Hochschule für Angewandte Wissenschaften Hamburg
*Hamburg University of Applied Sciences*

# Ausarbeitung Projekt 1

Marco Schneider

Border Gateway Protocol
Analyse & Topologie am BCIX Berlin

*Fakultät Technik und Informatik*
*Department Informatik*

*Faculty of Engineering and Computer Science*
*Department of Computer Science*

# Inhaltsverzeichnis

# 1 Einleitung

Das Border Gateway Protokoll (BGP) [13] ist der de-facto Standard [10] für das heutige Internet. Das bereits 1990 entwickelte Protokoll wird heute in der Version 4 [17] weltweit auf allen Autonomen Systemen (AS) eingesetzt. Es wird in der sog. „default-free-zone“ des Internets eingesetzt und bedingt, dass alle Router auf dieser Ebene zu jeder IP-Adresse einen expliziten Pfad zu dem AS kennen, das diese IP-Adresse „beherbergt“. Es arbeitet ausschließlich auf der Control-Plane, wird also nur zum Austausch der Routenpfade genutzt.

Es sind keinerlei Sicherheitsmechanismen im Protokoll implementiert, die einzige „Sicherheit“ für Fehlkonfigurationen sind die Policies (Filterregeln) der jeweiligen Betreiber. Dies wird oftmals jedoch nur in den unteren Hierarchieebenen praktiziert, da die größeren Carrier („uppertier's“) durch die große Anzahl an Kunden nicht für jeden Kunden eine eigene Filterregel anlegen können.

Dass dies nicht immer klappt, zeigte die Pakistan Telekom im Jahre 2008, als sie versuchte, YouTube zu sperren [11]. Die fehlerhafte Konfiguration führte dazu, dass der gesamte Traffic für YouTube über Pakistan geroutet wurde und so das AS innerhalb kürzester Zeit unnerreichbar wurde.

Doch es sind nicht nur falsche Konfigurationen: Eine gewollte Fehlleitung zeigte im März dieses Jahres die Türkei mit ihrer Sperre von Youtube & Twitter. Dort wurde politisch motiviert der Google-DNS mittels /32-Announce gekapert, um den Zugang zu besagten Diensten zu verwehren [18].

Durch den geheimdienstlichen Missbrauch („NSA-Skandal“, etc.) entwickelt sich immer mehr der Wunsch, dass der Übertragungsweg bekannt und transparent ist. Das Thema ist nicht neu, schon 2012 wurde in [21] das deutsche Internet Identifiziert und klassifiziert.

Die INET-Gruppe der HAW-Hamburg beschäftigt sich u.a. im Projekt Peeroskop mit der Analyse des deutschen Internets. Durch eine Kooperation mit einem Berliner IXP ergibt sich die Möglichkeit, die Topologie zu analysieren und diese mit öffentlichen Looking-Glass-Anbietern vergleichen zu können. Dieser Einblick und die daraus entstehenden Informationen sind unerlässlich, um z.B. Anomalien erkennen zu können, oder einfach ein sehr genaues Bild der Topologie zu erhalten.

In der folgenden Fragestellung wird eine kurze Übersicht zum aktuellen BGP-Routing und der verwendeten Topologie gegeben. Es folgt die Zielsetzung mit zwei Schwerpunkten für diese Projektarbeit. Abschließend werden die in Frage kommenden Datenquellen diskutiert.

Die Methodik in Kapitel 3 beschreibt, wie die Daten aufbereitet werden, um die Schwerpunkte dieser Arbeit beantworten zu können. Die daraus folgenden Ergebnisse werden im folgenden Kapitel 4 vorgestellt.

Eine kurze Zusammenfassung und ein Ausblick auf mögliche Folgeprojekte wird in Kapitel 5 gegeben.

# 2 Fragestellung

Eine vollständige Routingtabelle besteht gegenwärtig aus ca. 508.000 - 512.000 einzelnen Präfixen sowie ca. 47.550 einzelnen Autonomen Systemen (AS) [1] [19]. Diese AS bilden eine Hierarchie aus verschieden großen Providern.

Ganz oben in der Hierarchie sind die Tier-1-Provider (z.B. Level 3, Deutsche Telekom, TeliaSonera). Hauptmerkmal ist, dass diese Provider keinen Upstream-Provider haben, also niemanden dafür bezahlen, dass die Daten verteilt werden.

Darunter sind die Tier-2-Provider (z.B. Vodafone, Tele2), die einen Großteil ihrer Daten selber verteilen können, aber für einen gewissen Teil die Upstream-Provider brauchen.
In der untersten Hierarchieebene befinden sich die Tier-3-Provider, die Großteile, bzw. alle Daten über Upstreamprovider verteilen.

Tier-2- und Tier-3-Provider müssen die jeweils höhere Instanz für den Datenverkehr („IP-Transit“) bezahlen. Um die Kosten für den IP-Transit möglichst gering zu halten, versuchen die Provider, die Daten direkt untereinander (kostenfrei) auszutauschen. Dies ist eine sog. Peer-to-Peer-Verbindung zwischen den beiden Providern, oder kurz: ein Peering (vgl. Topologiemodell von Lixin Gao [9]).

Es sind jedoch nicht alle dieser Peerings für die Öffentlichkeit sichtbar. Diese sog. privaten Peerings sind von außen schwer zu erkennen, da sie u.a. nicht als Transit-Weg weiterverbreitet werden wollen. Dabei könnte eine Fehlkonfiguration entstehen, in der ein Provider einen Customer für seinen IP-Transit nutzen würde („valley-free-routing“).

An einem Internet Exchange Point (IXP) sind viele verschiedene Provider auf dem Layer 2 miteinander verbunden. Da jeder Provider eine vollständige Routingtabelle hat, stellt sich die Frage, ob man „mitten in der Topologie“ an einem IXP diese Peerings (besser) erkennen kann im Vergleich zu öffentlichen Looking-Glasses.

## 2.1 Zielsetzung

Die lokal bekannte Topologie des BCIX soll mit anderen Datenquellen verglichen werden. Dazu sollen verschiedene Daten von anderen Standpunkten in der Topologie genutzt werden, um so Unterschiede oder Fehler finden zu können.

Aktuell besteht das BCIX aus 58 Teilnehmern [6] mit eigenen AS', diese Topologie soll analysiert werden. Dazu soll das BCIX als Kern angesehen werden und zwei weitere Hops um das BCIX herum analysiert werden. Das BCIX ist in dieser Arbeit als Konglomerat aller Mitglieder zu verstehen, nicht als das einzelne AS des BCIX selber, da es nicht am Routing beteiligt ist.

Es sollen nur die Peerings, also die Verbindungen zwischen zwei AS', miteinander verglichen werden. Dabei soll es keine Rolle spielen, ob die Verbindung gerichtet[1] ist oder extrem wenig/viel genutzt wird. Das einfache Erkennen der Verbindung aus den Routingdaten ist ausreichend für den Vergleich der verschiedenen Quellen.

#### Schwerpunkt 1: Genauigkeit analysieren

Durch den direkten Zugriff auf einen IXP ist die lokale Topologie genau bekannt, es ist also eine „Ground-Truth[2]" vorhanden und kann mit den allgemein zugänglichen Quellen im Internet verglichen werden.
Die zusätzlich erkannten Peerings sollen analysiert und quantifiziert werden und hinsichtlich ihrer Relevanz zum BCIX bewertet werden.

