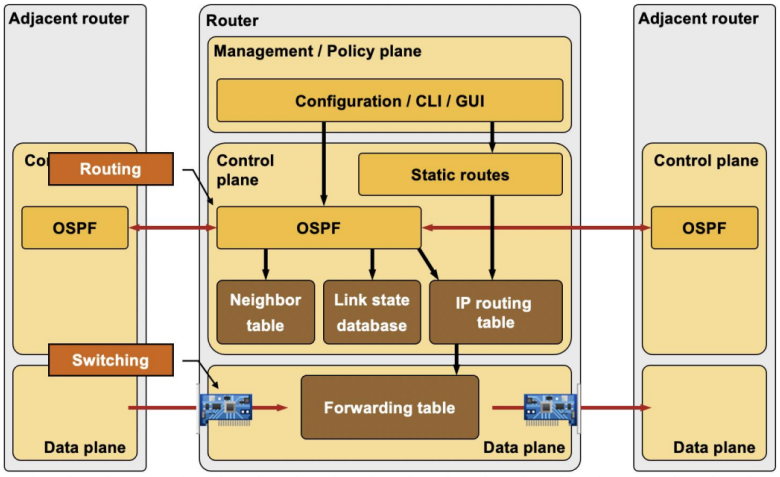
Computer Netzwerke 2 Zusammenfassung

1) Introduction

Network Planes



Slides goes into OSI Layers but nothing especially interesting.

Exercises:

- Spanning Tree → Mut zur Lücke
- Static ip route config

ip route <dst ip> <dst subnet> <ip to send to>

2) OSPF Open shortest path first

Protocol to exchange routing information. Widely used.

Each router does:

1) Establish neighbour adjacencies and exchanges LSAs (Link state advertisment)
2) Build a LSDB (link state database)
3) Run Dijkstra - SPF algorithm
4) Build its routing table

OSPF uses the following packets:

- Hello (1) (Type): Discover and maintain adjacencies → if no router is known → multicast to 224.0.0.5
  Sent every HelloInterval
  RouterDeadInterval to detect failed neighbors } default values?
- (DBD) Database Description (2): simplified version of LSA
  helps other routers detect if a new version of LSDB exists
- (LSR) Link state request (3): Sent after DBD, used to request a specific LSA
  used to update a LSDB entry, that the neighbor has a newer entry
- (LSU) Link state update (4): contains one or more LSA
  Ensuring all routers have the same view of the network
- Link state acknowledgment: Explicit ACK that a LSA was received

Flooding optimization DR/BDR/DROTHER

This does not apply if the connection is configured as point-to-point

The Designated Router does the LSA Forwarding and LSDB sync task on behalf of the routers in the OSPF domain → optimizes flooding → less packages → good

Designater Router elected based on priority

Backup designated router takes over if DR fails

all other routers are DROTHER

[eg. new network
Link state changed ? → LSU (includes LSA) to DR → ack back, floods LSU → ack back]

DR stays DR, even if router with lower prio connected. Also is not the DR again if the DR is turned off then on again

OSPF structure

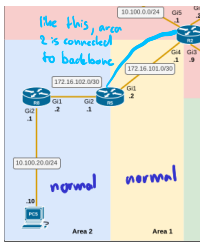
After every Link state change, OSPF needs to reevaluate its SPF calculation
This doesnt scale well at all → spread into areas

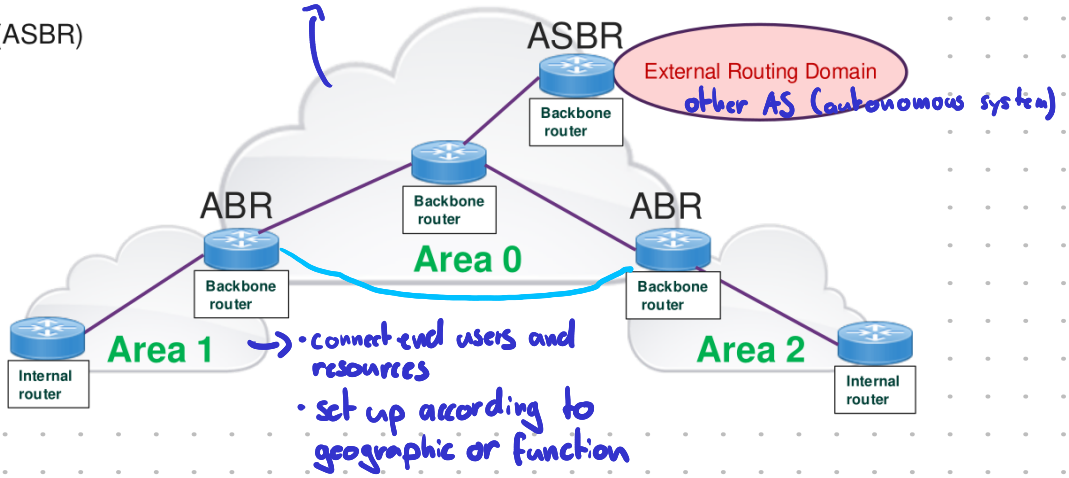
Area Border Router (ABR)

AS Boundary Router (ASBR)

Internal router

Backbone router

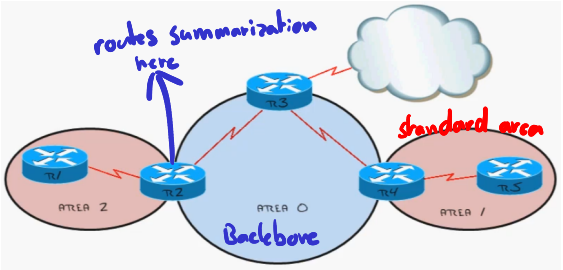




a virtual link could be connecting an area via backbone area
area 1 on left and right side in graphic above

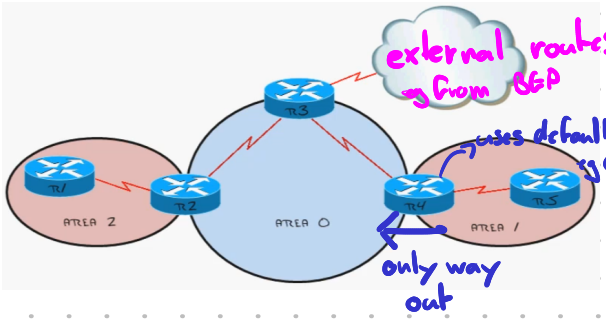
https://www.youtube.com/watch?v=nayaSIYkQp0

Areas and LSA types → YT video: CBT Nuggets Cisco OSPF areas and LSA Types



- standard area, all routes are redistributed
  can do route summarization → less routes
- Type 1 LSAs only flooded in the area the router belongs to

Type 2 LSA:
- also only within an area
- Sent by the designated router
- contains all routers from an area (slides say multi access network) ← to check what this is!
- Tells other routers who the DR is

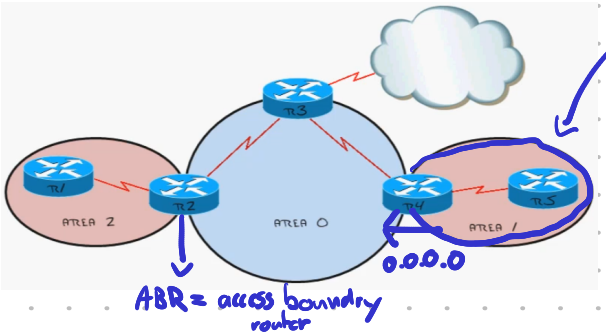


Stub area:
- limited in the routes it can receive
- doesn't want external routes
- only one way out anyway
- Blocks type 5 LSAs

Lab Exercise 3.4

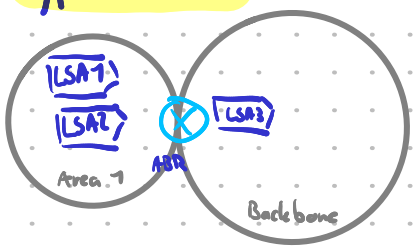
Under which circumstance does it make sense to use a stub area?

