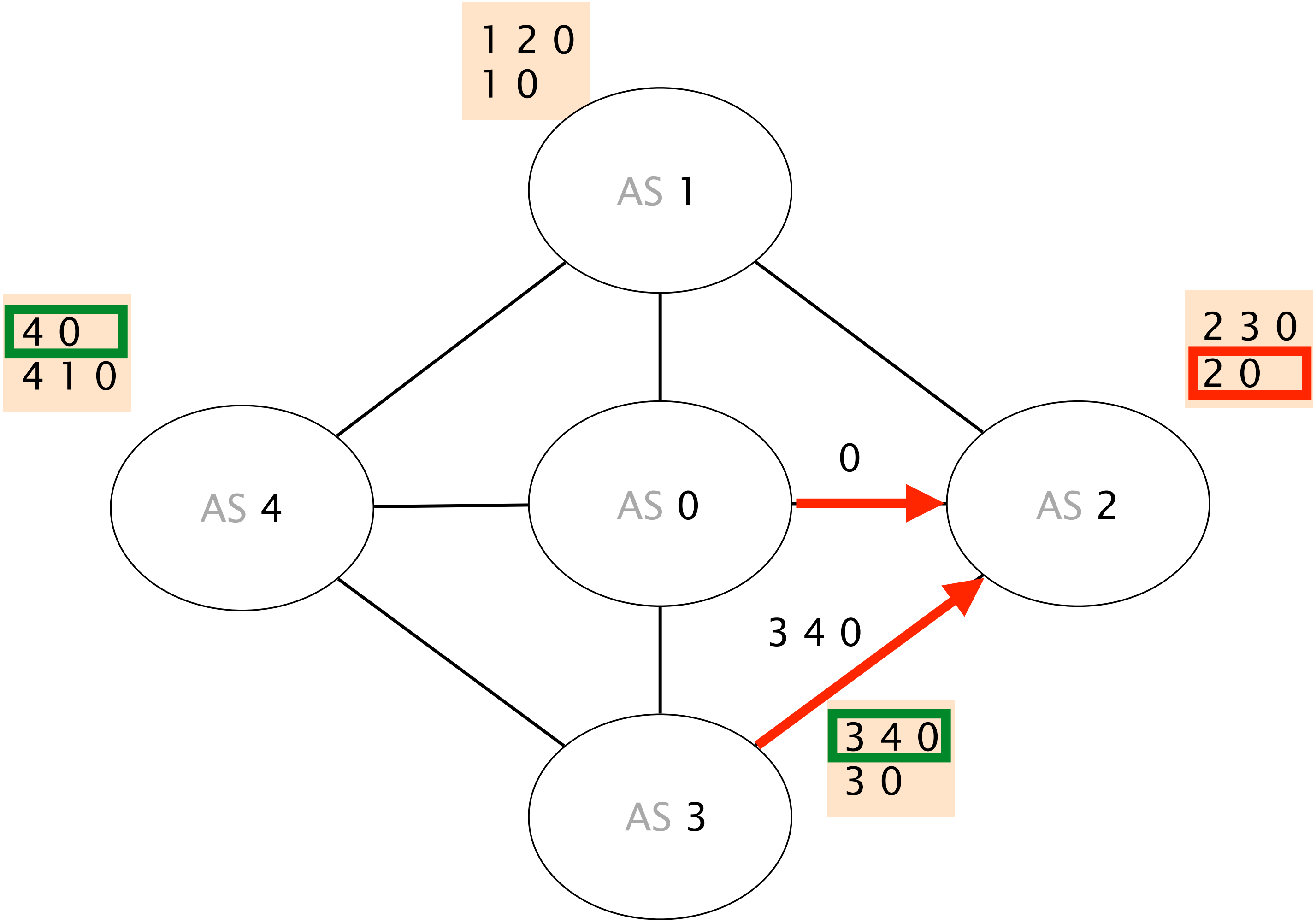
Exercise 6 - Convergence



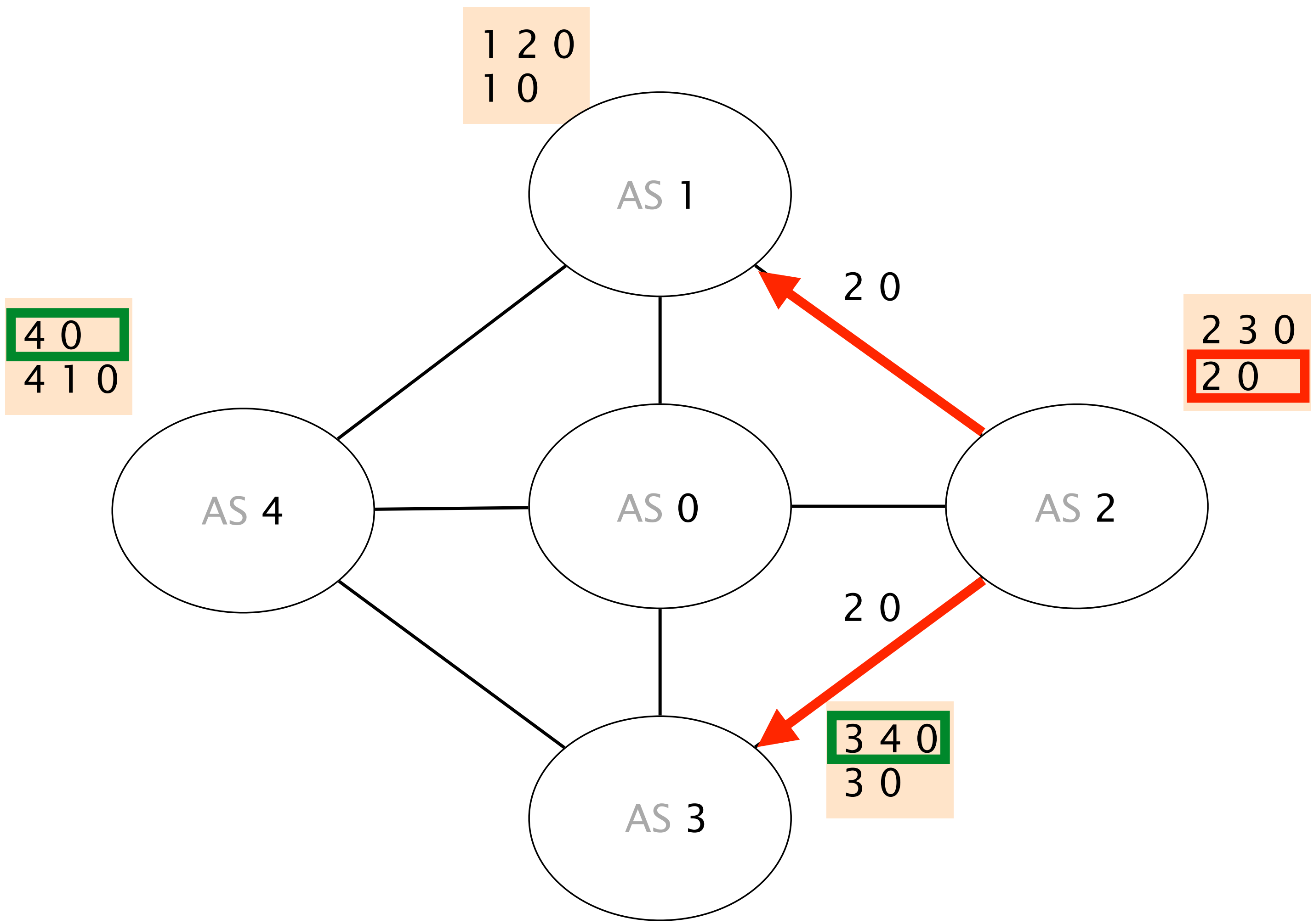