#### Schwerpunkt 2: Abweichungen erkennen

Die gefundenen Abweichungen sollen erkannt und analysiert werden. Dabei soll festgestellt werden, woher diese Abweichungen stammen und ebenfalls hinsichtlich der Relevanz zum BCIX bewertet werden.

## 2.2 Aktive und passive Messungen

Es gibt zwei unterschiedliche Messmethoden: einmal aktive, und einmal passive Messungen. Die passiven Messungen basieren auf Informationen von BGP Trace-Kollektoren, Route-Servern, Looking-Glasses und Internet Routing-Registry (IRR)-Datenbanken [22]. Bei aktiven Messungen hingegen werden Forwarding-Pfade mithilfe von Messpakete verifiziert.

In dieser Projektarbeit sollen rein passive Messungen zum Einsatz kommen. Zu einem späteren Zeitpunkt sollen die Ergebnisse mit dem aktiven Ansatz verglichen werden, da aktuell das Parallelprojekt der aktiven Messung [12] noch nicht zur Verfügung steht.

[1] Es wird i.d.R. beim BGP-Routing die Control-Plane als Richtung angenommen
[2] Aus der Geoinformatik entlehnt: Bezeichnet Fernerkundungsdaten, die mittels vor Ort gefundener Informationen verbessert werden

## 2.3 Datenquellen

Bei der Auswahl der verschiedenen Datenquellen sollte besonders Wert auf die Verfügbarkeit und die Zuverlässigkeit der Daten gelegt werden. Dies hat den Grund, dass z.B. Scripte bei Fehlern in den Rohdaten nicht korrekt funktionieren und es so zu einem erheblichen Mehraufwand kommen könnte. Weiterhin kann die Verifizierung der Daten aus einem wohlbekannten Anbieter geringer ausfallen, als bei einem sehr kleinen oder unbekannten Anbieter. Es gibt im Internet verschiedene Anbieter solcher Daten, die hier kurz vorgestellt und bewertet werden.

### 2.3.1 BCIX Route-Reflektor

BCIX [6] steht für „Berlin Internet Exchange and Peering Point“ und ist der drittgrößte[3] IXP in Deutschland. Durch eine Kooperation mit dem BCIX ist ein Zugriff auf den Route-Reflector (RR) möglich. Dort werden von vielen der angeschlossenen Mitgliedern die Routen auf Peering-Ebene ausgetauscht, d.h. diese Routen sind nicht notwendigerweise in der Hierarchie des Internets zu sehen (Public-Peering vs. Private-Peering).

Diese Daten sollen mit den anderen Datenquellen verglichen werden um so eine Aussage darüber treffen zu können, ob die Sicht mitten in der lokalen Topologie Vorteile bringt oder nicht.

### 2.3.2 Andere Quellen

#### RIPE-DB

Die RIPE [3] (Réseaux IP Européens, dt.: Europäische IP-Netze) ist eine RIR[4], koordiniert das europäische Internet und stellt eine Datenbank ([4]) zur Verfügung, die die Peerings zwischen den verschiedenen AS' beschränkt. Diese Daten werden von den Betreibern der AS' gepflegt und sind somit nicht immer aktuell. Dennoch sind diese Daten sehr wertvoll, da sie die von den Betreibern gewünschte Konfiguration anzeigen und somit Abweichungen besser festgestellt und nachvollzogen werden können.

[3] gemessen am Peak-Traffic; DE-CIX: 3,4 TBit/s; ECIX: 300 GBit/s; BCIX: 70 GBit/s
[4] Regional Internet Registry

### Route Views

Route Views [15] ist ein Projekt der University of Oregon, welches ursprünglich die Möglichkeit bieten sollte, Routeninformationen aus verschiedenen Quellen in Echtzeit anzeigen zu können. Aktuell stehen 18 verschiedene Looking-Glasses zur Verfügung, die eine öffentliche Nutzung erlauben. Ein besonders interessantes Angebot ist das Routing-Archiv, welches teilweise bis zum Jahre 1997 zurück die Routingtabellen zur Verfügung stellt.

Es gibt insgesamt zwei unterschiedliche Angebote, einmal OIX (entspricht der Ausgabe „sh ip bgp“ auf einem Cisco-Router) und einmal BGPDATA (entspricht allen BGP-Informationen im MRT-Format). Im Idealfall sollen diese Daten gleich sein, da aber verschiedene Router (Hardware vs. Software) an verschiedenen Standorten verwendet werden, werden beide Angebote untersucht.

### Packet Clearing House

Das PCH [16] ist eine non-profit Organisation, welche sich u.a. zum Ziel gesetzt hat, die Analyse des Internets zu fördern. Dazu gibt es eine Vielzahl von Routencollektoren, die u.a. auch archiviert werden. Leider war das Archiv im betreffenden Zeitraum oft nicht verfügbar, sodass diese Quelle als unzuverlässig eingeschätzt wurde und in diesem Teil der Arbeit keine Verwendung findet.

### UCLA - Internet Topology Collection

Das Internet Research Lab der UCLA[5] hat bis zum Juni 2013 das Projekt „Internet Topology Collection“ [20] betrieben. Dazu gehörte auch ein Script, welches verschiedene Quellen ausliest und aufbereitet zur Verfügung stellt. Glücklicherweise wurde dieses Script nicht mit dem Projekt eingestellt und liest auch heute noch 133 verschiedene Router aus und stellt eine Liste der AS-Verbindungen zur Verfügung. [22]

Dieses Projekt ist eine sehr wichtige Quelle, da bereits alle vorher genannten Quellen erfasst werden und so nicht mehr einzeln abgefragt werden müssen.

[5] University of California Los Angeles

# 3 Methodik

Um die Daten miteinander vergleichen zu können, ist es unerlässlich, dass die Daten auf ein gemeinsames Format gebracht werden. Dazu müssen im ersten Schritt die Rohdaten aus dem Maschinenformat in ein menschenlesbares Format konvertiert werden. Danach müssen die Daten analysiert werden auf Verbindungen zwischen den Autonomen Systemen. Wenn man einmal in die (menschenlesbaren) Daten schaut, erkennt man schnell, dass sich die gewünschten Informationen gut extrahieren lassen. Dazu ein Beispiel einer Update-Nachricht:

Listing 3.1: BGP v4 Update-Nachricht

```
FROM: 193.178.185.5 AS16374
TO: 193.178.185.26 AS51224
ORIGIN: IGP
ASPATH: 6939 30844 37545
NEXT_HOP: 193.178.185.34
MULTI_EXIT_DISC: 290
ANNOUNCE
   154.73.43.0/24
   154.73.42.0/24
   154.73.41.0/24
   154.73.40.0/24
```

Es ist beim passiven Ansatz ausreichend, den Erreichbarkeitsvektor „ASPATH“ zu analysieren. Es besteht zwar noch die Möglichkeit, dass der ASPATH manipuliert[1] wurde, die Wahrscheinlichkeit ist aber sehr gering.

Das von der UCLA genutzte Perl-Script [24] ist eine modifizierte Version des Tools bgpdump von Yu Zhang [23]. Es liest Dateien im MRT[2]-Format und generiert daraus eine Liste mit Verbindungen zwischen verschiedenen Autonomen Systemen. Da dieses Script schon hinreichend validiert wurde und exakt auf die Aufgabenstellung passt, wird es in dieser Projektarbeit verwendet.

[1] Route-Inflation gilt an dieser Stelle nicht als Manipulation
[2] Multi-Threaded Routing Toolkit [7]

Der Begriff „Links“ ist als Synonym zu einem Peering, also einer Verbindung zwischen zwei Autonomen Systemen zu verstehen. Es wird mit BGP-Dumps gearbeitet, sodass in dieser Arbeit nur die Control-Plane und nicht die Data-Plane betrachtet wird.

