INSTITUT FÜR INFORMATIK
DER LUDWIG–MAXIMILIANS–UNIVERSITÄT MÜNCHEN

Bachelorarbeit

Leitfaden für ein sicheres Inter-Domain-Routing am LRZ

Christian Simon

INSTITUT FÜR INFORMATIK
DER LUDWIG–MAXIMILIANS–UNIVERSITÄT MÜNCHEN

Bachelorarbeit

Leitfaden für ein sicheres Inter-Domain-Routing am LRZ

Christian Simon

Aufgabensteller:	PD Dr. Helmut Reiser
Betreuer:	Stefan Metzger Helmut Tröbs Bernhard Schmidt
Abgabetermin:	15. Dezember 2011

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

München, den 15. Dezember 2011

.....
(Christian Simon)

Abstract

Im Rahmen dieser Bachelorarbeit wird ein Leitfaden entwickelt, der die Verbesserung der Sicherheit des Inter-Domain Routing am Leibniz-Rechenzentrum (LRZ) zum Ziel hat. Bei diesem Routingdatenaustausch zwischen verschiedenen Autonomen Systemen (ASen) kommt das Border Gateway Protocol (BGP) zum Einsatz. Anhand einer Betrachtung der entsprechenden Spezifikationen werden daraus folgende Sicherheitsrisiken aufgeführt. Anschließend werden Techniken vorgestellt, die diese Risiken gezielt absichern.
Anhand einer Betrachtung der gegenwärtigen Situation am LRZ wird ein Leitfaden aufgestellt, der die Rahmenbedingungen, die durch das Münchener Wissenschaftsnetz (MWN) vorgegeben sind, berücksichtigt. In einem ersten Schritt wird ermöglicht, dass weitere ASe die Netzankündigungen des LRZ überprüfen können. Um Unregelmäßigkeiten im Routing rechtzeitig erkennen zu können, wird als weitere Maßnahme ein Monitoring der Netze des LRZ empfohlen. Zuletzt wird noch beschrieben, wie eingehende Routen überprüft werden können. Dies ist durch das technische Umfeld am LRZ nur bei IPv6 möglich.

Inhaltsverzeichnis

1 Einführung

In der heutigen Zeit nimmt das Internet als Kommunikationsmedium der Informationsgesellschaft, eine wesentliche Rolle im privaten wie auch im kommerziellen Datenaustausch ein. Diese enorme Bedeutung wird auch durch einen Blick auf die Nutzerzahlen des Internets bestätigt. Im 2. Quartal 2011 nutzten 74% der deutschen Erwachsenen das Internet [FW11]. Nach aktuellen Zahlen von Internet World Stats sind Ende März 2011 bereits 30,2% der Gesamtbevölkerung der Erde regelmäßig im Internet [IWS11].
Diese hohen Nutzerzahlen entstehen durch die vielfältigen Anwendungsfälle, die über das Internet abgewickelt werden können. Im privaten Bereich erfreuen sich vor allem soziale Netzwerke wie Facebook und Twitter steigender Beliebtheit. Zudem können z.B. Bankgeschäfte, Bestellungen, Behördengänge und viele weitere Vorgänge über das Internet zeitsparender und bequemer als über alternative Kommunikationswege erledigt werden. Ein weiterer Einsatzbereich ist die effektive Informationsbeschaffung, denn über Suchmaschinen kann innerhalb kürzester Zeit das Internet nach einer bestimmten Thematik durchsucht werden.
Die Bedeutung verstärkt sich nochmals, wenn man einen Blick auf den Nutzen für die Wirtschaft wirft. Es zeichnet sich dort der Trend ab, die eigene IT Umgebung durch Cloud Computing flexibler zu gestalten. Dadurch müssen die Verbindungen zu dieser Cloud, die in der Regel über das Internet hergestellt werden, zuverlässig funktionieren. Viele Unternehmen setzen auch virtuelle private Netzwerke (VPNs) ein, um etwa räumlich entfernte Unternehmensstandorte zu vernetzen oder um Mitarbeitern, die sich nicht im Unternehmen befinden, Zugriff auf das Intranet zu gewähren. Diese Netze nutzen das Internet um über einen Tunnel die jeweiligen Daten ins unternehmensinterne Netz zu transportieren.
Einer hohen Akzeptanz erfreut sich das Internet auch in der Wissenschaft. Schon lange bevor das Internet Anfang der 1990er für kommerzielle Zwecke geöffnet wurde, nutzten Forschungseinrichtungen das Netz zum weltweiten Austausch ihrer Ergebnisse und zur individuellen Kommunikation unter Wissenschaftlern. Für diese Zwecke wurden Protokolle und Dienste entwickelt, die bis heute in Verwendung sind. So betrieb die NSF (U.S. National Science Foundation) mit dem NSFNET das erste auf TCP/IP basierende WAN. Dienste aus dieser Zeit, die bis heute in Benutzung sind, sind z.B. E-Mail, Newsgruppen, Telnet und FTP[Tan03]. Die vermehrte Internetnutzung lässt sich auch am jährlich steigendem übertragenen Datenvolumen festhalten, denn schon ein Blick in die Statistiken der bayrischen Hochschulen zeigt innerhalb der letzten zehn Jahre ein Wachstum um mindestens den Faktor 40 [LRZ09].
Diese Verbreitung des Internets lässt erahnen, welche weitreichenden Folgen eine großflächige Störung hervorrufen würde. Daher wird dessen Zuverlässigkeit und Robustheit immer wichtiger. Um jedoch Aussagen darüber treffen zu können, muss ein genauerer Blick auf den technischen Unterbau geworfen werden. Im Wesentlichen besteht dieser aus den Protokollen der TCP/IP - Protokollfamilie. Während die Internet Protokolle (IP) der Versionen 4 und 6 sich um die Paketvermittlung kümmern, wird zum Herstellen einer Ende-zu-Ende-Verbindung zumeist auf die Protokolle TCP oder UDP zurückgegriffen. Nach ISO/OSI Architektur entspricht das den Schichten 3 für die Paketvermittlung und 4 für den Transport

in einer Ende-zu-Ende-Verbindung. Darauf setzen wiederum die Anwendungsprotokolle der Schichten 5-7 auf.
Ein solches ist auch das Border Gateway Protokoll (BGP). Es spielt dabei eine besondere Rolle für das Routing im Internet, denn es kommt dann zum Einsatz, wenn eine Kommunikation über das Netz des eigenen Provider hinaus stattfinden soll, dies wird auch als Inter-Domain Routing bezeichnet. Da BGP somit für den Betrieb des Internets in der heutigen Form notwendig ist, stellt jede Sicherheitsschwachstelle im Routingprotokoll BGP auch einen möglichen Angriffspunkt auf das Internet in seiner Gesamtheit dar.

1.1 Motivation

An mehreren realen Vorkommnissen kann man sehen, dass die Schwachstellen im BGP nicht nur von theoretischer Natur sind [Too11]. Grundsätzlich besteht die Schwierigkeit einen Vorfall überhaupt zu erkennen, da er unter Umständen nur wenige Netzbereiche des Internets betrifft und ggf. die Konnektivität gar nicht einschränkt, sondern die Pakete nur über einen anderen Pfad lenkt. Prinzipiell muss man zwischen einem gezielten Angriff auf ein Netz und einer versehentlichen Störung durch einen Konfigurationsfehler unterscheiden.
Ein bekannter Vorfall ereignete sich am 24. Februar 2008. Die pakistanischen Provider wurden von deren Regierung angewiesen, ein bestimmtes Video auf der Videoplattform YouTube.com für eigene Kunden zu zensieren. Durch einen Konfigurationsfehler des Providers Pakistan Telecommunication übernahm dieser fälschlicherweise das Netz von YouTube.com. Dies hatte zur Folge, dass für Internetbenutzer weltweit YouTube nicht erreichbar war. Der Normalzustand der beteiligten Netze ist in Abbildung 1.1 dargestellt. YouTube gibt dabei das Netz `208.65.152.0/22` bekannt und erhält dafür den kompletten Datenverkehr.

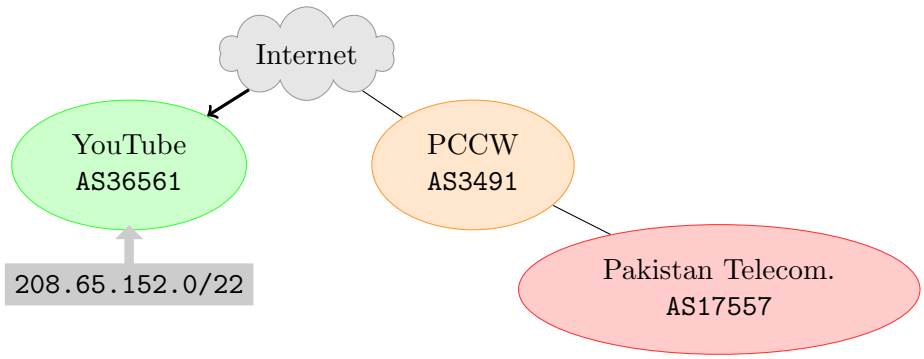

Abbildung 1.1: Normalzustand der beteiligten Provider

Der Provider Pakistan Telecommunication implementierte die geforderte Sperre dadurch, dass er alle Pakete, die an das Subnetz `208.65.153.0/24` gerichtet waren, verwarf. Fälschlicherweise machte er dieses Netz jedoch um 18:47 Uhr (UTC) über ein Announcement bei seinem Transitprovider (vgl. Abschnitt 2.1.1) PCCW bekannt. Da dieser eingehende Anfragen nicht ausreichend prüfte, fiel dort die Unzulässigkeit der Routinginformation nicht auf und diese konnte sich somit über das gesamte Internet weiterverbreiten. Da das von Pakis-

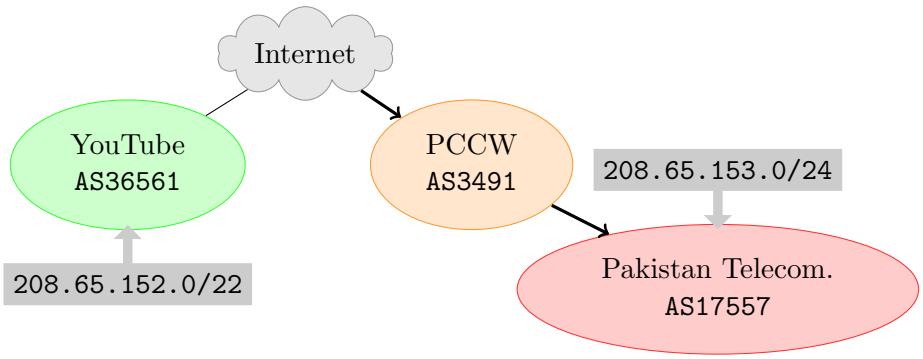

Abbildung 1.2: Phase 1: Pakistan Telefon gibt ein nicht zulässiges Netz bekannt

tan Telecommunication verbreitete Netz spezifischer ist, als das von YouTube, bevorzugten Router nach dem Longest Prefix Match-Prinzip (vgl. Glossar) das falsche Netz. Dadurch landeten alle Anfragen an das entsprechende Netz bei Pakistan Telecommunication, welche, wie in Abbildung 1.2 zu erkennen ist, die ankommenden Pakete verwarf. Für Besucher war YouTube somit nicht mehr zu erreichen.

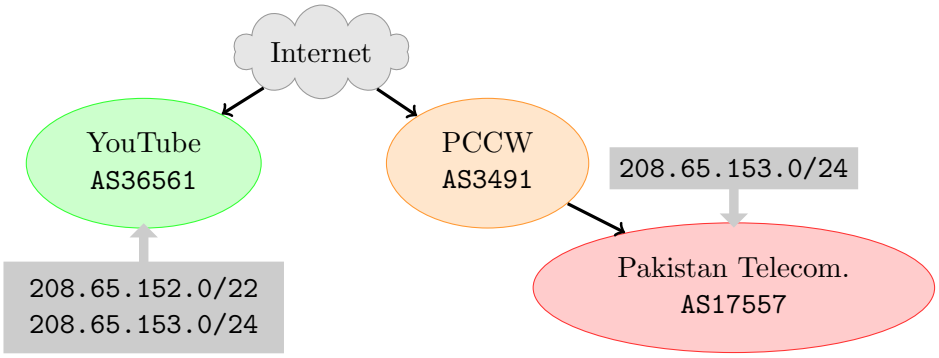

Abbildung 1.3: Phase 2: YouTube gibt das /24 Netz bekannt

YouTube reagierte um 20:07 Uhr (UTC), indem es dasselbe /24-Netz wie Pakistan Telecommunication bekannt gab. Das führte zu dem Effekt, dass die Router den Pfad wählten der kürzer war (vgl. Abbildung 1.3). Somit war für einen Teil der Nutzer die Seite wieder erreichbar, für einen anderen Teil jedoch, v.a. Nutzer die geographisch näher an Pakistan lagen, führte die Route weiterhin ins Leere.

Als YouTube um 20:51 Uhr (UTC) das angegriffene Netz in zwei Subnetze aufgeteilt hatte und diese /25-Netze bekannt gab, wurden alle Pakete wieder zu dem Netzwerk von YouTube geleitet, da diese spezifischeren Netze wiederum nach dem Longest Prefix Match-Prinzip gegenüber dem falschen Netz bevorzugt wurden (vgl. Abbildung 1.4). Um 21:02 Uhr (UTC) filterte PCCW schließlich sämtliche Netzbekanntmachungen von Pakistan Telecommunication und beendete somit den Vorfall [RIP08].

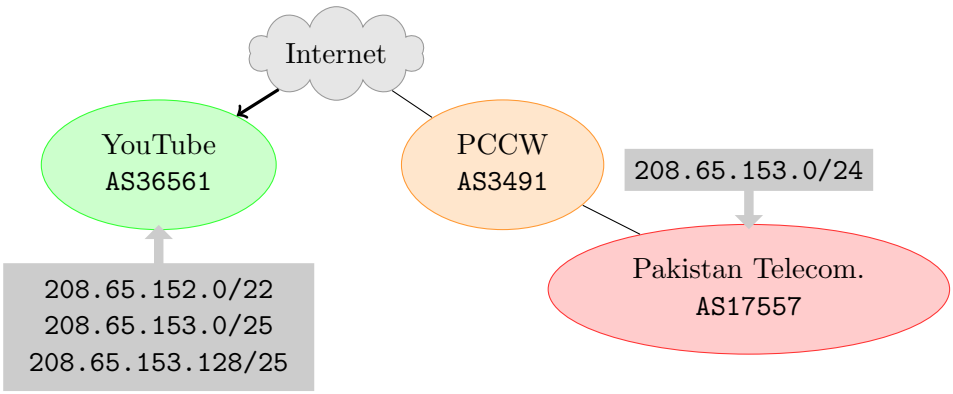

Abbildung 1.4: Phase 3: YouTube gibt zwei /25 Netze bekannt

Dieses Beispiel zeigt, dass sich selbst ein lokaler unbeabsichtigter Fehler beim Sperren eines Netzes global auswirken kann und so Teile des Internets unerreichbar werden. Begünstigt wurde dies dadurch, dass PCCW die eingehenden Netze ungeprüft übernommen hatte. Es ist daher nötig Erweiterungen für BGP auf den Weg zu bringen, so dass es resistenter gegenüber solchen Konfigurationsfehlern oder gezielten Angriffen wird. Denn in der jetzigen Situation kann eine berechtigte Netzankündigung nicht von einer unberechtigten unterschieden werden. Es ist nötig dass jeder Provider die eingehenden Netzankündigungen möglichst restriktiv filtert. Dadurch kann jedoch die Korrektheit der Routinginformationen schon durch wenige Provider, die nicht entsprechend filtern, gefährdet werden. Daher sollte die Korrektheit von Routinginformationen durch geeignete kryptographische Verfahren nachweisbar sein.

1.2 Zielsetzung

Diese Bachelorarbeit soll einen Leitfaden zur Verbesserung des externen Routings entwickeln, der die Bedürfnisse des Leibniz-Rechenzentrums (LRZ) speziell betrachtet und die Besonderheiten in dessen Infrastruktur berücksichtigt. Es sollen die Vor- und Nachteile einzelner Maßnahmen erläutert und gegenübergestellt werden. Geeignete Erweiterungen sollen in die lokale Routingpolicy einfließen, damit diese zuverlässig Anomalien im Routing erkennen und umgehen kann. Dadurch sollen Auswirkungen auf das gesamte Internet möglichst verhindert werden. Als praktische Aufgabe werden schließlich noch ausgewählte Vorkehrungen prototypisch implementiert. Es soll dabei insbesondere die Tabelle für das IPv6 Routing betrachtet werden, da das LRZ dort im Gegensatz zur Default Route bei IPv4 über eine vollständige Routingtabelle verfügt.

