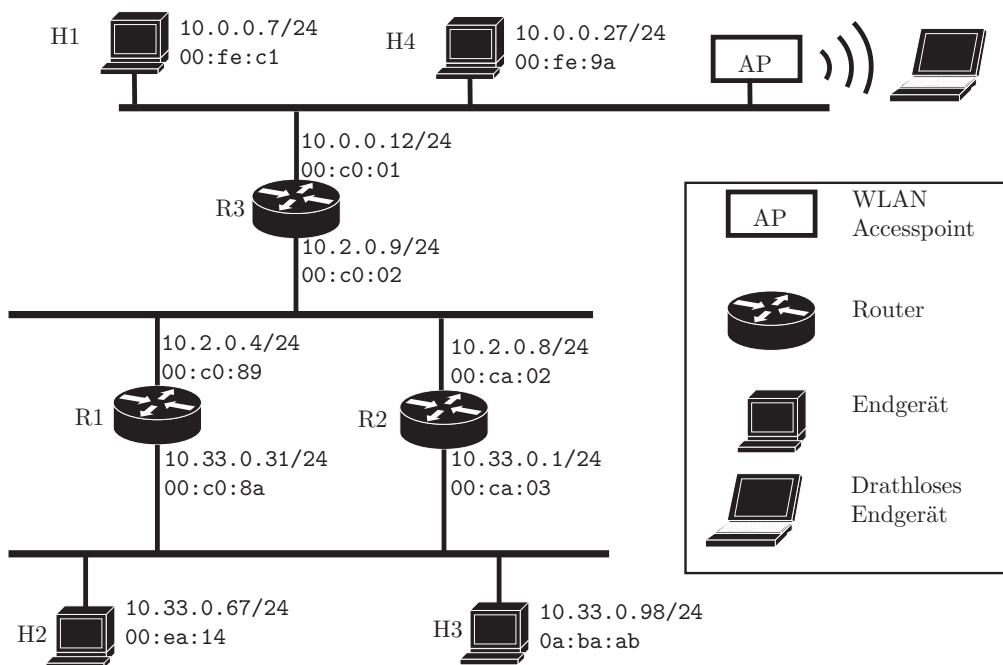


Übung 8: Forwarding, BGP, Peering

2019-12-12

1. Internetworking

Folgendes Netzwerk sei gegeben, die angeschlossenen Geräte haben eine MAC Adresse (hier auf drei Byte verkürzt) und eine IP Adresse.



- (a) Würde es Probleme geben, falls die Endgeräte H1 und H3 die gleiche MAC Adresse hätten? Erläutern Sie kurz! **Lösung:** Es gibt keine Probleme, da Hardware Adressen nur im lokalen Netzwerk gelten und H1 und H3 in unterschiedlichen lokalen Netzwerken sind
- (b) In der Vorlesung haben Sie verschiedene Rollen für Geräte die Netze verbinden kennen gelernt wie "Router", "Hub", "Repeater", "Bridge", "Switch".

Welche Rolle hat der Accesspoint AP? **Lösung:** Ethernet und WLAN verwenden verschiedene Paketformat (802.3 bzw. 802.11). WLAN emuliert zwar die Ethernet Semantik, ist aber genau genommen nicht Ethernet (z.B. drei bis vier Mac Adressen in jeden Paket, anstatt von zweien). Bei dem Accesspoint handelt es sich daher um eine Bridge.

- (c) Müssen beim Datenaustausch zwischen von Endgerät H1 und Endgerät H2 alle Pakete ausschließlich über Router R1 oder R2 weitergeleitet werden oder kann beispielsweise auch die Hinrichtung über R1 und Rückrichtung über R2 geleitet werden? **Lösung:** Nein, den oberen Layer ist der Weg auf der physikalischen Ebene völlig egal.
- (d) Das Endgerät H1 schickt ein Paket zu H2 (Echo Anfrage). H2 antwortet auf dieses Paket (Echo Antwort). Stellen Sie die dafür versandten Pakete in der korrekten Reihenfolge dar. Nutzen Sie dafür folgende Form:

	Quelle	Ziel
Ethernet	Mac Adresse Quelle	Mac Adresse Ziel
IP	Quell IP Adresse	Ziel IP Adresse

Lösung:

	Quelle	Ziel
Ethernet	00:fe:c1	00:c0:01
IP	10.0.0.7	10.33.0.67
	Quelle	Ziel
Ethernet	00:c0:02	00:ca:02
IP	10.0.0.7	10.33.0.67
	Quelle	Ziel
Ethernet	00:ca:03	00:ea:14
IP	10.0.0.7	10.33.0.67
	Quelle	Ziel
Ethernet	00:ea:14	00:c0:8a
IP	10.0.33.67	10.0.0.7
	Quelle	Ziel
Ethernet	00:c0:89	00:c0:02
IP	10.0.33.67	10.0.0.7
	Quelle	Ziel
Ethernet	00:c0:01	00:fe:c1
IP	10.0.33.67	10.0.0.7

- (e) Sie möchten in Ihren Hosts (H2, H3) ausnutzen, dass zwei redundante Router vorhanden sind. Sie überlegen sich folgendes Verfahren:
- Sie tragen als Router für H2 und H3 die IP 10.33.0.254 ein.
 - Sie modifizieren R1 und R2 so, dass diese auch auf die Anfrage für 10.33.0.254 antworten.

Beantworten Sie die folgenden Fragen zu dem Protokoll:

- Funktioniert dieses Verfahren?
- Welche Probleme sehen Sie hier?
- Was könnten Sie eventuell verbessern?
- Was passiert wenn ein Router ausfällt?

Lösung: Das Verfahren funktioniert. Wird auch in der Industrie so angewendet. Stichwort dazu ist First Hop Redundancy Protocol. Verbessern kann man, dass die Router sich koordinieren können. Wenn ein Router ausfällt, geht erstmal nix mehr für die entsprechenden Clients bis der Arp Cache gelöscht wird. Anstatt mit den eigenen Mac Adressen zu antworten können die Router auch virtuelle Mac Adressen übernehmen, die im Fehlerfall übernommen werden.

2. BGP-Routing

- (a) **Prefix** Ist 202.17.12.0/8 ein gültiger Prefix?

Lösung: Nein, bits ab der 9. Position sind gesetzt.

- (b) **Routing-Metriken und BGP** Angenommen, Sie betreiben ein Transit AS. Wie können Sie AS-Path Announcements nutzen, um Verkehr von Ihrem AS fernzuhalten?

Lösung: Indem man die eigene AS Number mehrfach in den AS-PATH schreibt.

- (c) **Lastbalanzierung** Angenommen, Sie betreiben ein AS mit einem 1.0.0.0/8 Prefix. Sie verbinden Ihr AS an zwei CIX-Punkten mit anderen AS-Betreibern. Wie können Sie AS-Path Announcements nutzen, um die Last zwischen diesen beiden Übergabepunkten zu balanzieren? Wie stellen Sie gleichzeitig Ausfallsicherheit her?

Lösung: Sie announcieren 1.0.0.0/9 und 1.0.0.0/8 von einem Gateway-Router; 1.128.0.0/9 und 1.0.0.0/8 vom anderen Gateway.

- (d) **Geographische Aufteilung** Angenommen, die beiden CIX sind geographisch weit voneinander entfernt. Hat das Einfluss auf Ihre Entscheidung oder auf andere Regeln, mit denen Sie Ihr AS betreiben?

Lösung: Die IP-Adressen bzw. Prefixe so im eigenen Netz verteilen, dass sie nahe den entsprechenden Prefixen liegen.

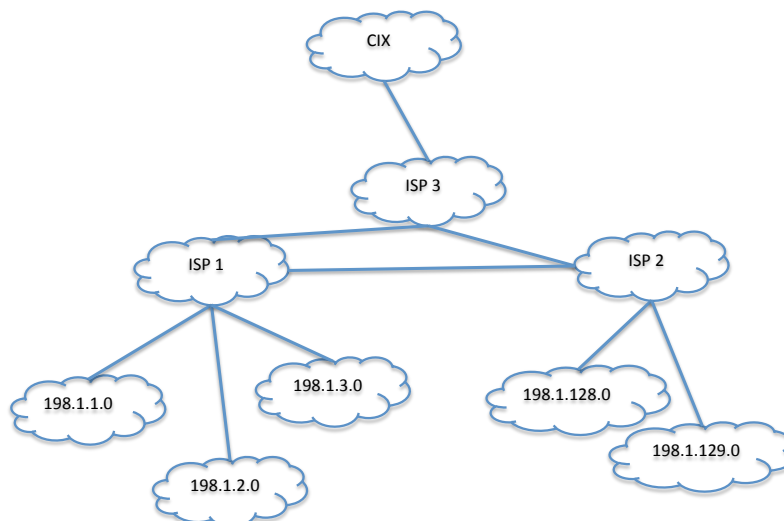
- (e) **Adresszuteilung** Sie betreiben ein Transit-AS, das Konnektivität an Endnutzer und Firmen verkauft. Ein Kunde möchte 1500 Rechner anschließen. Welchen Prefix wählen Sie?