Solution to Lab 3.4
In case there are no other border routers connected to the area which connect to external routing domains. In that case all external traffic can be sent to the same border router with a default route.



Totally stubby area:
- no external routes, incl areas you control
- only default route used 0.0.0.0/0
- one way out
- Blocks LSA Type 3, 4, 5 from entering
- Routing table really efficient

Type 3 LSA:

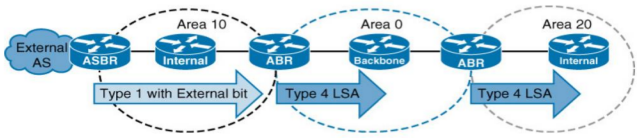


The ABR converts LSA 1 and 2 into an LSA Type 3 and forwards them to the other OSPF areas.

→ after the conversion we know the route is summarized

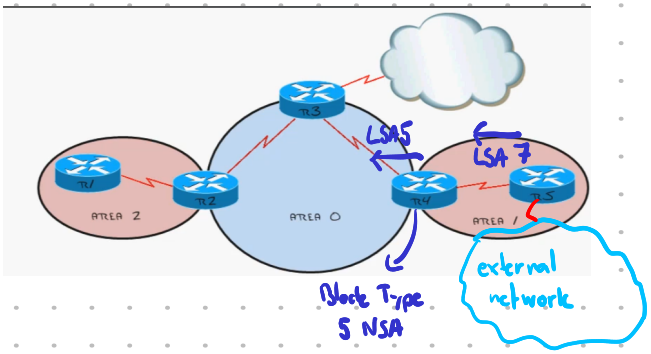
Type 4 LSA:
- This is a special use case.
- ASBR sends a Type 1 LSA with the external bit set.
- The Area Border Router converts this to a Type 4 LSA

⇒ used to find the ASBR



Type 5 LSA: Propagates external area routes in all OSPF areas. → routes from the ASBR

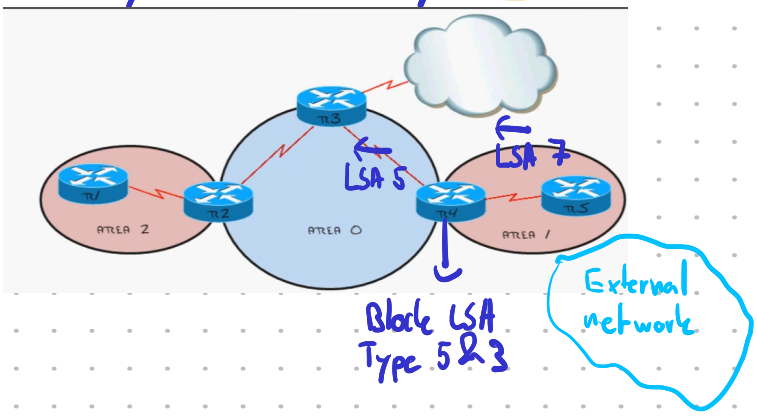
Not so stubby area (NSSA): → grew accidentally



- Same as a totally stuby area, but with an external network attached to it
- Default route not announced into NSSA area

Type 7 LSA: The Ninja LSAs
- Type 5 LSAs are blocked, so the are masked as Type 7. Once they are through the NSSA the Not so stubby area border router (NSSA ABR) converts them to Area 5.

Totally not so stubby area

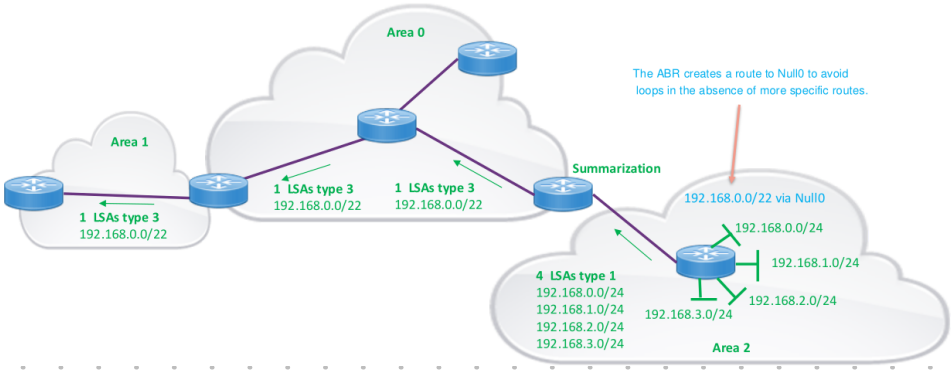


- also blocks Type 3 LSAs (and 5)
- creates a default route
- LSA Type 7 to 5 conversion as in NSSA

Route Summarization

Helps reduce two major Problems:
- Large Routing tables
- Frequent LSA flooding in an AS

Only Area Border routers and AS boundary routers do route summarization.



- The same works with Type 5 LSA.
- Route summarization requires good subnet planning.

Best Path calculation

In order to calculate the best path, all costs are added up. The route with the lowest cost is selected.

Cost = 10^8/Bandwith → 10M = 10^7 = $10^8/10^7$ = 10

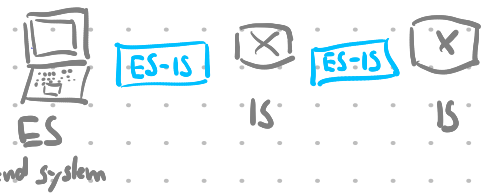
If costs are equal, the selection will work based on these attributes:

1) Intra-area, 2) Inter-Area (O), 3) External Type 1 (E1), 4) NSSA Type 1 (N1)
5) External Type 2 (E2), 6) NSSA Type 2 (N2) look at this again

I have read the lab again, unsure atm how to add this into the zsm/

3) IS-IS Intermediate system to Intermediate system

Similar to OSPF, widely used by ISPs



- also a link state protocol
- 2 level hierarchy
- simpler than OSPF
- well suited for IPv6

connectionless network protocol

No technical need for IP, other protocols work too: IPv4, IPv6, CLNP

TLV - Type Length Value

A theoretical concept of how to carry information in LSP (link state Packets). This is what makes IS-IS so adaptable, as new values can be added to value field

Type (one octet) | Length (one octet) | Value (of length L)

IS-IS adressing: Area ID (6 char). System ID (12 char). N-selector (2 char)

: 49.00A4.0000.0c11.1111.00

IS-IS PDU (protocol data units)

Hello: used to discover, build and maintain IS-IS adjacencies

Sent periodically

Mac address multicast Layer 2!

Layer 1 package sent to 01-80-C2-00-00-14
Layer 2 package sent to 01-80-C2-00-00-15

(LSP)
Link state protocol:
- Similar to OSPF LSA Type 1
- Contains TLV -> router, associated networks
- unicast on point-to-point
- multicast on broadcast media

CSNP: Complete Sequence Number PDU
summary of LSP in LSDB

PSNP: Partial Sequence Number PDU

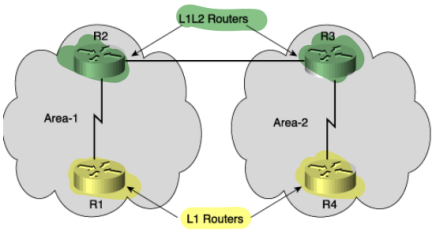
Request and ack missing LSDB from LSDB

Router levels / Areas

Level 1 router: only knows about its own area (inter-area)

Level 2 router: only knows about other areas

Level 1/2 routers: has two LSDB, on for L1 and a second one for L2.



