

BGP für IPv6

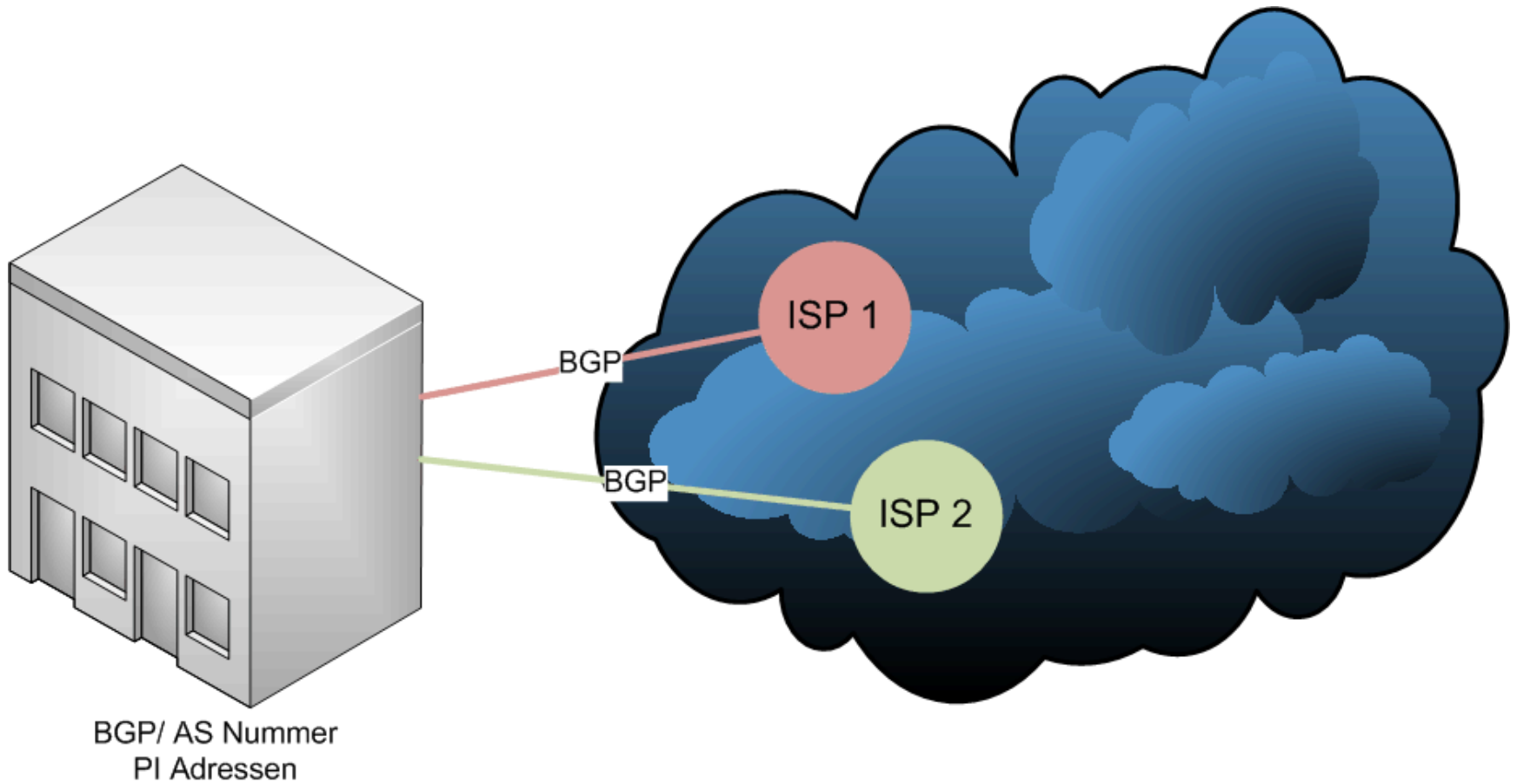
Wilhelm Boeddinghaus
Heise IPv6 Kongress 2014

Wer spricht?

- Dipl. Inf (FH) Wilhelm Boeddinghaus
- iubari GmbH
- 20 Jahre Netzwerk Erfahrung
- 11 Jahre Strato Netzwerkdesign
- Deutscher IPv6 Rat
- IPv6 Forum Gold Certified Engineer

- xing
- linked-in

Warum BGP?



Voraussetzungen

- Eigene AS Nummer
 - Eindeutige Nummer für Ihr Netz
- RIPE Mitgliedschaft = Local Internet Registry (LIR)
- Sponsoring LIR = ein anderer Provider

Voraussetzungen

- Eigene IP Nummern
 - IPv4 und/oder IPv6
- RIPE Mitgliedschaft = Local Internet Registry (LIR)
 - PA = Provider Aggregated
- Sponsoring LIR = ein anderer Provider
 - PI = Provider Independent

BGP

- TCP Applikation
- Port 179
- Beide Partner können Server oder Client sein
 - Firewall: SYN für Port 179 beidseitig offen
- Distance Vector Protocol

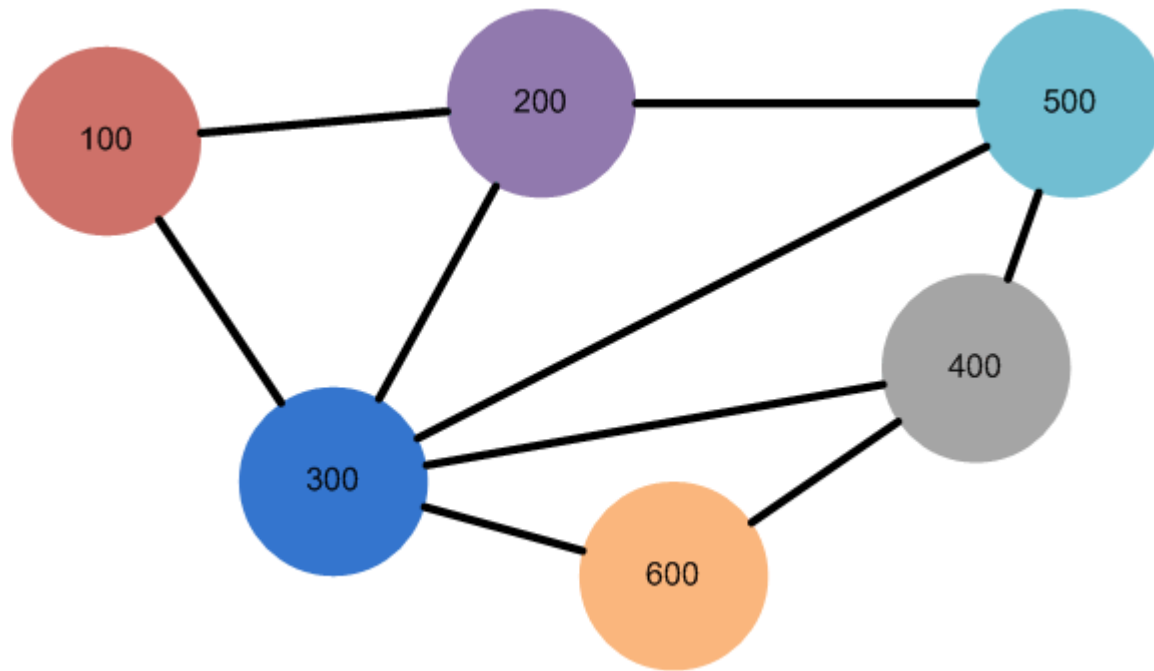
BGP

- Aufteilung des Internet in Autonome Systeme
- z.B. DTAG = AS 3320, Strato = AS 6724
- Jedes AS announced sein Netz an Nachbarn
- Erhält Informationen über Nachbarn und die Welt