## 3.1 Datenstruktur

Die Ausgabe des Perl-Scripts ist wie folgt vorgegeben:

IPv4/v6 daily : {ASN1} {ASN2}

Im ersten Schritt werden die verschiedenen Quellen mittels Script in die genannte Form umgewandelt und danach in eine Datenbank eingelesen. Die Datenbank speichert neben ASN1 und ASN2 zusätzlich noch die Herkunft der Daten (also von welchem Anbieter sie stammen) und einen Zeitstempel.

Die Entscheidung eine Datenbank zu nutzen wurde im Vorwege gefällt, da die recht großen Datenmengen (z.T. über 175.000 Einträge pro Quelle!) später performant nutzen zu können. Dabei wurde eine sqlite3-Datenbank verwendet, da kein extra Datenbankserver eingerichtet werden muss und die Performance mit einer SSD-Festplatte im zufriedenstellenden Bereich liegt.

Weiterhin können später komplexere Abfragen gestaltet werden um z.B. zwei Quellen direkt miteinander zu vergleichen. Dies war notwendig, um das Script der UCLA mit den anderen Quellen auf Korrektheit zu testen.

## 3.2 Vergleich der unterschiedlichen Daten

Ein einfacher Ansatz, um Daten miteinander zu vergleichen, wären z.B. zwei Arraylisten mit den zu vergleichenden Daten. Diese könnte man gegeneinander auswerten, welche Einträge gleich sind und diese dann aus beiden Arraylisten löschen. Diese Methode funktioniert zuverlässig, kostet aber auch sehr viel Rechenzeit (bei einem Test ca. 25 Minuten für einen Datensatz) und ist somit nicht günstig.

Eine weitere Idee ist die Nutzung von Python-Sets, welche sich einfach mittels einfachem „-“ im Quellcode abziehen lassen. So erreicht man die Differenz einfacher, leider jedoch auch mit sehr hohem Rechenaufwand. Diese Methode dauert ca. 10 Minuten pro Datensatz.

Bei der Aufgabenstellung fällt einem Informatiker schnell die Datenstruktur „Trie“ (Präfix-Baum) ein. Da die Daten bereits aber in einer Datenbank gespeichert werden, können auch SQL-Befehle benutzt werden. Dazu wird die Funktion „EXCEPT“ genutzt:

Listing 3.2: SQL-Abfrage

```
SELECT start, end FROM links WHERE sourceID=X EXCEPT
SELECT start, end FROM links WHERE sourceID=Y
```

Dabei stehen X und Y für die zu vergleichenden Quellen. Die SQL-Abfrage liefert alle Links aus Quelle X und zieht davon alle Links aus Quelle Y ab. Die Restmenge gibt dann Auskunft darüber, in welchen Punkten X genauer als Y ist (vgl. Venn-Diagramm in Bild 3.1). Die Laufzeit für Ausführung des Statements beträgt unter einer Sekunde bei aktuell 30 Datensätzen[3] in der Datenbank. Dies ist mehr als ausreichend schnell und stellt auch kein Problem bei größeren Datenmengen in der Datenbank dar.

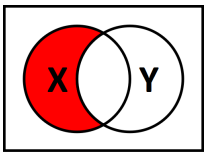


Abbildung 3.1: Ergebnis der SQL-Abfrage 3.2 (aus [14])

Mit dieser einfachen SQL-Abfrage wird die Frage beantwortet, in welchem Teil ein Datensatz besser/genauer ist, als ein anderer. Vergleichen wir nun unseren Ground-Truth-Datensatz mit den anderen Quellen, so erhalten wir (je nach Richtung des Vergleiches) die gewünschten Informationen über die Zusatzinformationen des jeweiligen Datensatzes. Einfach gesagt kann nach der Ausführung sofort eine Aussage getroffen werden, ob der Vergleichsdatensatz genau so gut ist (leere Menge) - oder ob der Referenzdatensatz mehr Informationen enthält (beliebige Menge).

Doch damit ist noch nicht die Frage auf das Gebiet um den BCIX beantwortet: nach der SQL-Abfrage müssen die Daten noch in Relation zum BCIX gesetzt werden, da Verbindungen am anderen Ende der Welt in diesem Kontext nicht wichtig sind. Dazu ist ein einfaches Aussortieren möglich, welches alle nicht am BCIX vorhandenen ASN löscht.

[3] Als Datensatz ist hier ein voller BGP-Dump eines Anbieters zu verstehen

# 4 Ergebnisse

Die Vergleichsdaten vom BCIX wurden am Route-Reflektor des BCIX über einen Zeitraum von ca. einem Monat[1] gesammelt. Dies entspricht einem MRT-Logfile von ca. 20 Megabyte.

Die verwendeten Daten der anderen Quellen stammen vom 28.07.2014.

Im ersten Schritt werden die verschiedenen Datenquellen miteinander verglichen in Bezug auf die „sichtbaren“ Links innerhalb des BCIX. Dabei wurden alle AS des BCIX auf Links untereinander untersucht. Die Ergebnisse sind in folgender Tabelle zu finden:

| Anbieter | Links | ASNs |
|---|---|---|
| BCIX | 20 | 58 |
| UCLA | 234 | 58 |
| Routeviews OIX | 121 | 58 |
| Routeviews BGPDATA | 125 | 58 |
| RIPE | 90 | 58 |

Es werden alle 58 vorhandenen AS am BCIX von allen Datenquellen erkannt. Dies wurde so erwartet und wäre auf eine Fehlkonfiguration zurückzuführen, falls ein AS nicht erkannt würde. Bei den Links fällt die sehr geringe Anzahl an Links am BCIX selber auf, wo die Sicht am genauesten sein sollte. Warum dies so ist, wird im folgenden Kapitel erklärt.

## 4.1 1-Hop-Problematik

Es erscheint paradox, dass im ’Zentrum des Wissens’, also am Routereflektor, nur 20 Links - und somit einen Bruchteil der anderen Quellen - erkannt werden. Zum einen liegt dies daran, dass ein sehr großer Teil der Peerings privat sind (vgl. [5]). Denken wir über das Peering an einem Routereflektor nach, so wird deutlich, dass alle Teilnehmer die Routen (gemäß ihrer Policies) mit dem RR austauschen (vgl. Bild 4.1). Dabei werden grundsätzlich keine Full-Tables übertragen, sondern nur die eigenen Netze (und natürlich die Netze der Customer) announced.

[1] 26.06.2014 - 28.07.2014

Dadurch kennen nun alle Peeringpartner des RR sämtliche lokalen Announcements der anderen AS und werden diese sehr wahrscheinlich in die lokale RIB übernehmen, da das entsprechende Netz direkt erreicht werden kann.

Dies hat aber auch zur Folge, dass der RR nur die beste Route zum Zielpräfix verteilt, die in diesem Falle nur aus einem Hop besteht. Durch den ASPATH von einem Element kann am RR des BCIX nicht festgestellt werden, welche Verbindungen tatsächlich existieren. Es ist aber davon auszugehen, dass alle Peeringpartner, die den RR nutzen, auch über sämtliche Routen verfügen (vgl. Bild 4.2).

Insgesamt sind 41 AS-Pfade mit einem Hop in der FIB des RR. Daraus lässt sich errechnen, dass mindestens 820 bidirektionale BGP-Verbindungen[2] am BCIX vorhanden sein müssen. Da insgesamt 58 Teilnehmer am BCIX angeschlossen sind, ist davon auszugehen, dass einige Teilnehmer privat peeren und somit die tatsächliche Anzahl noch höher ist.