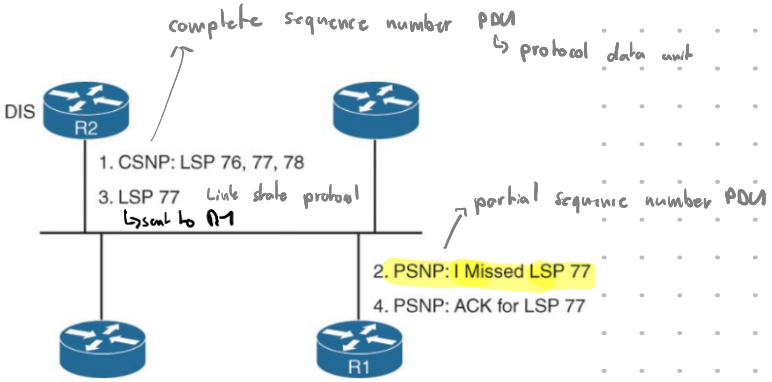
DIS - Designated Intermediate System

Similar to the OSPF designated router.

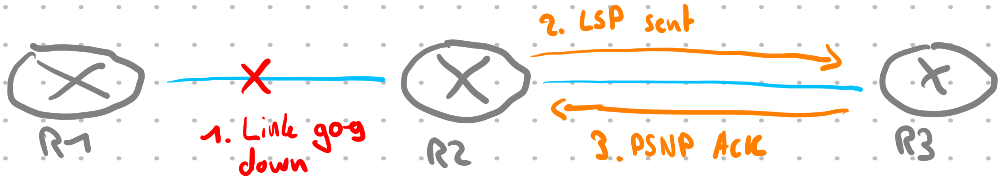
Responsibilities:
- Creating and updating a pseudonode ← why is this needed?
- Flooding LSPs over the LAN

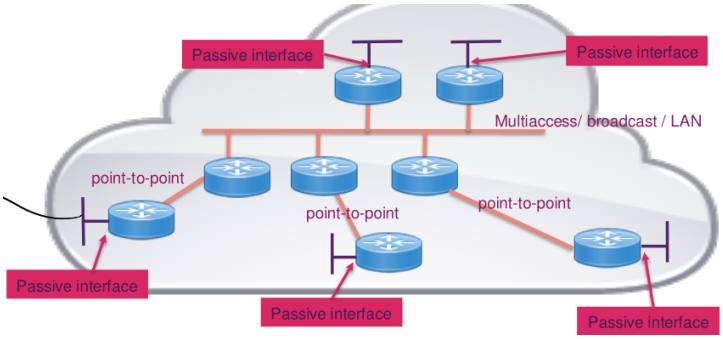
- Highest prio wins → if prio equal → highest Mac
- L1 and L2 DIS may not be same physical device
- No backup, higher priority router takes over DIS role

1. DIS R2 transmits CSNP periodically
   - Multicast MAC address
   - Ensure LSDB accuracy
2. Adjacent neighbors compare LSP within CSNP with LSDP
   1. R1 is missing LSP 77 and request it with PSNP
3. The DIS R2 sends the LSP 77 to R1
4. R1 use PSNPs to acknowledge the receipt of the LSP 77



Point-to-point has no DIS





Interfaces going outside of the IS-IS area are marked as passive

passive → no IS-IS traffic sent and incoming ignored.

This is how to lock down the IS-IS areas, otherwise somebody else can fiddle with the routes.

Areas

Why use areas ? Similar reasoning like for OSPF: makes administration easier, reduces size of LSDB, smaller SPF calculations and not every link state triggers a recalculation.

- L1 is routing intra-area and uses closest L1/L2 router to get out of area.
  ↳ default route
- ISIS L1 is equivalent to OSPF Totally Stubby area.

- L2 routing is inter-area and populates L1 routes into L2

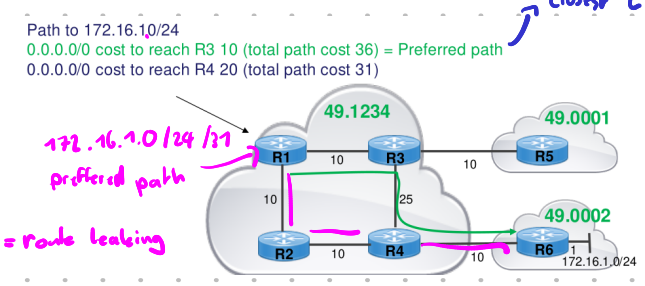
ISIS backbone can only consist of L1/L2 or L2 routers and will span into member routers in all areas.

↳ I understand this like there is one L2 backbone and every L1/L2 router connected to a separate area is part of the backbone too → to double check

Best Path selection

1) L1 intra-area routes
2) L2 intra-area routes
3) L2->L1 leaked routes
4) L1 inter-area route with external metrics ← clarify what this is exactly
5) L2 inter-area route with external metrics
6) L2->L1 leaked routes with external metrics

Route leaking

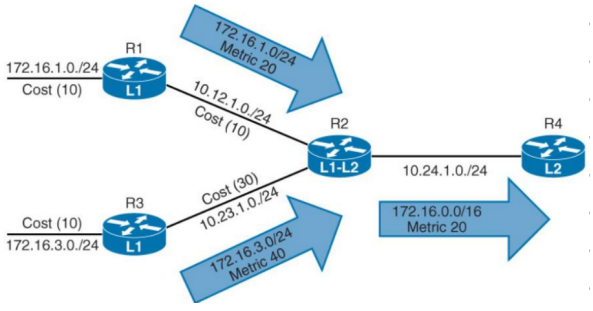


pink route more optimal

IP Summarization

as shown in picture (similar to OSPF)

lowest metric kept.

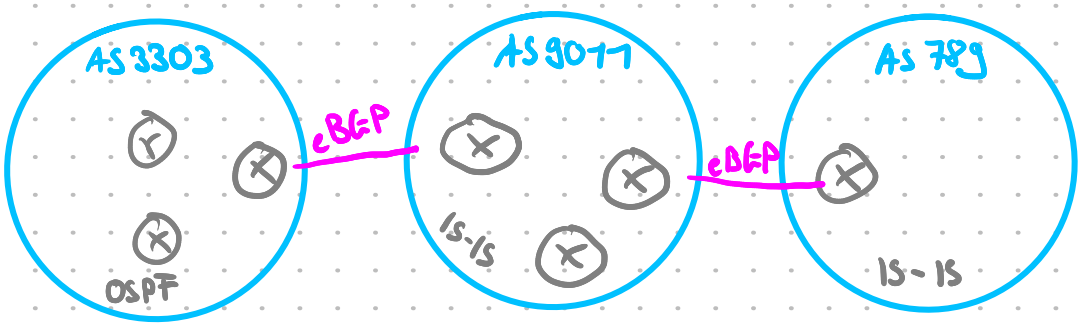


4-5) BGP

Border Gateway Protocol advertises, learns and chooses the best path inside the global internet.

Uses TCP on port 179.

BGP connects multiple Autonomous Systems together. That can be different ISPs or customer AS. AS64512-65535 are private and not routed.



BGP Messages

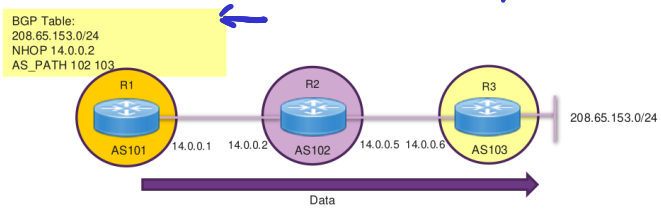
Open: Sent to establish a BGP peer
contains AS number, router ID and hold time
↳ we can detect if we are in the same AS

BGP Update: Transfer routing information between peers
is composed of:
↳ Withdraws: Routes to remove
↳ Attributes: AS path, many others
↳ Prefixes: new routes to add

→ distributed into:
↳ well-known mandatory: AS path, origin, nexthop
↳ well-known discretionary: local preference
↳ optional - transitive
↳ optional non-transitive

Mut. zur Lücke, dass kann ich mir eh nicht merken.