Exercise 6 - Convergence



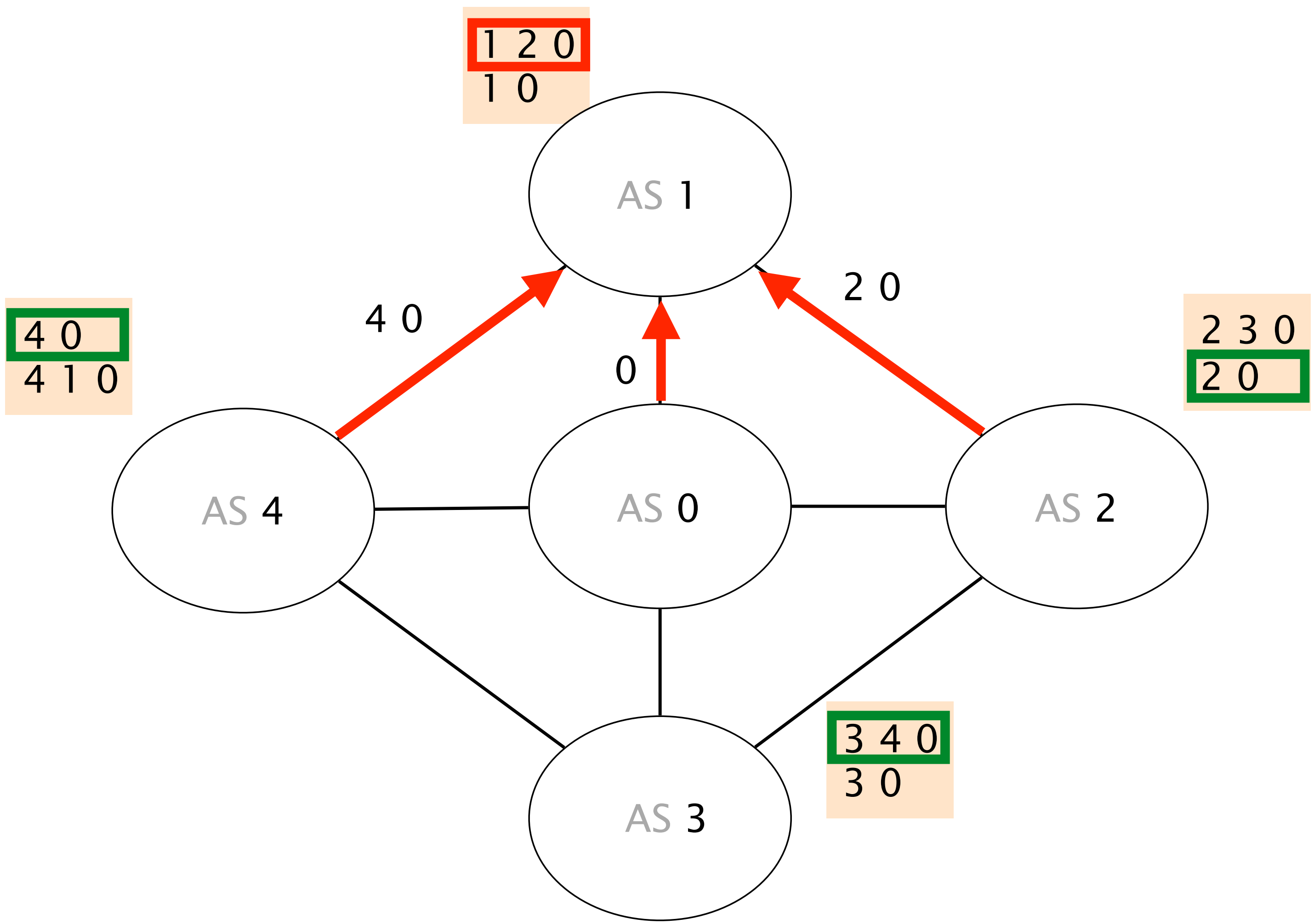
Exercise 6 - Convergence