Die 20 erkannten Routen sind Customer-Routen, die vom Upstreamprovider (beide am BCIX) an den RR weitergegeben werden. Damit existiert ein AS-Pfad aus zwei Elementen und kann somit vom Script erkannt werden. Somit sind nicht nur 20 Verbindungen am BCIX vorhanden, sondern es gibt insgesamt 20 Peerings am BCIX, wo der Customer Mitglied des BCIX ist, aber durch einen Provider (ebenfalls Mitglied des BCIX) verbunden ist.

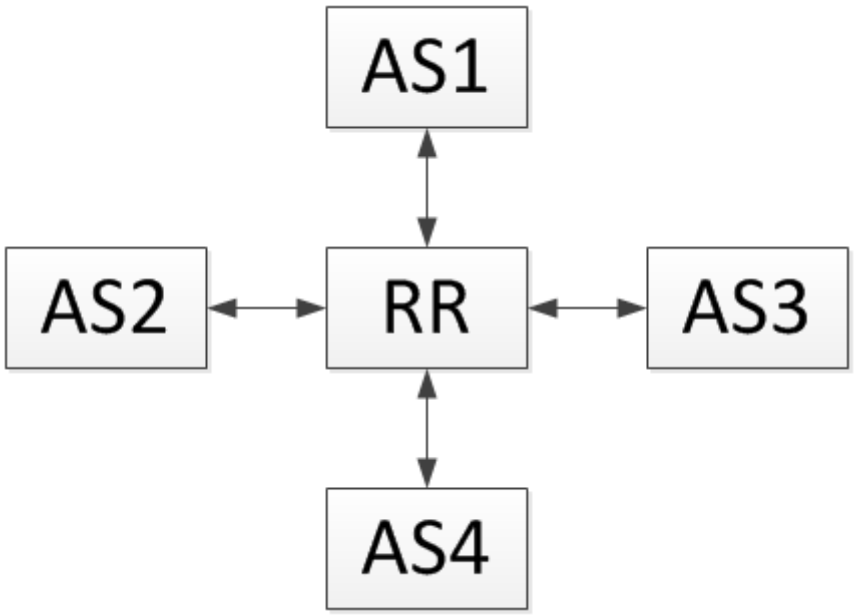


Abbildung 4.1: schematisches Peering mit RR

[2] $\frac{n*(n-1)}{2}$

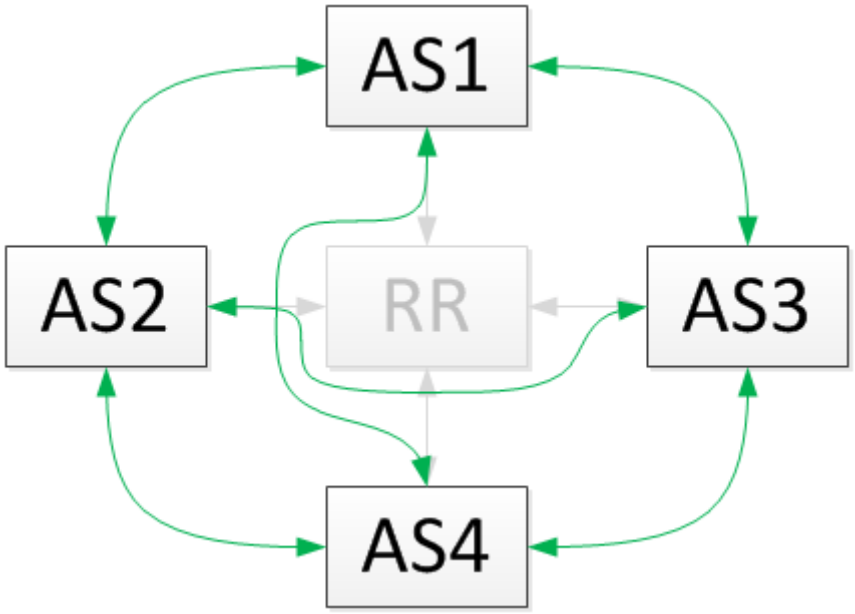


Abbildung 4.2: schematisch resultierendes Peering

## 4.2 BCIX-Topologie

In Kapitel 2.1 sollte ursprünglich das BCIX '+ zwei Hops' analysiert werden. Zwei Hops um das BCIX klingen erstmal nicht viel, da aber einige Teilnehmer des BCIX mit Tier-1-Providern (z.B. TeliaSonera - AS 1299) bzw. mit Autonomen Systemen peeren, die sehr viele Peeringpartner haben (z.B. Hurricane Electric - AS 6939), werden mit diesen zwei Hops schon knapp 71%[3] aller AS weltweit erreicht.
In Bild 4.3 ist als Übersicht die Topologie aller BCIX-Teilnehmer (insg. 58) dargestellt. Es zeigt den Aufbau des BCIX um zu verdeutlichen, das selbst diese 58 Teilnehmer schon so komplex vernetzt sind, dass eine sinnvolle Darstellung kaum mehr möglich ist.

Um die Bilder übersichtlich zu machen, wurden die Bilder in Hierarchieebenen eingeteilt. Als einziger Tier-1-Provider steht TeliaSonera ganz oben, darunter folgen die Tier-2-, und die Tier-3-Provider. Die Klassifikation erfolgte über die AS-Rank Abfrage [2] der CAIDA [8]. Tier-1-Provider wurden klassifiziert durch das Merkmal, keine Upstreamprovider zu haben. Tier-2-Provider hingegeben haben Peerings und Upstreamprovider, aber auch Customer. Tier-3-Provider und Stub's wurden in der untersten Kategorie zusammengefasst: Keine Customer, einen oder mehrere Upstreamprovider und vereinzelt Peerings.

[3] 2-Hop-AS: 33.774, Gesamt lt. CIDR-Report: 47.649

In den Bildern 4.7, 4.8 und 4.9 wurden der besseren Übersichtlichkeit wegen die Bilder abgeschnitten und untereinander fortgeführt. Die dargestellten Ebenen sind von oben gesehen T2-T3 T2-T3.

Die AS am BCIX wurden alle farblich markiert, um in der Topologie erkennen zu können, wie das BCIX involviert ist.

Aus diesem Grund und der Tatsache, dass der Genauigkeitsvorteil lokal begrenzt ist, wurden nur die Routen im BCIX selber und zusätzlich einem Hop analysiert. Es werden im Folgenden nur Graphen mit den direkt involvierten AS gezeigt, da schon die Gesamttopologie des BCIX (vgl. Bild 4.3) zu unübersichtlich erscheint, um sinnvolle Darstellungen zu ermöglichen.

