

# **Aufgabenstellung BGP-Versuch**

Praktikum Rechnernetze SS 2008

25. Juni 2008

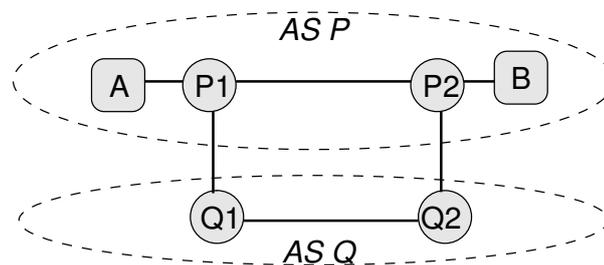
# 1 Theoretischer Teil

## 1.1 Intra-Domain-Routing vs. Inter-Domain-Routing

1. Warum werden für Inter- und Intra-AS-Routing unterschiedliche Protokolle verwendet?
2. Warum sind bei Intra-AS-Routingprotokollen, z.B. bei OSPF, Policies nur von begrenzter Bedeutung?
3. Warum spielen Policies dagegen im Inter-AS-Verkehr eine wichtigere Rolle als eine globale Optimierung des Netzverkehrs?
4. Wie umgeht BGP als Pfad-Vektor-Protokoll (im Gegensatz zu einem Distanz-Vektor-Protokoll) das Problem von Routing-Schleifen?
5. Während es viele Intra-AS-Routingprotokolle gibt, gibt es als Inter-AS-Protokoll quasi nur BGP als de-facto-Standard. Warum können sich alternative Protokolle hier nur schwer durchsetzen?

## 1.2 BGP und sein Entscheidungsprozess

1. Wiederholen Sie folgende für BGP wichtige Begriffe: AS, Policy, Peer, Peering, Pfad, Route Advertisement/Route Update, Routing Information Base, Transit Network, Stub Network, Internet Exchange Point, Exterior/Interior Gateway Routing Protocol, MED, Local Preference. Sie brauchen dazu nichts aufschreiben.
2. Stellen Sie den Entscheidungsprozess von BGP dar! Beschreiben Sie jeweils an einem Beispiel, wie die folgenden Einflussmöglichkeiten durch Admins wirken: Ändern der Local-Preference, künstliche AS-Path-Verlängerung, Ändern des MED.



3. Betrachten Sie die in der Skizze gegebene Topologie bestehend aus zwei autonomen Systemen  $P$  und  $Q$ , den Routern  $P1, P2, Q1, Q2$  und den Hosts  $A$  und  $B$ . Nehmen Sie für

alle Kanten ein Kantengewicht von 5 an und dass BGP als Inter-AS Routingprotokoll eingesetzt wird.

- a) Nennen Sie die kostenoptimale Route von *A* nach *B*.
  - b) Nehmen Sie nun an, dass sich das Kantengewicht zwischen *P1* und *P2* auf 20 erhöht. Nennen Sie für diesen Fall die kostenoptimale Route von *A* nach *B*. Welchen Weg nimmt ein Paket tatsächlich zwischen *A* und *B*? Begründen Sie kurz Ihre Antwort!
  - c) Nehmen Sie nun an, dass die Verbindung zwischen *P1* und *P2* komplett ausfällt. Wie wird jetzt der Verkehr von *A* nach *B* geroutet?
4. Lesen Sie den Artikel "Geordnetes Chaos - Wie Routing dem Internet seine Selbstheilungskräfte verleiht" aus c't 03/2006, S. 130.
- a) Nennen Sie zwei plausible Gründe, warum ein ISP nicht mit einem anderen ISP ein Peering eingehen möchte.
  - b) Schauen Sie sich das im Artikel gezeigte Bild der autonomen Systeme in Deutschland an. Zählen Sie für jeden der "Content Provider" und der "Business Customers" auf, mit welchen anderen AS-Betreibern sie verbunden sind. Interpretieren Sie das Ergebnis hinsichtlich der möglichen Strategie der einzelnen Teilnehmer-AS's.

### 1.3 Routing im realen Internet: BGP-Routen zum Web-Server der UC Berkeley

Nun wollen wir mal BGP etwas praktischer erkunden. Betrachten Sie dazu folgendes Beispiel, in dem wir versuchen, herauszubekommen, wie IP-Pakete von hier zum Web Server der Uni Berkeley geroutet werden.