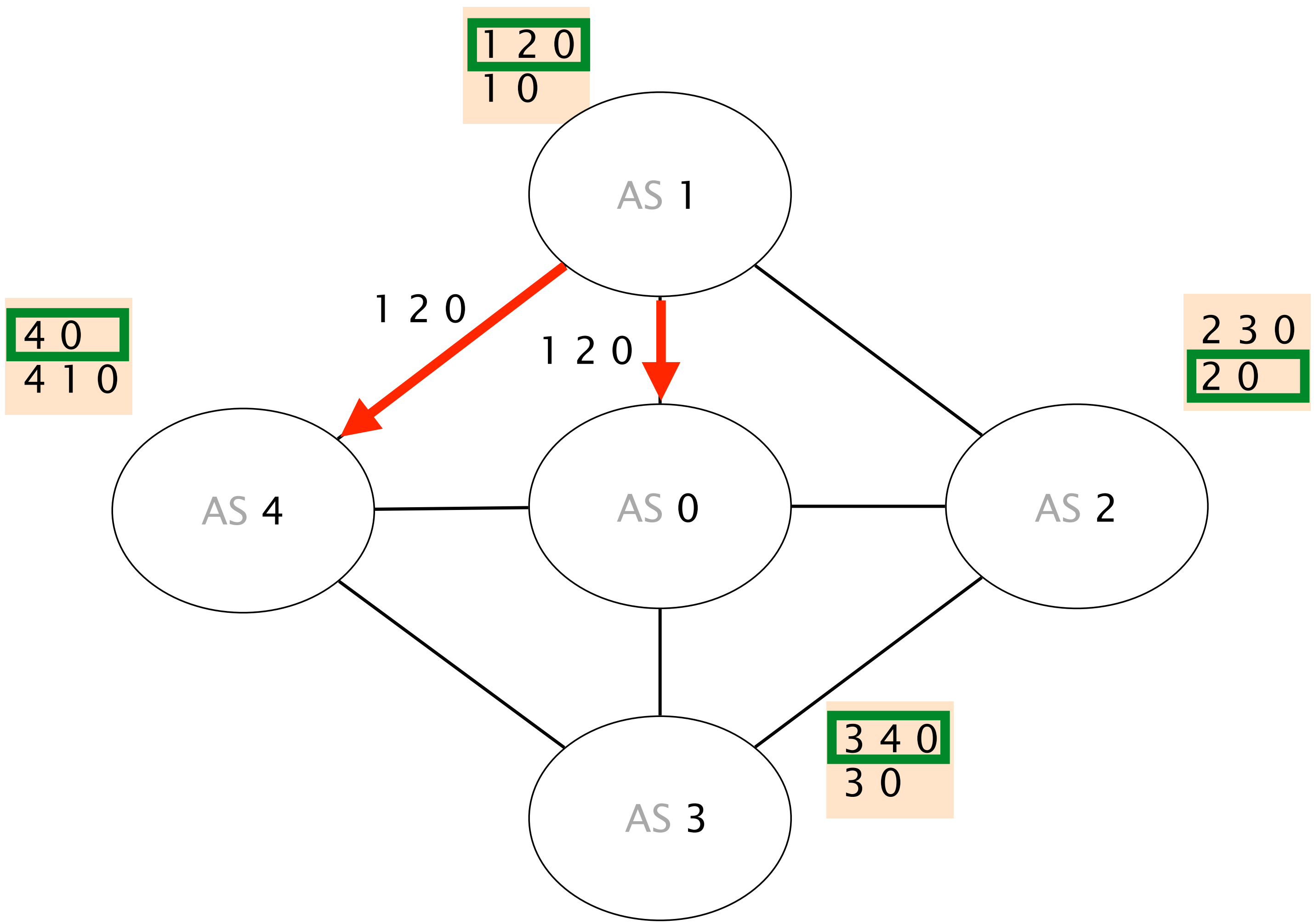
Exercise 6 - Convergence



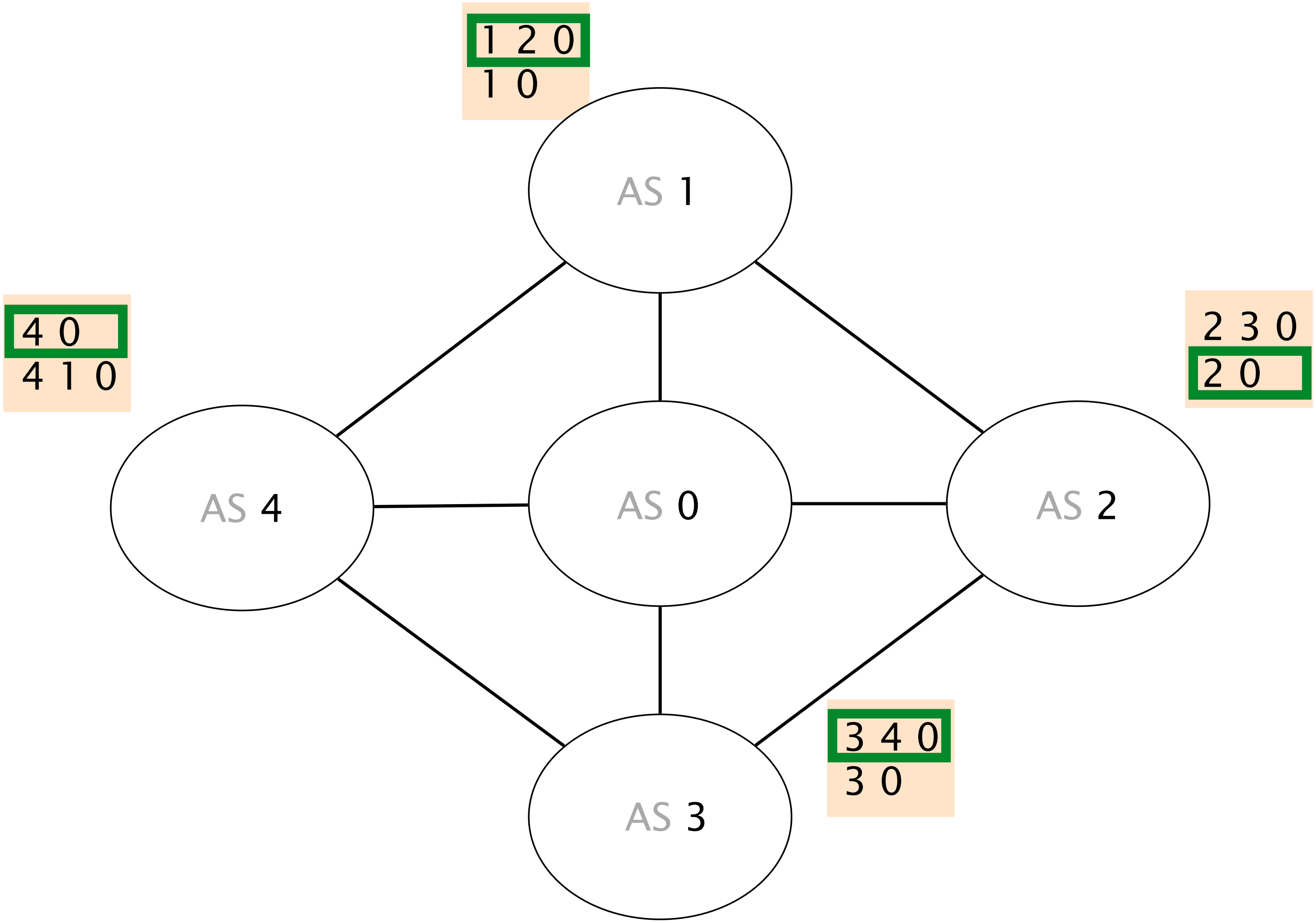
Exercise 6 - Convergence



Exercise 6 - Convergence



Exercise 6 - Convergence



Important deadline

Use Slack or email to ask your last questions until 14.08.2019

Exam: 21.08.2019 – HIL D 15

Individual Questions