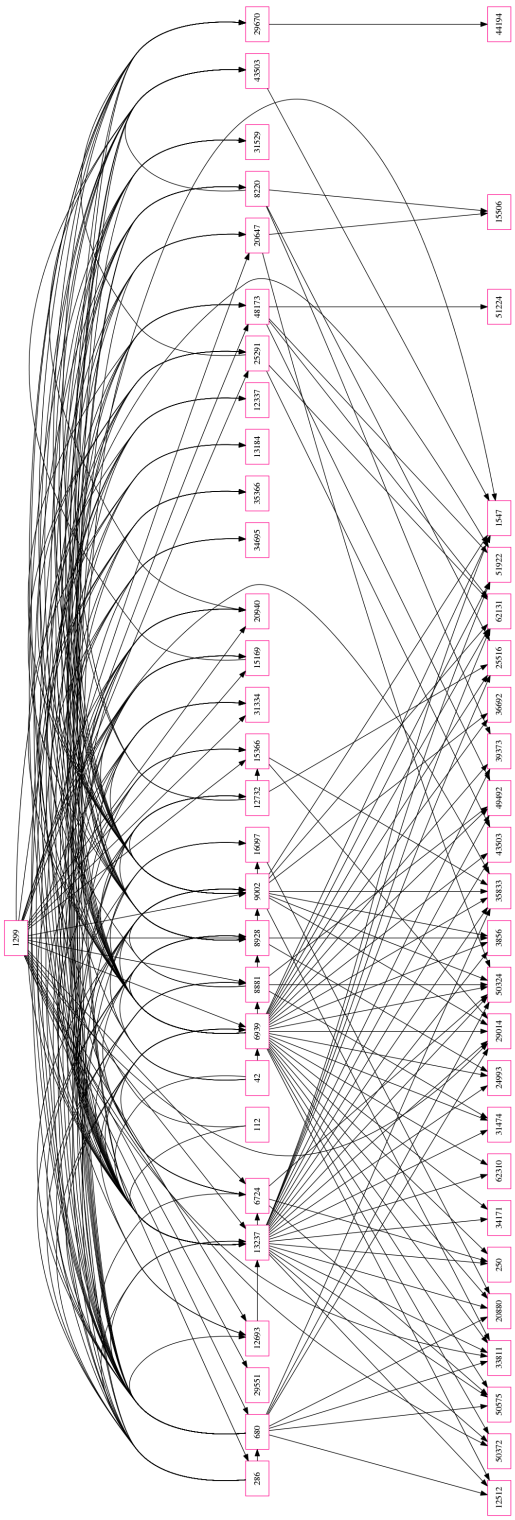


Abbildung 4.3: Topologie des BCIX, Tiers gemäß [2]

## 4.3 Analyse der Ergebnisse

Im folgenden Abschnitt werden die Ergebnisse dargestellt. Dazu wurden gemäß des Kapitels 3.2 die Daten verglichen und werden hier kurz erläutert. Als erstes werden die Ergebnisse vorgestellt und am Ende des Kapitels noch einmal diskutiert.

### 4.3.1 Statistiken der Quellen

Um eine kurze Übersicht über die Quellen zu bekommen, folgt hier eine kurze Zusammenfassung in Tabellenform. Dabei stehen „Links gesamt“ für alle erkannten Verbindungen der Datenquelle und „Links BCIX“ für alle Verbindungen innerhalb des BCIX.

| Anbieter | Datenstand | Links gesamt | Links BCIX |
|---|---|---|---|
| BCIX | 2014-07-28 14:07:38 | 4.788 | 20 |
| UCLA | 2014-07-28 17:58:51 | 176.268 | 234 |
| RIPE | 2014-07-28 14:28:00 | 100.271 | 90 |
| RouteViews BGPDATA | 2014-07-28 14:23:32 | 100.502 | 125 |
| RouteViews OIX | 2014-07-28 14:15:20 | 100.071 | 121 |

### 4.3.2 Sichtbarkeit der Links

In Bild 4.4 werden typischen Kategorien von Links gezeigt. Dazu zählen Downstream (Datenfluss Richtung Customer), Upstream (Datenfluss Richtung Provider), Peering (Datenaustausch zwischen zwei Providern), Stubs (Endknoten) und Siblings (Coexistenz zweier verschiedener AS, die logisch zusammen gehören, z.B. durch Firmenübernahme). Dabei gilt: der Kunde (dies kann auch ein Provider sein) bezahlt den Provider für die Verteilung seiner Daten im Internet (s. 4.5). Der Provider stellt sicher, dass die Kunden-Announcements an alle Peeringpartner weitergeleitet werden. Weiterhin versorgt er all seine Kunden mit seiner gesamten Routingtabelle.

Downstreams sind systembedingt durch die Advertisements auf der Controlplane zu erkennen: Ein Provider muss alle Netze seines Kunden (ungefiltert) weitergeben, um ihn so für andere AS erreichbar zu machen (vgl. Bild 4.5). Schwieriger sind alle anderen Kategorien aufzuspüren, da jeweils anhand der Policies des jeweiligen Betreibers nicht alle Routen weitergegeben werden, um z.B. das Valley-Modell (multihomed) nicht zu verletzen und Kosten zu sparen (Peerings).

Upstream-Links können hierarchiebedingt nur sehr schwierig auf der Controlplane erkannt werden. Es war nicht möglich, einen „Rückweg“ einer BGP-Verbindung zu verifizieren. Dies wäre

ohnehin nur in wenigen Szenarien (z.B. Multihomed Stub, ein nahes Peering oder eine Fehlkonfiguration) oder bei direktem Zugriff auf einen entsprechenden Router möglich.

Daher ist die Frage interessant, wie viele Links am BCIX von den verschiedenen Quellen gemeinsam erkannt werden, um so Rückschlüsse auf die Kategorie schließen zu können. Insgesamt werden von allen Quellen 238 verschiedene Links erkannt.

In der folgenden Tabelle wird die Anzahl der erkannten Links pro Quelle gezeigt. Dabei ist zu erwarten, dass bei steigender Quellenzahl (n Quellen können den Link verifizieren), die Wahrscheinlichkeit eines erkannten Peerings abnimmt.

| Anzahl Quellen | Summe Links | UCLA | RIPE | RV-BGPDATA | RV-OIX | BCIX |
|---|---|---|---|---|---|---|
| 1 | 112 | 108 | 3 | 1 | 0 | 0 |
| 2 | 4 | 4 | 0 | 3 | 0 | 1 |
| 3 | 35 | 35 | 0 | 35 | 33 | 2 |
| 4 | 70 | 70 | 70 | 70 | 70 | 0 |
| 5 | 17 | 17 | 17 | 17 | 17 | 17 |

Die Auswertung der jeweiligen Quellenzahl erfolgt im nächsten Abschnitt detailliert.

#### Sichtbar von einer Quelle

Die meisten Links werden von nur einer Quelle erkannt. Dazu macht die UCLA erwartungsgemäß den größten Anteil mit 108 Links aus, drei Links werden von der RIPE erkannt und die BGPDATA der RouteViews erkennt einen Link.

Wird ein Link nur von einer Quelle erkannt, so ist die Wahrscheinlichkeit am Höchsten, dass es sich dabei um ein Peering handelt. Das hängt mit dem Peering selber zusammen: Die Announcements von Peeringpartnern werden i.d.R. nur an die eigenen Customer weitergegeben, um keinen unnötigen Traffic von anderen Peeringpartnern auf sich zu ziehen und möglichst viel des eigenen Upstream-Traffics einzusparen.

Der Großteil der Informationen kommt von der UCLA, die mit einigen, großen Anbietern (z.B. DFN, RETN, Lambdanet) peert. Die zusätzlich gewonnenen Informationen sind ausschließlich Peerings[4]. Damit ist in diesem Fall die These bestätigt, dass bei Sichtbarkeit von einer Quelle es sich höchstwahrscheinlich um ein Peering handelt, das von anderen AS nicht erkannt werden kann.