Lösung: Z.B. einen /21-Prefix oder einen /20 und einen /19 Prefix.

- (f) **CIDR vs. NAT** Sowohl Classless Interdomain Routing (CIDR) als auch Network Address Translation (NAT) wurden als Lösung für das Problem knapper IP-Adressen angesehen. Stimmen Sie dem zu? Wo sind die Unterschiede?

Lösung: Lösen unterschiedliche Probleme. CIDR das Aggregationsproblem; NAT für private Netze geeignet

- (g) **CIDR oder Classfull Interdomain Routing?** Betrachten Sie unten stehende Abbildung. Welche Prefixe werden durch die einzelnen ISPs annoncirt? Machen Sie sinnvolle Annahmen über Transit/Customer und Peer-Beziehungen. Geben Sie jeweils eine Lösung für Classless und Classfull Interdomain Routing an.



Lösung: ISP 1 und 2 als Peers aufgefasst, sonst als Client/Transit-Beziehung. Die folgende Lösung geht davon aus, dass ISP 1 Adressen bis 198.1.127.255 besitzt und entsprechend annoncirt; ISP 2 entsprechend. Laut Aufgabenstellung (und wie in der Tafelübung besprochen) müsste man das entsprechend zu 198.1.0/22 ändern. Das Prinzip bleibt aber das gleiche.

CF ISP 1 nach ISP 3: 198.1.1/B, 198.1.2/B, 198.1.3/B; ISP 2 nach ISP 3: 198.1.128/B, 198.1.129/B
ISP 3 nach CIX: alle 5 Prefixe einzeln.

Mit CIDR

- ISP 1 nach ISP 3: 198.1.0/17

- ISP 1 nach ISP 2: 198.1.0/17
- ISP 1 zu seinen Kunden: 198.1.0/17 und 198.1.128/17 (und viele andere)
- ISP 2 nach ISP 3: 198.1.128/17
- ISP 2 nach ISP 1: 198.1.128/17
- ISP 1 zu seinen Kunden: 198.1.0/17 und 198.1.128/17 (und viele andere)
- ISP 3 nach CIX, ISP 1, 2: 198.0.0/15 (und viele andere)

Anmerkung: Das würde ohne CIDR ähnlich passieren, aber eben mit allen Prefixen einzeln aufgezählt.

(h) **BGP und Link-Fehler** Betrachten Sie folgende Situation:

- BlaNet ist Kunde von ISP-Transit1 und hat Prefix 198.32.1.0/24 erhalten; BlaNet ist ein reines Stub-Netz ohne Multihoming.
- ISP-Transit1 verkauft viele Prefixe unter 198.32/13 und annouciert daher diesen Prefix 198.32.0.0/13 an Backbone-AS.
- ISP-Transit1 benutzt 0.0.0.0/0 mit Ziel Backbone-AS in allen Routern. (Erläutern Sie, warum!)
- Beschreiben Sie was geschieht, wenn ein Packet mit Ziel 198.32.1.1 bei Backbone-AS eintrifft und die Verbindung zu BlaNet ausgefallen ist.
- Wie lösen Sie das Problem?

Lösung:

- Verbindung zu BlaNet ist ausgefallen, Paket für 198.32.1.1 kommt bei Backbone AS an. Dann passiert:
 - i. Das Paket trifft bei ISP-Transit1 ein. Dort gibt es keine Route zu 198.32.1/24 (wegen detektiertem Linkausfall wurde diese Route entfernt). Aber es gibt die allgemeine Route 0/0. Das Paket wird gemäß dieser Route zurück zu Backbone-AS geschickt.
 - ii. Von dort wird das Paket gemäß der annoucierten Route 198.32/13 zurück an ISP-Transit1 geschickt. Resultat: Routing-Loop!
- Lösung: Man muss in ISP-Transit1 eine "Null-Route" für alle selbst annoucierten Präfixe installieren. Also innerhalb von ISP-Transit1 gibt es einen Präfix 198.32/13, der alle Pakete verwirft. Bei intakten Links zu Kunden-AS gibt es einen spezifischeren Präfix, der dafür sorgt, dass das Paket richtig ankommt. Bei ausgefallenem Link und entferntem spezifischen Präfix führt diese Null-Route dazu, dass Pakete für das eigene Netz nicht zurück in das Backbone-Netz gespeist werden.