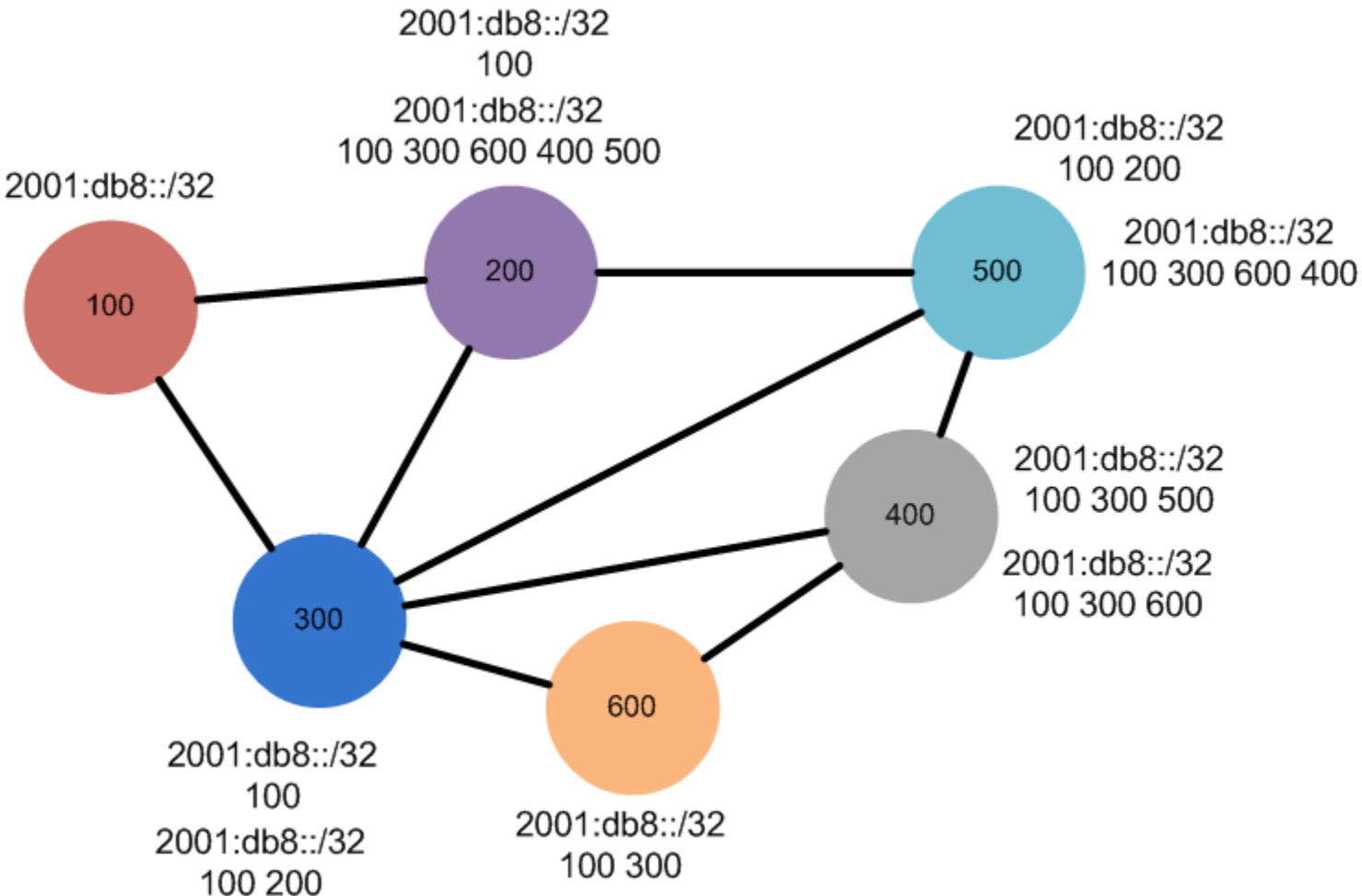
eBGP vs. iBGP

- External BGP zwischen AS
- Internal BGP innerhalb eines AS
- Ähnlich, aber doch anders