Dafür brauchen wir zunächst die IP-Adresse des Zielhosts.

```
$ host www.berkeley.edu
www.berkeley.edu is an alias for www.w3.berkeley.edu.
www.w3.berkeley.edu has address 169.229.131.81
```

Anschließend schauen wir uns mal die Route von einem Host (im Beispiel ist die Quelle ein Rechner im Netz der LMU) zum Zielhost an. Wie Sie sehen, bekommen wir hier zwar die DNS-Namen und IP-Adressen der dazwischen liegenden Router, aber eigentlich würden uns die AS-Nummern interessieren. Dazu gleich mehr.

```
$ /usr/sbin/traceroute 169.229.131.81
traceroute to 169.229.131.81 (169.229.131.81), 30 hops max, 40 byte packets
 1  gongo.nm.ifi.lmu.de (141.84.218.125)  0.172 ms  0.141 ms  0.139 ms
 2  brouter-129-187-214.informatik.uni-muenchen.de (129.187.214.254)  1.029 ms  0.879 ms  1.645 ms
 3  vl-3003.csr1-kb1.lrz-muenchen.de (129.187.0.137)  1.571 ms  1.026 ms  0.791 ms
 4  vl-3002.csr1-2wr.lrz-muenchen.de (129.187.0.133)  0.789 ms  0.770 ms  0.789 ms
 5  xr-garl-te1-3-147.x-win.dfn.de (188.1.37.89)  0.766 ms  0.828 ms  0.789 ms
 6  zr-fra1-te0-7-0-1.x-win.dfn.de (188.1.145.53)  8.432 ms  8.413 ms  9.027 ms
 7  dfn.rtl1.fra.de.geant2.net (62.40.124.33)  8.285 ms  8.219 ms  8.157 ms
 8  abilene-wash-gw.rtl1.fra.de.geant2.net (62.40.125.18)  100.936 ms  111.088 ms  101.048 ms
 9  so-0-0-0.0.rtr.atla.net.internet2.edu (64.57.28.6)  114.482 ms  121.175 ms  115.143 ms
10  so-0-2-0.0.rtr.hous.net.internet2.edu (64.57.28.43)  139.294 ms  138.587 ms  138.266 ms
```

```

11 so-3-0-0.0.rtr.losa.net.internet2.edu (64.57.28.44) 169.918 ms 170.242 ms 175.492 ms
12 hpr-lax-hpr--i2-newnet.cenic.net (137.164.26.132) 171.797 ms 172.340 ms 171.837 ms
13 svl-hpr--lax-hpr-10ge.cenic.net (137.164.25.13) 179.161 ms 179.282 ms 179.137 ms
14 oak-hpr--svl-hpr-10ge.cenic.net (137.164.25.9) 180.076 ms 180.275 ms 180.193 ms
15 hpr-ucb-ge--oak-hpr.cenic.net (137.164.27.130) 180.115 ms 180.079 ms 180.085 ms
16 t2-3.inr-202-reccev.Berkeley.EDU (128.32.0.39) 180.230 ms 180.270 ms 180.216 ms
17 t1-1.inr-211-srb.Berkeley.EDU (128.32.255.43) 180.183 ms 180.134 ms 180.189 ms
18 * * *

```

Es ist nicht schlimm, dass uns der WWW-Server selbst nicht antwortet, er dürfte unmittelbar hinter dem letzten Hop sein, bis Berkeley sind wir ja schon gekommen.

Verbinden Sie sich nun per telnet mit `route-server.belwue.de`. Hinweis: Das Belwue-Netz ist das Landeshochschulnetz des Landes Baden-Württemberg. Dort können Sie in einer Emulation einer Cisco IOS-Shell die BGP-Routen zu beliebigen IP-Adressen erkunden, die ein dortiger Router in die Welt hat.

Geben Sie am Prompt folgendes Kommando ein. Es erscheint (zumindest aktuell):

```

route-server>show ip bgp 169.229.131.81
BGP routing table entry for 169.229.0.0/16, version 128725904
Paths: (1 available, best #1, table Default-IP-Routing-Table)
  Not advertised to any peer
  680 20965 11537 2153 25
    188.1.233.229 (metric 18) from 193.196.190.132 (193.196.190.132)
      Origin IGP, metric 0, localpref 450, valid, internal, best

```

1. Interpretieren Sie diese Antwort. Über welche AS'e führt die BGP-Route von hier nach `www.berkeley.edu`? Übernehmen Sie das Ergebnis in Ihre Lösung.

Hinweis: Die Zeile, die mit 680 anfängt, ist die BGP-Route. 680 ist die AS-Nummer des Deutschen Forschungsnetzes (DFN), durch welches der Verkehr als erstes geleitet würde, wenn er das Belwue-Netz verlassen hat. Dahinter kommen die Nummern der anderen AS's auf dieser BGP-Route. Über diese weiteren AS'es können Sie sich auf <http://ws.arin.net/whois> (Amerika) oder <http://www.ripe.net/whois> (Europa) informieren.