1.3 Überblick

In Kapitel 2 wird der technische Unterbau des Internet allgemein behandelt. Es wird darin auf dessen Struktur eingegangen und die verwendeten Protokolle kurz eingeführt. Dabei werden die Routingprotokolle speziell betrachtet. Schließlich wird der Fokus auf die Verwaltungsebene des Internets gelegt.
Das Kapitel 3 beschäftigt sich detailliert mit BGP. Es wird zuerst anhand entsprechender

RFC Spezifikationen eingeführt und somit dessen Verhalten aufgezeigt. Im abschließenden Abschnitt wird auf die potentiellen Schwächen von BGP eingegangen.
Mögliche Gegenmaßnahmen, um diese Schwächen zu beseitigen, werden in Kapitel 4 behandelt. Für den einfacheren Überblick erfolgt eine Gegenüberstellung von den zuvor erläuterten Sicherheitsproblemen und möglichen Lösungsansätzen.
In Kapitel 5 wird die gegenwärtige Situation am LRZ beschrieben. Ausgehend von diesen Rahmenbedingungen wurde ein Leitfaden zur Verbesserung der Sicherheit des Inter-Domain Routing entwickelt, der in Kapitel 6 vorgestellt wird. Es werden für das LRZ geeignete Maßnahmen ausgewählt und Empfehlungen gegeben, wie diese eingesetzt werden können. Diese werden schließlich in Kapitel 7 prototypisch implementiert.
In Kapitel 8 werden die Ergebnisse der Arbeit zusammengefasst und es wird ein Ausblick gewagt, welche Technologien zukünftig die Sicherheit von BGP verbessern könnten.

2 Grundlagen

Dieses Kapitel befasst sich mit dem Internet an sich, das auch oftmals als Netz der Netze bezeichnet wird. Dass diese Bezeichnung genau die Struktur des heutigen Internets beschreibt, wird nach Abschnitt 2.1 deutlich. Daraufhin werden die dem Internet zugrunde liegenden Protokolle der TCP/IP Familie genauer erläutert. Diese spielen bei den späteren Untersuchungen eine wichtige Rolle und werden daher kurz eingeführt. Schließlich wird ein kurzer Einblick in die Verwaltung des Internets durch die ICANN geworfen. Damit die Zuständigkeiten bei der Vergabe und Verwaltung von Internet Ressourcen klarer werden.

2.1 Netzstruktur des Internets

Das Internet besteht im Wesentlichen aus dem Zusammenschluss der Netze einer Vielzahl von Netzbetreibern. Mehrere Netze, die untereinander eine einheitliche Routingpolicy einsetzen, können zu einem Autonomen System (AS) zusammengefasst werden. Alle Router eines ASes tauschen über interne Routing Protokolle (vgl. Abschnitt 2.2.3) die Erreichbarkeiten von Zielen innerhalb dieses Verwaltungsbereichs aus[HB96]. Solche ASe werden in der Regel von Internet Service Provider (ISP), größeren Unternehmen oder auch von wissenschaftlichen Einrichtungen betrieben. Daher kann sich die geographische Ausdehnung sehr unterscheiden. So erstreckt sich das AS der Deutschen Telekom über weite Teile Europas und wird dabei durch mehrfache Anbindungen nach Nordamerika und Asien ergänzt. Das AS des LRZ ist dagegen bis auf einige wenige Institute auf den Münchner Großraum beschränkt [Tel11] [LRZ11].
Verfügen zwei Router über eine Verbindung untereinander, so wird diese als Intra-AS bezeichnet, wenn beide Router dem gleichen AS angehören. Unterscheiden sich dagegen die ASe der beiden Router so spricht man von einer Inter-AS-Verbindung. Darauf aufbauend können Router in zwei Klassen eingeteilt werden: Wenn ein Router ausschließlich über Intra-AS-Verbindungen verfügt, wird er auch Inner Router genannt. Besitzt ein Router dagegen mindestens eine Inter-AS Verbindung, so wird er als der Edge Router bezeichnet (vgl. dazu Abbildung 2.1). Verschiedene ASe tauschen also über Inter-AS-Verbindungen Routinginformationen und Datenverkehr aus. Da ein AS auch als Verwaltungsdomain eines Betreibers gesehen werden kann, bezeichnet man diesen Austausch auch als Inter-Domain-Routing. Dabei unterscheidet man grundsätzlich zwei Varianten: Transit und Peering.

2.1.1 Transit

Bei einer Transitverbindung zwischen zwei ASen tritt ein AS, welches typischerweise das größere ist, als Provider, das andere als Kunde auf. Dabei ermöglicht der Provider dem Kunden, gegen Zahlung entsprechender Gebühren, den kompletten Datenverkehr über ihn abzuwickeln[vdB08]. Der Kunde erhält wahlweise eine Default Route oder die komplette Routingtabelle des Providers. Ein AS das zu mindestens zwei Providern Transitverbindungen

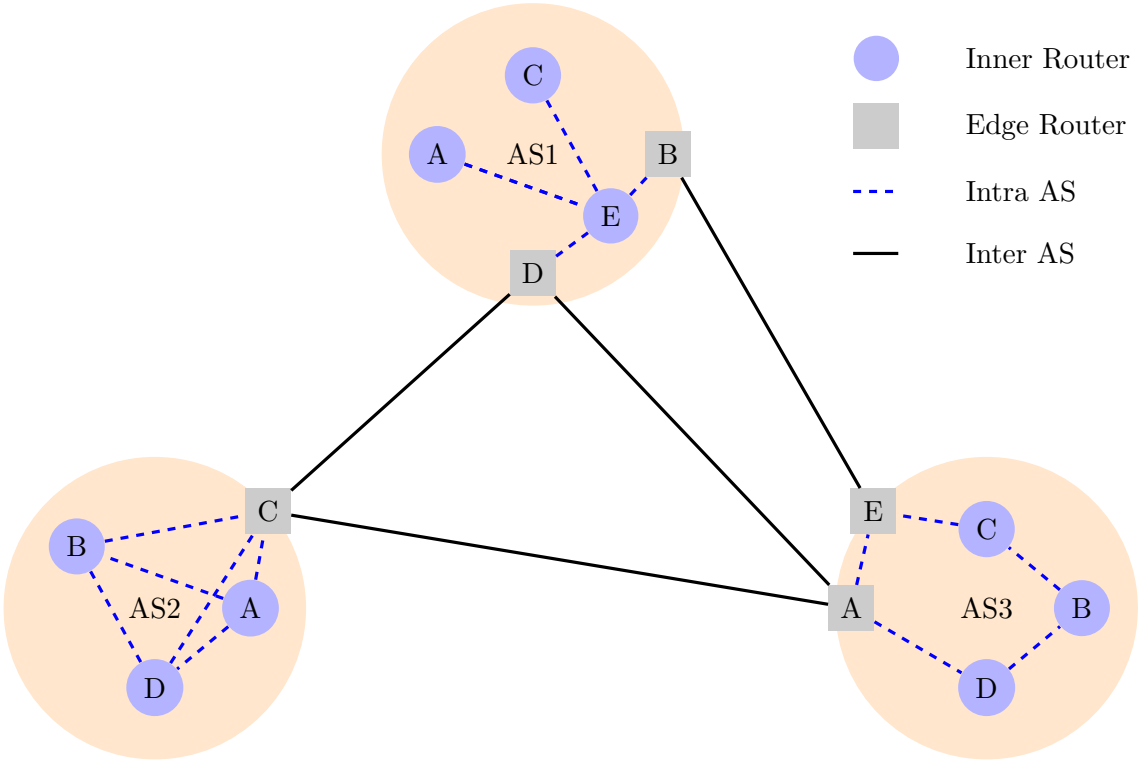

Abbildung 2.1: Beispieltopologie zur Unterscheidung Intra AS vs. Inter AS

hält, wird als Multihomed AS bezeichnet. Dies sorgt für Redundanz bei der Internetanbindung, denn bei Ausfall eines Providers ist die Erreichbarkeit über die zusätzliche Anbindung gesichert. Da der Kunde Einfluss darauf hat, über welchen Transit Provider er welche Zielnetze kontaktiert, kann der Datenverkehr über beide Provider verteilt werden (Load Balancing).
Allerdings hat Multihoming negative Auswirkungen: Der Kunde hat grundsätzlich höhere Kosten, denn es sind mindestens zwei Transitverbindungen nötig. Weiterhin wird nun der Einsatz von BGP in diesem AS nötig, da es nicht mehr ausreicht den Datenverkehr über eine Default Route schlicht an einen Provider weiterzuleiten. Er benötigt eine eigene lokale Routingpolicy um zu entscheiden, welche Anbindung für welche Fälle den Vorzug hat. Ein weiterer Nachteil, der nicht den Kunden selbst betrifft, sondern sich auf das gesamte Internet auswirkt, ist, dass durch Multihoming eine große Anzahl relativ kleiner Netze in die Routingtabellen weltweit gelangt und somit die Anforderungen an die Router erhöht.[Mü10]

2.1.2 Peering

Neben dem Transit gibt es noch die Möglichkeit ein Peering zwischen zwei oder mehreren ASen auszuhandeln. Diese Art ist in der Regel für die beteiligten ASe (Peers) kostenfrei, da die Peers eine ähnliche Größe haben und sich so deren Betreiber durch das Peering Vorteile erhoffen. Denn durch geeignete Peerings kann der Datenverkehr an Transitprovider gesenkt werden, da die Partner Datenverkehr zwischen ihren beiden ASen direkt abwickeln. Außerdem kann eine direkte Verbindung zweier Netze sowohl geringe Latenzen als auch höhere Bandbreiten ermöglichen, im Vergleich zum Weg über Transit-ASe. Deshalb versucht

ein Netzbetreiber stets den meisten Traffic über Peerings abzuwickeln. Der Datenaustausch zwischen den ASen wird typischerweise an Internetknoten (wie z.B. DE-CIX, AMS-IX) vollzogen, denn dort sind viele ASe mit einem Point of Presence (PoP) vertreten.[vdB08]

2.1.3 Einordnung der Provider

Grundsätzlich lassen sich alle Provider in eine Hierarchie einordnen. Ein Netzbetreiber, dessen AS ausschließlich über Peeringabkommen verfügt und er darüber das gesamte Internet erreicht, wird auch als Tier-1-Provider bezeichnet. So ein Provider verfügt über ein interkontinental ausgebautes Netz, welches ihm ermöglicht hohe Bandbreiten über weite Strecken zu transportieren. Weltweit beschränkt sich die Zahl der Tier-1-Provider auf gut ein Dutzend. Tier-2-Netzbetreiber dagegen definieren sich dadurch, dass sie mindestens eine Transitverbindung mit einem Tier-1-Provider benötigen, um das gesamte Internet zu erreichen. Tier-2 Provider sind meist nur regional bzw. national tätig. Ein Tier-3-Provider dagegen betreibt nur Transitverbindungen zu Tier-2 bzw. Tier-3-Providern, jedoch nicht zu Tier-1-Providern. Zusätzlich sorgen bei Tier-2- bzw. Tier-3-Providern Peerings mit ähnlich gearteten Partner-ASen für direkten Austausch untereinander [Bei02]. Ein kurzer Auszug über ausgewählte ASe im Umfeld des Leibniz Rechenzentrums sowie deren Einordnung in die Providerhierarchie sind in Abbildung 2.2 dargestellt.

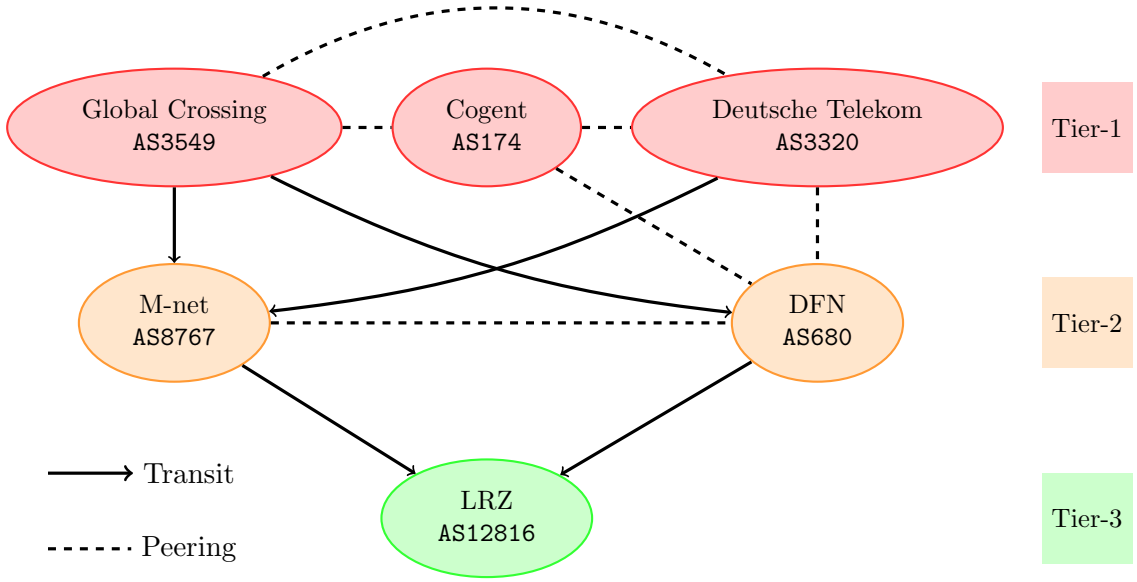

Abbildung 2.2: Einige ASe und deren Beziehungen zueinander

2.2 Internetprotokollfamilie TCP/IP

Damit die Kommunikation im Internet geregelt abläuft, sind verbindliche Protokolle nötig. Diese werden in der Internetprotokollfamilie TCP/IP zusammengefasst. Im Folgenden sol-

len die zwei für die weitere Betrachtung des Themas relevanten Protokolle kurz eingeführt werden. Es handelt sich dabei um die namensgebenden Protokolle Internet Protokoll (IP), zuständig für die Vermittlung und das Transmission Control Protocol (TCP), zuständig für den Datentransport.

2.2.1 Internet Protokoll

Das Internet Protokoll ermöglicht den Paketaustausch zwischen zwei Endsystemen im Internet. Das Internet Protokoll ist im ISO/OSI Modell Schicht 3 zuzuordnen und stellt somit die erste Schicht dar, die komplett unabhängig von der physikalischen Verbindung ist. Da die Systeme zumeist nicht über eine direkte Verbindung untereinander verfügen, werden die Pakete von entsprechenden Transitsystemen, die als Router bezeichnet werden, weitergeleitet. Ein IP Paket setzt sich aus zwei Teilen zusammen: einem Header (vgl. Abbildung 2.3 bzw. 2.4), der die für die Übertragung nötigen Informationen beinhaltet, gefolgt von einem Payload. Quell- und Zieladresse sind obligatorische Bestandteile des Headers, daher besitzt jede Netzwerkschnittstelle eines Systems eine eigene sogenannte IP-Adresse. Damit ein System weiß, über welche nächsten Stationen ein Paket mit bestimmten Ziel geschickt werden muss, ist eine Routingtabelle von Nöten. Im Internet finden heute die Versionen 4 und 6 des Internet Protokolls Verwendung, deren Unterschiede kurz erläutert werden.