Zwei von der RIPE stammenden Links existieren nicht in den Routingtabellen der anderen Anbieter, bei dem dritten Link handelt es sich um einen gewöhnlichen Upstream-Link. Da es

[4] diese Information stammt von [2] und wurde für jeden Link geprüft

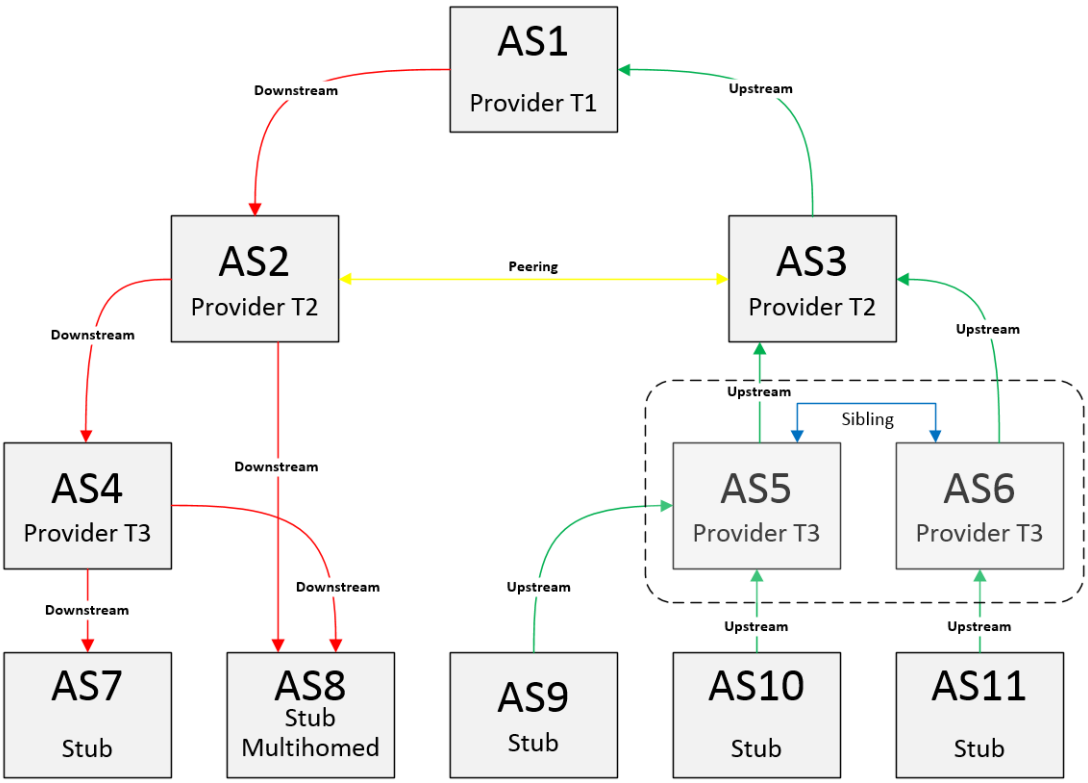


Abbildung 4.4: Schematischer Aufbau der BGP-Topologie

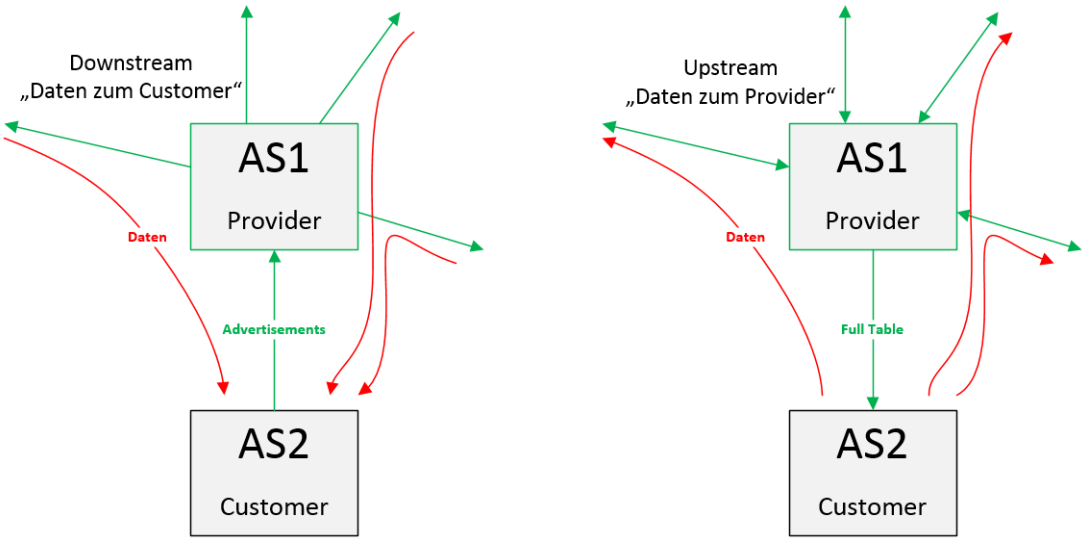


Abbildung 4.5: Schematische Darstellung von Up- & Downstreams

sich bei den nicht existierenden Verbindungen um alte bzw. falsche Informationen handelt, wurde eine Klassifikation nicht mehr vorgenommen.

#### Sichtbar von zwei Quellen

Drei der vier Links werden von der UCLA und RouteViews gesehen, es handelt sich bei zweien um Upstreamlinks und um einen nicht mehr existenten Link, der aber nach Klassifikation ein Peering hätte sein sollen.

Der verbleibende Link von der UCLA und dem BCIX ist ebenfalls ein Upstreamlink, sodass die Vermutung nahe liegt, dass ab zwei Quellen nur noch Upstreamlinks sichtbar sein werden.

Damit ist auch bestätigt, dass im Falle einer Sichtbarkeit von mehreren Quellen zwar noch die Chance eines Peerings besteht, wahrscheinlicher (und bestätigt) jedoch ist ein gewöhnlicher Upstream.

#### Sichtbar von drei Quellen

33 der 35 Links werden von den Quellen UCLA und beiden RouteViews-Daten erkannt. Es handelt sich in allen Fällen um Upstreamlinks. Die zwei verbleibenden Links werden vom BCIX,

der UCLA und RouteViews-BGPDATA erkannt. Auch hier handelt es sich erwartunsgemäß um Upstreamlinks.

#### Sichtbar von vier Quellen

Alle Links, die von vier Quellen erkannt werden, sind von den gleichen Quellen: Es handelt sich immer um UCLA, RV-BGPDATA, RV-OIX und RIPE.

Je mehr Quellen eine Route sehen, desto 'unspektakulärer' ist sie für die Analyse: es handelt sich auch nur (größere) Upstreamlinks.

#### Sichtbar von fünf Quellen

Von allen Datenquellen gesehene Links müssen auch wie bei vier Quellen genannt, sehr große und bekannte Upstreamlinks gemäß des Valley-Modells sein. Anders ist nicht zu erklären, warum diese Links weltumspannend erkannt werden können. Es handelt sich dabei hauptsächlich um Upstreams in lokaler Nähe vom BCIX, z.B. sind drei Links der insgesamt 17 vorhandenen Kunden von AS 48173 (unbelieveable Machine Company) - einem Mitglied des BCIX.

Es wurde gezeigt, dass bei der Sichtbarkeit von einer Quelle die Chance auf ein entdecktes Peering am Größten ist. Dies wurde anhand der Daten in diesem Fall bestätigt, denn alle Links (bis auf eine Ausnahme der RIPE), die von einer Quelle erkannt wurden, waren Peerings zwischen zwei Anbietern.