BGP Update ←, data flows the other direction



Loop prevention: if his own AS number is in the AS path, the update will be ignored.

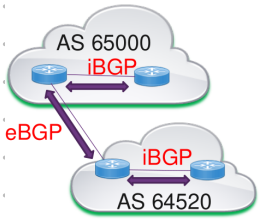
BGP route selection:

More specific route wins!

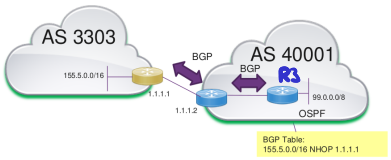
- BGP uses 11 steps to determine the best path:
1. Prefer highest weight (local to router)
2. Prefer highest local preference (global within AS) if network is directly attached → preferred
3. Prefer routes that the router originated
4. Prefer shorter AS paths (only length is compared) route with less AS numbers
5. Prefer lowest origin code (IGP < EGP < Incomplete)
6. Prefer lowest MED (also called metric)
7. Prefer external (EBGP) paths over internal (IBGP)
8. For IBGP paths, prefer path through closest IGP neighbor
9. For EBGP paths, prefer oldest (most stable) path
10. Prefer paths from router with the lower BGP router-ID
11. Prefer the path that comes from the lowest neighbor address

Put this on sheet to take to exam

eBGP and iBGP



- eBGP refers to BGP between different AS
  - ↳ next hop automatically updated
- iBGP refers to BGP in an AS
  - ↳ next hop not updated by default.
    - ↳ turn on next-hop self feature

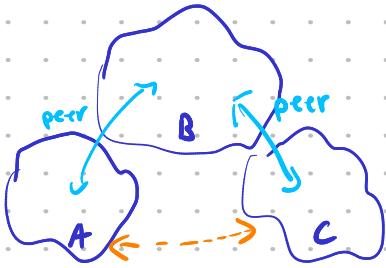


next hop for R3 not updated → not reachable

Peering and Transit

Peering: · ISPs provide reachability to eachother
· peering partners considered equal

Transit: ISPs provides reachability to all destinations
· smaller ISPs transit through bigger ISPs
· smaller ISPs pay the bigger ISPs



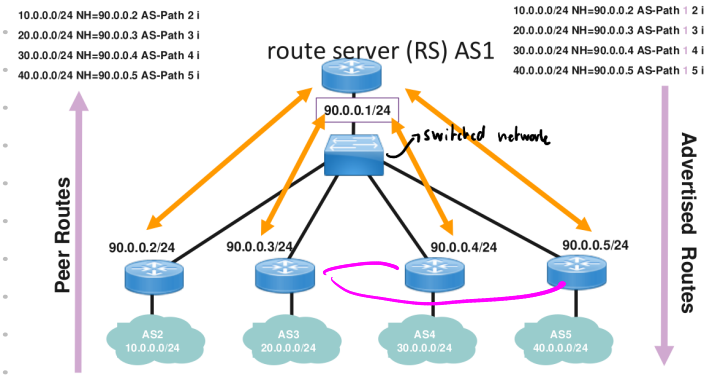
ISPs usually don't like being used as the transit AS. If A were to connect to C, via B → b would be the transit AS.

ip as-path access-list 10 permit ^$ → This regex only allows local originated routes and disallows the a to c peering via B

These ISP peering/transit settings are a big and complicated task.

IXP - Internet Exchange point

an IXP allows CDN and ISPs to connect via a switche network.



private peering also possible, but requires contract with each peer.

Look into this again! Slides don't offer too much

Start BGP advanced slides...

BGP pre pending feature → to influence incoming traffic ← How?

AS 2579 3257 559 559 559

Enterprise Internet Connectivity Options

Single-Homed

No BGP:
- most home setups
- customer has default route 0.0.0.0 to ISP
- static IP from ISP to customer network

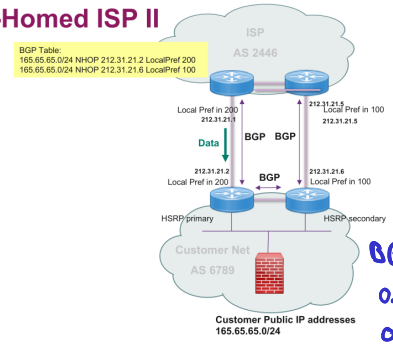
BGP: client runs his own AS

BGP update sent to ISP

customer BGP Table: 0.0.0.0/0 NHOP x.x.x.x

ISP BGP table: exact IP NHOP y.y.y.y

Dual-Homed ISP II



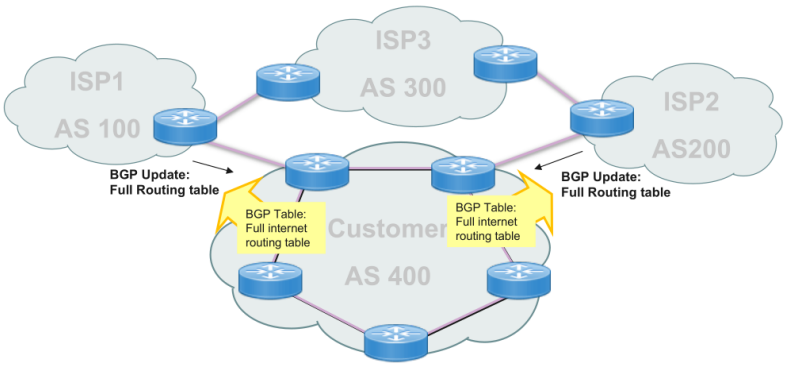
- customer runs bgp and uses localPref to determine which link is used
- ISP/customer can use MED multi-exit descriminator to influence how traffic leaves the network

BGP Table:
123.2.2.0/24 NHOP R1 MED 50
123.2.2.0/24 NHOP R2 MED 100

BGP Table:
0.0.0.0/24 NHOP 212.31.21.5 localpref 100
0.0.0.0/24 200

- also possible to use multiple metrics

Multi-Homed



- used in an active/active design
- AS400 doesn't want to be transit, only routes from his AS advertised to ISPs

Routing Security

BGP is not secure by default. Eg Youtube had 208.65.153.0/22 announced, and the pakistani announced 208.65.153.0/24 -> more specific. Which led to all yt traffic being rerouted to Pakistan.

Other cases happened too.

Solution: RPKI - Resource public key infrastructure that is essentially a big PKI where a router can look up if an address is valid

Prevents BGP route highjacking and accidental misconfigurations

a certificate authority in RPKI is called trust anchor. Typical chain of trust implemented. a ROA is the object signed, containing Prefix [208.65.153.0], maximum prefix length [/24], AS Number [AS 17285]

6) Network Design

Pillars of networking: } Cost
- Scalability
- Speed
- Availability
- Security
- Managability

calculation example ↗ network design übungen

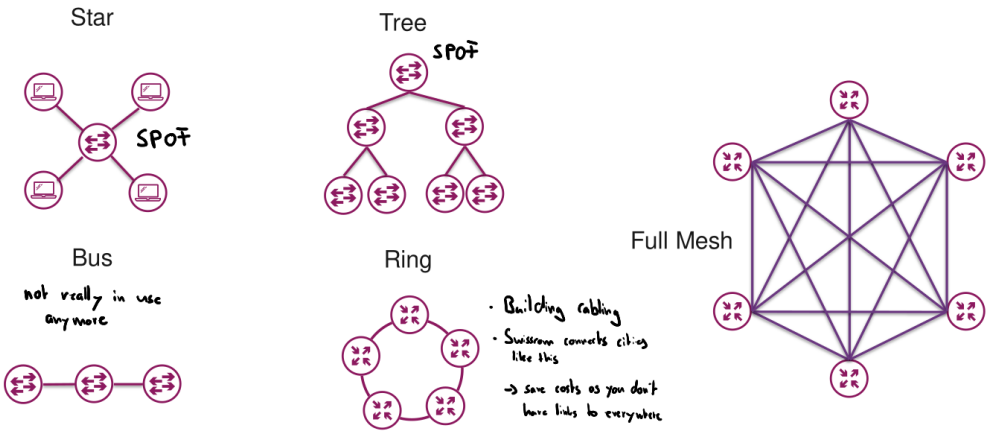
$$MTBF_{combined} = \sum\left(\frac{1}{MTBF_n}\right)^{-1}$$

$$MTBF_{parallel} = \sum\frac{1}{n}\left(\frac{1}{MTBF}\right)^{-1}$$

$$Availability = \frac{\text{Mean time between failure}}{MTBF + \text{Mean time to repair}}$$

Reasons for downtime: 1/2 hardware failure, 1/4 human error, 1/4 software failure, natural disasters, ...