(i) **Übereifrige Prefix-Aggregation** Betrachten Sie die Situation aus unten stehenden Tabelle: BlaNet ist wie oben Kunde eines einzelnen ISPs,

EinsNetz und ZweiNetz haben ebenfalls ihre IP-Adressen von ISP1 bezogen, sind aber zusätzlich multi-homed mit ISP2.

- Zeichnen Sie das auf!
- Nehmen Sie an, die beiden ISPs announzieren aggressiv aggregierte Prefixe an einen gemeinsamen CIX. Dort kommt ein Paket für 198.24.17.1 an. Erläutern Sie, was mit dem Paket geschieht?
- Was sind korrekt Aggregierungs- bzw. Announzierungsregeln für diesen Fall? Was solle insbes. auch ISP1 verbreiten?

AS	Adressbereich	Prefix	Provider
ISP1	198.24.0 – 198.31.255.255	198.24.0.0/13	—
EinsNetz	198.24.0.0 – 198.24.15.255	198.24.0.0/20	ISP1, ISP2
BlaNet	198.24.16.0 – 198.24.23.255	198.24.16.0/21	ISP1
ZweiNetz	198.24.56.0 – 198.24.63.255	198.24.56.0/21	ISP1, ISP2
ISP2	198.32.0.0 –	198.32.0.0/13	—

Lösung:

- ISP1 an CIX: 198.24.0.0/13; ISP2 an CIX: 198.32.0.0/13, 198.24.0.0/18(!). Ein Paket für 198.24.17.1 wird per longest prefix matching an ISP2 verschickt. Der kennt aber keine Route dorthin. Ein sog. black hole ist entstanden.
- ISP2 darf keine Aggregation 198.24.0.0/18 verbreiten. Er muss die beiden Prefixe für EinsNetz und ZweiNetz (198.24.56.0/21 und 198.24.0.0/20) separat announzieren. ISP1 sollte diese beiden ebenfalls verbreiten, sonst wird alle Verkehr für EinsNetz und ZweiNetz über ISP2 fließen (longest prefix matching in CIX!). Das mag für ISP1 gut oder schlecht sein, je nach Vertrag. Allgemein ist beim Aggregieren von Prefixen, die nicht aus dem eigenen Adressbereich stammen, große Vorsicht geboten.

3. **The Art of Peering** Nehmen Sie an, Sie sind ein Content-Provider AS A (z.B. web hoster). Wir betrachten weiter

- einen Access-Provider B (z.B. Privatkunden-ISP),
- einen Transit-Provider AS X, mit dem AS B ein Peering-Agreement hat, AS A aber nur Transit kauft,
- einen weiteren Transit-Provider AS Y, mit dem AS A und AS B ein Transit-Agreement haben.

In dieser Situation sendet B an A nur wenig Verkehr (z.B. kleine HTTP-Anfragen), aber A muss an B große Datenmengen senden und dafür seinen

Transit-Provider bezahlen. Typischerweise wird der Verkehr über X fließen (wegen Peering zwischen X und B).

A möchten ein Peering-Agreement mit B eingehen, d.h., kostenlos Traffic mit A und B tauschen. Das ist für A attraktiv, da dann die Kosten für A wegfallen. Für B ist das nicht interessant, da B den Verkehr von A ja schon kostenlos erhält.

Wie kann A ein solches Peering-Agreement erreichen?

Lösung: Eine unmittelbare Verhandlung von A mit B wird voraussichtlich scheitern.

A sendet seine Antworten nicht über X, sondern über Y an B. Dadurch wird der Verkehr von A nach B teurer; B muss transit bezahlen. Nach einiger Zeit kann dann A mit B verhandeln und vermutlich Erfolg haben. Ggf. kann A auch aus B heraus Daten von A anfordern (z.B. durch Web Crawler o.ä.) und so den Datenumfang und damit die Kosten von B erhöhen.

Im "Peering Playbook" ist dies als Traffic Manipulation beschrieben. Dort finden sich noch viele weitere Möglichkeiten. <http://drpeering.net/white-papers/Art-Of-Peering-The-Peering-Playbook.html>