2.2.1.1 IPv4

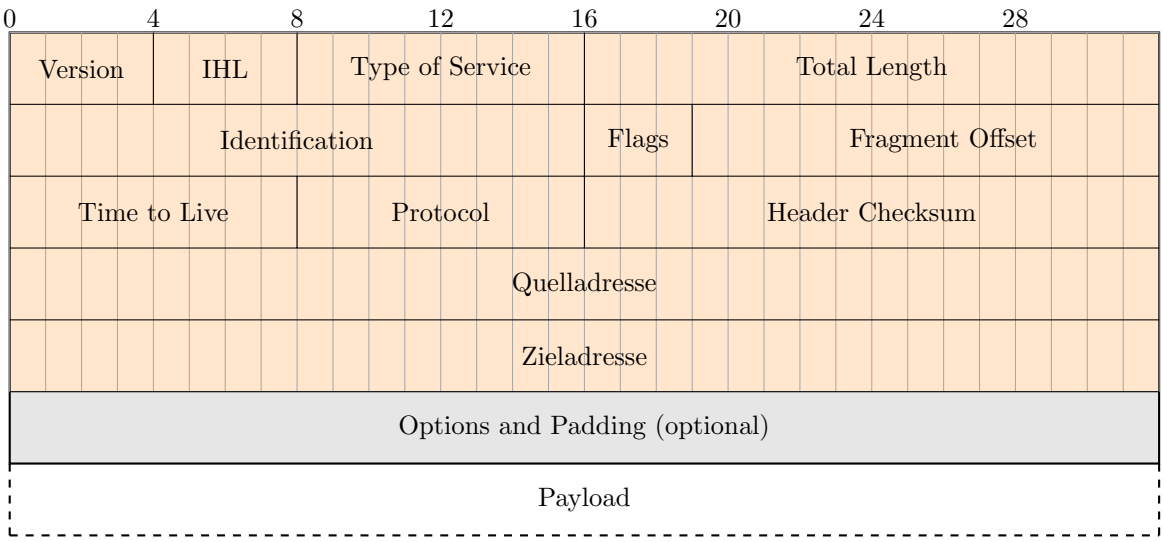

Abbildung 2.3: IPv4 Header mit anschließendem Payload

In der Version 4, wie in RFC 791 definiert, werden zur Adressierung der Hosts Adressen mit einer Länge von 32 Bit verwendet [Pos81a]. Die Adressen werden meist in der gepunkteten Dezimaldarstellung visualisiert. Dabei wird die Adresse in vier Octets aufgeteilt und jeweils deren Dezimalwert getrennt durch je einen Punkt angegeben. Die Adresse selbst kann in einen Netzteil und einen Hostteil getrennt werden. Wo sich die Grenze befindet, wird durch das Netzsuffix angeben. In der Classless Inter-Domain Routing Notation (CIDR)

wird dieses Suffix hinter der Adresse aufgeführt[FLYV93]. Die Angabe `192.168.1.200/24` hat beispielsweise ein Netzsuffix von 24. Der Netzteil nimmt also die ersten 24 Bit ein, somit bleiben für den Hostteil die restlichen acht Bits. Als Faustregel gilt: Je kleiner das Suffix ist, desto größer ist das jeweilige Netz. Insbesondere bezeichnet `/0` den Gesamtadressraum und `/32` einen einzelnen Host.
Der IPv4 Header ist in der Regel 20 Byte lang, kann aber durch optionale Parameter um 40 Byte verlängert werden. Dessen detaillierter Aufbau ist in Abbildung 2.3 dargestellt. Neben den Quell- und Zieladressen findet man noch weitere Informationen, die der Paketverarbeitung dienen, im Header. Im Feld Header Checksum wird eine Prüfsumme des Headers aufgeführt. Erwähnenswert ist auch noch das Feld Time-to-Live (TTL), das an jedem Router um 1 dekrementiert wird. Sobald das Feld den Wert 0 erreicht, wird das entsprechende Paket verworfen. Dies sorgt dafür, dass Pakete nicht endlos in Routingschleifen kreisen können. Im Feld Protocol wird der Typ des Nutzdatenprotokolls angegeben. Der Bereich Options and Padding wurde bei der Konzeption von IPv4 für etwaige Erweiterungen vorgesehen. Er wird im praktischen Einsatz jedoch nicht benutzt.
Durch das rasante Wachstum des Internets stehen in einigen Regionen der Welt schon heute keine neuen IPv4 Adressen mehr zur Verfügung. Weltweit betrachtet wird aktuellen Schätzungen nach spätestens im Jahr 2014 der letzte IPv4-Adressbereich durch eine Regional Internet Registry (RIR, vgl. 2.3) zugeteilt[Hus11]. Ein weiteres Problem an dem zu kleinen Adressraum besteht darin, dass jedes AS anstatt eines einzigen größeren mehrere kleine Adressblöcke zugewiesen bekommt. Dies führt zu einer unnötigen Vergrößerung der Routingtabellen. Um diese Probleme zu beseitigen, musste ein Nachfolgeprotokoll mit größerem Adressraum entwickelt werden.

2.2.1.2 IPv6

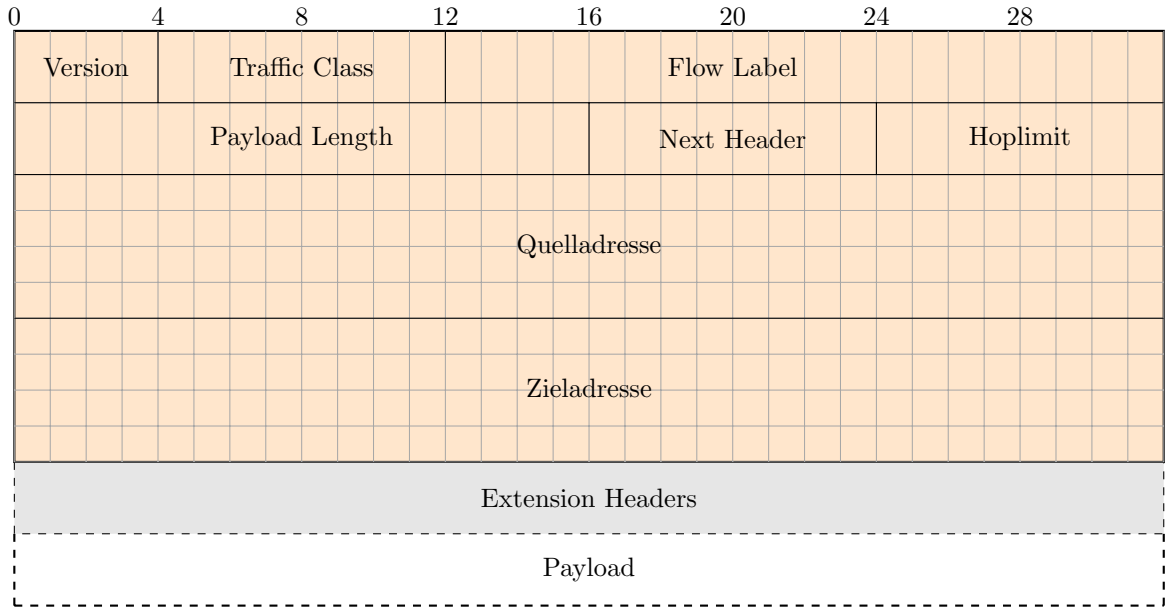

Abbildung 2.4: IPv6 Header mit anschließendem Payload

Im Jahr 1995 wurde ein solcher Nachfolger unter dem Namen IPv6 in RFC 1883 spezifiziert. Die Versionsnummer 6 ergibt sich dadurch, dass als Version 5 bereits das Stream Protokoll der Version 2 (ST2) in RFC 1819 spezifiziert wurde. Da ST2 jedoch die experimentelle Phase nie beendet hat, gilt IPv6 als Nachfolger (aktuellste Spezifikation in RFC 2460). Es wurden viele Veränderungen in dieser neuen Version implementiert, die durch die langjährigen Erfahrungen mit dem Vorgänger als sinnvoll erachtet wurden.
Um bei IPv6 einer künftigen Adressknappheit vorzubeugen, wurde der Adressraum mit 128 Bit Adresslänge, deutlich vergrößert. Als Standardnotation für Adressen wurde eine hexadezimale Darstellung ausgewählt, die in 16-Bit-Blöcken, jeweils getrennt durch einen Doppelpunkt, notiert wird. Zur Verkürzung können führende Nullen weggelassen werden. Als zusätzliche Verkürzung darf einmal je Adresse eine Folge von Blöcken mit Wert 0 durch zwei Doppelpunkte (`::`) ersetzt werden. Die Angabe der Netzgröße erfolgt bei IPv6 analog zu IPv4 in der CIDR Notation. Zudem wurde der IPv6-Header aufgeräumt, er hat nun eine feste Länge von 40 Byte. Er kann durch Extension Header bei Bedarf erweitert werden, die sich zwischen dem Header und der Payload des Pakets befinden. Dadurch, dass die heutigen Netze, für welche IPv6 entwickelt wurde, zuverlässiger sind als diejenigen zu Zeiten der IPv4 Entwicklung, wird auf die Header Checksum verzichtet. Diese musste, da sich die TTL an jedem Router ändert, jeweils neu berechnet werden. Dadurch wird deren CPU entlastet. Durch Umbenennung von TTL in Hoplimit und von Protocol in Next Header soll sich der Zweck, der sich seit Version 4 nicht geändert hat, leichter erschließen lassen.[DH98]
Der großzügige Adressraum ermöglicht eine Reduzierung der Routingtabellen im Vergleich zu IPv4. Denn dadurch kann jedes AS ein ausreichend großes Netz erhalten, so dass weitere Zuweisungen möglichst verhindert werden. Somit kann die Zahl der Netze, die je AS in die globalen Routingtabellen wandern, gering gehalten werden. Dies ist ein essentieller Faktor, wenn die Skalierung des Internets verbessert werden soll. [WDMC07]

2.2.2 Transmission Control Protocol

Auf der Grundlage, die das Internet Protokoll schafft, setzen die Protokolle der Transportschicht auf. Für diese Arbeit ist dabei das Transmission Control Protocol (TCP) maßgeblich. Ein TCP Paket, wie in Abbildung 2.5 dargestellt, unterteilt sich in einen Header, der die Verbindungsinformationen, und einen Nutzdatenteil, der den zu übertragenden Bitstrom beinhaltet. Es ermöglicht eine verbindungsorientierte, bidirektionale, und zuverlässige Kommunikation zwischen zwei Endsystemen. Diese Eigenschaften werden nun kurz erläutert.
Um die Zuverlässigkeit der Paketübermittlung sicherzustellen, werden gesendete Pakete von der Gegenstelle nach erfolgreicher Übertragung mit einem `ACK` bestätigt. Die Integrität des Pakets kann anhand einer Prüfsumme über deren Inhalt gewährleistet werden. Sollte der Absender kein `ACK` erhalten, wird das Paket nach Ablauf einer gewissen Zeitspanne (Retransmission Timeout) erneut versendet [Hol02].

Zu Beginn eines Datenaustausches steht der Verbindungsaufbau. Dieser erfolgt bei TCP durch einen Three Way Handshake, im Zuge dessen werden nötige Verbindungsinformationen ausgetauscht. In Abbildung 2.6 wird der Ablauf eines Verbindungsaufbaus von Host A zu Host B skizziert. Das erste Paket von A zu B erhält eine `SYN` Kennzeichnung (sog. Flag) im Header, die darauf hinweist, dass es sich bei diesem Paket um eine Verbindungsanforderung handelt. Zudem generiert A eine Sequenznummer X, die ebenfalls im Header mit übertragen wird. Jetzt wird das Paket von B empfangen, dieser will die Verbindung annehmen. Dazu sen-

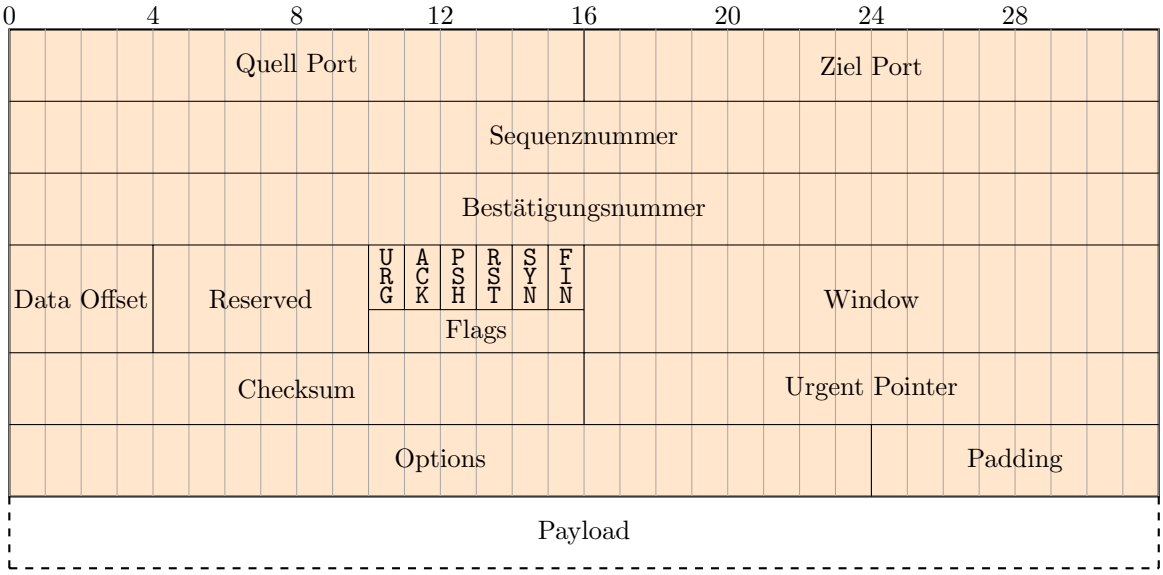

Abbildung 2.5: TCP Header mit anschließendem Payload

det er ein Paket zurück, welches sowohl mit `ACK` als auch mit `SYN` gekennzeichnet ist. Damit bestätigt er den Empfang des ersten Pakets (`ACK`) und stimmt der Verbindungsanforderung zu (`SYN`). Damit A weiß, welches Paket von B bestätigt wird, gibt B die Sequenznummer X des ersten Pakets von A addiert mit 1 als Bestätigungsnummer an. Auch B generiert für dieses Paket eine zufällige Sequenznummer Y. Sobald A nun die Antwort von B empfängt, bestätigt A mit einem ACK Paket. Dieses trägt als Sequenznummer die Bestätigungsnummer des vorhergehenden Antwortpakets, also X+1 und als Bestätigungsnummer die letzte von B empfangene Sequenznummer Y, wiederum addiert mit 1.
Die Verbindung ist nun aufgebaut, somit ist jetzt eine Übertragung von Nutzdaten möglich. Dabei wird je Paket die Sequenznummer immer um die Länge der Payload erhöht. Dies hält die Reihenfolge der Pakete bei der Versendung fest, so dass sie beim Empfänger wieder in der richtigen Reihenfolge zusammengesetzt werden können. Ein `ACK` für eine Sequenznummer wird erst versendet, wenn alle vorhergehenden Pakete fehlerfrei empfangen wurden.[Pos81b]
Der Verbindungsabbau erfolgt wiederum mit einem eigenen Kennzeichen `FIN` im Header. Sendet A ein solches Paket zu B, erklärt A gegenüber B, dass von A keinerlei Datenpakete mehr folgen werden. Dies bestätigt B mit einem `ACK`. B kann nun noch verbleibende Pakete übertragen und dann schließlich auch ein `FIN` Paket senden. Nachdem dieses von A noch bestätigt worden ist, ist die Verbindung abgebaut.[Tan03]
Wird ein Paket mit der `RST` Kennzeichnung empfangen, muss die Verbindung sofort abgebrochen werden. Solche Pakete werden unter anderem versandt, wenn technische Probleme auftreten oder ein Verbindungsaufbau nicht erwünscht ist.

2.2.3 Routingprotokolle

Nachdem die grundlegenden Protokolle IP und TCP eingeführt sind, werden nun die Routingprotokolle betrachtet. Aus Abschnitt 2.2.1 ist bekannt, dass jedes über IP kommunizierende System eine Routingtabelle besitzt. Während in einem kleinen lokalen Netzwerk

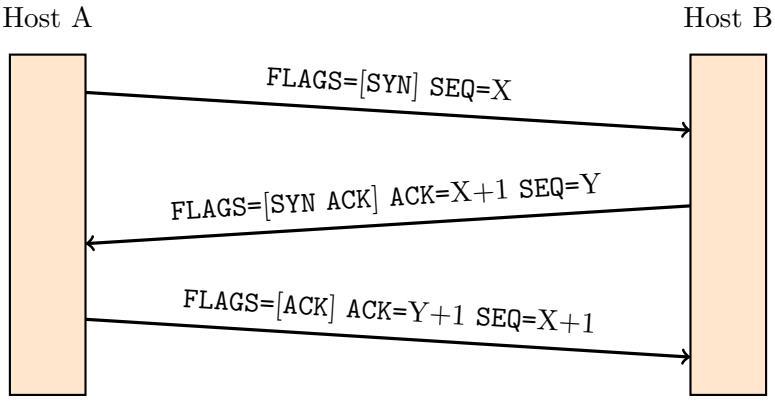

Abbildung 2.6: TCP Verbindungsaufbau von Host A zu Host B: Three Way Handshake

diese Tabelle ohne großen Aufwand statisch eingerichtet werden kann, ist es bei großen und dynamischen Netzen, wie dem Internet, nötig diese Tabellen über Routingprotokolle aktuell zu halten.
Je nachdem, ob es sich um Netze aus dem Intra-AS oder dem Inter-AS Bereich (vgl. Abschnitt 2.1) handelt, unterscheidet man zwischen zwei Einsatzbereichen von Routingprotokollen: Für Routen innerhalb des eigenen ASes (Intra-AS) werden Interior Gateway Protocols (IGPs) und für die darüber hinaus (Inter-AS) Exterior Gateway Protocols (EGPs) verwendet.
Es existieren bei den IGPs zwei technische Ansätze: Link State Protokolle und Distanzvektorprotokolle. Eine spezielle Form von Distanzvektorprotokollen stellen Pfadvektorprotokolle dar. Diese Varianten werden in den folgenden Abschnitten kurz beschrieben.