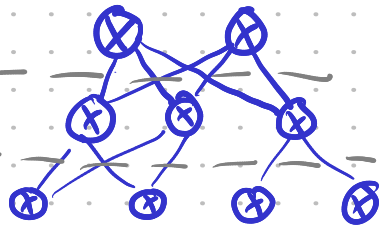
Well-Known Topologies



A campus network usually connects multiple buildings

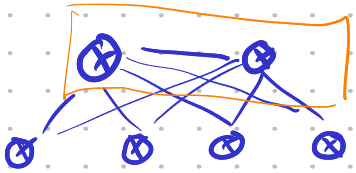
In a hierarchical design, we have:
- core layer
- distribution layer[2]
- connection layer

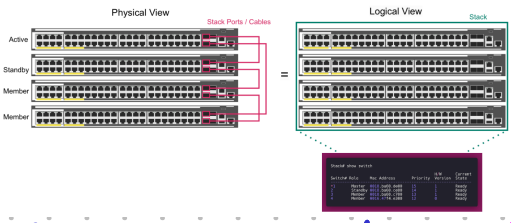
↳ still most used
- easy to scale up, by simply add more hardware here
- clients connect -> connection layer



[2] aggregates data from multiple switches and goes to core

Collapsed core for smaller networks:
↳ get rid off distribution layer





Switch stacking is aggregates switching locally to make management easier.

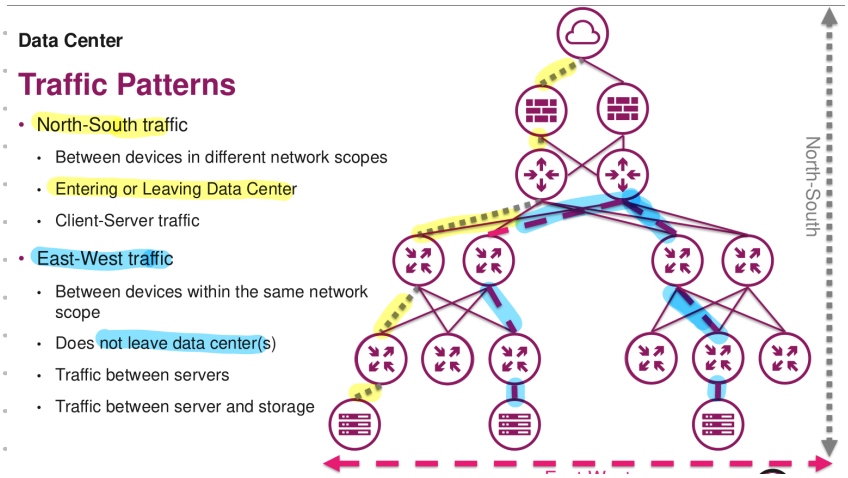
First Hop redundancy: Todo

Virtual router redundancy protocol (VRRP): multiple routers (↗ one active, the other hot standby) sharing one IP (↗ two different Macs), set this IP as default gateway -> redundancy

The better version of VRRP is Gateway Load Balancing Protocol GLBP

Same IP and the same virtual Mac is used . Default gateway is load balanced

Layer 2 is bad , layer 3 preffered.



Exercises have a massive diagram about how to design a campus network

This is probably a subject that is worth discussing with Ramon.

7) Multicast

Unicast: Sender sends N unicast packages to N receivers (even those that don't want it)

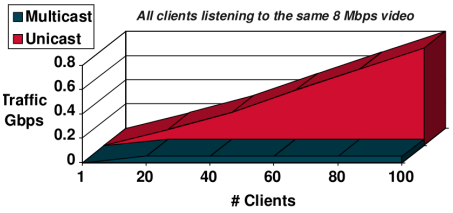
Broadcast: One broadcast package sent per network to N receivers. Can't use 255.255.255.255 as this adress is not routed -> 139.1.1.255

Multicast: One multicast package sent and received by everyone that listens to it.

- only received by those who want the package
- best balanced solution
- routers must actively participate in multicast.

Used in one to many streams:

- Iptv
- live video streaming
- live radio



Multicast is UDP based -> best effort, no congestion avoidance, recipient must expect duplicated packages

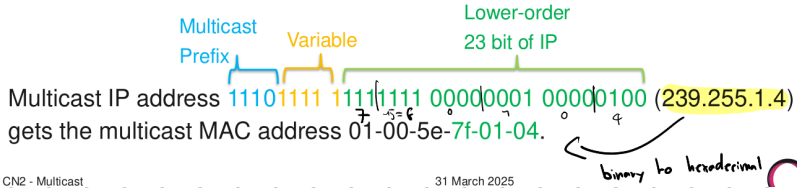
Multicast sender: sending the packages

Multicast receiver: must be multicast capeable, raises interest in multicast group -> will receive the multicast traffic.

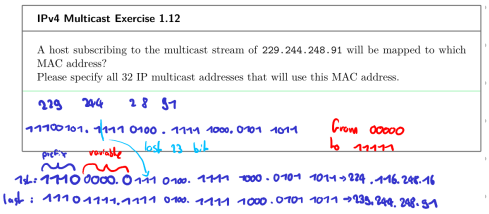
Reserved Multicast Addressing

- **Reserved link-local addresses**
  - 224.0.0.0–224.0.0.255
  - Transmitted with TTL = 1
  - Examples
    - 224.0.0.1 All systems on this subnet
    - 224.0.0.5 OSPF routers
    - 224.0.0.13 PIMv2 routers
    - 224.0.0.22 IGMPv3
- **Other IANA reserved addresses**
  - 224.0.1.0–224.0.1.255
  - Not local in scope (transmitted with TTL > 1)
  - Examples
    - 224.0.1.1 NTP (Network Time Protocol)
- **Administratively scoped addresses** ← our own application
  - 239.0.0.0–239.255.255.255
  - Private address space
  - Similar to RFC1918 unicast addresses
- **SSM (Source Specific Multicast) range**
  - 232.0.0.0–232.255.255.255

IP multicast Map address mapping



→ 32 IP addresses get the same MAC!