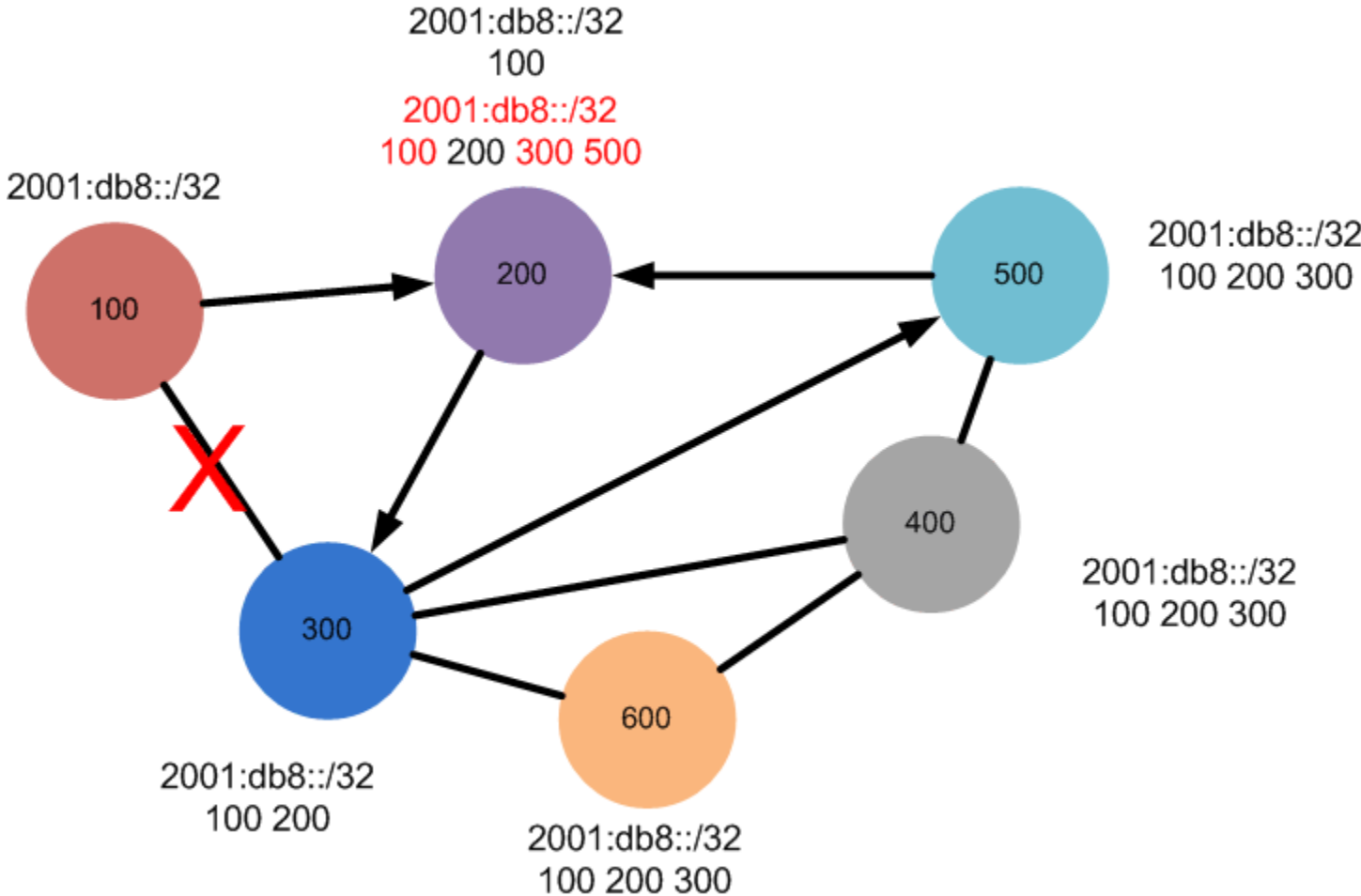
eBGP



eBGP



eBGP



eBGP

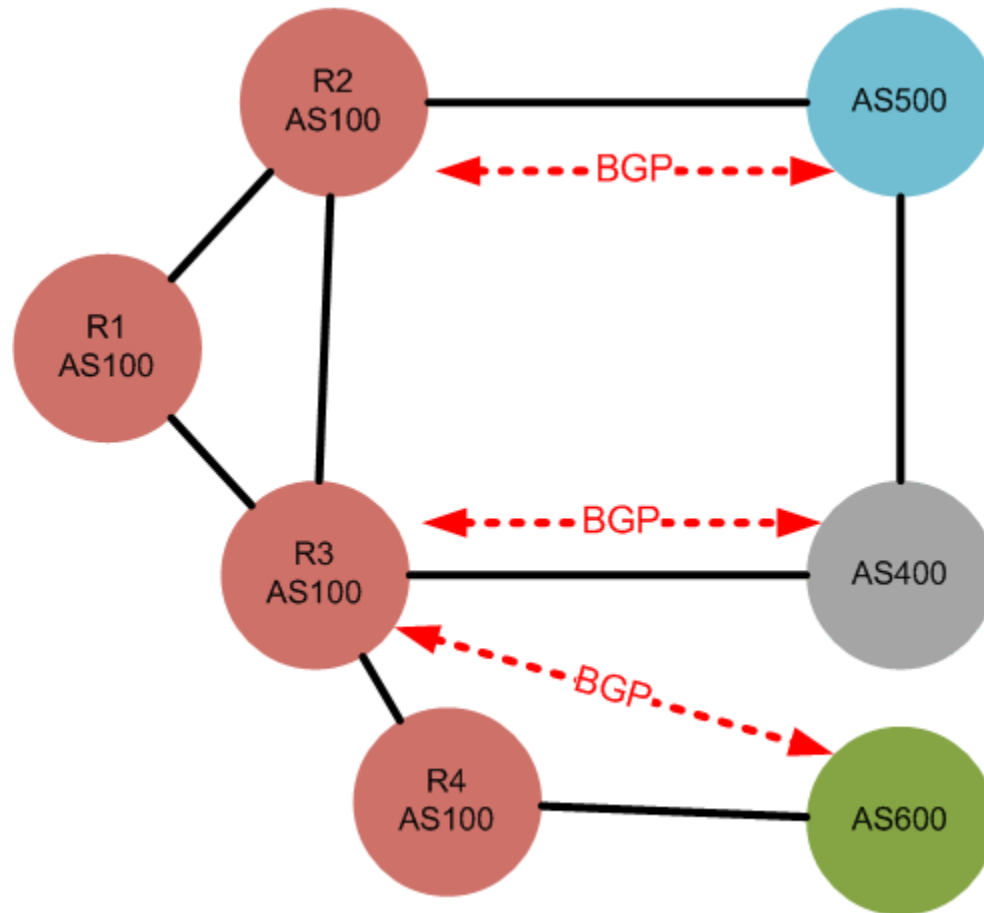
- Eigene AS Nummer in AS-Path
 - Announcement verwerfen
 - Ausnahme: accept-AS Option ist konfiguriert
 - Nur für Migration
 - Vorsicht !
- Loop Prevention bei externem BGP

eBGP

- eBGP immer nur zwischen direkten Partnern
- Keine Loops
- TTL oder Hop Limit = 1
- Kein Weiterleiten

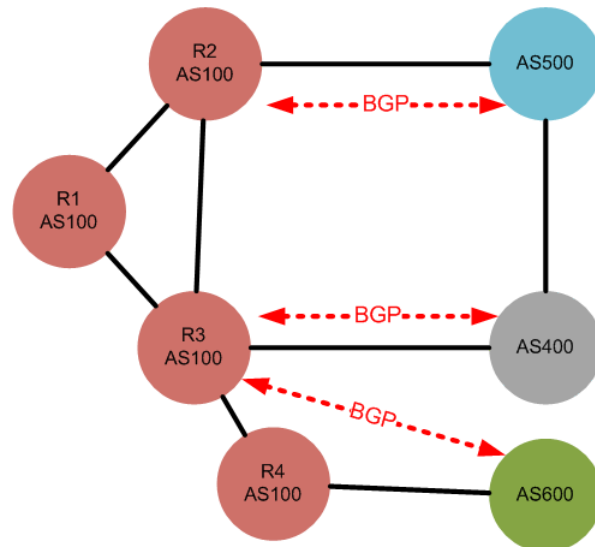
- Ausnahme: ebgp Multihop
 - TTL wird größer
 - Selten gebraucht

eBGP Multihop



eBGP Multihop

- Hot Potato Routing
 - Wenn ein externer Pfad da ist
 - Paket raus aus eigenem AS
 - Umweg über zusätzlichen Router vermeiden



eBGP

- TTL nur auf den nötigen Wert legen
- Gut dokumentieren
- Vermeiden, wenn möglich
- Dekonfigurieren, wenn möglich

- Migration, keine Dauerlösung

BGP

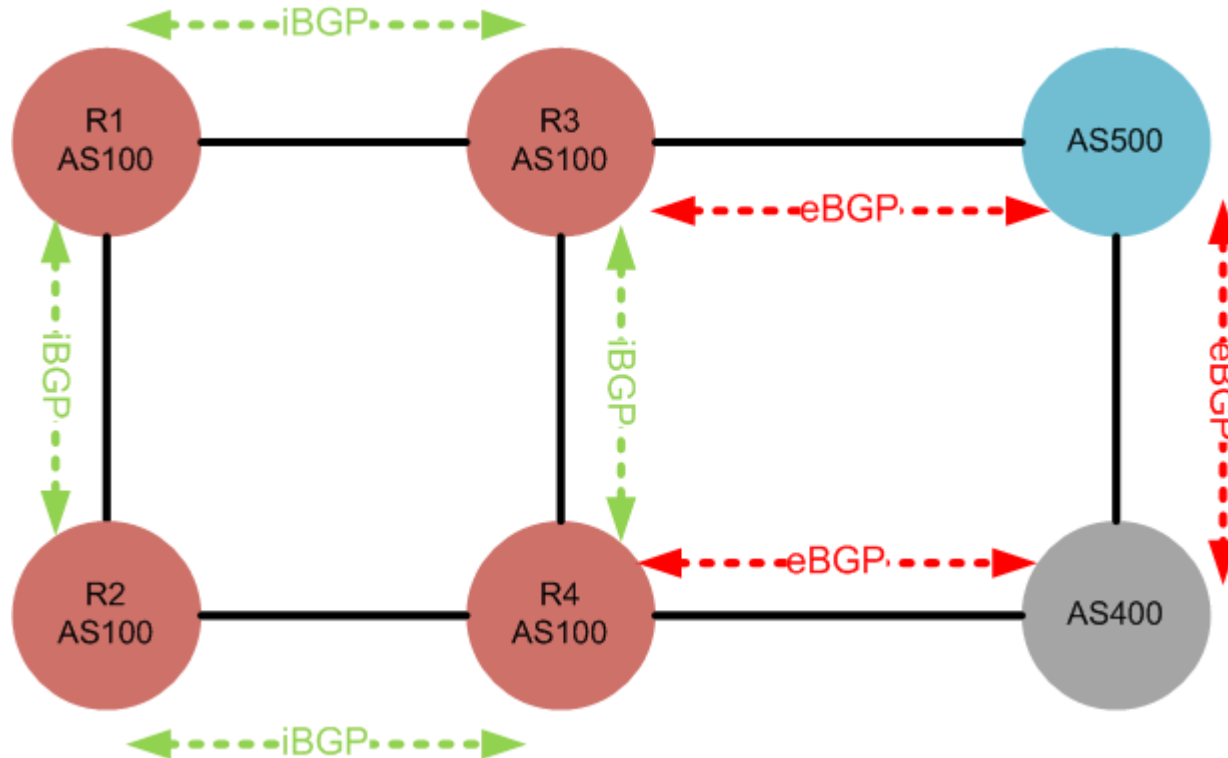
- TCP Applikation
- Port 179
- Beide Partner können Server oder Client sein
 - Firewall: SYN für Port 179 beidseitig offen
- Distance Vector Protocol