Auf `route-server.belwue.de` gibt es auch ein `traceroute`-Kommando, das den Routern gleich die passende AS-Nummer zuordnet.

```
route-server> traceroute 169.229.131.81
```

2. Visualisieren Sie das Ergebnis in einer Zeichnung, stellen Sie also die AS's als gestrichelte Ellipsen dar, die Router mit einem kleinen Kreis, die Verbindungen als Linien. Beziehen Sie die IP-Adressen und AS-Nummern mit ein, sowie die Ortsinformationen, die man mit etwas Phantasie aus den Routernamen ableiten kann.

Hinweise:

- HPR steht für "high performance router", also keinen Namen einer Stadt.
- "wash, atla, hstn, losang, lax, svl, oak, ucb" sind Abkürzungen für Städte. "svl" steht dabei für Sunnyvale, CA, die anderen sind leicht herauszubekommen.

- Statt die IP-Adressen immer wieder auszuschreiben, können Sie sie auch zusammenfassen. Dafür schreiben Sie einmal an eine geeignete Stelle einen Prefix hin, z.B. 188.1.0.0/16 und dann an die Router nur noch den verbleibenden Teil der IP-Adresse, also z.B. 37.13 und 18.26.

3. Analysieren Sie auf die gleiche Art die Routen zu drei Zielen in Australien:

- a) www.uwa.edu.au (University of Western Australia Perth)
- b) www.wesfarmers.com.au (Wesfarmers ist ein großer australischer Handelskonzern)
- c) www.bayweb.com.au (die Internetpräsenz des kleinen Badeortes Byron Bay)

Welche Hops/AS'e haben die Routen gemeinsam? Wo weichen Sie ab? Stellen Sie das Ergebnis näherungsweise in einer stilisierten Weltkarte dar. Beachten Sie dabei die Antwortzeiten, die über die Ausbreitungsverzögerung Rückschlüsse auf Entfernungen zulassen! Wer sind die Betreiber der einzelnen AS'e? Erklären Sie Gemeinsamkeiten und Unterschiede zwischen den Routen!

Hinweis: Sie dürfen hier ruhig ein bisschen spekulieren, wenn Sie Ihre Vermutung einigermaßen plausibel begründen.

Was wäre jeweils die optimale geringste Entfernung auf der Erdoberfläche und die damit verbundene Ausbreitungsverzögerung (round-trip!)? Wieso können wir real nicht unter den Wert kommen? Wie weit sind die realen vom optimalen Wert entfernt und warum?

## 2 Praktischer Teil

### 2.1 Szenario I: BGP kennenlernen in einem einfachen Netz von zwei Autonomen Systemen

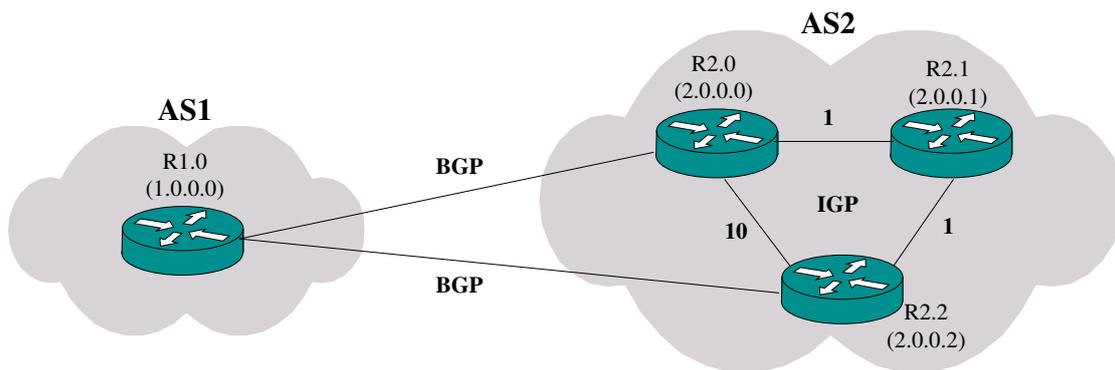


Abbildung 2.1: Ein Einfaches Netz

Abbildung 2.1 zeigt ein einfaches Netz aus zwei Autonomen Systemen AS1 und AS2. AS1 hat einen Router (BGP Speaker) R1.0, wobei AS2 drei Router R2.0, R2.1 und R2.2 beinhaltet. Innerhalb von AS2 wird ein IGP verwendet, die IGP-Gewichte der Verbindungen innerhalb AS2 sind in der Abbildung ersichtlich. R1.0 hält eBGP-Sessions mit R2.0 und R2.2 in AS2.