2.2.3.1 Link State Protokolle

Bei den Link State Protokollen kennen alle Router die komplette Netzwerktopologie. Dazu ist es nötig, dass jeder Router eine Liste mit allen direkten Verbindungen, die er zu anderen Router hält, erstellt. Zu dieser Liste fügt er die Netzwerke hinzu, die er direkt erreicht. Diese Liste sendet der jeweilige Router an alle seine direkten Nachbarn. Diese wiederum senden die Liste weiter, so dass jeder Router sie erhält. Dieser Vorgang wird auch als Flooding bezeichnet. Sollte sich an den Daten etwas ändern, z.B. der Ausfall einer Verbindung, wird dieses Ereignis über dasselbe Verfahren allen Routern mitgeteilt. Durch diese Form der Informationsweitergabe vergeht nach einer Änderung nur eine sehr kurze Zeitspanne bis alle Router dieselbe Netzwerktopologie vorliegen haben. Diese Zeitspanne wird auch als Konvergenzzeit bezeichnet.
Aus der Gesamttopologie muss nun der kürzeste Weg zu einem gewissen Ziel ermittelt werden. Bei Open Shortest Path First (OSPF), welches eine verbreitete Implementierung eines Link State Protokolls ist, geschieht dies anhand des Dijkstra-Algorithmus [Dom02].
Ein Einsatz von Link State Protokollen als EGP wird dadurch verhindert, dass es nötig ist in jedem Router die gesamte Topologie zu speichern. Durch die Größe des Internet ist dies schlicht nicht möglich.

2.2.3.2 Distanzvektorprotokolle

Einen anderen Weg gehen die Distanzvektorprotokolle. Sie betrachten je Zielroute (Vektor) die entsprechende Entfernung (Distanz), dabei wird die jeweils geringste Entfernung bevorzugt. Mögliche Entfernungen ergeben sich aus der Summe der Kantengewichte, der Anzahl der Router auf dem jeweiligen Pfad (Hops) oder auch der Latenz zwischen den Systemen. Wir wollen nun das Routing Information Protocol (RIP) betrachten, es setzt als Maß der Entfernung auf die Anzahl der Hops. Zu Beginn besitzt ein Router nur die Routen seiner direkt verbundenen Netze, diese haben dadurch eine Distanz von 0 Hops. In regelmäßigen Abständen, bei RIP alle 30 Sekunden, teilt er nun seinen Nachbarn die über ihn erreichbaren Netze samt deren Distanzen mit. Gleichzeitig ist er empfangsbereit für die Mitteilungen seiner Nachbarn. Wenn er nun eine neue Route oder eine mit niedrigerer Distanz empfängt, nimmt er diese in seine Routingtabelle auf. Neue Routinginformationen verteilen sich deshalb rasch auf den teilnehmenden Routern. Fällt dagegen ein Router aus und sorgt damit für eine Unerreichbarkeit eines Teilnetzes, dauert es lange bis die Route von allen Tabellen der Router entfernt wurde und die Tabellen aller Router konvergent sind. Beim Hinzufügen von Routen ist also die Konvergenzzeit wesentlich geringer als beim Entfernen. Ein weiteres Problem ist, dass es keine Erkennung von Routingschleifen, die durch spezielle Topologien auftreten können, gibt.

2.2.3.3 Pfadvektorprotokolle

Pfadvektorprotokolle versuchen diese Schwächen der Distanzvektorprotokolle zu umgehen, indem sie zusätzlich zur Zieladresse und dem Next-Hop, den gesamten Pfad zum Zielsystem austauschen. Somit können unter anderem Routingschleifen effektiv verhindert werden, da lediglich der Pfad daraufhin untersucht werden muss, ob das eigene System darin vorkommt. Details dazu finden sich in Kapitel 3, in dem das Pfadvektorprotokoll Border Gateway Protocol (BGP) ausführlich behandelt wird.

2.3 Verwaltung des Internets

Neben der Spezifikation der technischen Grundlagen ist für den Datenaustausch eine Instanz nötig, die sich um die Belange der Internetverwaltung kümmert. Dies wird durch die gemeinnützige Organisation Internet Corporation for Assigned Names and Numbers (ICANN) verwirklicht. Gegründet wurde die ICANN im Jahre 1998, nachdem sich zuvor die U.S. Regierung um deren Aufgaben gekümmert hatte. Noch bis 2009 war die ICANN direkt dem Wirtschaftsministerium der USA unterstellt, nun ist sie jedoch von diesem unabhängig. Dadurch, dass ihr Sitz in den USA liegt, ist sie aber weiterhin den dortigen Gesetzen unterstellt. Zu den Hauptaufgaben der ICANN gehören laut deren Statuten die eindeutige Zuweisung und Delegierung von IP Adressen und AS Nummern [ICA11]. Diese Verwaltung wird durch die Internet Assigned Numbers Authority (IANA) vorgenommen. Die IANA ist eine Unterabteilung der ICANN.

Vergabe von IP Adressen und AS Nummern

Bei der Verteilung der Ressourcen geht die IANA hierarchisch vor. Zuerst wird der gesamte Bereich in relativ große Blöcke aufgeteilt. Diese Blöcke werden nun den Regional Internet

Registries (RIR) zugeteilt. Es gibt insgesamt fünf RIRs, wobei diese jeweils für gewisse Regionen zuständig sind (vgl. Tabelle 2.1).

RIR	Region
AfriNIC	Afrika
APNIC	Asien, Pazifik
ARIN	Nordamerika
RIPE NCC	Europa
LACNIC	Lateinamerika, Karibik

Tabelle 2.1: Überblick über die Zuständigkeiten der 5 RIRs

Die RIRs teilen nun Bereiche aus ihren Ressourcen an ihre Mitglieder, den Local Internet Registries (LIR), zu. Diese LIRs sind meist Provider, die schließlich einzelne Adressbereiche an ihre Kunden weitergeben. Diese Vergabehierarchie wird in Abbildung 2.7 noch verdeutlicht. Dies ist das Standardverfahren zur IP Adresszuteilung von Provider Aggregated (PA) Space. Dass heißt also, dass der Endkunde ein Netz aus dem größeren Netz seines LIRs erhält. AS Nummern und Provider Independent (PI) Space Adressen werden an den Endkunden direkt von den jeweiligen RIRs vergeben.
Eine Besonderheit stellt der Legacy Space dar. Dieser enthält Adressbereiche, die bereits zugeteilt worden sind, bevor die heutige Vergabestruktur existierte. So haben Endkunden die Adressen direkt zugewiesen bekommen und unterstehen keiner RIR. Diese Netze sind meist sehr groß, durch die Fragmentierung benutzter Netze können diese, aber nur mit großen Aufwand verkleinert und ggf. neu verteilt werden.

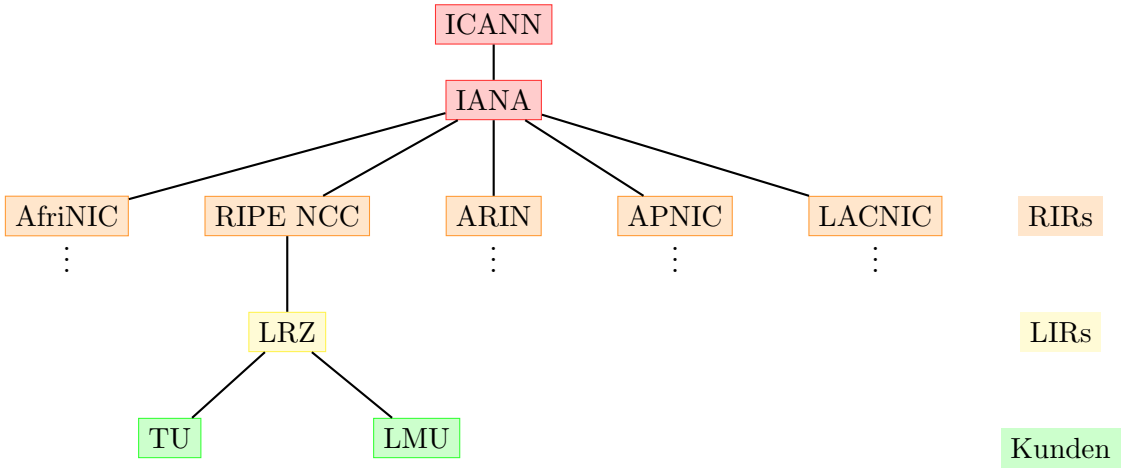

Abbildung 2.7: Vergabehierarchie von IP Adressen / AS Nummern

3 Border Gateway Protocol (BGP)

Dieses Kapitel behandelt das im Internet verwendete Routingprotokoll BGP. Zuerst soll durch einen Rückblick über die bisherigen Protokolle der Weg zur heute verwendeten Version 4 von BGP aufgezeigt werden. Der nachfolgende Abschnitt erläutert die Protokolleigenschaften und -spezifikationen, die in den entsprechenden RFCs festgehalten sind. Schließlich werden die Sicherheitsprobleme, die sich daraus ergeben, benannt und anhand der zu erwartenden Gefahren eingeschätzt.

3.1 Historische Entwicklung von BGP

Eine wichtige Rolle bei der Entwicklung des Internets spielte dessen Vorgänger, das ARPANET. Ursprünglich wurde es 1975 zum Austausch zwischen Forschungszentren und dem Militär in den Vereinigten Staaten gegründet. Als sich die militärischen Standorte 1983 durch ein eigenes Netz abspalteten, wurde das Netz nur noch von Wissenschaft und Forschung genutzt. Zu dieser Zeit gab es einen zentralen Kernbereich des Internets, der allein von dem Internet Network Operations Center betrieben wurde [IBM95].
Er bestand aus mehreren Core Gateways, die untereinander verbunden waren, und so das Backbone des ARPANETs bildeten. Diese Router tauschten die Routinginformationen untereinander mit dem Distanzvektorprotokoll Gateway-to-Gateway Protocol (GGP) aus[HS82]. An diese Core Router wurden dann die ASe angebunden. Der Routinginformationsaustausch zwischen den Edge Router der ASe und den Core Gateways erfolgte dann über ein EGP. Dafür kam das gleichnamige Exterior Gateway Protocol (EGP), welches in RFC 904 spezifiziert wurde, zum Einsatz. Es konnte bereits zwischen Intra-AS und Inter-AS-Bereichen unterscheiden und war somit in der Lage zusammen mit einem IGP betrieben zu werden. Eine wesentliche Einschränkung war jedoch, dass es nur auf Topologien eingesetzt werden konnte, die verbundene ASe in eine hierarchische Ordnung brachten [Mil84]. Diese Ordnung zeichnete sich auch in der Konfiguration von EGP ab. Denn dort musste bei den Peers angegeben werden, ob sie näher am Kernbereich oder weiter entfernt als das eigene System eingeordnet sind.
Die beiden Protokolle EGP und GGP hatten jedoch einen entscheidenden Nachteil: Sie richteten sich auf eine einzige zentralisierte Infrastruktur aus. Als ab 1986 mit dem NSFNET ein weiteres alternatives Backbone zur Verfügung stand, wurde dieses Problem offensichtlich. Dies motivierte die Entwicklung des Nachfolgeprotokolls Border Gateway Protocol (BGP). Die erste Spezifikation erschien 1989 als RFC 1105 [LR89]. Diese enthielt bereits, die grundlegenden Aspekte, die später in diesem Kapitel Erwähnung finden werden. Ab 1989 wurde diese erste BGP Version im NSFNET produktiv eingesetzt. Aus dem Betrieb ergaben sich viele Verbesserungen, die dann in BGP-2 implementiert wurden. Da es nun nicht mehr nur einen einzigen Kernbereich gab, wurde die relative Notation aufgegeben, denn die Systeme waren jetzt nicht mehr nach der Entfernung zum Kernbereich einzuordnen. Außerdem wurden die Nachrichtentypen verändert, unter anderem wurden die Pfadattribute eingeführt

[LR90]. BGP-3 führte dann einen BGP Identifier ein, der eine Verbindungskollision verhindert [LR91].
Die heute verwendete Version BGP-4 unterstützt schließlich die CIDR-Notation (vgl. Abschnitt 2.2.1), die eine feinere Netzaufteilung ermöglicht [Koz05]. Im Jahr 2006 wurde eine aktualisierte Version von BGP-4 in RFC 4271 herausgegeben. Dort wurden einige erweiternde RFCs, die seitdem erschienen waren, in die offizielle Spezifikation integriert. So wurde beispielsweise die Nutzung des Route Reflector, der in RFC 4456 eingeführt wurde, in die BGP-4 Spezifikation aufgenommen. Es wurde im Zuge dessen auch die TCP-MD5 Authentifizierung (wird in Abschnitt 4.1.2 behandelt) als verpflichtende Erweiterung aufgenommen. Dieser RFC 4271 ist die Grundlage für die weiteren Untersuchungen von BGP in dieser Arbeit.

3.2 Protokollbeschreibung

Es wird nun genau auf die Einzelheiten der Spezifikation von BGP eingegangen. Als Transportprotokoll setzt BGP auf TCP auf und nutzt dazu den Port 179. Eine solche Verbindung zwischen zwei BGP Routern wird auch als BGP Session bezeichnet. Diese Sessions befinden sich immer in einem von sechs festgelegten Zuständen. Diese Zustände werden in Abschnitt 3.2.2 definiert und beschrieben. Ist eine BGP Session hergestellt, tauschen die beiden Router Informationen über deren Erreichbarkeiten aus. Diese werden über zu diesem Zweck festgelegte Nachrichten ausgetauscht. Deren Aufbau wird in Abschnitt 3.2.3 erklärt. Im Routeauswahlprozess, den Abschnitt 3.2.4 vorstellt, werden die empfangen Routinginformationen ausgewertet und eine Route je Ziel ausgewählt. Dies findet nach verschiedenen Kriterien statt, wobei der Betreiber des Routers diese durch Konfigurationsoptionen beeinflussen kann. Da beim Design von BGP versucht wurde, das Protokoll möglichst flexibel zu definieren, sind einige Protokollerweiterungen im Einsatz, die in Abschnitt 3.2.5 vorgestellt werden sollen. Je nachdem, wo BGP eingesetzt wird, kann man zwischen iBGP und eBGP unterscheiden.

3.2.1 iBGP vs. eBGP

Als iBGP wird eine BGP Session zwischen zwei Routern desselben ASes bezeichnet. Um alle Informationen in einem AS auszutauschen, muss jeder Inner Router zu jedem Edge Router eine BGP Session offen halten. Zudem müssen die Edge Router untereinander ebenfalls iBGP Sessions aufbauen. Diese Topologie wird dann als „fully meshed“ bezeichnet. Um diese Vielzahl an Verbindungen zu reduzieren, empfiehlt RFC 4456 die Verwendung eines ggf. mehrerer Route Reflektor(en). Diese bauen zu jedem BGP Router (im eigenen AS) genau eine iBGP Verbindung auf. Über diese(n) können sich die Routinginformationen weiterverbreiten [BCC06]. Ein Beispiel eines Route Reflectors stellt Router E im AS1 in Abbildung 2.1 dar, während die Router im AS2 „fully meshed“ verbunden sind.
eBGP dagegen bezeichnet eine BGP Session, die zwischen zwei Routern verschiedener ASe aufgebaut wird.

