

# Internet - wie funktioniert das eigentlich?

Heino Gutschmidt, [heino.gutschmidt@ines-consulting.de](mailto:heino.gutschmidt@ines-consulting.de)

INES Consulting GmbH

16.02.06

## 1 Grundlagen BGP-basierten Routings

- autonome Systeme (AS)
- Wegewahl im Internet
- BGP Routing Information Base - Theorie
- Informationsquellen

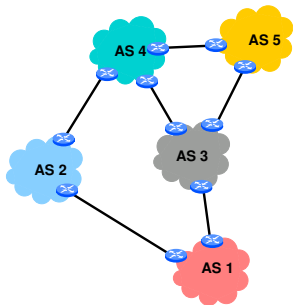
## 2 Demonstration

- Mini Internet
- BGP-Routing Information Base/Routing Tabelle - Praxis
- Auswirkungen eines Linkausfalls
- Software Zebra

# Fragen ?

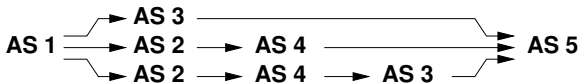


# autonome Systeme



- Internet besteht aus einer Kopplung von autonomen Systemen
- vergleichbar mit Autobahnen/Bundesstraßen
- die Border-Router “sprechen“ Border Gateway Protocol (BGP)
- Austausch von Prefix-Informationen (z.B. 213.187.64.0/19) verbunden mit der AS-Nummer (z.B. 16272)
- zwischen zwei AS existieren mindestens zwei unterschiedliche Wege

# Wegewahl im Internet basierend auf AS-Pfaden



- Definition Weg zwischen zwei AS als AS-Pfad
- Länge des AS-Pfades ist wichtiges Kriterium für die Wegewahl (insgesamt ca. 10 Parameter, je nach Implementation)

# BGP Routing Information Base

- logische Representation der BGP Routing Table
- besteht aus den Komponenten:
  - Input Policy Engine → empfangene Routing Informationen
  - Output Policy Engine → zu sendende Routing Informationen
  - Local Policy Engine → Routing Tabelle
- ca. 180.000 (!) Einträge für gesamtes Internet

- Beispiel (BGP-RIB):

	Network	Next Hop	Metric	LocPrf	Weight	Path
*>	12.6.208.0/20	80.81.192.63	0	800	0	174 1742 i
*		208.175.236.25	10	499	0	1273 174 1742 i
*		62.67.38.1	10	500	0	3356 209 1742 i

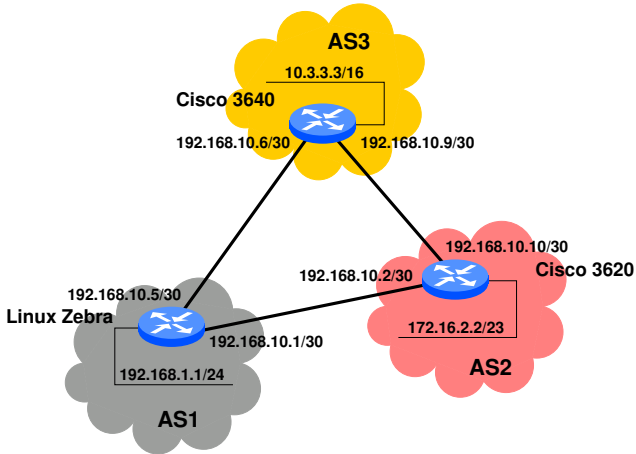
- Beispiel (Routing Table):

```
B          12.6.208.0/20 [20/0] via 80.81.192.63, 18:54:32
```

# Informationsquellen

- Route Servers (<http://www.traceroute.org>)
  - meist Cisco Command Line Interface
  - `show ip bgp regexp <ASNUMBER>$`
  - `show ip route | incl <PREFIX>`
- RIPE/whois (<http://www.ripe.net>)
- IRRToolSet-4.8.2 ⇒ `prtraceroute`

# Mini Internet



# BGP-Routing Information Base/Routing Tabelle - Praxis

- RIB:

	Network	Next Hop	Metric	LocPrf	Weight	Path
*	10.3.0.0/16	192.168.10.2				0 2 3 i
*>		192.168.10.6	0			0 3 i
*>	172.16.2.0/23	192.168.10.2	0			0 2 i
*		192.168.10.6				0 3 2 i
*>	192.168.1.0	0.0.0.0	0		32768	i

- Routing Tabelle:

```

B>* 10.3.0.0/16 [20/0] via 192.168.10.6, eth1, 03:31:03
B>* 172.16.2.0/23 [20/0] via 192.168.10.2, eth0, 03:06:54
C>* 192.168.1.0/24 is directly connected, lo
C>* 192.168.10.0/30 is directly connected, eth0
C>* 192.168.10.4/30 is directly connected, eth1

```

## Auswirkungen eines Linkausfalls

- ping zwischen AS1 und AS2 (192.168.1.1/172.16.2.2)
- Nutzung des kurzen AS-Pfades AS1 - AS2
- Linkausfall zwischen AS1 und AS2:

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 10.3.0.0/16	192.168.10.6	0			0 3 i
*> 172.16.2.0/23	192.168.10.6				0 3 2 i
*> 192.168.1.0	0.0.0.0	0		32768	i

- Nutzung des längeren AS-Pfades AS1 - AS3 - AS2
- Praxis: hold time 180s, keepalive 60s (Minderung Route-Flapping)

# Software

- GNU Zebra unterstützt BGP-4, RIPv1, RIPv2, RIPv6, OSPFv2, OSPFv3
- <http://www.zebra.org/>
- sehr modular - Zebra-Daemon und Protocol-Daemon
- Cisco CLI
- Lauzeitkonfiguration über CLI auf TCP-Port 2601 (zebrad) und 2605 (bgpd)
- Erfahrungen:
  - sehr stabil
  - preiswerte Alternative bei geringeren Bandbreiten (PC-Architektur)
  - Einsatzbeschränkung durch Linux-unterstützte WAN-Hardware