Layer 2 Multicast / IGMP

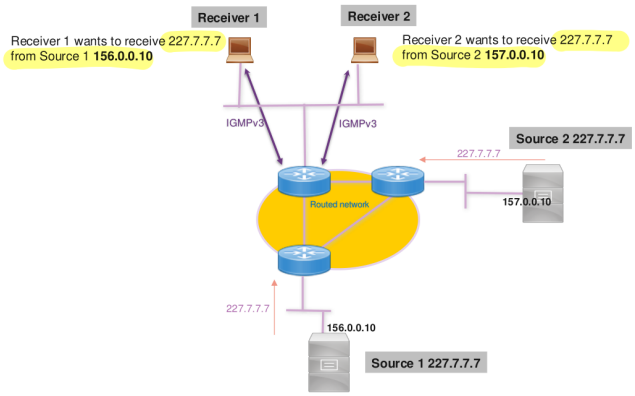
IGMP → Internet Group Management Protocol

A client sends a IGMP membership report to 224.0.0.2 (all routers) to join a multicast group

IGMPv1 has no mechanism to leave a group → v not used anymore

In IGMPv2 a user can request to leave a group → also sending to 224.0.0.2

Host-Router Signaling: IGMPv3



IGMPv3 has an exclude list

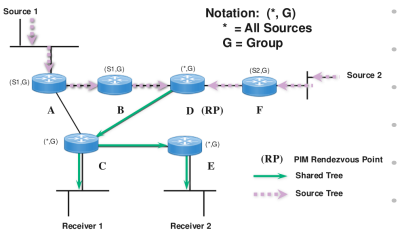
IGMP snooping on switches:

- a switch builds an IGMP table and can forward on the correct port
- If IGMP snooping is disabled, a switch forwards multicast frames to all except receiving port.

IP Multicast on Layer III

- Multicast group adresses are not in the unicast route table → seperate table

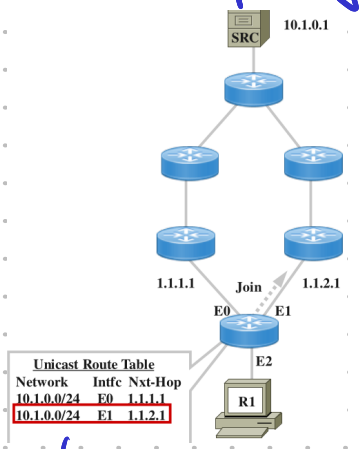
Source based Tree (S, G) S= source, g= group, * = all sources



- Multicast builds a tree from dst to src

Reverse Path Forwarding

Multicast packages are dropped, if it isn't the shortest path



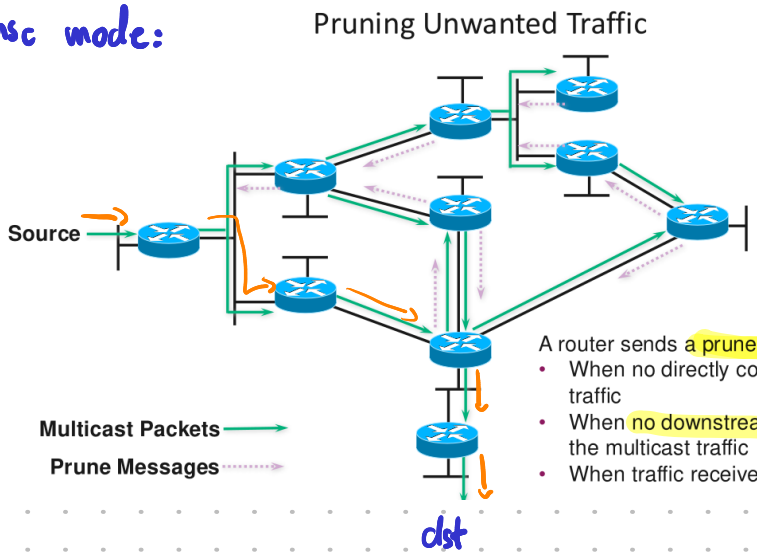
- Tree built from dst to src
- Incoming multicast traffic must match the unicast route interface, otherwise it isn't the shortest path and the multicast packet is dropped.

↳ if equal cost, higher IP wins

PIM: Protocol Independent Multicast

IGMP v2 and v1

Dense mode:



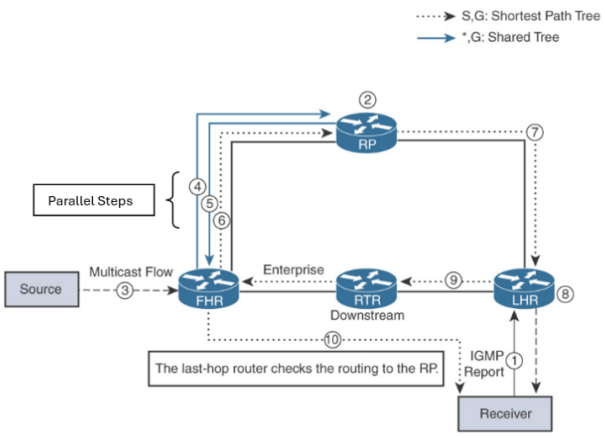
Final tree

- Flood and prune repeats every 3min

A router sends a prune message:

- When no directly connected hosts receiving the multicast traffic
- When no downstream routers are interested in receiving the multicast traffic
- When traffic received on a non-RPF interface

PIM sparse mode - Any source multicast



IGMPv1

1. Receiver sends an IGMP request to the last-hop router (LHR), also known as the gateway. The LHR reviews its configuration, containing RP information, for the requested group and sends a request in PIM control protocol towards the RP.
2. The (*, G) update is used to provide the RP information of an active receiver for the multicast group.
3. The source sends the multicast stream to the first-hop router (FHR).
4. The FHR, configured with PIM sparse-mode and relevant RP information, encapsulates the multicast stream in a unicast packet called the register message and sends it to the RP.
5. The RP already has a receiver that wants to join the multicast group using the (*, G) entry. Consequently, the RP sends an (S,G) PIM join towards the FHR. It also sends a register stop message directly to the FHR.
6. After receiving the register stop message, the FHR sends the multicast stream to the RP as an (S,G) flow.
7. Based on the interested receiver, the outgoing interface list is populated, and the (S,G) flow is sent to the LHR via the PIM-enabled interfaces aligned to the unicast routing table. From the LHR, the multicast packet flows to the receiver.
8. At the LHR, the router checks its unicast routing table to reach the source.
9. If the check verifies an alternate path is more optimal based on the unicast routing protocol, the (S,G) join is sent to the source from the LHR.
10. The source sends the multicast flow to the LHR via the (S,G) state and is created in the multicast routing table. When the data flow has migrated to the shortest path tree (S,G), the multicast LHR sends an RP prune message towards the RP.

Source specific multicast SSM

- This is the protocol you want to use!
- It seems so simple, I don't know why the other versions are so complicated

How Does SSM Work?

Here's how the components interact in SSM:

- **Source (S)**: The device that sends multicast data.
- **Group (G)**: A multicast IP address (in the range 232.0.0.0/8 for IPv4).
- **Receiver**: Subscribes to a specific **(S,G)** channel.
- **Router**: Forwards only traffic from S to receivers that explicitly request (S,G).

SSM uses the **IGMPv3 (IPv4)** or **MLDv2 (IPv6)** protocols, which allow receivers to specify the source address in their join requests.

Example

Imagine a live football stream:

- Source IP: `192.0.2.10`
- Multicast Group: `232.1.1.1`

A receiver that wants to watch this match would send an IGMPv3 "Include" request:

> "I want to receive traffic from 192.0.2.10 to group 232.1.1.1"

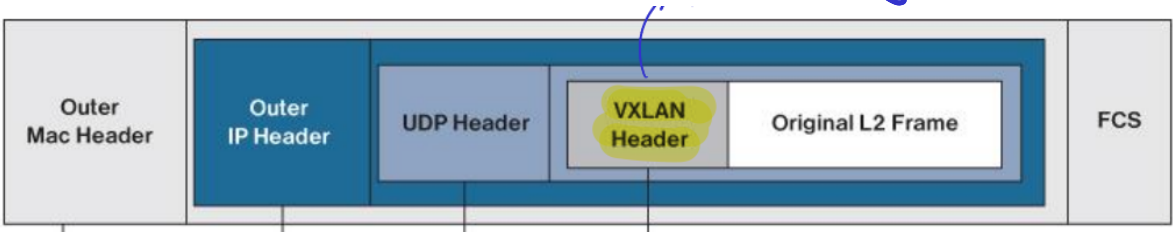
Routers will then only forward packets from that source to that group.

