

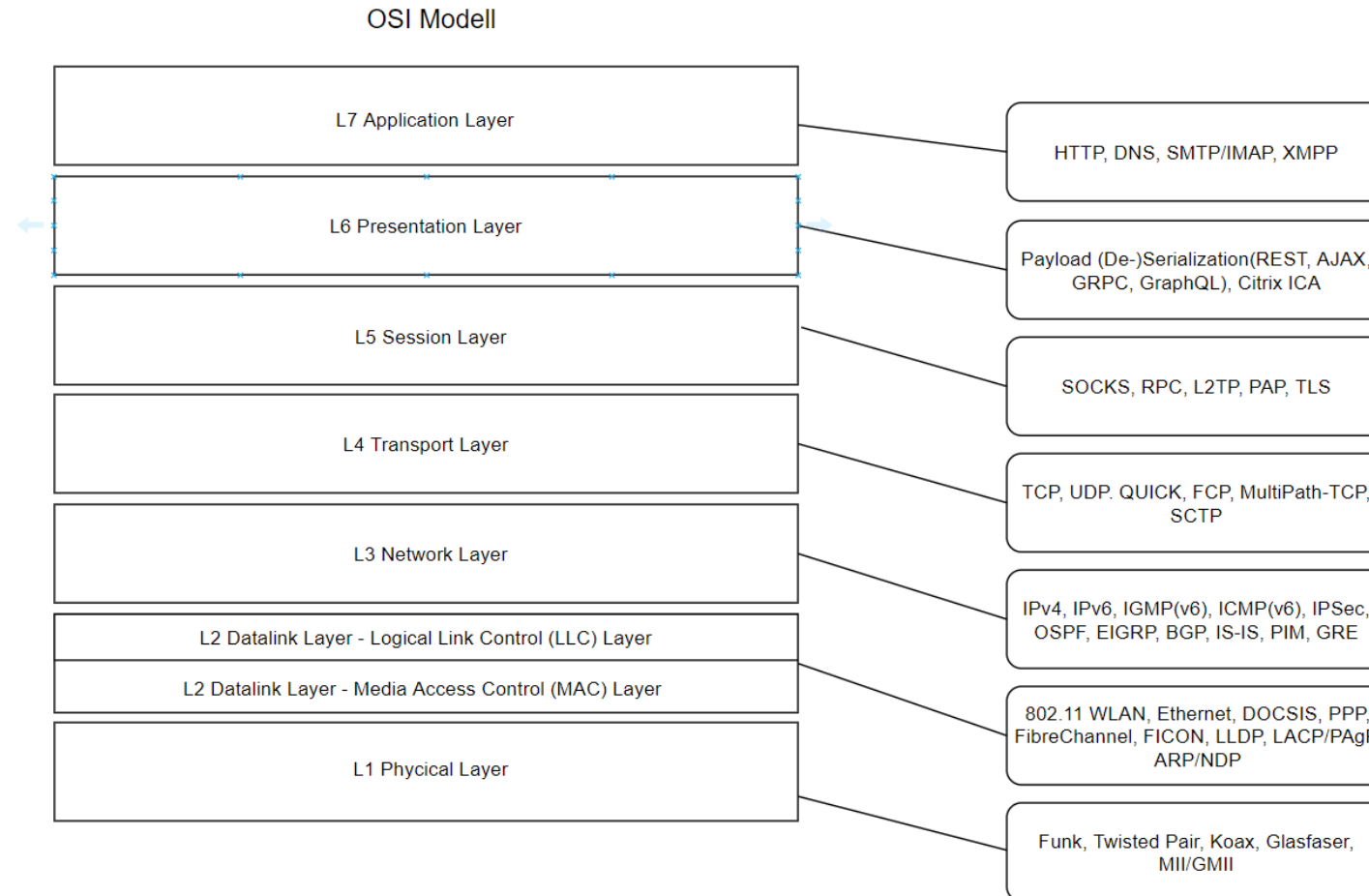
IP Routing

**Frankfurter Linux User
Group (FraLUG) e.V.**



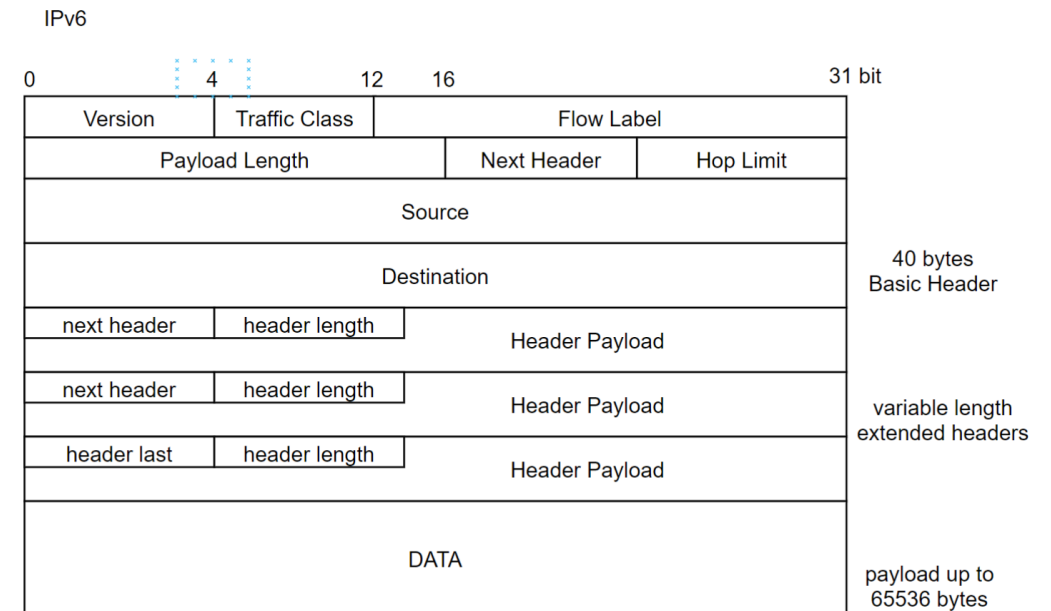
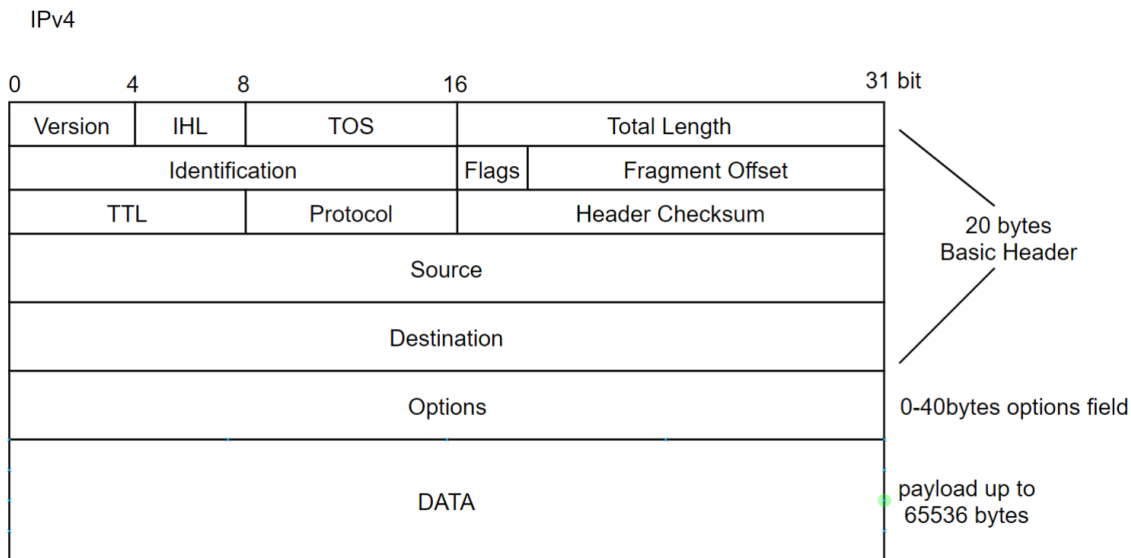
BASICS

Schichten und Kapselung



Internet Layer IPv4/IPv6

- PDUs auf Layer 3 werden Pakete genannt.
- Sie enthalten Header und Payload
- unidirektional



Layer 4 - Segments & Datagrams

TCP Segment

Source Port		Destination Port					
Sequence Number							
Acknowledgement Number							
offset	U	A	P	R	S	F	Window Size
Checksum		Urgent Pointer					
Options							
Payload							

UDP

Source Port	Destination Port
length	checksum
Data	

ROUTING

Routing

- Findet auf Layer 3 statt
- kann statisch oder dynamisch (via Routingprotokoll) stattfinden.
- verwendet immer den Longest Match
- verwendet je nach Protocol verschiedene Attribute zur Auswahl bei multiplen Routen

address space management

- Oberste Instanz ICANN (Internet Corporation for Assigned Names and Numbers)
- Vergibt Netzbereiche an RIRs
- verschiedene RIRs(RIR - Region Internet Registry) verwalten lokale Teile des Adressraums selbstständig (RIPE, ARIN, APNIC, AFRINIC)
- RIRs weisen Address-Ranges einzelnen Entitäten in ihrem Zuständigkeitsbereich zu.