iBGP

- 2 Router im gleichen AS
 - Internes BGP
 - Administrative Distance 200

- 2 Router in verschiedenen AS
 - Externes BGP
 - Administrative Distance 20

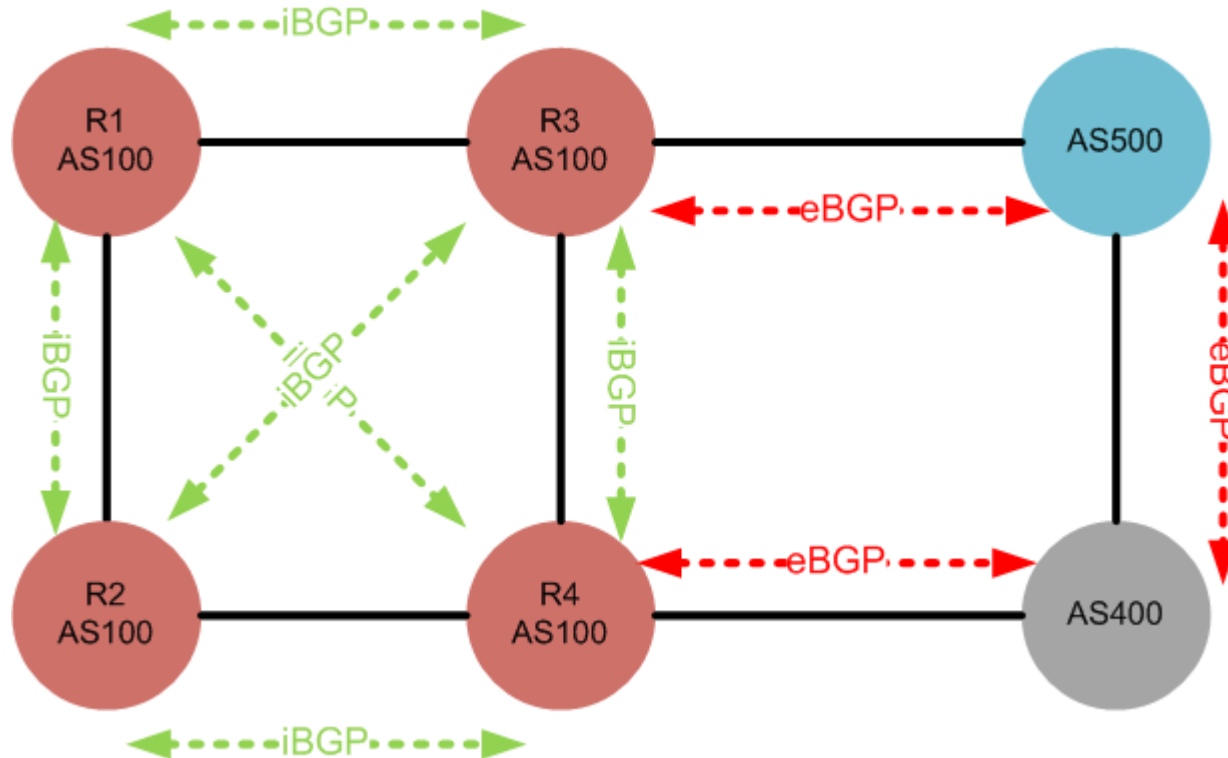
iBGP



iBGP

- Jeder iBGP Router muß mit jedem iBGP Router eine BGP Session haben
- Full Mesh
- Viele Verbindungen von Loopback zu Loopback
- Folgen nicht unbedingt den Kabeln

iBGP



iBGP

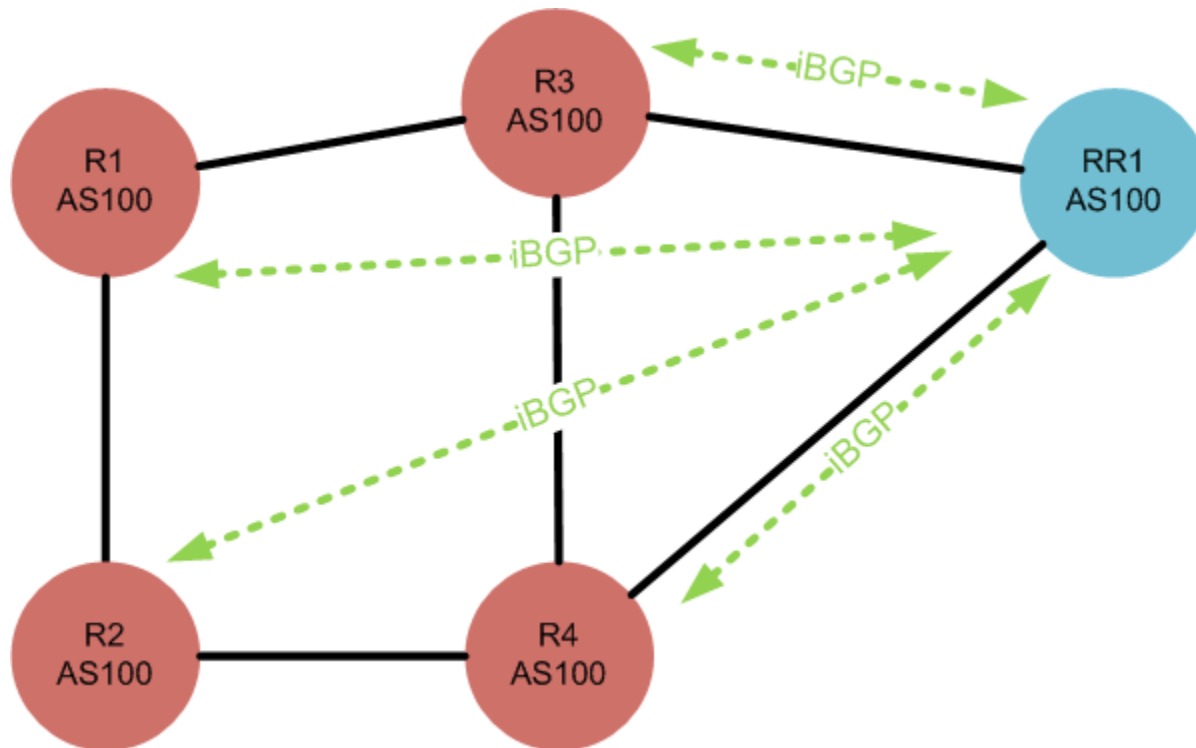
- Router gibt keine Prefixe an Nachbarn weiter, die er von iBGP Nachbarn gelernt hat
- Gibt eBGP Prefixe weiter
- iBGP ändert den Next Hop nicht
 - Nexthop self überschreibt das
- eBGP ändert den Nexthop