08) EVPN / VxLAN

Provides enterprises a solution to offer efficient L2/L3 connectivity across a campus network!

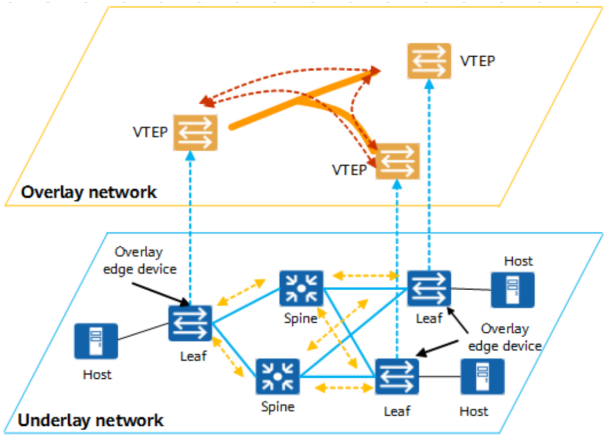
VxLan: Virtual extensible Local area network

- UDP encapsulation on port 4789
- VTEP virtual tunnel end points
- 24 bits VNI fields → ~16 Mio segments



Encapsulated VxLAN header

Overlay / Underlay:



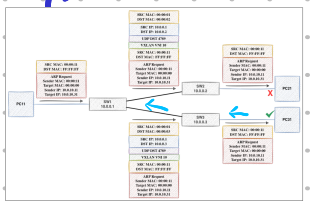
- Overlay is a virtual L2 network that relies on the underlay for packet forwarding
- does encapsulation

- Provides L3 IP connectivity as the underlay network.
- OSPF, IS-IS, BGP

BUM Traffic = Broadcast, unicast and multicast traffic

There are different ways to learn how to learn Mac-adresses

Flood and learn



- The learned MACs are then sent back ←

Static

Flood and learn with multicast: · multicast groups used, either statically configured or via control plane

[more details in exercises

Exercise 2.13 Flood and learn

Briefly summarize the learning process of the switches involved when srv-03 connects to srv-06 the first time.

1) For srv03, to learn Mac of srv06 sends ARP request to leaf-01
2) a) Forward ARP broadcast to all vlan30, local interfaces
   b) Forward as multicast to 239.0.0.30 on VNI30030
3) all Leafs of VNI30030 learn Mac of SRV03
4) leaf 9 sends traffic into VLAN 30
5) srv6 replies with its Mac to leaf09
6) leaf 9 sends unicast packet to leaf01 containing info from 5
7) leaf 1 learns MAC of srv6 and maps it to leaf 9
8) leaf 1 can now encapsulate frames to srv07 as VXLAN datagrams

EVPN (Ethernet VPN)

· Overcome flood-and-learn limitations, doesn't rely on data plane learning

· Uses MP-BGB to distribute VTEPs

Slides have little information → exercises done

↳ Read Lab and add information here!

09/10) MPLS Multi-protocol label switching

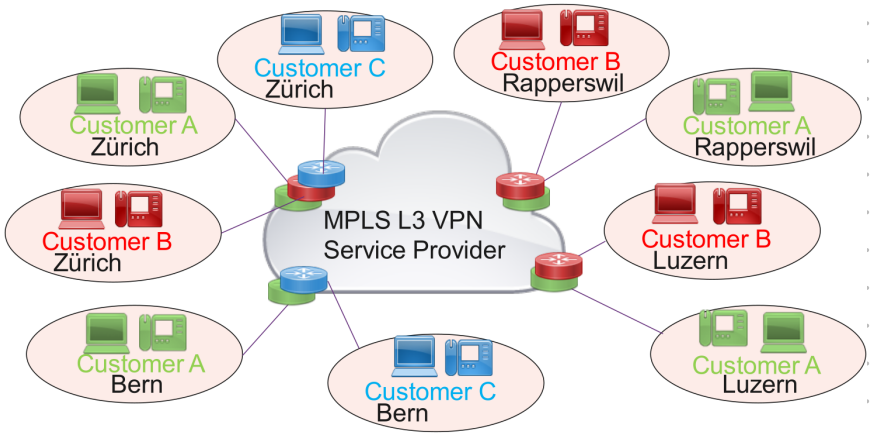
Overview of WAN technologies:

- point-to-point → leased private line
  - Dark fiber → own physical cable
  - CWDM (Coarse wavelength division multiplexing): max distance 120 km, different wavelengths only optics, separation
  - DWDM (Dense wavelength division multiplexing): 80 different channels, max distance 1000 km, wavelength difference 0.8 nm
    ↳ used in all modern submarin cables

- Circuit switching: necessary to establish a connection dynamically and before communication can happen. → eg. phone call and alike

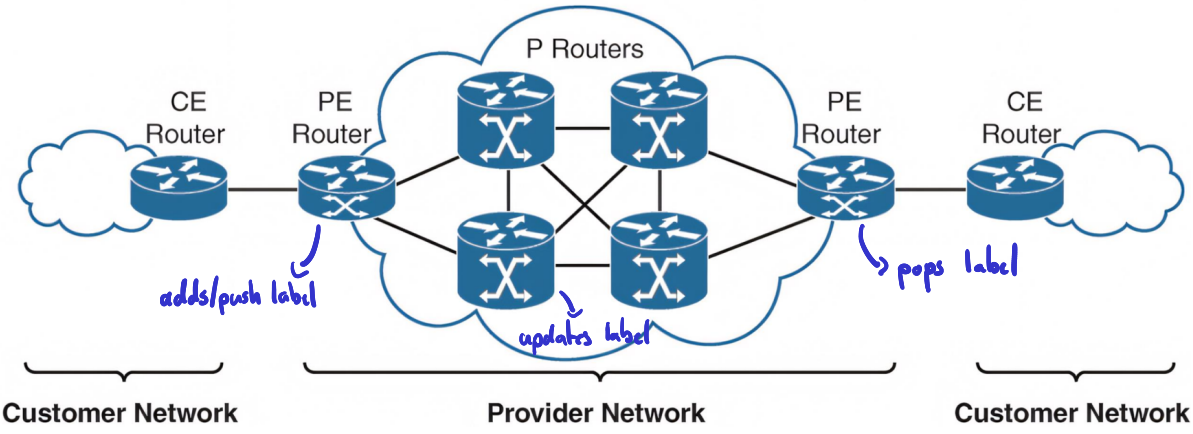
Packet switching: splits traffic data into packets that are then routed => MPLS

Modern WAN of choice



MPLS is multiprotocol → works with L2/L3. Switching happens based on labels instead of IP adresses

P: provider router, PE: provider edge router, CE: customer edge router



Label information is added between Layer 2 and Layer 3 packet information.
↳ contains Label and TTL

LDP - Label distribution protocol

LDP establishes a session by performing: periodically send MPLS hello messages
respond to received hello messages

UDP for hello message, TCP to establish session

control plane + RIB (Routing information Base) from IP protocol

LIB: Label Information Base → label to prefix bindings

FIB: Forwarding information base → used for IP forwarding, contains IP prefixes → next hops

LFIB: Label Forwarding information base → MPLS forwarding, contains inlabel → { out-label, interface, next-hop

↳ show mpls forwarding (IP)

Data plane

Label allocation:

1) IP protocol builds the IP routing table
2) Each LSR (label switched router) assigns a label to IP individually → LIB
3) LSR announces labels to all other LSR
4) LFIB and FIB built based on received information

Penultimate Hop popping: second to last router removes MPLS label as CE router can route directly

VRF - Virtual Router and Forwarding

Separate Virtual Routers on a physical device.