IPv4 - Special Addresses

- Loopback Adressen 127.0.0.0/8
- RFC 1918 IPv4 Private Address Ranges:
 - 10.0.0.0/8 (10.0.0.0 - 10.255.255.255)
 - 172.16.0.0/12 (172.16.0.0 - 172.31.255.255)
 - 192.168.0.0/16 (192.168.0.0 - 192.168.255.255)
- CGN - carrier grade NAT: 100.64.0.0/10 (100.127.0.0 - 100.127.255.255)
- IPv4 Link Local 169.254.0.0/16
- IETF reserved 192.0.0.0/24; 192.0.2.0/24; 198.51.100.0/24, 203.0.113.0/24; 233.252.0.0/24; 240.0.0.0/4
- IPv6/IPv4 Relay 192.88.99.0/24&2002::/16
- Communication Testrange 198.18.0.0/15
- IP Multicast 224.0.0.0/4
- Broadcast 255.255.255.255/32

IPv6 - Special Addresses

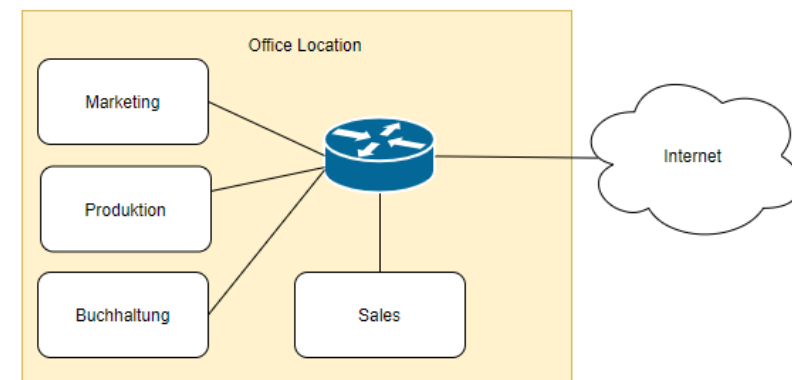
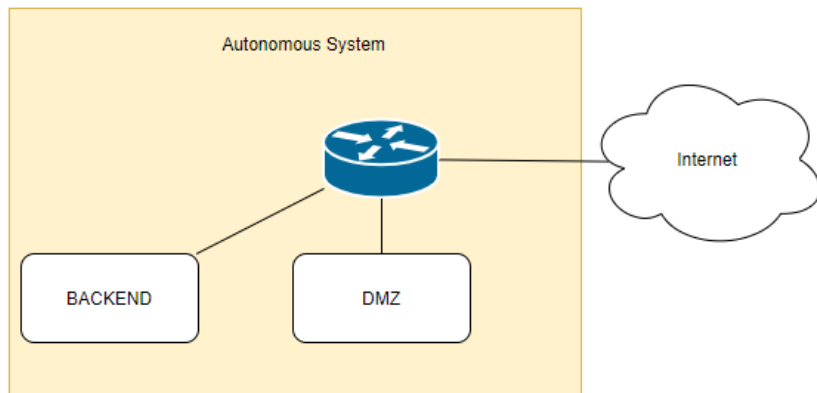
- Loopback Adresse `::1/128`
- IPv6 Link Local Unicast `FE80::/10`
- `::ffff:0:0/96` IPv4-mapped Address
- `64:ff9b::/96` IPv4-IPv6 Translation
- `64:ff9b:1::/48` IPv4-IPv6 Translation
- `100::/64` Discard-Only Address Block
- `2001::/23` IETF Protocol Assignments
- `2001::/32` TEREDO Tunneling
- `2001:1::1/128` Port Control Protocol Anycast
- `2001:1::2/128` NAT anycast traversal
- `2001:2::/48` Benchmarking
- `2001:3::/32` Automatic Multicast Tunneling

IPv6 - Special Addresses

- 2001:4:112::/48 Anycast DNS Reverse Lookup
- 2001:20::/28 ORCHIDv2 (Overlay Routable Cryptographic Hash Identifiers)
- 2001:db8::/32 Documentation
- 2002::/16 6to4 Tunneling
- 2620:4f:8000::/48 Anycast DNS Reverse Lookup
- fc00::/7 Unique-Local

Subnetting

- Subnet ist Splitting des IPv4/v6-Adressraumes in Segmente und erlaubt die weitestgehend unabhängiger Verwaltung verschiedener Segmente durch unterschiedliche Entitäten oder die Zuordnung einzelner Segmente für bestimmte Zwecke. (Zoning)
- Verwendet VLSM (variable length subnet masks)



Subnetting

- Jede IP Adress Konfiguration besteht aus Adresse und Netzmaske
Angegeben wird das im Format Adresse/<Netmask|CIDR>
- Bsp: 192.168.1.1/24 ist eine IPv4 Adresse und hat 32bit.
Der Netprefix ist 24bit lang. Alle Adressen im Hostanteil(blau) sind lokal erreichbar und brauchen keinen Router.
- Adressen die eine Änderung im grünen Bereich(Netz) erfordern müssen geroutet werden.

11111111.11111111.11111111.00000000

11000000.10101000.00000001.00000001

PBR und VRFs

- PBR(Policy Based Routing) verwendet Route-Maps oder Diskriminatoren für die Routing-Entscheidung.
 - Routetags
 - Source Address
 - Source und/oder Destinationports
- VRFs (Virtual Route Forwarder) erlauben das betreiben verschiedener von einander unabhängiger Routing Tables auf einem System.