#### Aufgaben:

1. Schreiben Sie ein C-BGP Skript zu der angegebenen Topologie. Beachten Sie, dass unterschiedliche Routing-Verfahren zwischen den AS'en und innerhalb der AS'e angewendet werden.
2. Ermitteln Sie mit dem C-BGP Tools die RIB des Routers R2.1!
3. Ermitteln Sie die RIB des Routers R2.1 für den Fall, dass der Link zwischen R1.0 und R2.2 unterbrochen wird!
4. Ermitteln Sie die RIB des Routers R2.1 für den Fall, dass der Link zwischen R2.0 und R2.1 unterbrochen wird! (Hinweis: In diesem Fall ist die Verbindung zwischen R1.0 und R2.2 wieder hergestellt.)

## 2.2 Szenario II: BGP-Filter zwischen drei AS'es

Zu dieser Aufgabe ist eine Topologie-Datei `topologie.topo` vorbereitet. Diese Datei beschreibt die Peering-Beziehungen zwischen verschiedenen Autonomen Systemen. Wir betrachten dazu nur AS21(0.21.0.0), AS3300(12.228.0.0) und AS1239(4.215.0.0).

### Aufgaben:

1. Machen Sie sich mit der Topologie vertraut, z.B. mit `graphviz` oder per Textanalyse. Die Topologie ist sehr groß, betrachten Sie vor allem die AS'e, die von AS21 nur wenige AS-Hops entfernt sind. Beachten: Hier in diesem Beispiel mappen AS-Nummern auf IP-Prefixes, z.B. AS3300 enthält den Router 12.228.0.0 ( $12 \times 256 + 228 = 3300$ ).
2. Schreiben Sie ein `cbgp`-Skript, das folgende drei Dinge tut: Laden Sie die Topologie-Datei in das C-BGP Tool und starten Sie damit die Simulation! Propagieren Sie das Prefix 255/8 von Stub-Netz AS21 (multihomed an AS1239 und AS3300)! Lassen Sie alle AS-Pfade zum Prefix 255/8 in eine Datei schreiben (`bgp topology record-route`). Speichern Sie das Skript unter `scen-2-task-1.cli`.
3. Kopieren Sie `scen-2-task-1.cli` nach `scen-2-task-2.cli`. Verändern Sie `scen-2-task-2.cli`, sodass folgende Routing-Policy umgesetzt wird: "Bei allen BGP-Routen, die an AS 1239 (4.215.0.0) gesendet werden, soll AS 21 sechs Mal zusätzlich vorn angefügt ("prepending") werden."

Zeigen Sie den Erfolg dieser Maßnahme an einem Beispiel!

## 2.3 Szenario III: BGP in einem kleinen Internet

### 2.3.1 Topologie

Schauen Sie sich das Szenario “Fig. 1 Network topology of the Tiny-Internet.” an, lesen Sie den Text dazu, um alle gewünschten Eigenschaften des Routings zu erfassen.

1. Bilden Sie die Topologie nach!
2. Setzen Sie in den AS'es Intradomain-Routing auf sowie ggf. benötigte statische Routen. Begründen Sie dabei ihre Design-Entscheidungen, z.B. über die IGP-Gewichte!
3. Konfigurieren Sie die BGP-Router so, dass eine Interdomain-Konnektivität erreicht wird! Setzen Sie zunächst keine BGP-Routing-Policies ein!

Hinweis: Bitte halten Sie sich kurz bei der Beschreibung. Mehr als 6 Seiten müssen da nicht rauskommen.

### 2.3.2 BGP-Filter

1. Setzen Sie die Geschäftsbeziehungen der Domains (siehe Text) durch BGP-Routing-Policies um, die die Verbreitung der Routen entsprechend begrenzen!

### 2.3.3 Abgabe der Ergebnisse

1. Am Ende des Versuchs erwarten wir von Ihnen ein c-bgp-Skript, das das Szenario entsprechend umsetzt. Fügen Sie dieses Skript Ihrer Ausarbeitung bei.
2. Beschreiben Sie, wie sie Ihren Aufbau validiert haben! D.h. wie stellen Sie sicher, dass das Routing in ihrem modellierten Szenario den Anforderungen gerecht wird? Ggf. müssen hier mit einer Test-Automatisierung z.B. in einer Skriptsprache arbeiten.
3. In der letzten Veranstaltung sollen Sie Ihre Lösung den anderen präsentieren. Zeigen Sie dazu zunächst den “Normalfall”, d.h. dass alle Router und Links entsprechend den Anforderungen angeordnet sind und funktionieren. An ausgewählten Beispielen sollen Sie dabei das Routing begründen.
4. Wir werden Ihnen daraufhin Router und/oder Links nennen, die Sie in der Simulation deaktivieren sollen. Stellen Sie dann dar, wie sich das Routing der Situation “anpasst”, trotzdem aber die Geschäftsbeziehungen eingehalten werden!