iBGP Route Reflector

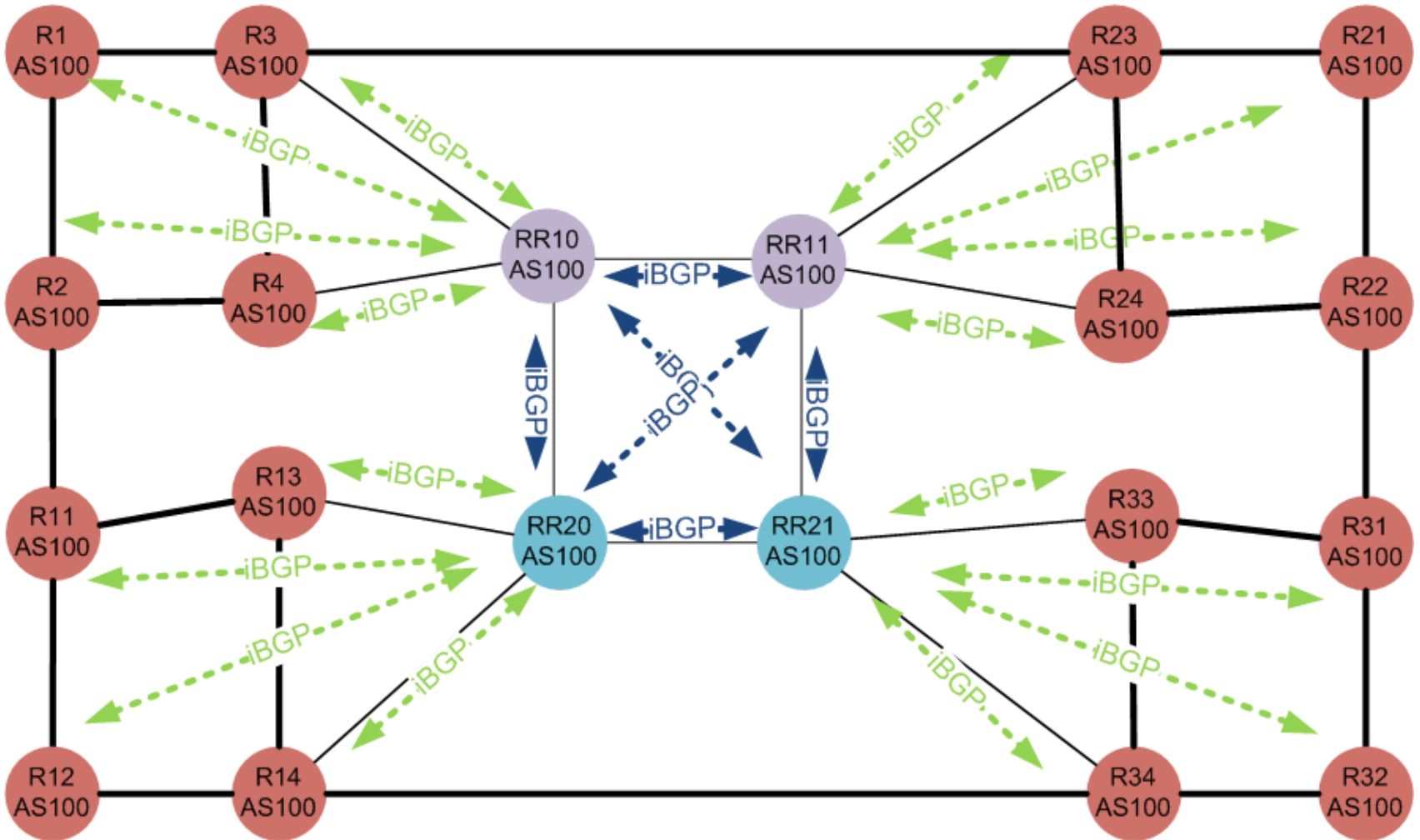
- Gibt interne Routen weiter
- Kein Full Mesh mehr nötig
- Der RR Client merkt nichts davon
- Alle Konfiguration auf Route Reflektor