3.2.2 Zustandsmodell einer Session

Für eine BGP Session sind genau sechs Zustände definiert. Ein Übergang zwischen diesen Zuständen ist nur anhand wohldefinierter Transitionen möglich, diese sind in dem endlichen

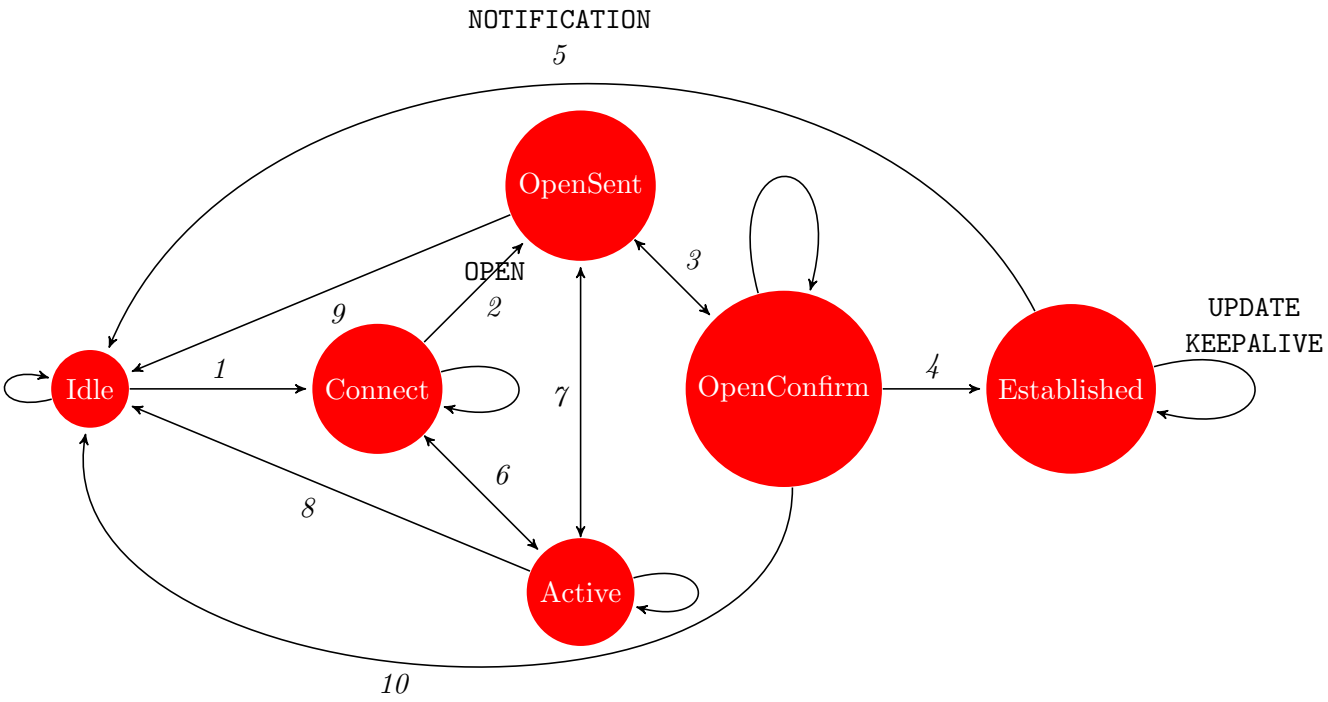

Abbildung 3.1: Endlicher Automat der Zustände einer BGP Session

Automaten in Abbildung 3.1 genauer aufgezeigt.
Die erwähnten Nachrichtentypen werden im nachfolgenden Abschnitt ausführlich besprochen. Die zulässigen Zustände sind [Bei02][RLH06]:

Idle:

Während des Idle Zustandes ist die BGP Funktionalität zwischen den beiden Routern deaktiviert. Dies kann daran liegen, dass das entsprechende Interface noch nicht aktiviert ist bzw. BGP generell auf diesem Pfad deaktiviert ist. Eine weitere Ursache kann auch eine abgebrochenen Session sein, die sich nach dem Zurücksetzen im Zustand Idle befindet. Sobald eine Verbindung hergestellt werden soll, wechselt die entsprechende Session in den Zustand Connect (*1*) und versucht sich zu dem Peer zu verbinden. Dazu wartet der Router auf TCP Verbindungen von dessen Peer und versucht parallel dazu selbst eine Verbindung aufzubauen.

Connect:

Die Session bleibt so lange im Connect Zustand, bis entweder eine TCP Verbindung erfolgreich aufgebaut ist oder kein TCP Sitzungsaufbau möglich ist. Im Falle einer erfolgreichen Verbindung wird eine `OPEN` Nachricht versendet und in den Zustand OpenSent gewechselt (*2*). Ist keine Verbindung möglich, wird in den Zustand Active gewechselt (*6*).

Active:

In diesem Zustand werden abgelaufene Timer der Zustände Connect und OpenSent behandelt. Zuerst wird versucht mittels einer Transition in den vorhergehenden Zustand (*6* bzw. *7*) doch noch einen erfolgreichen Verbindungsaufbau zu ermöglichen. Sollte dies wiederholt nicht möglich sein, wird nach Ablauf von Timern die Session in den Zustand Idle zurückgesetzt (*8*).

OpenSent:

Es wird nun auf die `OPEN` Nachricht vom Gegenüber gewartet. Es wird in den Zustand ...

- ... OpenConfirm (*3*) gewechselt, falls die Daten der `OPEN` Nachricht eine Verbindung erlauben. Dazu wird ein `KEEPALIVE` Nachricht versendet.
- ... Active (*7*) gewechselt, falls keine Nachricht eingeht.
- ... Idle (*9*) gewechselt und die Session somit zurückgesetzt, falls die Daten der `OPEN` Nachricht dies erforderlich machen. Mögliche Ursachen dafür sind eine unterschiedliche Protokollversion oder eine falsch konfigurierte AS Nummer.

OpenConfirm:

In diesem Zustand wird auf die erste `KEEPALIVE` Nachricht vom Gegenüber gewartet. Es wird in den Zustand Established (*4*) gewechselt, sobald die `KEEPALIVE` Nachricht eingeht. Falls keine Nachricht eingeht, wird zuerst in den Zustand OpenSent (*3*) gewechselt. Falls wiederholt keine Nachricht eintrifft, wird in den Zustand Idle (*10*) gewechselt und die Session somit zurückgesetzt.

Established:

Die Session ist nun soweit aufgebaut, dass `UPDATE` und `KEEPALIVE` Nachrichten ausgetauscht werden können. Tritt ein Fehler auf, so wird eine `NOTIFICATION` gesendet bzw. empfangen. In diesem Fall wird die Verbindung durch einen Wechsel in Zustand Idle (*5*) zurückgesetzt.

Der Verlauf bis zum Aufbau einer BGP Session passiert also im Normalfall diese Zustände: Idle $\xrightarrow{1}$ Connect $\xrightarrow{2}$ OpenSent $\xrightarrow{3}$ OpenConfirm $\xrightarrow{4}$ Established.

3.2.3 Nachrichten

BGP unterteilt die TCP Kommunikation in Nachrichten. Diese werden erst ausgewertet, sobald sie komplett eingetroffen sind. Eine Nachricht besteht aus einem 19 Byte großen Header und ggf. einer Payload. Die `KEEPALIVE` Nachricht verzichtet ihrerseits auf eine Payload, während die restlichen hier behandelten Nachrichtentypen `OPEN`,`UPDATE` und `NOTIFICATION` über ein Payload verfügen. Der Aufbau einer Nachricht ist in Abbildung 3.2 dargestellt.
Wie bei allen Abbildungen von Nachrichten sind orange Felder mit fester Länge definiert, rote Felder hingegen besitzen eine variable Länge. Die entsprechende Länge ist mit Hilfe des Rasters sehr einfach abzulesen. Sind Felder mit einem grauen Bereich hinterlegt, können diese ggf. komplett weggelassen werden.
Der Header unterteilt sich in drei Felder: Ein 16 Byte langes Marker Feld, welches nur noch

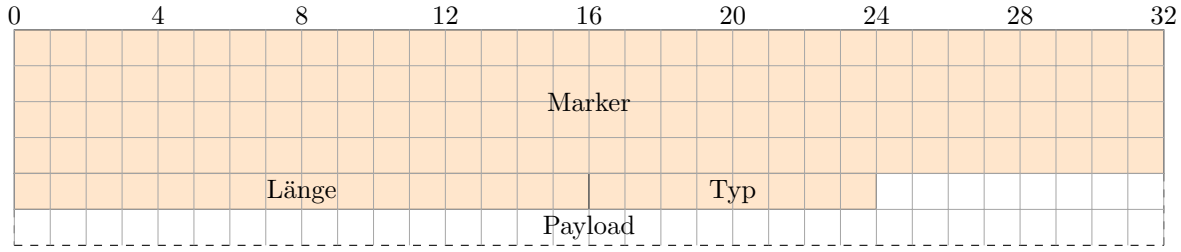

Abbildung 3.2: BGP Header mit anschließendem Payload

aus Kompatibilitätsgründen existiert und nur Einsen enthält, ein 2 Byte Feld für die Länge der Nachricht, welches Werte zwischen 19 (keine Nutzdaten) und 4096 (Maximallänge) annehmen darf, sowie zum Schluss noch eine Angabe des Nachrichtentyps. Die verschiedenen Typen werden nun genauer behandelt:

3.2.3.1 `OPEN` Nachricht

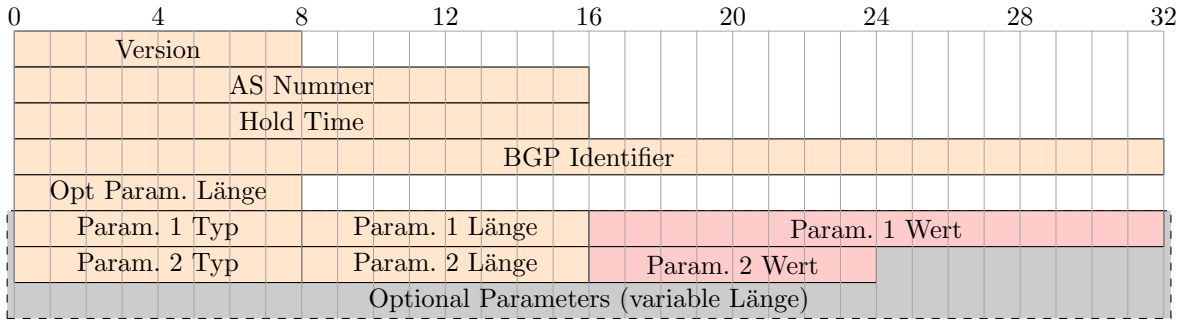

Abbildung 3.3: BGP `OPEN` Message

Bei der Verbindungsinitialisierung (vgl. Schritt *2* in Abbildung 3.1) tauschen die BGP Peers `OPEN` Nachrichten aus. Der Aufbau einer solchen Nachricht ist in Abbildung 3.3 dargestellt.
Der Absender gibt in dem Feld Version, die BGP Version seiner Implementierung an. Anschließend folgt das Feld AS Nummer. Dort gibt er seine eigene AS Nummer an. Beide Felder werden von der Gegenstelle geprüft und für den Fall, dass die Angaben nicht kompatibel sind, wird die Verbindung getrennt.
Die Hold Time gibt den Vorschlag des Senders für den gemeinsamen Hold Timer an. Als Wert für den Hold Timer wird die kleinere der beiden Hold Times gesetzt. Der Hold Timer gibt den Zeitabstand an, nach dieser ein Peer entweder eine `UPDATE` oder `NOTIFICATION` Nachricht senden muss, weil ansonsten sein Gegenüber annimmt, dass keine Erreichbarkeit mehr gegeben ist. Der Wert muss größer als 3 Sekunden sein oder 0 betragen, was einem unendlichen Hold Timer entspricht. Damit der Hold Timer deaktiviert wird müssen beide Peers eine Hold Time von 0 angeben. Von einer Deaktivierung des Hold Timer wird jedoch abgeraten, da es dann nicht mehr möglich ist, ausgefallene Verbindungen zu erkennen.
Im Feld BGP Identifier wird eine IP Adresse des Quellsystems übertragen, sie gilt jedoch für alle Sessions und Schnittstellen des Systems und dient zu dessen Identifizierung. Sie wird

dazu benutzt um mehrfache Verbindungen zwischen zwei Routern zu erkennen (sog. Kollisionen).
Durch die Felder Opt Param. Länge und Optional Parameters kann die Nachricht durch Parameter erweitert werden. Dies hält das BGP Protokoll flexibel, da darüber beispielsweise verwendete Erweiterungen ausgehandelt werden könnten. Die zusätzlichen Parameter haben diese Form: Parametertyp, Parameterlänge, Parameterwert. In früheren Spezifikationen wurde hier eine optionale Authentication Information definiert, diese ist jedoch im RFC 4271 als veraltet gekennzeichnet. Diese Parameter werden u.a. von der Multiprotocol Extension verwendet (vgl. Abschnitt 3.2.5)

3.2.3.2 `UPDATE` Nachricht

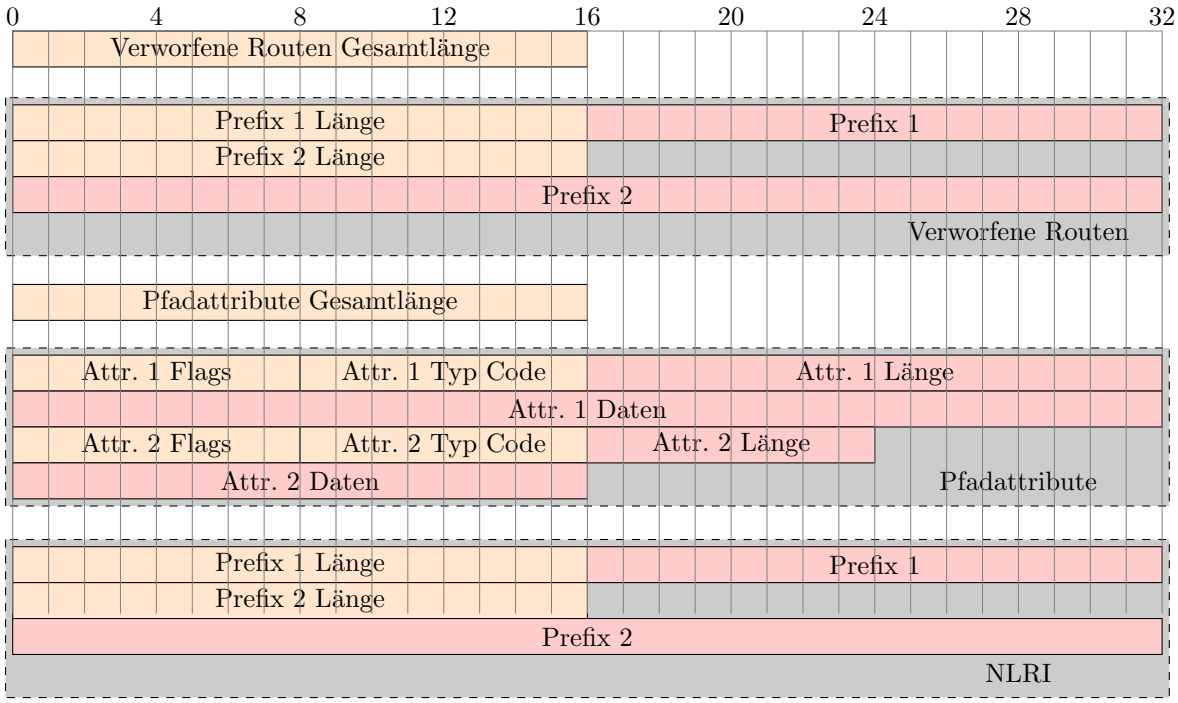

Abbildung 3.4: BGP `UPDATE` Message

Über `UPDATE` Nachrichten werden die Routinginformationen übertragen. Sie enthalten im Wesentlichen drei Bereiche, die eine variable Länge besitzen: die verworfenen Routen, die Pfadattribute und die Network Layer Reachability Information NLRI. Durch Angabe der Längen zweier Felder, ergibt sich die Länge des dritten Feldes über die Nachrichtenlänge aus dem Header.
Der Aufbau des Inhalts des Verworfenen-Routen-Feldes ist identisch mit dem des NLRI: Er besteht jeweils aus 2-Tupel der Form Prefix Länge und Prefix. Dabei entspricht die Prefixlänge dem Netzsuffix und das Prefix dem Netzteil der Route. Prefix Felder werden mit beliebigen Bits zur nächsten Bytegrenze aufgefüllt, sollten diese nicht an einer solchen enden. Das Verworfene-Routen-Feld gibt dabei die Netze an, die nun nicht mehr erreichbar sind. Im NLRI dagegen stehen die Netze, die bekanntgegeben werden sollen.
Die Informationen im Bereich Pfadattribute beziehen sich auf die Netze, die im NLRI an-

gegeben wurden. Ein einzelnes Pfadattribut besteht dabei aus einem 4-Tupel: Flags, Typ Code, Länge und dem Wert. Die Flags geben die Art eines Attributs an. Über die Flags werden u.a. die Attribute in vier Kategorien eingeteilt:

Bekannt und obligatorisch (A)

Diese Attribute sind allen BGP Implementationen bekannt. Sobald ein NLRI angegeben wird müssen diese Attribute auch mit geschickt werden.