Customers can be separated by using virtual routers → same IPs allowed :)

Route distinguisher: unique value to identify a routing table on a router

Route Target = route distinguisher + IP → makes the route globally unique

↳ eg: 1:100:192.168.1.0/24

Based on the route target, it is decided which routes are routed to which VRF

Route reflector: used to reflect routes across AS/network with PE routers
↳ check this again!

VPNv4 · control plane: redistribute IP route with iBGP
· forwarding plane: does the forwarding using MPLS label
↳ clarify what this is

Summary: MPLS L3VPN Technology Components

- PE-CE link
  - CE configured to route IP traffic to/from adjacent PE router
    - Variety of routing options; static routes, eBGP, OSPF, IS-IS
- MPLS L3VPN Control Plane
  - Separation of customer routing via virtual VPN routing table
    - In PE router: customer interfaces connected to virtual routing table
    - Between PE routers: customer routes exchanged via BGP (using Route Reflectors)
- MPLS L3VPN Forwarding Plane
  - Separation of customer VPN traffic via additional VPN label
    - VPN label used by receiving PE to identify VPN routing table

MPLS configuration

202506-june.md 9+, M | P3.conf

CN2 > MPLS > P3.conf

```
1   router ospf 1
2
3   interface GigabitEthernet1
4     ip ospf 1 area 0
5     mpls ip
6
7   interface GigabitEthernet2
8     ip ospf 1 area 0
9     mpls ip
10
11  interface GigabitEthernet3
12    ip ospf 1 area 0
13    mpls ip
14
15  interface Lo0
16      ip ospf 1 area 0
17
18  mpls label range 3000 3099
```

- Underlay IP connectivity using OSPF (or another IGP)
- Configuration of MPLS with command MPLS IP

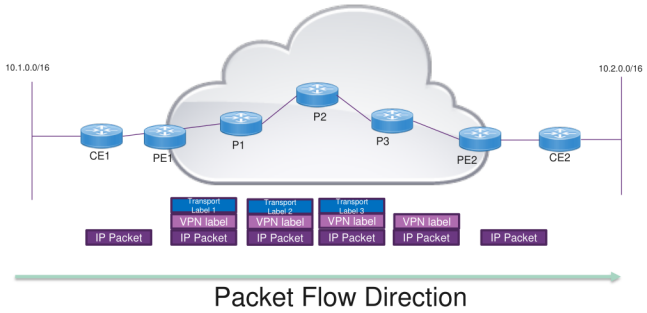
```
SW04
=====

conf t
ip routing
ip multicast-routing
router ospf 1
ip pim rp-address 10.10.10.3

interface vlan 1
no shutdown
ip address 10.10.4.1 255.255.255.0
ip ospf 1 area 0
ip pim sparse-mode
```

more notes for this?

Route Propagation

1) route from CE router advertised to other PE via OSPF/...
2) PE adds it to VRF based on Route Target
3) PE advertises route to other PEs via iBGP
4) Other PE routers do step 2 again.

MPLS L3 VPN Forwarding



- PE1 learned Transport Label 1 learned from P1 via LDP
- P1 learned Transport Label 2 learned from P2 via LDP
- P2 learned Transport Label 3 learned from P3 via LDP
- P3 learned that he has to performed Penultimate Hop Popping (PHP) from PE2 via LDP
- PE1 received the VPN label from PE2 via iBGP

Layer 2 MPLS

Works very similar to Layer 3 MPLS -> multiprotocol :)

- looks like a big switch from the customer POV

Continue here with MPLS exercises and Lab.

unicast Reverse Path Forwarding uRPF
Filters out packets with spoofed src adresses

12) QoS Quality of Service

Networks are built to meet the needs of end users. This is the goal

QoS based routing:
- multiple paths not just shortest one
- path pinning → don't instantly switch path when a better one is found
  ↳ otherwise two routes might be switched constantly

Typical bottlenecks: speed mismatches or buffers

Buffers: incoming → when device can't handle all packets at once
outgoing → waiting for transmission, as not to overload devices e.g. router → rasberry pi

! Network Metrics !

Latency / Delay (ms): delay in data transfers (ping time)

Jitter (ms): variation of ping time, difference of highest and lowest delay

Packet loss (%): data packets not arriving

Bandwidth (Gbit/s): data transfer rate
↳ [Throughput (Gbit/s): data that succesfully dist]

Queuing Mechanism

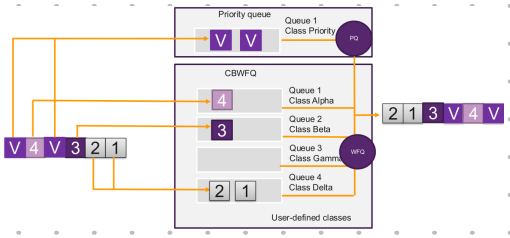
FIFO: first in first out, no decision thus the fastest

Priority Queuing: different queus that are prioritized, prio 1 cleared first → queus below might starve

(weighted) round robin: Multiple queus that are dispatched 1 by 1. Ratio might change if weighted

CBWFQ: class-based weighted fair queuing → extends weighted round robin by supporting user-defined classes.

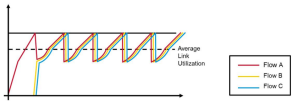
LLQ: low latency queuing → used most often today
strict prio queuing for low delay traffic such as voice calls
↳ dispatched instantly



Tail dropping: traffic that is dropped 'cause queue is full
it is not known what traffic is dropped

⇒ RED: Randomly Early detection → drops packeages so the queue never gets full → avoid tail dropping
↳ Exists as weighted RED too → lower prio packets dropped first

TCP global synchronization
1. Multiple TCP sessions start at different times
2. TCP window sizes are increased
3. Tail drops cause many packets of many sessions to be dropped at the same time
4. TCP sessions restart at the same time (synchronized)

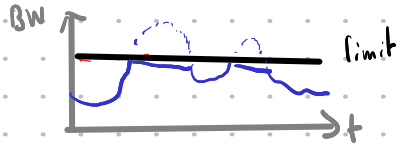


TCP throttles sending to avoid network congestion

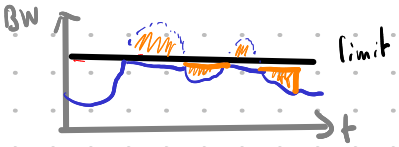
looked at in Asys too

TCP starvation: UDP has no such mechanisms and simply sends. Thus UDP packages can fill up queues and prevent TCP packages from being sent → solution use RED random early detection to drop (mainly) UDP packeges

Policing: hard limit the bandwith



Shaping: delay traffic to remain within BW (bandwidth) limits

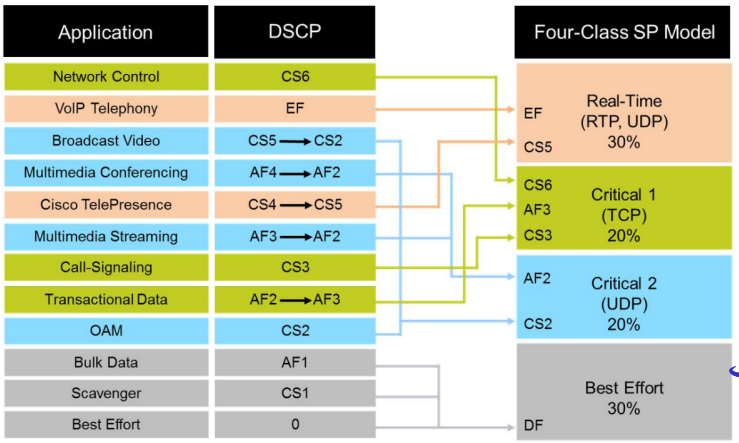


Qos models

Best effort: internet is best effort → net neutrality

Integrated service: Nut zur Lücke

Differential Services: de facto standard, simple and scalable mechanism for classifying and managing network traffic to provide QoS



Marking of QoS at Layer 2 in dot1q header or layer 3 in Tos field

1st CE router does the marking

→ different models exist.