Static Routing/Policy based Routing

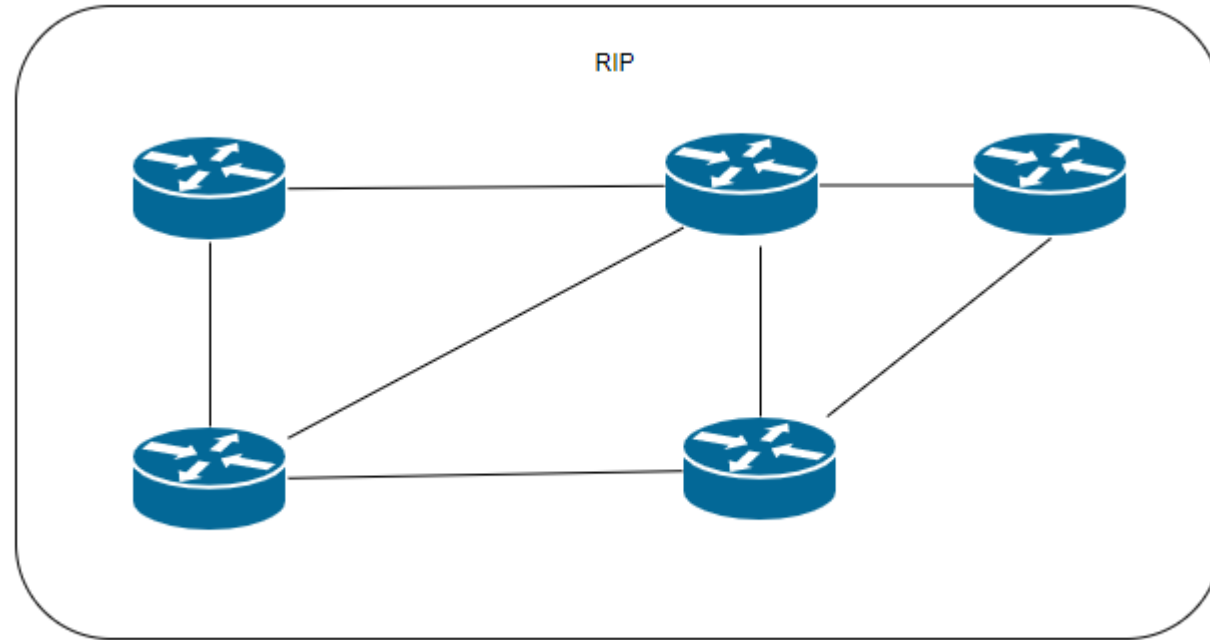
- Für jedes Ziel(Subnet oder Host-Adresse) wird eine fixe Route je Gerät auf der Strecke installiert.
- Routingentscheidungen werden anhand der Routen und manuell konfigurieren Diskriminatoren getroffen.
- Nicht handhabbar für komplexe Installationen durch extremen Verwaltungsaufwand
- anfällig für Human Error/Route-Poisoning, Split-Horizon Szenarien, und unreachable Route Issues

Routing Information Protocol

RIP - Routing Information Protocol

- Interior Gateway Protocol
- Distance Vector Protocol
- heutige Inkarnationen RIPv2(IPv4) und RIPv6 (IPv6)
- max 15 HOPs
- verwendet Hopcount als Metrics
- überträgt immer die Komplette Routingtabelle
- Übertragung erfolgt via Multicast (RIPv1 nutzte Broadcast)
RIPv2 224.0.0.9 UDP/520
RIPv6 ff02::9 UDP/521
- Verwendung KMU/SoHo