Bekannt und optional (B)

Diese Art von Attributen sind auch allen BGP sprechenden Routern bekannt, allerdings muss nicht jedes `UPDATE`, das NLRI enthält, diese auch verwenden.

Zusätzlich und transitiv (C) bzw nicht-transitiv(D)

Diese Attribute müssen nicht in allen BGP Router implementiert sein. Es können für Erweiterungen eigene Attribute definiert werden. Je nachdem, ob das Attribut als transitiv markiert ist oder nicht, werden die Attribute über weitere BGP-Router verteilt oder nicht.

Zudem geben die Flags an, ob das nachfolgende Längenfeld ein oder zwei Bytes lang ist. Dieses Feld gibt wiederum die Länge des Wertfeldes an.
Die Tabelle 3.1 listet die nach RFC 4271 definierten Pfadattribute mit ihrer Funktion und der jeweiligen Einteilung in die vorher definierten Kategorien auf.

Attributsbezeichnung	Typcode	Inhalt	Kategorie
`ORIGIN`	1	Gibt die Herkunft der Route an (0⇒IGP, 1⇒EGP, 2⇒Unbekannt)	A
`AS_PATH`	2	Gibt den AS Pfad zum NLRI an	A
`NEXT_HOP`	3	Gibt den Next-Hop zur NLRI an	A
`MULTI_EXIT_DISC`	4	Findet Verwendung bei mehreren Links mit einem bestimmten AS, um einen Link zu bevorzugen	D
`LOCAL_PREF`	5	Findet Verwendung bei iBGP, um Routen im eigenen AS zu gewichten, darf das eigene AS nicht verlassen	A nur bei iBGP
`ATOMIC_AGGREGATE`	6	Wenn dieses Flag vorhanden ist, wurde bei der Aggregation beim AS_PATH mindestens ein AS entfernt	B
`AGGREGATOR`	7	Bezeichnet den Router mit IP und ASN, der das aggregiert Prefix erstellt hat	C

Tabelle 3.1: Übersicht über Pfadattribute der `UPDATE` Nachrichten nach RFC 4271 [RLH06]

3.2.3.3 `NOTIFICATION` Nachricht

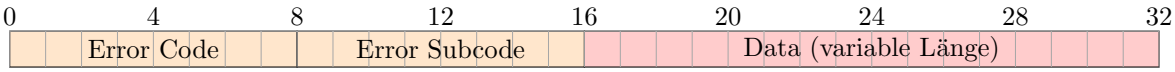

Abbildung 3.5: BGP `NOTIFICATION` Message

`NOTIFICATION` Nachrichten kündigen dem Peer immer einen Fehler an, der entdeckt wurde. Die BGP Session wird nach dieser Nachricht immer sofort durch Wechseln in den Zustand Idle (vgl. *5*, *8*, *9* und *10* in Abbildung 3.1) zurückgesetzt. Genauere Informationen finden sich dann in den Feldern Error Code, Error Subcode und ggf. Data. Im Data Feld können über die in RFC 4271 definierten Error (Sub-)Codes hinaus Informationen mitgeteilt werden.

3.2.3.4 `KEEPALIVE` Nachricht

Eine `KEEPALIVE` Nachricht dient dazu ein Timeout des Hold Timers zu verhindern. Sie besteht nur aus einem BGP Header und wird geschickt, sobald der Hold Timer auszulaufen droht. Außerdem wird diese Nachricht als Bestätigung der `OPEN` Nachricht beim Sitzungsaufbau verwendet.

3.2.4 Routeauswahlprozess

Nachdem jetzt der Routinginformationsaustausch zwischen zwei Routern behandelt wurde, erläutert dieser Abschnitt die Weiterverarbeitung dieser Daten. Dazu hat jeder BGP Router drei verschiedene Arten von Routingtabellen, diese werden auch als Routing Information Bases (RIBs) bezeichnet. In der `Loc-RIB` speichert der Router die Einträge, die durch den Routeauswahlprozess, als beste Route für ein gewisses Ziel bestimmt worden sind. Dies ist auch die Routingtabelle, die der Router zur Paketvermittlung einsetzt.
Daneben gibt es noch je BGP Session im Zustand Established eine `Adj-RIB-In` und eine `Adj-RIB-In`. In die `Adj-RIB-In` werden die vom Peer eingehenden Routen aus den `UPDATE` Nachrichten direkt nach dem Empfang übernommen. Die `Adj-RIB-In` enthält dagegen die Routen, die dem Peer über `UPDATE` Nachrichten mitgeteilt werden sollen.
Der Routeauswahlprozess soll nicht nur die beste Route für ein Ziel suchen, sondern auch diejenigen auswählen, die an die Peers weitergegeben werden sollen. Zudem soll die Menge der Routinginformationen durch Aggregation reduziert werden. Dabei werden mehrere Routen die Teil eines größeren Netzes sind, durch eine einzige Route von diesem größeren Netz ersetzt. Dieser Prozess kann in drei eigenständige Phasen aufgeteilt werden:

3.2.4.1 Bestimmung Präferenz der Routen (Phase 1)

Diese Phase wird immer ausgeführt, sobald der BGP Router eine `UPDATE` Nachricht empfängt, die eine neue, veränderte oder verworfene Route enthält. Während dieser Phase sind Zugriffe durch andere parallel laufende Prozesse auf die `Adj-RIB-In` durch einen Lock gesperrt.
Nun muss für die neuen Routen die Präferenz bestimmt werden. Bei Routen die über eBGP eingegangen sind, wird die Präferenz nur aus der Konfiguration des Routers berechnet. Dagegen wird bei iBGP die Präferenz aus dem LOCAL_PREV-Attribut (vgl. Tabelle 3.1) verwendet, wobei sich diese ggf. durch die Konfiguration des Routers noch verändern lässt. Phase 1 ist nun beendet und es wird Phase 2 gestartet.

3.2.4.2 Auswahl der besten Route (Phase 2)

Solange diese Phase aktiv ist, wird ebenfalls der Zugriff auf die `Adj-RIB-In` Tabellen durch andere gesperrt. Phase 1 und Phase 2 können daher niemals gleichzeitig aktiv sein.
Es wird jetzt jede Route daraufhin untersucht, dass dessen NEXT_HOP Attribut (vgl. Tabelle 3.1) auflösbar ist und sich dies nicht durch die Informationen der letzten `UPDATE` Nachrichten verändert hat. Falls die Route nicht mehr auflösbar ist, wird sie im Prozess nicht weiter betrachtet. Ebenso werden Routen bei denen das eigene AS bereits im Pfad vorkommt von der weiteren Betrachtung ausgeschlossen, da solche Routen Routingschleifen verursachen können (vgl. Abschnitt 2.2.3.3).
Nun werden alle verbleibenden Routen nach Ziel betrachtet, falls für ein Ziel nur eine Route existiert, wird diese ausgewählt. Existieren dagegen mehrere Routen für dasselbe Ziel, wird diejenige ausgewählt, welche die höchste Präferenz (nach Phase 1) hat. Sollten hier mehrere Routen in Frage kommen, wird die verwendete Route nach den folgenden, sog. Tie-Break Regeln bestimmt:

- Entferne alle Routen, die nicht den kürzesten AS Path haben.
- Entferne alle Routen, die nicht das niedrigste ORIGIN Attribut besitzen.
- Entferne alle Routen, die vom selben AS empfangen wurden, und das niedrigere MULTI_EXIT_DISC Attribut haben.
- Falls eine Route existiert, die über eBGP empfangen wurde, entferne alle Routen die über iBGP empfangen wurden.
- Entferne alle Routen, deren NEXT_HOP Attribut nicht die kleinste Metrik hat.
- Entferne alle Routen bis auf die mit dem geringsten BGP Identifier.
- Entferne alle Routen bis auf die mit der niedrigsten Peer IP Adresse.

Nach diesen Regeln wurde für ein Ziel genau eine Route übrig gelassen. Nun werden die ausgewählten Routen in die `Loc-RIB` übernommen.

3.2.4.3 Bestimmung der Weitergabe der Routen (Phase 3)

Phase 3 wird ausgeführt, wenn sich entweder Routen in der `Loc-RIB` geändert haben, lokal verbundene eingefügte Routen sich verändert haben oder eine neue Session erstellt wurde.
Sie blockiert bis ein evtl. laufender Phase-2-Prozess beendet wurde. Es werden alle Routen, die in der Tabelle `Loc-RIB` stehen, in die jeweilige `Adj-RIB-In` anhand der lokal konfigurierten Policy übernommen. Eine Route, deren NEXT_HOP nicht weitergegeben wird, darf ebenfalls nicht weitergeleitet werden. Sollte eine vormals weitergegebene Route nicht mehr zu Verfügung stehen, so muss diese im Verworfenen-Routen-Feld einer entsprechenden `UPDATE` Nachricht dem Peer gegenüber als ungültig erklärt werden.
Zum Abschluss dieser Phase kann die Gesamtzahl der Routinginformationen nun dadurch verkleinert werden, dass mehrere Routen zu kleineren Netzbereichen, durch eine Route zu einem größeren Bereich, der die anderen Netze zusammenfasst, ersetzt werden.

3.2.5 Protokollerweiterungen

Die Protokollspezifikation von BGP legt großen Wert auf eine spätere Erweiterbarkeit. Dies wird vor allem durch die zusätzlichen Pfadattribute erreicht, die, obwohl sie einer Implementierung nicht bekannt sein müssen, von ihr trotzdem weiterverarbeitet werden können. Daneben unterstützt die `OPEN` Nachricht ebenfalls optionale Parameter, über welche die jeweiligen Erweiterungen ausgehandelt werden können. Einige wichtige Erweiterungen werden nun kurz erläutert:

4 Byte AS Nummern (RFC 4893):

Da der AS Nummernraum, der bei BGP-4 zwei Byte groß ist und in absehbarer Zeit aufgebraucht sein wird, kann über diese Erweiterung der Raum auf 4 Byte vergrößert werden. Damit ein stufenweiser Umstieg möglich wird, wurde darauf geachtet, dass BGP Router, die lediglich 2 Byte AS Nummern unterstützen, möglichst kompatibel sind. Dazu wird für diese Router anstatt der jeweiligen 4 Byte AS Nummern die reservierte AS Nummer `AS23456`, die auch als `AS_TRANS` bezeichnet wird, verwendet. Die AS Pfadattribute werden, zum einen kompatibel für Implementierungen ohne Erweiterungen, zum anderen durch zusätzliche Attribute übertragen.

Multiprotocol Extensions (RFC 4760):

Da die Spezifikation von BGP-4 lediglich IPv4 Unicast Routing behandelt, ist es erst durch eine Erweiterung wie diese möglich IPv6 und Multicast Routing zu betreiben. Die IPv4 spezifischen Felder werden dabei durch die zusätzlichen Pfadattribute `MP_REACH_NLRI` und `MP_UNREACH_NLRI` ersetzt. Diese Attribute sind der Kategorie D also zusätzlich und nichttransitiv zuzuordnen (vgl. Abschnitt 3.2.3).

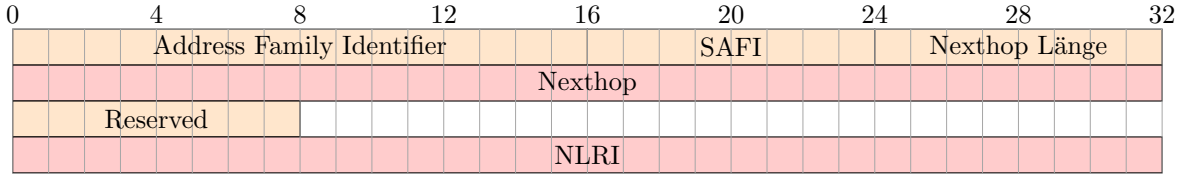

Abbildung 3.6: Pfadattribut `MP_REACH_NLRI` der Multiprotocol Extensions

Das Attribut `MP_REACH_NLRI` übernimmt die Aufgaben des Feldes NLRI der `UPDATE` Nachricht. Dessen Aufbau ist Abbildung 3.6 zu entnehmen. Das Feld Address Family Identifier (AFI) gibt dabei eine festgelegte Nummer für das jeweils verwendete Protokoll an. Das Feld Subsequent Address Family Identifiers (SAFI) unterteilt die AFI noch genauer. Die Werte von 1 für Unicast und 2 für Multicast sind bereits in RFC 4760 festgelegt. Das Feld Reserved dient der Kompatibilität zu der Vorgängerspezifikation RFC 2858 und sollte immer den Wert 0 haben. Die weiteren Felder sind analog zu den jeweiligen Feldern der `UPDATE` Nachricht.

Im `MP_UNREACH_NLRI` Attribut werden die nicht mehr erreichbaren Netze angegeben. Dazu existieren wieder die Felder AFI und SAFI, um den Protokolltyp festzulegen. Im Feld Verworfene-Routen werden dann wie bei dem gleichnamigen Feld der `UPDATE` Nachricht die fortan nicht mehr gültigen Netze übertragen[BCKR07].
Während des Sitzungsaufbaus handeln die beiden BGP Peers in der `OPEN` Nachricht die

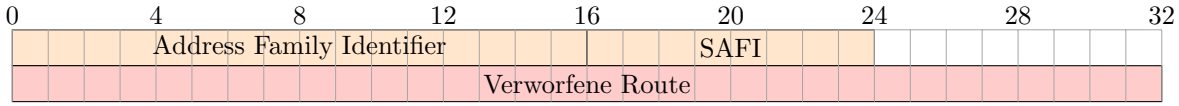

Abbildung 3.7: Pfadattribut `MP_UNREACH_NLRI` der Multiprotocol Extensions

möglichen Protokolle durch Angabe der AFI und SAFI Nummern aus. Diese Nummern werden übrigens durch die IANA den entsprechenden Protokollen zugeteilt [RP94]. BGP mit dieser Erweiterung wird auch als Multiprotocol BGP (MBGP) bezeichnet.

3.3 Sicherheitsschwachstellen BGP

Da BGP zu einer Zeit entwickelt wurde, als das Internet im Vergleich mit der heutigen Ausdehnung ziemlich klein war, spielten Sicherheitsüberlegungen damals eine nachgeordnete Rolle. Durch das stetige Wachstum des Internets wird auch die Verlässlichkeit des BGP Routings immer wichtiger. In diesem Abschnitt sollen die bestehenden Gefahren aufgezeigt und klassifiziert werden.
Die möglichen Auswirkungen auf BGP, die sich aus den Gefahren ergeben werden, in dem nachfolgendem Abschnitt behandelt. Man kann nach der Quelle der Gefahr unterscheiden. Zum einem wirken sich Schwachstellen in IP bzw. TCP direkt auf das BGP Routing aus. Diese Gefahren werden in 3.3.2 untersucht. Zum anderen können Gefahren durch Sicherheitsprobleme des BGP Protokolls selbst verursacht werden. Diese Gefahren finden in Abschnitt 3.3.3 Erwähnung.

3.3.1 Auswirkungen auf BGP

Schwachstellen können sowohl durch vorsätzliche Manipulationen, also gezielte Angriffe, als auch durch unbeabsichtigte Konfigurationsfehler, wie zum Beispiel simplen Tippfehlern, ausgenutzt werden. Deren Auswirkungen lassen sich grundsätzlich in diese verschiedenen Kategorien einteilen:

3.3.1.1 Verfälschter Ursprung

Als verfälschter Ursprung wird eine Netzbekanntmachung bezeichnet, die von einem Peer ausgeht, der nicht berechtigt ist, diesen Netzbereich in einer Bekanntmachung zu veröffentlichen. Dabei kann die Netzbereich-AS Beziehung verletzt sein oder der Peer eine fremde AS Nummer verwenden.
Für den entsprechenden Traffic der betroffenen Netze kann dies zur Folge haben, dass der unberechtigte Empfänger den Datenverkehr verwirft (sog. Blackholing) oder den Datenverkehr mitliest und anschließend evtl. verändert an den richtigen Empfänger weiterleitet (sog. Eavesdropping).