Weiterhin wurde angenommen, dass noch Peerings nachgewiesen werden können, die von mehr als einer Quelle erkannt werden. Da bereits aber bei zwei unterschiedlichen Quellen kein Peering mehr vorhanden ist, kann man davon ausgehen, dass auch bei weiter zunehmender Quellenzahl die Chance, ein Peering zu entdecken, gegen null geht. Diese These wurde für diese Studienarbeit bestätigt: es wurden bei zwei oder mehr unterschiedlichen Quellen keine Peerings mehr entdeckt.

Dies liegt hauptsächlich daran, dass ein Peering lokal sehr begrenzt, d.h. maximal an die Peeringpartner und deren Customer weitergegeben wird. Das liegt in der Natur des Peerings: die Partner wollen die Daten kostenneutral untereinander austauschen, um so Transit-Gebühren zu sparen. Es würde für sie nur nachteilig sein, die vom Peer gelernten Routen an

den Upstream-Provider weiterzugeben, da unnötiger Traffic entsteht und möglicherweise das Valley-Modell verletzt wird.

## 4.4 Differenztopologien

In diesem Abschnitt werden die Differenztopologien grafisch dargestellt. Dabei handelt es sich immer um einen Vergleich der Daten des BCIX mit je einem der verschiedenen Datenquellen. Dabei gilt {Datenquelle}\{BCIX}, also die Differenztopologie zeigt, welche Peerings zusätzlich zu den BCIX-Daten von der jeweiligen Quelle vorhanden sind. Die Topologie ist begrenzt auf alle AS, die maximal 1 Hop von einem BCIX-Mitglied entfernt sind ('BCIX + 1 Hop'-Topologie).

Die AS des BCIX wurden farblich markiert, um diese besser von den umliegenden AS unterscheiden zu können und so auch Links innerhalb des BCIX verdeutlichen zu können.

### 4.4.1 BCIX - UCLA

In Bild 4.6 ist die Differenz zwischen BCIX und UCLA dargestellt. Dabei handelt es sich also um all die Links, die die UCLA zusätzlich zum BCIX erkennt. Insgesamt werden 152 Links erkannt, davon starten und enden am BCIX jedoch keine. Auf die BCIX + 1 Hop-Topologie reduziert bleiben 15 zusätzliche Links, die hier erkannt werden: Allerdings handelt es sich nur um Upstreams.

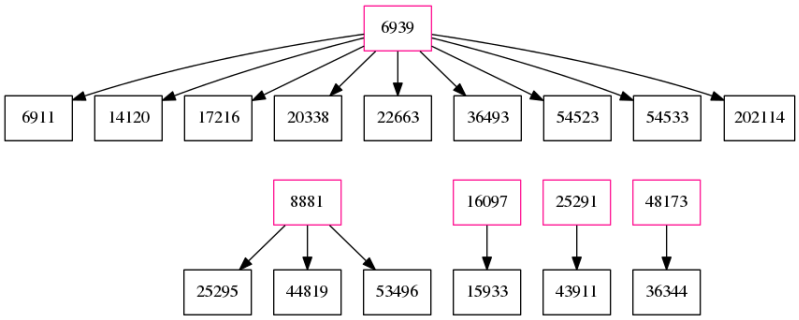


Abbildung 4.6: Differenztopologie BCIX - UCLA

### 4.4.2 BCIX - RIPE

In Bild 4.7 (Seite 22) werden insgesamt 27 zusätzliche Links innerhalb des BCIX erkannt, drei davon starten und enden am BCIX selber. Es wird in diesem Beispiel besonders eindrucksvoll

sichtbar, dass die RIPE in Bezug auf die tatsächlichen Links von Hurricane Electric (AS 6939) nur schlechte Informationen hat. Außerdem ist HE Mitglied des BCIX, daher liegen sehr genaue Informationen über den Ist-Zustand vor.

### 4.4.3 BCIX - RouteViews BGPDATA

Auch bei dem Vergleich der BGPDATA des RouteViews Projekt gibt es insgesamt 16 Links am BCIX, die zusätzlich erkannt werden. Einer davon startet und endet am BCIX, drei weitere starten am BCIX. In Bild 4.8 ist die Differenz zu sehen.

### 4.4.4 BCIX - RouteViews OIX

In Bild 4.9 auf Seite 23 starten insgesamt 31 Links am BCIX und es enden sechs Stück dort. Drei verlaufen innerhalb des BCIX.

Bei allen zusätzlich erkannten Routen handelt es sich um P2C-Beziehungen, d.h. alle zusätzlich erkannten Routen sind Customer Routen aus Sicht des jeweiligen Providers. Dieses Verhalten lässt sich durch verschiedene Policies erklären, bzw. durch multiple Upstream-Provider der entsprechenden Customer (Links innerhalb des BCIX, die gar nicht erkannt wurden). Dabei kommt es auf den Standpunkt des Looking-Glasses in der Topologie an, vereinfacht gesagt gilt: topologisch näher dran und/oder viele Messpunkte = bessere Ergebnisse.

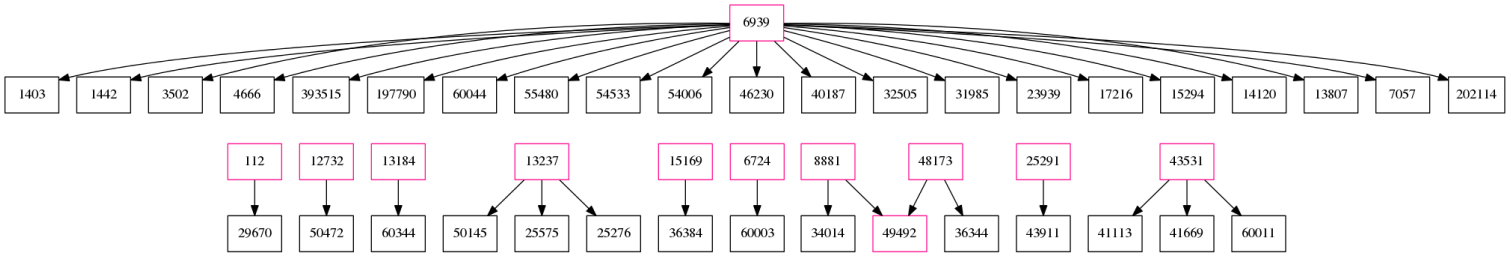


Abbildung 4.7: Differenztopologie BCIX - RIPE

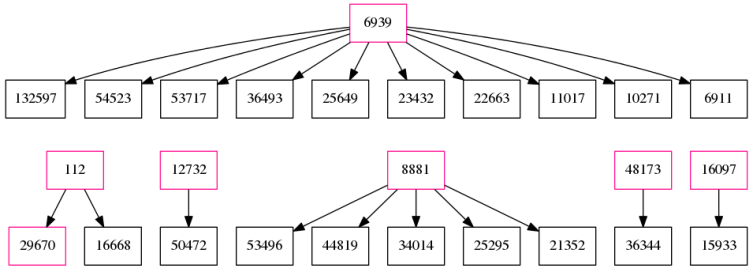


Abbildung 4.8: Differenztopologie BCIX - RouteViews BGPDATA

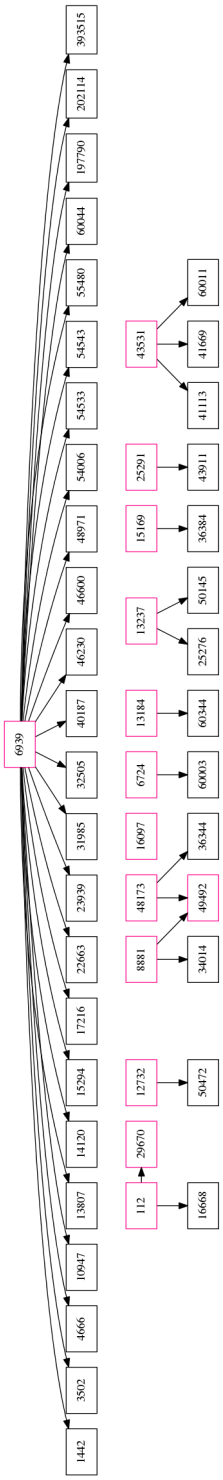


Abbildung 4.9: Differenztopologie BCIX - RouteViews OIX

# 5 Zusammenfassung & Ausblick

In dieser Arbeit wurde ein Vergleich von einer lokal bekannten Topologie um das BCIX in Berlin mit verschiedenen, öffentlich zugänglichen Datenquellen erstellt.