RIP - Routing Information Protocol

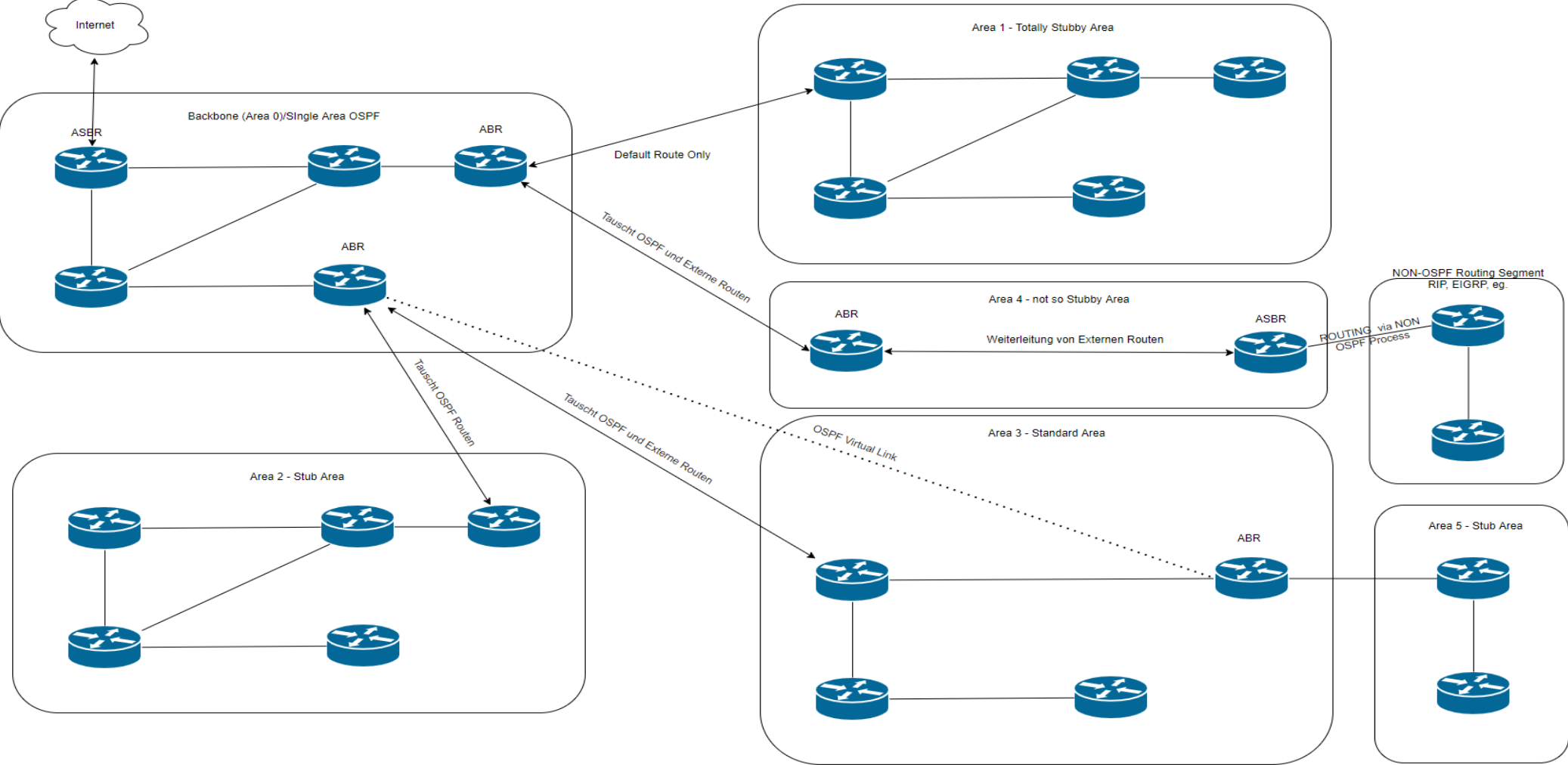


Open Path Shortest First

OSPF - Open Shortest Path First

- Interior Gateway Protocol
- Linkstate Protocol, benutzt Link State Advertisements(LSAs)
- unterteilt Routing Segmente in Areas (Backbone, Standard, Stub, Totally Stub, not so stubby), two-level hierarchy
- Differenziert zwischen Router-Typen (ABR, ASBR, IR, BR)
- Formt Nachbarschaft durch Hello advertisements auf Links
- Router ist die Boundary
- Benutzt DR und BDR als zentrale Trackinginstanz der LinkState Table
- Dijkstra Algorithmus
- Verwendung in komplexen Netzen (Datacenter, große Firmennetze)