3.3.1.2 Verfälschter Pfad

Von einem verfälschten Pfad spricht man, wenn das AS_PATH Attribut einer `UPDATE` Nachricht durch einen Router unerlaubt manipuliert wird. Bei der Weitergabe einer Routinginformation darf die AS_PATH Information nicht verändert werden, stattdessen muss lediglich

die eigene AS Nummer angehängt werden.
Wird dagegen dieser AS_PATH durch ein AS, das böse Absichten hegt, unerlaubt manipuliert, ist es möglich dadurch die Routeauswahl von nahen ASen zu beeinflussen. Es kann z.B. durch eine Verkürzung des AS_PATH Attributs so Datenverkehr für ein bestimmtes Zielnetz über sich selbst umlenken. Es ist diesem AS nun wiederum möglich den Traffic zu verwerfen bzw. ihn mitzulesen.

3.3.1.3 Provozierte Instabilität

Dabei wird eine BGP Session in den Zustand Idle zurückgesetzt und diese somit beendet. Dies veranlasst die beiden betroffenen Router alle Routinginformationen erneut auszutauschen. Da Routen, die ursprünglich über die nun unterbrochene Session weitergeleitet wurden, nun nicht mehr gültig sind, muss dies den benachbarten Routern mitgeteilt werden. Dadurch wird eine erhebliche Menge an `UPDATE` Nachrichten erzeugt, die sich über weite Teile des BGP-Netzes verbreiten. Ändert sich der Zustand einer Session mehrfach kurz aufeinander folgend, so wird dies auch als Session Flapping bezeichnet.

3.3.1.4 Verlust der Verfügbarkeit

Dabei wird durch Unbefugte die Erreichbarkeit des BGP Dienstes gestört. Solange ein BGP Peer keinen Zugriff auf den BGP Dienst seines Partners hat, ist weder ein Nutzdaten- noch ein Routinginformationsaustausch möglich. Eine mögliche Ursache für den Verlust der Verfügbarkeit kann ein (D)DoS Angriff sein, der in Abschnitt 3.3.2.3 speziell behandelt wird.

3.3.1.5 Verlust der Vertraulichkeit

Der Schutz der Vertraulichkeit der übertragenen Routinginformationen ist in BGP nicht vorgesehen. Entsprechend erfolgt die Übertragung im Klartext. Ein denkbarer Anwendungsfall, wo Vertraulichkeit gewünscht werden kann, ist bei Verbindungen zwischen Transitprovider und deren Kunden. Um die Kunden ASe zuverlässig zu verbergen müsste die Vertraulichkeit der ausgetauschten Routinginformationen auf dieser Verbindung sichergestellt werden.
Allerdings ist Vertraulichkeit der Daten keine gängige Anforderung, die durch Provider gestellt wird [Mur06]. Daher wird diese im Rahmen der Arbeit nicht weiter betrachtet.

3.3.1.6 Verlust der Integrität

Die Integrität der Pakete, kann bei der Übertragung nicht garantiert werden. Diese könnten also von Dritten manipuliert worden sein. Für eine derartige Manipulation bieten sich sich drei Stellen an: Der IP-Header, der TCP-Header und die Nachricht selbst.
Diese Schwäche wird durch die Attacken TCP-Reset (vgl. Abschnitt 3.3.2.1) und Session Hijacking (vgl. Abschnitt 3.3.2.2) aktiv ausgenutzt:

3.3.2 Schwachstellen in Basisprotokollen (TCP/IP)

Diese Schwachstellen entstehen durch mangelnde Sicherheitseigenschaften der Protokolle TCP und IP. Sie werden von beliebigen Dritten nicht am BGP teilnehmenden Systemen ausgenutzt. Dabei sind die folgenden Sicherheitsrisiken zu betrachten:

3.3.2.1 TCP-Reset

Ein Zurücksetzen der TCP-Verbindung unterbricht ebenso die dazugehörige BGP Session. Es ist also nötig, dass diese erneut aufgebaut werden muss. In der Zeitspanne bis alle Routinginformationen wieder ausgetauscht worden sind, ist kein Nutzdatenverkehr über den entsprechenden Link möglich.

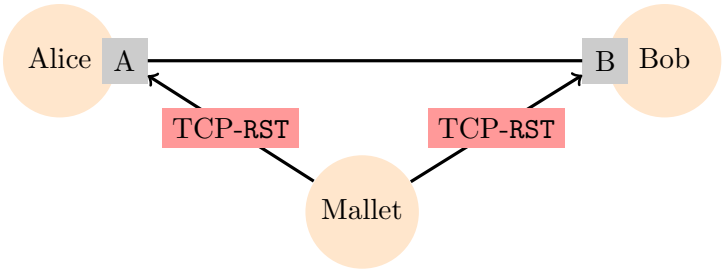

Abbildung 3.8: TCP-Reset einer BGP Session zwischen Alice und Bob

Eine TCP Verbindung kann durch Versand eines TCP Pakets mit dem Flag `RST` bzw. `FIN` beendet werden. Damit dieses Paket beim Empfänger den Eindruck erweckt, es komme von einem BGP Peer, wird es von einem Dritten mit gefälschter Quelladresse versandt (sog. Spoofing). In Abbildung 3.8 wird dieser Vorgang dargestellt. Zwischen den Routern von Alice und Bob ist eine BGP Session aufgebaut. Mallet versucht nun durch gefälschte TCP-`RST` Pakete an Alice und Bob, die BGP Session zu unterbrechen. Damit dies erfolgreich ist, müssen die gefälschten Pakete von Mallet die richtigen Sequenznummern und Quellports tragen. Um an diese Daten zu gelangen muss entweder der Datenverkehr zwischen den beiden Peers belauscht oder die entsprechenden Werte erraten werden. Bei modernen TCP Implementierungen sind die initialen Sequenznummern, bzw. der Quellport so gut randomisiert, dass diese nur schwer erraten werden können. Dies erschwert einen derartigen Angriff, so dass er ohne weitere Informationen nicht möglich ist.
Allerdings kann durch Verwendung von Internet Control Message Protocol (ICMP) Paketen die Attacke ermöglicht werden. Denn bei diesen Paketen ist keine Kenntnis der TCP Sequenznummern bzw. des Quellports nötig. Das Paket wird nun wieder mit gefälschter Quelladresse versehen und teilt dem Zielsystem der Attacke mit, dass der entsprechende Peer nicht zu erreichen ist, welches daraufhin die TCP Verbindung mit dem Peer trennt.

3.3.2.2 Session Hijacking

Befindet sich ein Angreifer Mallet zwischen zwei BGP-Systemen Alice & Bob und kann er damit alle ausgetauschten Pakete mitlesen und deren Zustellung unterbinden, dann wird dies auch als Man-in-the-middle-Angriff (MITM) bezeichnet. Diese Konstellation ist in Abbildung 3.9 skizziert. Während Alice & Bob der Meinung sind, dass sie Pakete direkt miteinander austauschen, schaltet Mallet sich in die Kommunikation mit ein und kann Pakete manipulieren. Mallet kann deshalb Einfluss auf die BGP Session der beiden Router ausüben. Dies ermöglicht dem Angreifer Mallet beliebige BGP Nachrichten an beide Router zu schicken. Der Angreifer kann sich nun wie ein BGP System verhalten. Dies führt dazu, dass er Atta-

cken, die im folgenden Abschnitt 3.3.3 beschrieben sind, in die Tat umsetzen kann.[KSM07]

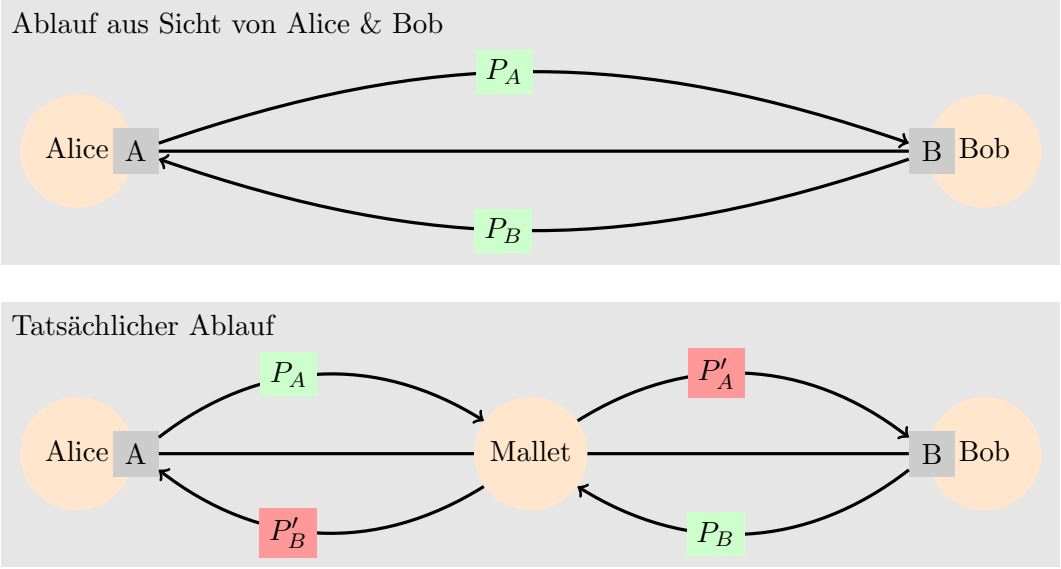

Abbildung 3.9: Man-in-the-Middle Attacke einer BGP Session zwischen Alice und Bob

3.3.2.3 (D)DoS

Unter einem Denial of Service (DoS) versteht man eine Dienstunterbrechung aufgrund einer Überlastsituation. Wird diese Überlast von einen größeren Anzahl von verteilten Systemen erzeugt, spricht man auch von einem Distributed Denial of Service (DDoS) [Wik11a]. Ein Angreifer kann dabei zwischen zwei Wegen wählen:

Überlast der Verbindung

Bei dieser Form wird durch eine große Menge an Nutzdatenverkehr die Verbindung der beiden Peers überlastet. Da der BGP Austausch über dieselbe Verbindung stattfindet, kann es passieren, dass die beiden Peers keine `KEEPALIVE` Nachrichten mehr austauschen können. Als Folge dessen wird die BGP Session in den Zustand Idle gesetzt und muss erneut aufgebaut werden. Die führt zu einer provozierten Instabilität (vgl Abschnitt 3.3.1.3).

Überlast des BGP Dienstes

Der zweite Weg führt über eine hohe Anzahl an Verbindungen direkt zum BGP Dienst des Routers. Für jede Verbindung muss die BGP Instanz Ressourcen reservieren und benötigt für die Verbindungsinitialisierung Rechenzeit. Dies wird solange durchgeführt, bis der Router so viele Ressourcen verbraucht hat, dass er den regulären BGP Sessions nicht nachkommen kann und diese abbrechen. Dadurch wird die Verfügbarkeit des BGP Dienstes eingeschränkt (vgl. Abschnitt 3.3.1.4).

3.3.3 Schwachstellen in BGP

In diesem Abschnitt werden die Gefahren vorgestellt, die von anderen BGP Systemen ausgehen können. Diese entstehen durch mangelnde Sicherheitsmechanismen im BGP Protokoll selbst. Es werden mögliche Routinganomalien, die sich auf das globale BGP System auswirken, vorgestellt. Zur vereinfachten Betrachtung wird die Beispieltopologie aus Abbildung 3.10 verwendet.

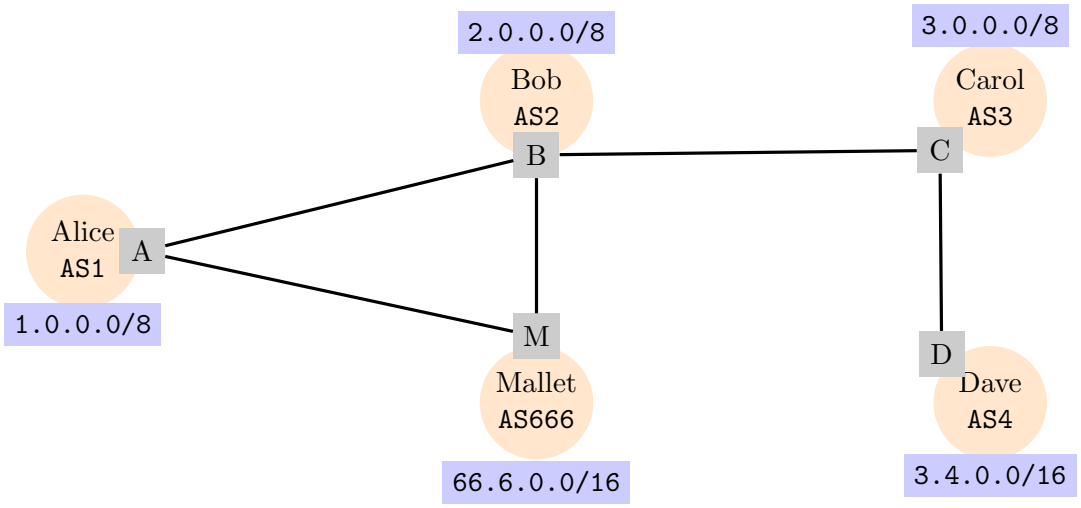

Abbildung 3.10: Beispieltopologie zur Veranschaulichung von BGP Schwächen

3.3.3.1 Umlenkung von Datenverkehr

Bei den folgenden Mechanismen wird Einfluss auf das Routing benachbarter ASe genommen. Dazu ist es nötig, dass diese unberechtigte `UPDATE` Nachricht sich im Routeauswahlprozess (vgl.3.2.4) gegenüber der authentischen durchsetzen kann. Eine Umlenkung von Nutzdatenverkehr kann über verschiedene Wege erreicht werden:

Prefix Hijacking

Beim Prefix Hijacking erklärt sich ein BGP Router eines ASes zuständig für einen Netzbereich, der eigentlich einem anderen AS zugewiesen ist. Er gibt also eine `UPDATE` Nachricht heraus mit einem NLRI für den er als Ursprung sein eigenes AS angibt (vgl. Abschnitt 3.3.1.1). Folglich konkurriert jetzt diese verfälschte Routinginformation mit der richtigen des authentischen ASes.
Die weiteren BGP Router des Internets müssen nun eine der beiden Routen auswählen. Dies geschieht anhand der Tie-Break Regeln in Abschnitt 3.2.4. Da beide Routen sich in der Regel erst bei Betrachtung der Länge des AS_PATH Attributs unterscheiden, wird dann diejenige Route bevorzugt, die sich näher (bezogen auf die Anzahl der durchlaufenen ASe) am jeweiligen Ursprung der Routeinformation befindet. Genau diese Situation lag während Phase 2 beim YouTube Vorfall aus Abschnitt 1.1 vor.
Um einen Prefix Hijack durchzuführen würde Mallet in der Beispieltopologie aus Abbildung

3.10 das Netz `3.4.0.0/16` von Dave zusätzlich bekanntgeben. Dadurch würde der Datenverkehr von den ASen von Alice und Bob bei Mallet empfangen werden, da die falsche Ankündigung den kürzeren AS_PATH als die valide hat. Carol würde jedoch weiterhin die richtige Route nutzen und Dave Pakete schicken können.[Hor09] [BFMR10]

Deaggregation Hijacking

Als Deaggregation bezeichnet man das Bekanntgeben einer Route, die Teilmenge einer bereits veröffentlichten Route ist. Dadurch kann ein Provider dem Kunden ein Teilnetz eines seiner Netze zur Verfügung stellen. Der Kunde kann dann für dieses Teilnetz Netzankündigungen herausgeben. Er deaggregiert somit das Netz seines Providers. Nach dem Longest Prefix Match Prinzip wird das speziellere Netz, das der Kunde bekannt gibt, bevorzugt. In der Beispieltopologie deaggregiert Dave mit dem Teilnetz `3.4.0.0/16`, das übergeordnete Netz `3.0.0.0/8` von Carol. Da dies mit dem Einverständnis von Carol passiert ist diese Deaggregation legitim.
Das Deaggregation Hijacking bezeichnet, die Deaggregation ohne entsprechende Zuweisung des Betreibers vom übergeordneten Netz. Bei der Deaggregation wird also genau umgekehrt vorgegangen wie bei der Aggreagtion (vgl. Abschnitt 3.2.4). Deaggregation Hijacking enführt also den entsprechende Netzbereich ähnlich wie beim Prefix Hijacking. Allerdings wird dadurch der gesamte Datenverkehr für das Teilnetz zum Entführer umgelenkt, da dieser die spezifischere Route bekanntgibt [KSM07]. Dies wiederum entspricht Phase 1 beim YouTube Vorfall aus Abschnitt 1.1.
Will Mallet aus der Beispieltopologie das Netz von Dave mittels Deaggregation Hijacking entführen, könnte er die beiden Netze `3.4.0.0/17` und `3.4.128.0/17` ankündigen. Dadurch hält er in den Routingtabellen von Alice, Bob und Carol die spezifischste Route und der gesamte Datenverkehr zu diesem Netz erreicht Mallet.