- Cluster möglich
 - Redundanz

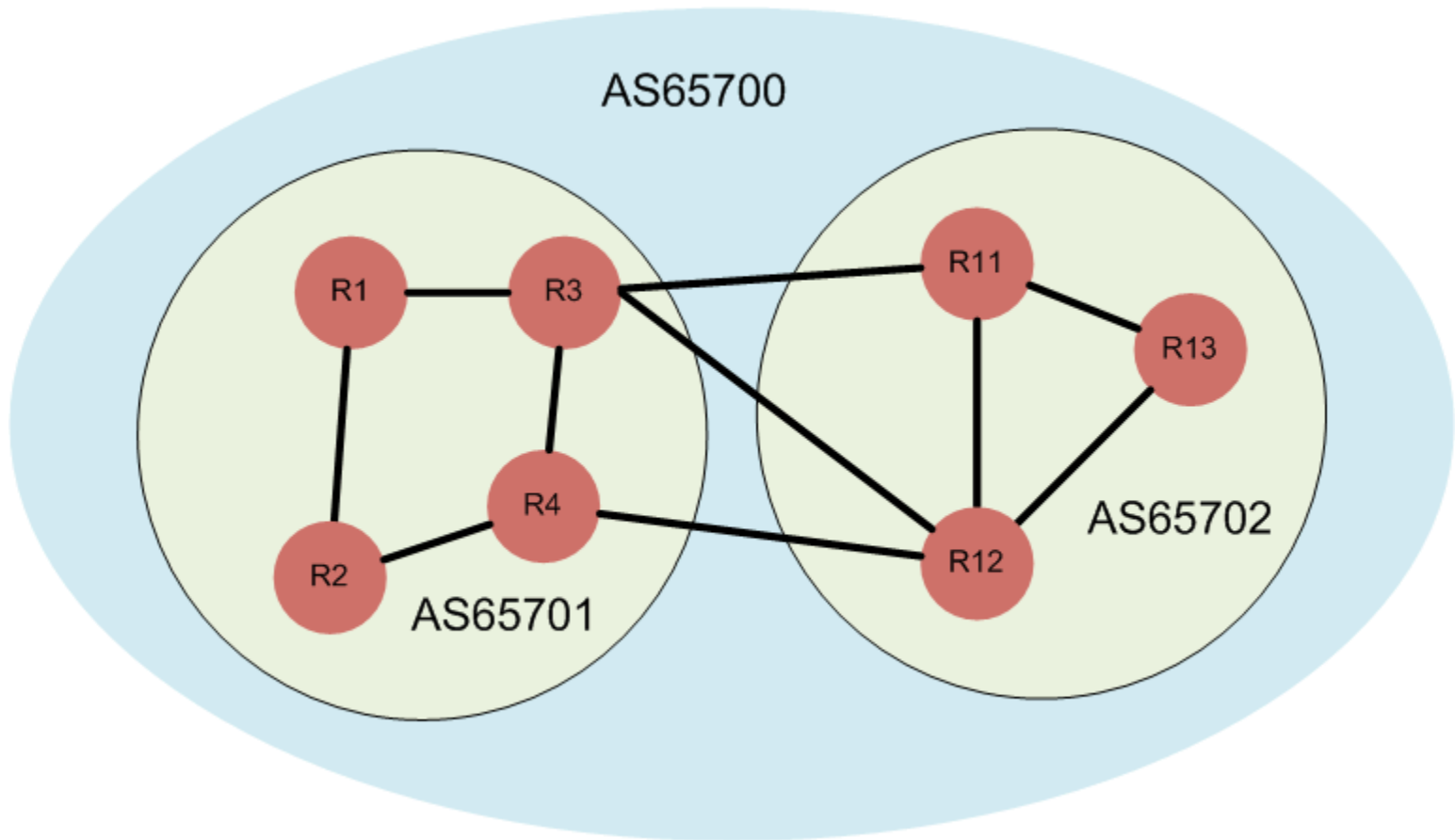
iBGP Route Reflector



iBGP Route Reflector



iBGP Confederation



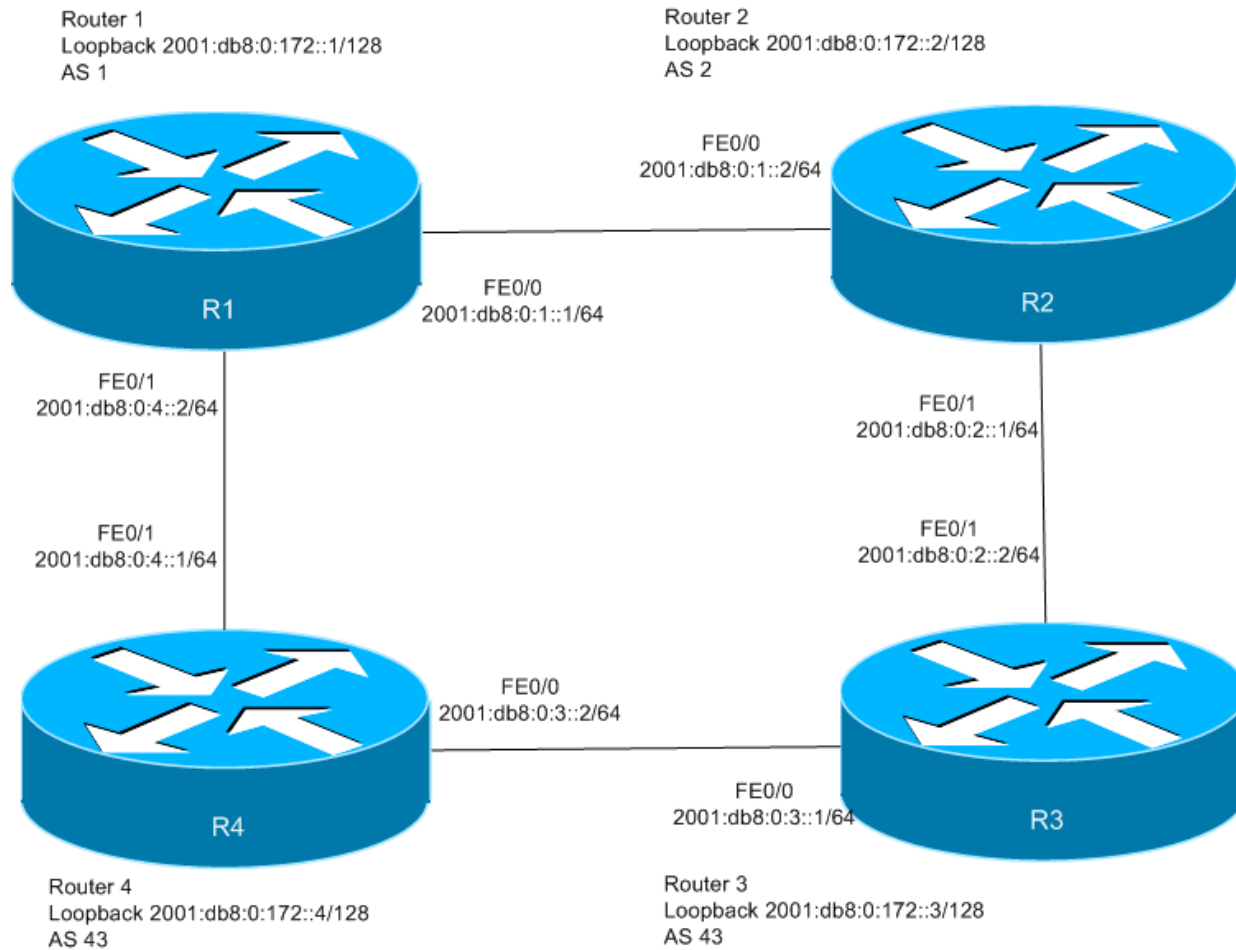
BGP Konfiguration

- Beispiel Konfiguration Cisco IOS
- Gebräuchlich
- Viel genutzt
- Cisco ist stark im Enterprise Markt

BGP Testlab

- GNS3 Cisco Simulator
 - Ältere Cisco Systeme werden simuliert
 - Image ist original
 - Großer Server notwendig
 - Cloud Services nutzen, z.B. AWS
- Cisco Cloud Services Router
 - Frei runterladen
 - Premium Lizenz 60 Tage gültig

BGP Lab

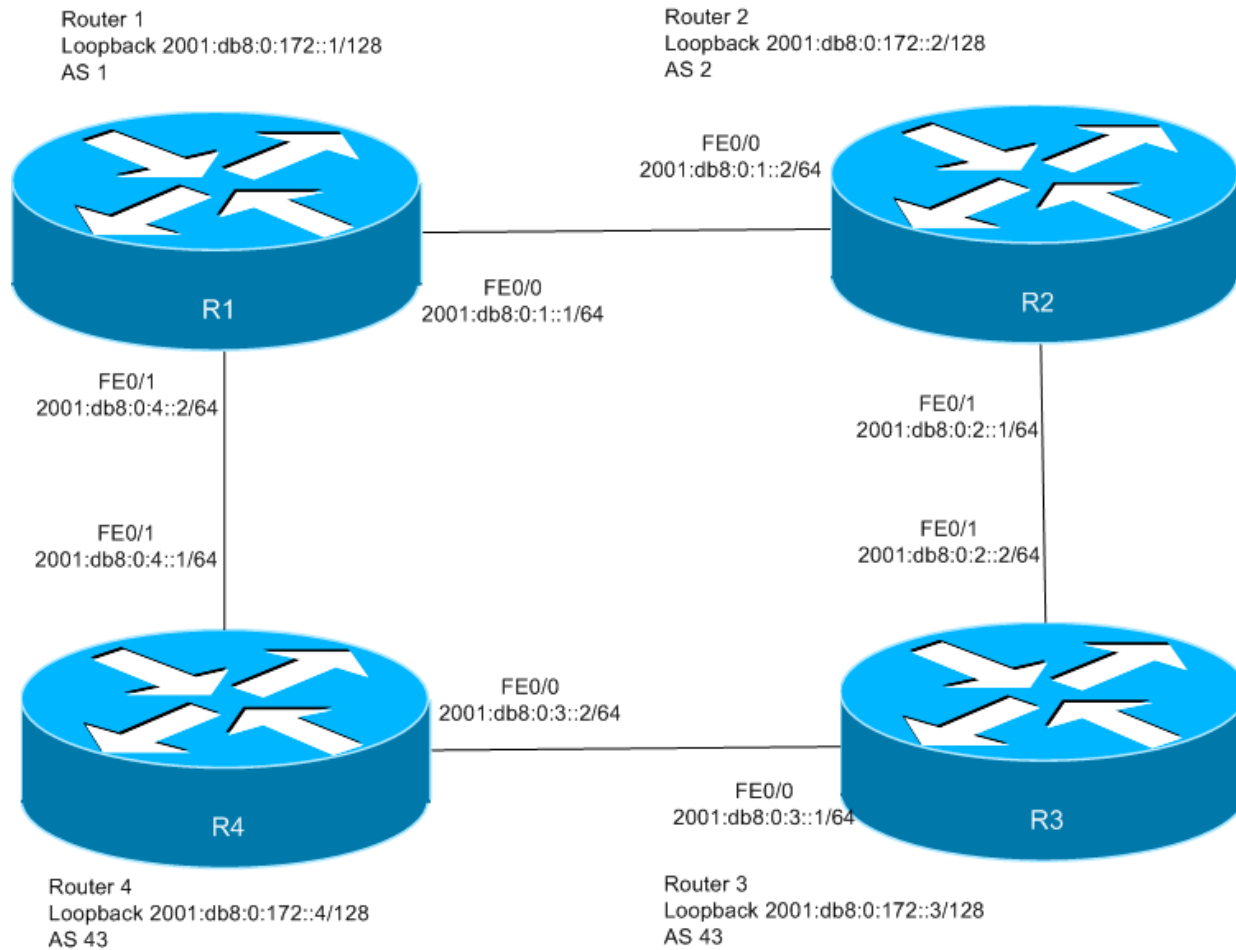


BGP Lab

- Externes BGP zwischen R1 und R2
- Externes BGP zwischen R1 und R4
- Externes BGP zwischen R2 und R3