OSPF - Open Shortest Path First



Intermediate System to Intermediate System

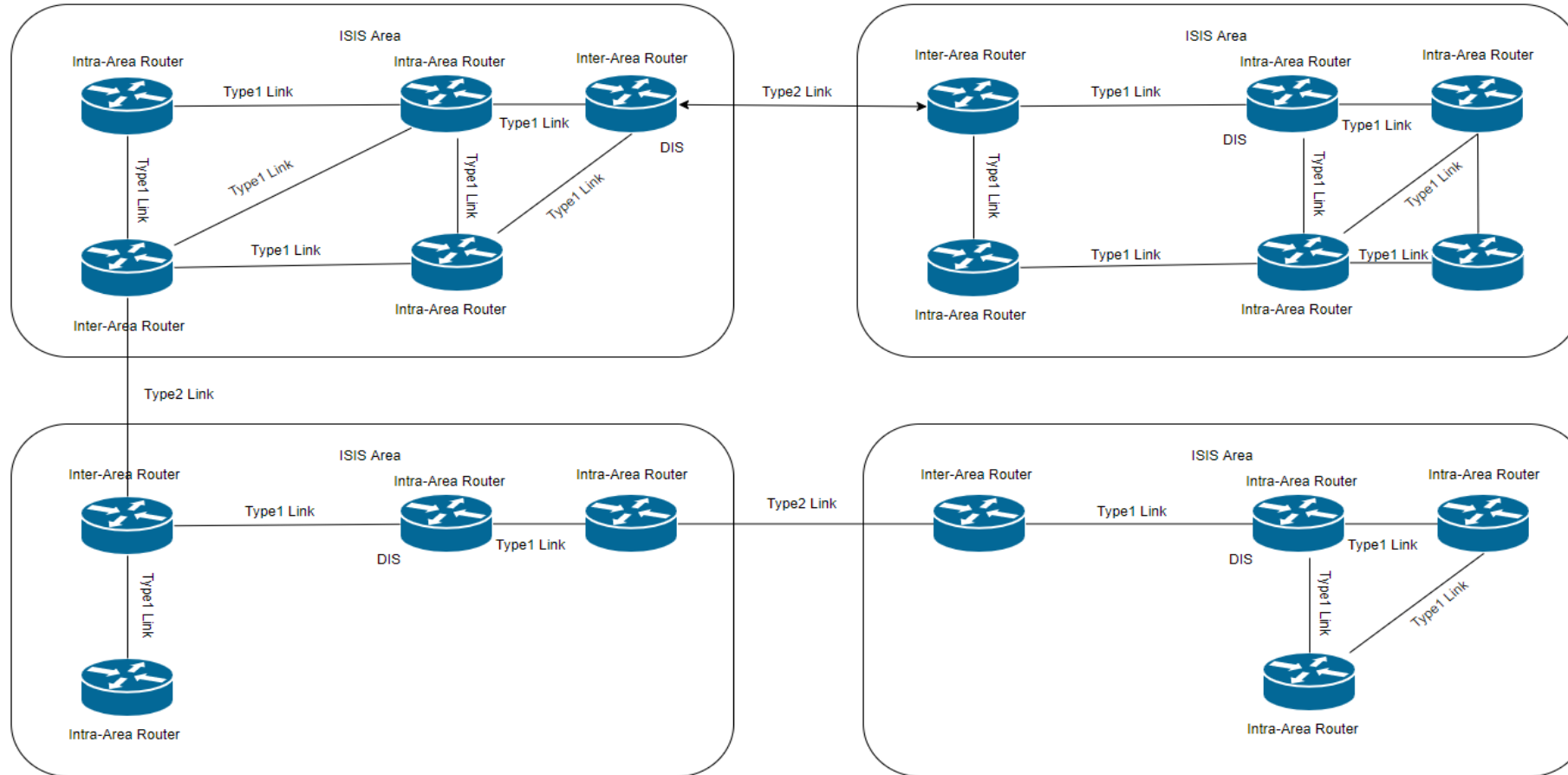
IS-IS

Intermediate System to Intermediate System

- Interior Gateway Protocol
- Link State Protocol
- Benutzt intern IP-unabhängige Adressen (OSI CLNS Addresses)
- Unterteilt in Areas (Intra-Area und Inter-Area)
- Link ist die Boundary
- Dijkstra Algorithmus
- Verwendung Carrier-Backbones, Network Fabrics (SPB/FSPF)
- Verwendung von einem Designated Router (DIS) zum Tracking des Status der Link State DB
- Request statt Linkstate Flooding
- Erweiterbar durch TLVs (Type Length Values)

IS-IS

Intermediate System to Intermediate System



Border Gateway Protocol

- "Autonomous System - A set of Routers under a single technical administration, using an Interior Gateway Protocol and common metrics to determine how to route packets within the autonomous system and using an inter-autonomous system routing protocol to determine how to route packets to other autonomous systems." RFC4271

BGP - Border Gateway Protocol

- Exterior Gateway Protocol zur Verbindung von Autonomen Systemen
- Path Vector Protocol
- AS-Nummern werden durch die RIRs verwaltet.
- Wird benutzt um Autonome Systeme zu verbinden
- Benutzt TCP/179
- iBGP Fully Meshed per default
- Benutzt Confederations und Route Reflectors
- ASN 32bit(bis 2011 16bit)
- Multihoming tauglich
- Entscheidet anhand von Path Vectors (Anzahl von AS auf dem Pfad, local preference, MED, Policymaps etc. pp.)
- Verwendung ISPs und Große Organisationen

BGP - Border Gateway Protocol

- Verwaltet Pfade in der BGP Topology Table
- Verwendet eine Neighbor Table zur Verwaltung von Nachbarschaften
- Benutzt Attribute in Update und Notification Messages.
- nur aktive Routen werden zwischen den AS ausgetauscht.
- BGP Router akzeptieren keine Routen von anderen AS mit ihrer eigenen AS Nummer im Pfad (loopfree AS-PATH)
- Benutzt Local Preference und MED Attribute zur Bestimmung des Übergabepunkts bei Mehrfachverbindungen zwischen 2 Autonomen Systemen
- Erlaubt die Nutzung von Tags zum Filtern

BGP - Border Gateway Protocol

mandatory attributes:

- AS-PATH
- Next-Hop
- Origin

discretionary attributes:

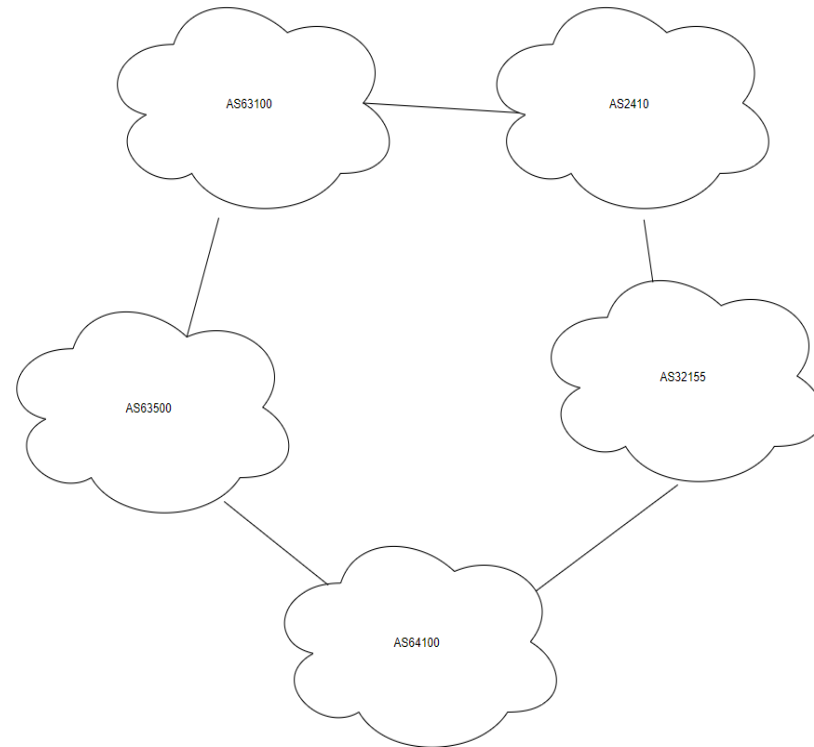
- local preference
- atomic aggregate

transitive attributes:

- community(tags)
- aggregator

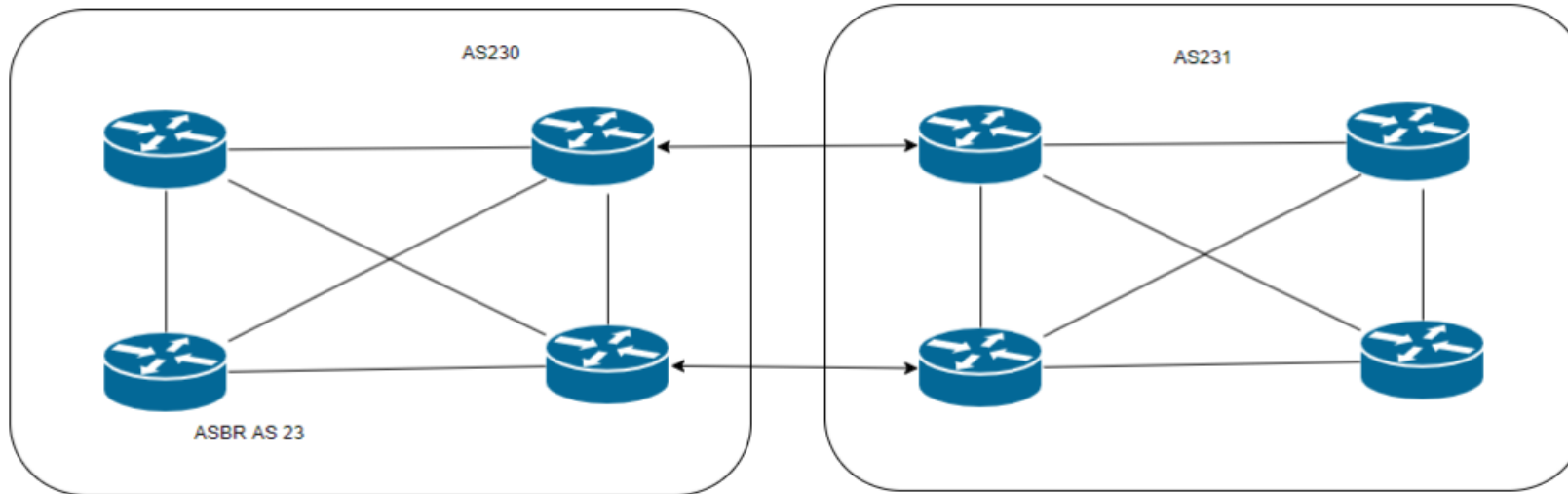
optional nontransitive:

- Multi-Exit-Discriminator(MED)