Das Ziel war die Analyse und Quantifizierung der lokalen BCIX-Topologie - also die Frage, ob aus dem Zugriff mitten in der zu untersuchenden Topologie mehr Informationen bringt, als eine umfassende Analyse fremder Quellen.

Da die Anzahl der verwendeten Route-Collectoren der verschiedenen Anbieter stark voneinander abweicht, war zu erwarten, dass mehr verwendete Collectoren auch ein genaueres Bild erreichen. Dies wurde in dieser Arbeit bestätigt, da die UCLA mit insgesamt 133 Collectoren immerhin 234 der Links innerhalb des BCIX erkennen kann, was gut einem Drittel aller vorhandenen Links entspricht.

Die anderen Datenquellen mit deutlich weniger Collectoren im Einsatz, haben somit je ca. 100 Links erkennen können. Es lässt sich somit zusammenfassen, dass die große Anzahl der Collectoren der UCLA für ein möglichst genaues Topologiebild unerlässlich sind. Als einzige Alternative zu einer sehr großen Anzahl von Route-Collektoren ist der Zugriff auf die zu untersuchende Topologie. Nur damit und in unmittelbarer Nähe entsteht ein genaues Bild.

Bei der Klassifikation der einzelnen Quellen zeigte sich abhängig von der Quantität der erkannten Links pro Quelle ein recht genaues Bild: Wurde ein Link nur von einer Quelle erkannt, so war es (bis auf eine Ausnahme) ein Peering. Ab zwei Quellen waren die zusätzlich erkannten Links nur noch gewöhnliche P2C-Upstreams. Damit wurde die zugehörige These in dieser Arbeit bestätigt.

Es wurde gezeigt, dass es immer von Vorteil ist, einen lokalen Zugriff auf die zu untersuchende Topologie zu haben. Falls diese Möglichkeit nicht besteht, so kann mit einer Vielzahl von Route-Collectoren immer noch ein sehr genaues Bild außerhalb der Topologie erstellt werden.

Für die Zukunft bleibt nach Fertigstellung des Projektes mit dem aktiven Messen der Vergleich beider Ergebnisse, um so noch genauere Aussagen treffen zu können. Diese Flussmessung können Hinweise über die in dieser Arbeit nicht erkannten lokalen Routen liefern und auch die Richtung der Peerings und Policies zeigen.

# Literaturverzeichnis


[1] : *Asia Pacific Network Information Centre (APNIC), Usage Statistics on 26th August 2014*. – URL http://thyme.apnic.net/ap-data/2014/08/26/0400/mail-global

[2] CAIDA AS Rank: Information for a single AS: AS Relationship Table. . – URL http://as-rank.caida.org/?mode0=as-info&mode1=as-table&as=

[3] : *Réseaux IP Européens (RIPE)*. – URL http://www.ripe.net/

[4] : *RIPE Routing Information Service (RIS)*. – URL http://www.ripe.net/projects/ris/rawdata.html

[5] AGER, Bernhard ; CHATZIS, Nikolaos ; FELDMANN, Anja ; SARRAR, Nadi ; UHLIG, Steve ; WILLINGER, Walter: Anatomy of a large European IXP. In: *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication* ACM (Veranst.), 2012, S. 163–174

[6] BCIX e.V.: Berlin Commercial Internet Exchange Point, URL http://www.bcix.de

[7] BLUNK, L. ; KARIR, M. ; LABOVITZ, C.: Multi-Threaded Routing Toolkit (MRT) Routing Information Export Format / IETF. October 2011 (6396). – RFC

[8] : *The Cooperative Association for Internet Data Analysis homepage*. http://www.caida.org/home/. 2010

[9] GAO, Lixin: On Inferring Autonomous System Relationships in the Internet. In: *IEEE/ACM Trans. Netw.* 9, Nr. 6, S. 733–745. – URL http://dx.doi.org/10.1109/90.974527. – ISSN 1063-6692

[10] HAWKINSON, John ; BATES, Tony: Guidelines for creation, selection, and registration of an Autonomous System (AS) / IETF. March 1996 (1930). – RFC

[11] HIJACKING, YouTube: *A RIPE NCC RIS case study*. 2008. – URL http://www.ripe.net/internet-coordination/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study

[12] KROHN, Andreas: Aktive Messverfahren zur Topologievalidierung im Routing-Atlas, URL http://users.informatik.haw-hamburg.de/~ubicomp/projekte/master12-13-seminar/krohn/folien.pdf, 2012

[13] LOUGHEED, Kirk ; REKHTER, Yakov: Border Gateway Protocol (BGP) / IETF. June 1990 (1163). – RFC

[14] MATE2CODE, Wikimedia: *Venn Diagram*. 2009. – URL http://de.wikipedia.org/wiki/Datei:Venn0100.svg

[15] MEYER, David u. a.: *University of oregon route views project*. 2005. – URL http://www.routeviews.org

[16] PCH: Packet Clearing House, URL https://www.pch.net

[17] REKHTER, Y. ; LI, T. ; HARES, S.: A Border Gateway Protocol 4 (BGP-4) / IETF. January 2006 (4271). – RFC

[18] TOONK, Andree: Turkey Hijacking IP addresses for popular Global DNS providers. . – URL http://www.bgpmon.net/turkey-hijacking-ip-addresses-for-popular-global-dns-providers

[19] TOONK, Andree: What caused today's Internet hiccup. . – URL http://www.bgpmon.net/what-caused-todays-internet-hiccup

[20] UCLA, IRL: Internet topology collection. . – URL http://irl.cs.ucla.edu/topology

[21] WÄHLISCH, Matthias ; SCHMIDT, Thomas C. ; BRÜN, Markus de ; HÄBERLEN, Thomas: Exposing a Nation-Centric View on the German Internet – A Change in Perspective on the AS Level. In: *Proc. of the 13th Passive and Active Measurement Conference (PAM)* Bd. 7192. Berlin Heidelberg : Springer-Verlag, 2012, S. 200–210

[22] ZHANG, Beichuan ; LIU, Raymond ; MASSEY, Daniel ; ZHANG, Lixia: Collecting the Internet AS-level Topology. In: *SIGCOMM Comput. Commun. Rev.* 35, Nr. 1, S. 53–61. – URL http://doi.acm.org/10.1145/1052812.1052825. – ISSN 0146-4833

[23] ZHANG, Yu: BGP-dump Repository. . – URL https://github.com/YuZhang/bgpdump-zy

[24] ZHANG, Yu: extractor.pl Repository. . – URL https://github.com/YuZhang/topo-lite/blob/master/extractor.pl