- Internes BGP zwischen R3 und R4

BGP Lab



BGP Vorarbeiten

- ipv6 unicast-routing
 - IPv6 Routing ist nicht per default an
 - Ändert sich vielleicht in neuen IOS Versionen
- BGP Router ID
 - 32 bit Zahl in Verkleidung einer IPv4 Adresse
 - Wird aus IPv4 Adresse gewonnen
 - IPv6 only: Von Hand setzen

BGP Address Family

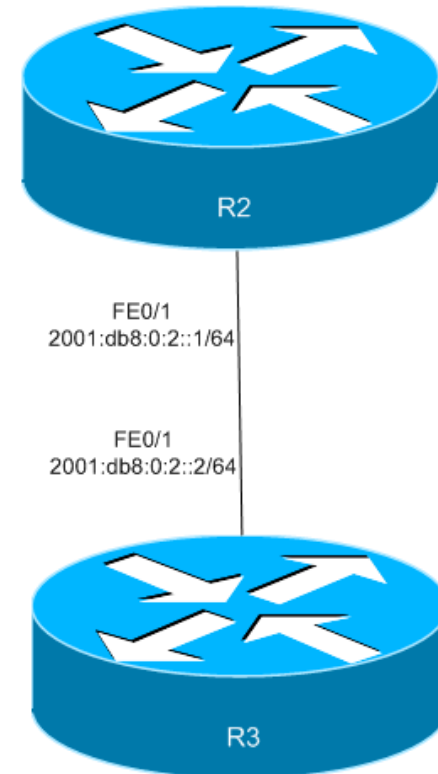
- ipv4
- ipv6
- vpnv4
 - MPLS Layer 3 VPN
- vpnv6
 - MPLS Layer 3 VPN
- eVPN
 - Für MAC Adressen, Layer2 über Layer 3

R2 – R3 eBGP

```
router bgp 2
no synchronization
bgp router-id 2.2.2.2
bgp log-neighbor-changes
neighbor 2001:DB8:0:1::1 remote-as 1
neighbor 2001:DB8:0:2::2 remote-as 43
no auto-summary
!
address-family ipv6
neighbor 2001:DB8:0:1::1 activate
neighbor 2001:DB8:0:2::2 activate
network 2001:DB8:888::/48
exit-address-family
```

```
router bgp 43
no synchronization
bgp router-id 3.3.3.3
bgp log-neighbor-changes
neighbor 2001:DB8:0:2::1 remote-as 2
neighbor 2001:DB8:0:172::4 remote-as 43
neighbor 2001:DB8:0:172::4 update-source Loopback0
no auto-summary
!
address-family ipv6
neighbor 2001:DB8:0:2::1 activate
neighbor 2001:DB8:0:172::4 activate
neighbor 2001:DB8:0:172::4 next-hop-self
exit-address-family
```

Router 2
Loopback 2001:db8:0:172::2/128
AS 2



Router 3
Loopback 2001:db8:0:172::3/128
AS 43

R3 – R4 iBGP

- Internes BGP
- Full Mesh, jeder mit jedem
- Muss auch bei Leitungsausfall gehen

- iBGP hängt an Loopback Interfaces
- iBGP benötigt dafür IGP (OSPFv3, IS-IS)

OSPFv3

```
ipv6 router ospf 200
```

- router-id 3.3.3.3

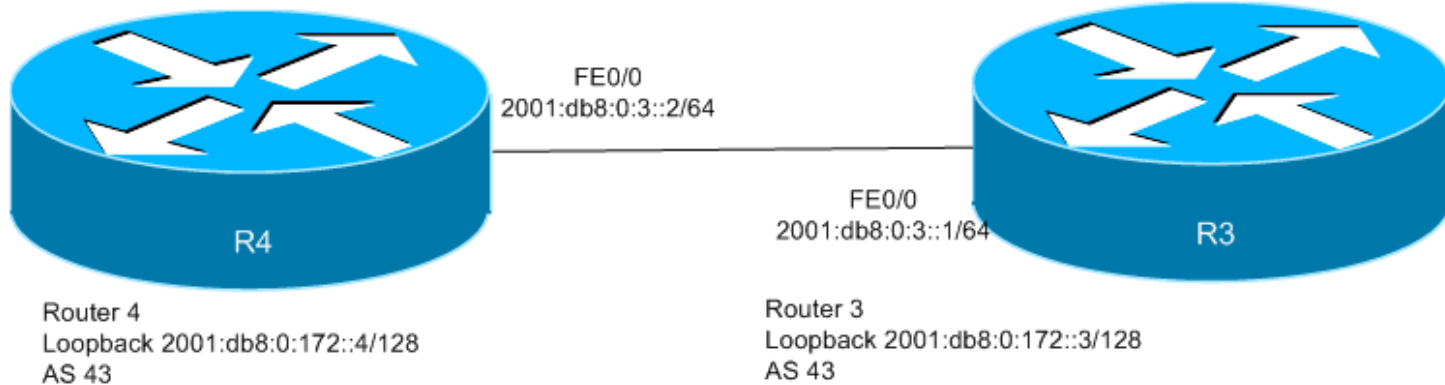
```
interface Loopback0
```

```
no ip address
```

```
ipv6 address 2001:DB8:0:172::3/128
```

```
ipv6 ospf 200 area 0
```

iBGP Konfiguration



```
router bgp 43
no synchronization
bgp router-id 4.4.4.4
bgp log-neighbor-changes
neighbor 2001:DB8:0:4::2 remote-as 1
neighbor 2001:DB8:0:172::3 remote-as 43
neighbor 2001:DB8:0:172::3 update-source Loopback0
no auto-summary
!
address-family ipv6
neighbor 2001:DB8:0:4::2 activate
neighbor 2001:DB8:0:172::3 activate
neighbor 2001:DB8:0:172::3 next-hop-self
exit-address-family
```

```
router bgp 43
no synchronization
bgp router-id 3.3.3.3
bgp log-neighbor-changes
neighbor 2001:DB8:0:172::4 remote-as 43
neighbor 2001:DB8:0:172::4 update-source Loopback0
no auto-summary
!
address-family ipv6
neighbor 2001:DB8:0:172::4 activate
neighbor 2001:DB8:0:172::4 next-hop-self
exit-address-family
```

Router 3 BGP Tabelle

```
R3#show bgp ipv6 unicast
```

```
BGP table version is 2, local router ID is 3.3.3.3
```

```
Status codes: s suppressed, d damped, h history, * valid, > best, i -  
internal, r RIB-failure, S Stale
```

```
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
---------	----------	--------	--------	--------	------