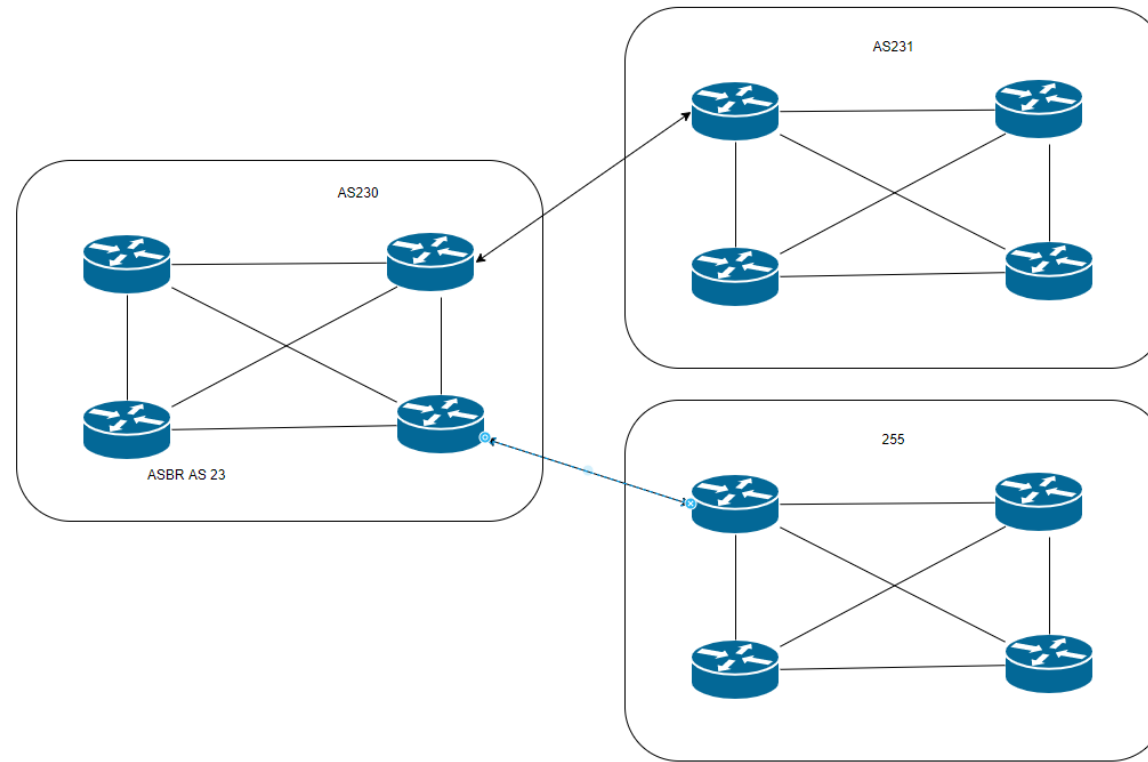
BGP - Border Gateway Protocol

- Multihome BGP für multiple Externe Anbindungen - multihomed single provider



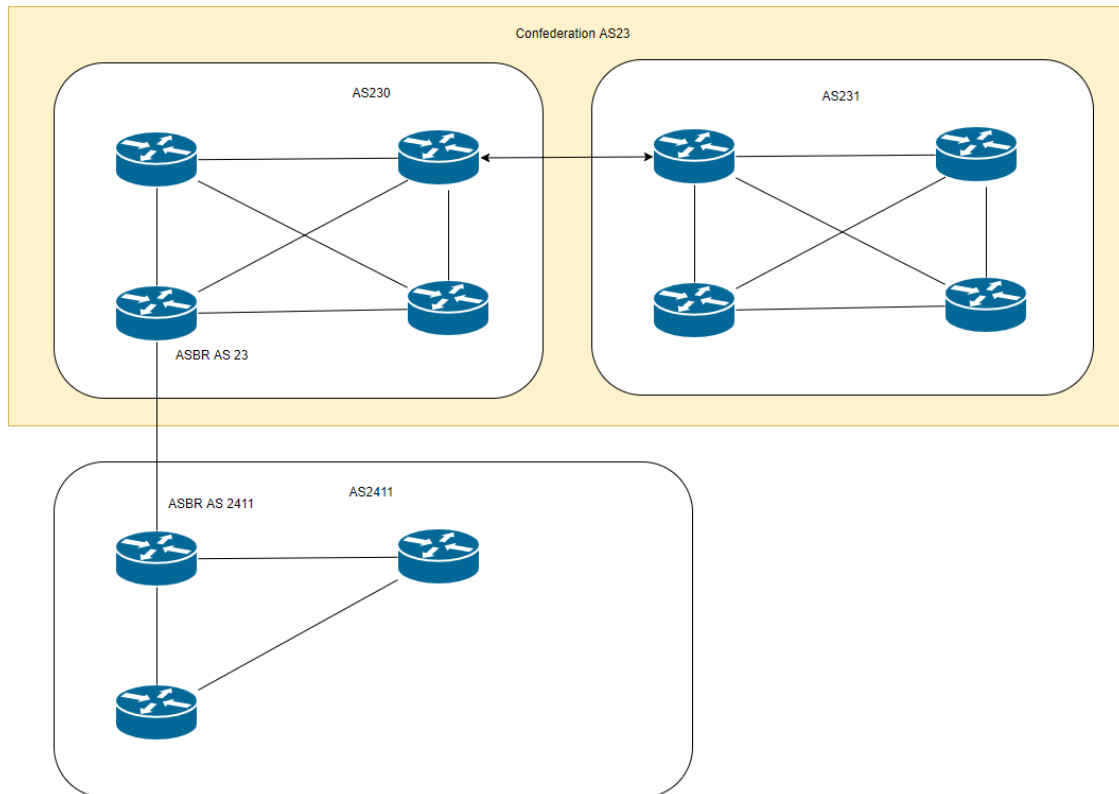
BGP - Border Gateway Protocol

- Multihome BGP für multiple Externe Anbindungen - multihomed multiple provider



BGP - Border Gateway Protocol

- BGP Confederations erlauben es verschiedene AS nach aussen hin als ein AS darzustellen.



DANKE