```
*> 2001:DB8:888::/48
```

2001:DB8:0:2::1	0	0	2	i
-----------------	---	---	---	---

Router 4 BGP Tabelle

R4#show bgp ipv6 unicast

BGP table version is 2, local router ID is 4.4.4.4

Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, r RIB-failure, S Stale

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*>i2001:DB8:888::/48	2001:DB8:0:172::3	0	100	0	2 i

Warum ist BGP so komplex?

- Filtern ist wichtig
- Filtern, welche BGP Prefixe reinkommen
- Filtern, welche BGP Prefixe rausgehen

- Filter-list
- Prefix-list
- Route-map

Warum ist BGP so komplex?

- Muss ich alles können?
- Nein, für die meisten Unternehmen reicht Basiswissen

Warum ist BGP so komplex?

- Muss ich alles können?
- Nein, für die meisten Unternehmen reicht Basiswissen
- Bei BGP kann vieles eingestellt werden
- Für Service Provider wichtig

BGP Begriffe

- BGP Decision Path
 - Welches Prefix von welchem Nachbarn nehme ich?
 - BGP lernt viel doppelt
 - BGP muss eine Auswahl treffen
 - Diese Auswahl kann beeinflusst werden

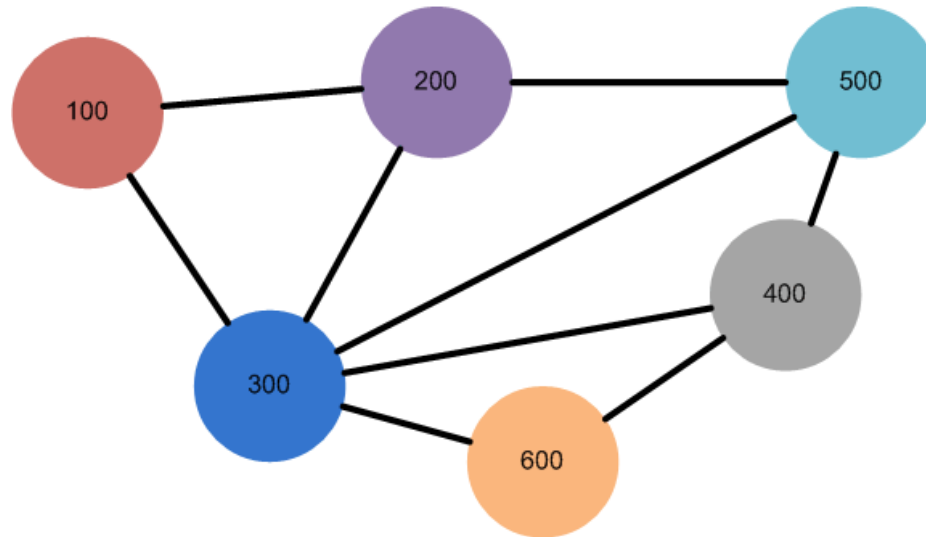
BGP Decision Path

- Local Preference
 - Beeinflusst eingehende Announcements
 - Beeinflusst ausgehenden Traffic

 - Legt fest, über welchen Uplink Traffic fließt
 - Gilt für das ganze Autonome System

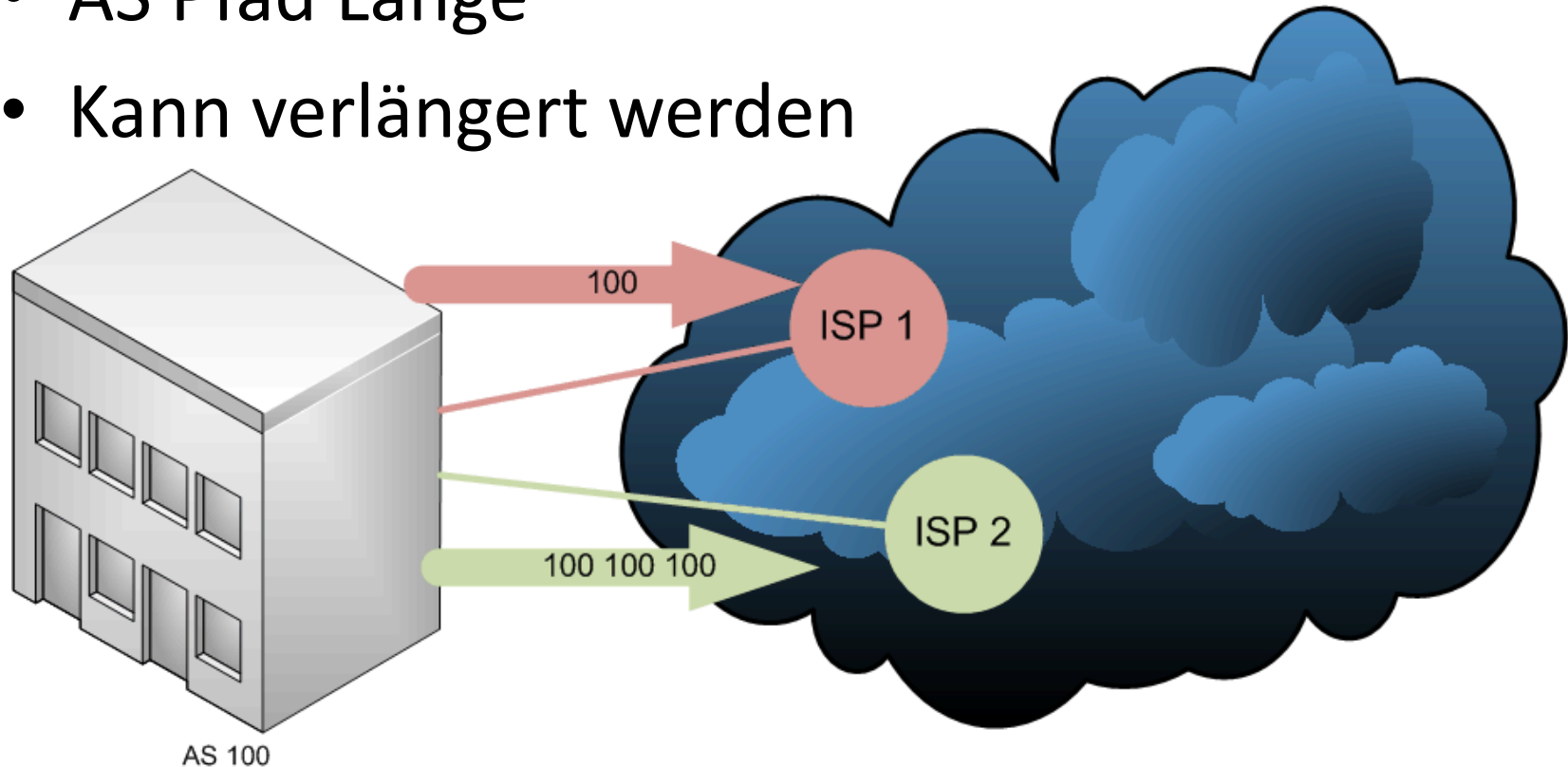
BGP Decision Path

- AS Pfad Länge
- Weg, den das Prefix genommen hat
- Sagt nichts über die wirkliche Länge des Weges aus



BGP Decision Path

- AS Pfad Länge
- Kann verlängert werden



BGP Fazit

- Multihoming
- Eigene AS Nummer
- Eigene IP Nummern
- Unabhängigkeit

- Komplexität
- Knowhow
- Hardware

BGP Fazit

- Komplexität
- Knowhow
- Hardware

- IPv6 macht unabhängig
 - Heute 16:30 Uhr, nach der Kaffeepause