Gefälschte Pfadinformationen

Um eine gefälschte Pfadinformation in einer `UPDATE` Nachricht handelt es sich, wenn diese Information von einem BGP Router entgegen den Spezifikationen erzeugt bzw. manipuliert worden ist (vgl. Abschnitt 3.3.1.2). Denn ein Empfänger kann nicht erkennen, ob die Angaben in den Pfadattributen, denen der jeweiligen Ursprungssysteme entsprechen oder von einen BGP Router auf dem Pfad unzulässig verändert wurden. Ebenso ist nicht gesichert, dass der AS_PATH korrekt ist.

Wir betrachten dazu lediglich die Ankündigung des Netzes `3.4.0.0/16` von Dave aus der Beispieltopologie von Abbildung 3.10. Diese Ankündigung verbreitet sich über die grün markierten `UPDATE` Nachrichten aus Abbildung 3.11. Mallet will nun den Datenverkehr belauschen den Alice an Dave sendet. Dazu verkürzt Mallet den AS_PATH der für das entsprechende Netz empfangenen `UPDATE` Nachricht und leitet diese an Alice weiter. Alice hat nun zwei Routinginformationen für das Zielnetz `3.4.0.0/16`: Nämlich die aus der dritten und vierten Nachricht von Abbildung 3.11. Da Mallets Pfad kürzer ist als Bobs, entscheidet sich Alice für die Route von Mallet. Nun empfängt Mallet den Datenverkehr den Alice an Dave richtet. Aus dem Blickwinkel von Alice scheint alles in Ordnung zu sein, denn sogar die Beziehung von Netz und Ursprung entspricht der Realität

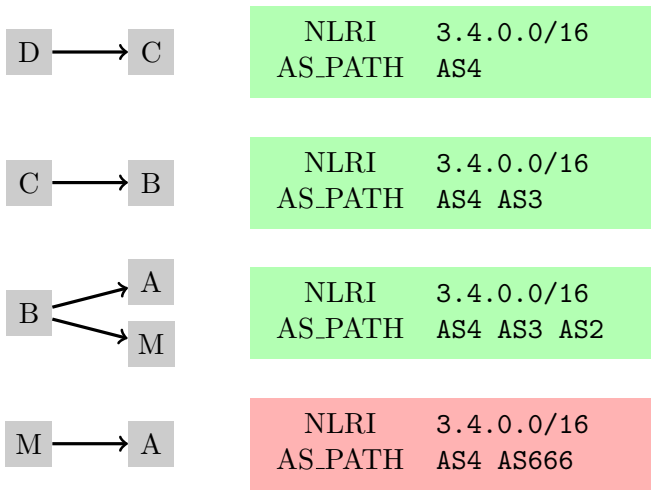

Abbildung 3.11: `UPDATE` Nachrichten mit teilweise modifizierten Pfad

3.3.3.2 Bogon Ankündigung

Bei einem Bogon Announcement kündigt ein AS einen Netzbereich an, der dem AS nicht zugewiesen ist. Da dieser Netzbereich jedoch nicht vergeben bzw. für andere Zwecke reserviert ist, gibt es in diesem Fall kein AS, das als Opfer bezeichnet werden kann. Solche Netzbereiche werden auch Bogons genannt und sollten, sofern alle BGP Instanzen korrekt konfiguriert sind, nicht in der globalen Routingtabelle auftreten. Immer wieder können solche Netze beobachtet werden. Da diese Netzbereiche keinem Provider zugeordnet sind, werden solche Bogons häufig für illegale Zwecke verwendet. Der entsprechende Angreifer erhofft sich dadurch Anonymität. [Cym11]

3.3.3.3 Update Flapping

Als Update Flapping wird ein im schnellen Wechsel durchgeführtes Bekanntgeben und wieder Verwerfen von Netzen bezeichnet. Die Problematik hierbei ist, dass diese Updates ggf. sogar global weitergereicht werden müssen. Wird dies mit einer entsprechend hohen Frequenz und einer Mehrzahl an Netzbereichen und Routern durchgeführt, kann dadurch eine erhebliche Last für BGP Router weltweit erzeugt werden. Dadurch besteht die Gefahr, dass für die betroffenen Netze in den Routingtabellen kein konvergenter Zustand mehr erreicht werden kann.

3.3.4 Zusammenfassende Risikobewertung

Zusammenfassend lässt sich sagen, dass die Spezifikation von BGP-4 selbst keine Mechanismen enthält, die einen sicheren Betrieb gewährleisten können. Zum einen ist die Integrität von BGP Nachrichten durch Sicherheitsgefahren aus den Basisprotokollen gefährdet. Zum anderen existiert keine Möglichkeit die Authentizität von BGP Peers bzw. die Korrektheit der Inhalte von `UPDATE` Nachrichten zu überprüfen [Mur06]. Daher sind für einen zuverlässigen Betrieb von BGP weitere Maßnahmen notwendig, die in Kapitel 4 detailliert erläutert werden.

4 Lösungsansätze für diese Sicherheitsprobleme

Dieses Kapitel stellt verschiedene Sicherheitsmaßnahmen vor, über die das BGP Protokoll bzw. dessen Basisprotokolle TCP und IP abgesichert werden können. In Abschnitt 4.1 werden Gegenmaßnahmen zu den Schwächen in den Internetprotokollen TCP und IP behandelt. Darauf aufbauend wird in Abschnitt 4.2 das Anwendungsprotokoll BGP selbst betrachtet. In Abschnitt 4.3 werden die vorgestellten Sicherheitserweiterungen den im vorhergehenden Kapitel ermittelten Sicherheitsrisiken gegenübergestellt. Dies bietet einen guten Überblick, auf welche Sicherheitsgefahren sich welcher Lösungsansatz bezieht.

4.1 Lösungen für Schwachstellen in den Basisprotokollen

In diesem Abschnitt werden Lösungsansätze für die Sicherheitsprobleme von BGP vorgestellt, die sich durch die Basisprotokolle TCP und IP ergeben. Dadurch sollen unerwünschte Effekte durch Systeme, die nicht regulär am BGP teilnehmen, verhindert werden.

4.1.1 Generalised TTL security mechanism

Der Generalized TTL Security Mechanism (GTSM) macht sich zu Nutze, dass BGP Sessions zumeist zwischen Routern aufgebaut werden, die direkt miteinander verbunden sind. Die TTL im IP-Header verringert sich auf dem Weg zum Empfänger also genau um den Wert 1. Bei GTSM werden nun alle ankommenden BGP Pakete, die eine TTL aufweisen, die kleiner als 254 ist, verworfen. Da der maximale Wert der TTL 255 beträgt, können so nur noch BGP Pakete von Systemen, die direkt mit dem jeweiligen Router mit GTSM verbunden sind, den Filter passieren. Dazu müssen diese Systeme alle BGP Pakete mit maximaler TTL von 255 versenden [GHM+07]. Weil bei IPv6 lediglich das Feld TTL in Hoplimit umbenannt wurde funktioniert es dort analog dazu (vgl. Abschnitt 2.2.1.2) .
Dieser Mechanismus behebt zwar nicht alle Sicherheitsprobleme, die durch IP und TCP hervorgerufen werden. Allerdings schränkt er die Anzahl der Systeme, die diese Lücken ausnutzen können erheblich ein, da diese maximal einen Hop entfernt sein dürfen. Außerdem ist keine Möglichkeit bekannt, die TTL so zu verändern, dass es scheint man sei nur einen Hop entfernt, obwohl man weiter entfernt ist. Um einen Angriff mit den erwähnten Schwachstellen zu führen, benötigt man daher zwingend Zugriff auf ein System, das einen Hop entfernt ist. Angreifer, die weiter als einen Hop entfernt sind, stellen mit diesem Mechanismus keine Gefahr mehr für die Integrität der Pakete dar (vgl. Abschnitt 3.3.1.6). Ebenso kann ein solcher Angreifer keinen (D)DoS mehr auf den BGP Dienst durchführen(vgl. Abschnitt 3.3.2.3).
Der GTSM verringert die Bedrohungen also nicht, sondern er schränkt lediglich den Kreis der Systeme ein, die diese ausnutzen können. Dies zeigt auch dessen Benennung während der Entwicklung: Generalized TTL Security Hack

4.1.2 TCP-MD5

Eine weitere Möglichkeit eine BGP Sitzung durch Angriffe auf TCP bzw. IP Ebene zu schützen, besteht durch die TCP-MD5 Signatur Erweiterung, die in RFC 2385 spezifiziert ist. Dazu müssen die beiden Peers zuvor einen sog. Pre Shared Key (PSK) vereinbaren. Der Schlüsselaustausch muss Out-of-Band erfolgen, damit eine Kompromittierung erschwert wird. Um die Datenintegrität eines Pakets sicherzustellen wird jetzt ein Hashwert aus Quell-/Zieladresse, Protokolltyp, Paketlänge, TCP-Header, TCP-Payload und diesem Pre Shared Key gebildet und dieser zwischen Header und Payload des TCP Pakets versandt. Da es sich bei MD5 um eine kryptographische Hashfunktion handelt, erlaubt das Ergebnis keinen Rückschluss auf den ursprünglichen Inhalt, der gehasht wurde. Wird nun das Paket empfangen, kann das Zielsystem ebenfalls diesen Hash bilden, da es über alle Eingabewerte insbesondere den Schlüssel verfügt. Stimmt der Hash Wert aus dem Paket mit dem Ergebnis der Hashfunktion auf dem Zielsystem überein, so wurde das Paket während der Übertragung nicht manipuliert. Stimmt im Gegensatz dazu der Hashwert nicht überein, wird das Paket verworfen. Dies würde das Quellsystem dazu veranlassen, das Paket erneut zu versenden[Hef98].
Diese Erweiterung wird bei einer Vielzahl von BGP Sessions verwendet, sie muss nach RFC 4271 von einer BGP Implementierung auch zwingend unterstützt werden. Sie sichert die Integrität der übertragenen Daten und authentifiziert durch die Kenntnis des Schlüssels den Peer. Allerdings hängt die Sicherheit des Verfahrens stark von der Wahl des Schlüssels ab. Denn wird an dieser Stelle ein unsicheres Passwort verwendet, lässt sich das mit modernen Rechnern in kurzer Zeit über eine Bruteforce Attacke ermitteln. Diese berechnet für alle möglichen Passwortkombinationen den Hashwert. Sobald dieser mit dem eines vorgegeben Datenpakets übereinstimmt, ist das Passwort gefunden.
Im RFC 3562 werden Empfehlungen abgegeben, wie ein sicherer Schlüssel gewählt werden kann. Dies sind die Anforderungen, die an solche Schlüssel gestellt werden [Lee03]:

- Die Länge des Schlüssels sollte zwischen 12-24 Bytes sein.
- Je BGP Session sollte ein eigener Schlüssel verwendet werden.
- Spätestens nach 90 Tagen sollte der Schlüssel erneuert werden.

Jedoch, auch wenn diese Empfehlungen aus dem Jahr 2003 berücksichtigt werden, dürfte auf modernen Rechnern der Schlüssel vor Ablauf der 90 Tage berechnet werden können. Es wurde daher mit TCP-Authentication Option (TCP-AO) im RFC 5925 ein Nachfolger für TCP-MD5 vorgestellt. Bei diesem Mechanismus wurde der verwendete Hashalgorithmus generalisiert, sowie Möglichkeiten geschaffen bei einer bestehenden TCP Verbindung sowohl den Schlüssel als auch den Hashalgorithmus zu wechseln. Somit ist es möglich einen Schlüsseltausch ohne Abbruch einer BGP Session durchzuführen. Allerdings empfehlen die Entwickler von TCP-AO ausdrücklich, dort, wo es möglich ist, IPSec zu verwenden [TMB10]. Denn ein Einsatz von IPSec kann an der fehlenden Unterstützung von Router Hard- bzw. Software scheitern. Außerdem gibt es IPSec Implementierungen, die nicht kompatibel zueiander sind. Mehr über IPSec ist im folgenden Abschnitt zu finden.
Sowohl TCP-MD5 als auch TCP-AO schützen vor Angriffen auf die Integrität der BGP Pakete (vgl. Abschnitt 3.3.1.6). Außerdem verhindern beide Mechanismen auch eine Überlast des BGP Dienstes durch (D)DoS (vgl. Abschnitt 3.3.2.3), da unberechtigt versandte Pakete

keine gültige Prüfsumme aufweisen und so bereits von der TCP-Implementierung verworfen werden.

4.1.3 IPSec

Mit Internet Protocol Security (IPSec) lässt sich ebenfalls die Integrität der ausgetauschten Pakete verifizieren. Es ist als Erweiterung für das IP Protokoll direkt in der Vermittlungsschicht angesiedelt. Die Spezifikation von IPSec sieht in RFC 4301 die eigenständigen Unterprotokolle Internet Key Exchange (IKE), Authentication Header AH und Encapsulated Security Payload (ESP) vor. Während IKE für den Schlüsselaustausch zuständig ist, sichern AH und ESP die Nutzdatenkommunikation ab. Um den BGP Datenaustausch abzusichern, muss IPSec im Transportmodus für den entsprechenden TCP Dienstport 179 aktiviert werden. Zusätzlich kann über IPSec Nutzdatenverkehr abgesichert werden. Dafür bietet sich der Tunnelmodus an. Dies wird in dem Abschnitt nicht betrachtet, da es keine Einfluss auf die Sicherheit von BGP selbst hat.
Im Allgemeinen wird zu Beginn einer IPSec Kommunikation über IKE die Verbindung initialisiert. Dabei authentifizieren sich beide Kommunikationspartner gegenseitig entweder anhand von Zertifikaten (mit Hilfe einer PKI vgl. Abschnitt 4.2.1) oder einem PSK. Anschließend werden die Verschlüsselungs- bzw. Integritätssicherungsalgorithmen ausgehandelt, sowie die entsprechenden Schlüssel und dessen Gültigkeit. Der Schlüsselaustausch wird dabei durch das Diffie-Hellman Verfahren vollzogen. Dieses garantiert sichere Schlüsselübertragung über einen ungesicherten Kanal. Darüber hinaus wird damit noch festgelegt, ob der Nutzdatenverkehr mit AH, ESP oder beidem gesichert werden soll. Diese ausgehandelten Parameter werden in der sog. Security Association (SA) gespeichert.[Rei11]
Mit dem AH wird die Integrität sowie die Authentizität des Ursprungs von Paketen abgesichert. Dazu wird für jedes IP-Paket ein Message Authentication Code (MAC) nach dem mittels IKE ausgehandelten Algorithmus erzeugt. Eine IPSec-Implementation muss dabei zumindest den HMAC-SHA1-96 unterstützen[Man07]. Optional können auch AES-XCBC-MAC-96 oder HMAC-MD5-96 unterstützt werden. Als Nachricht werden die statischen Felder des Headers sowie die Payload des Pakets gesichert. Dieser MAC wird nun im IP-Paket nach den IP-Header in einem Feld des AH eingefügt und anschließend verschickt. Der Empfänger des Pakets kann damit durch Überprüfung sicherstellen, dass die Integrität und Authentizität des Pakets gewahrt wurde. Der AH kann jedoch keine Vertraulichkeit gewährleisten, da die Nutzdaten nach wie vor im Klartext übertragen werden.[Ken05a]
Um Vertraulichkeit der übertragenen Daten zu gewährleisten, ist im Rahmen von IPSec der Einsatz von ESP nötig. Dort werden die Nutzdaten des IP-Pakets durch ein Verschlüsselungsverfahren, das von IKE ausgewählt wurde, geschützt. Dabei muss jede Implementation von IPSec AES-CBC und TripleDES-CBC unterstützen [Man07]. Dies sichert ebenfalls die Authentizität und die Integrität des Pakets. Der IP-Header kann damit jedoch nicht verschlüsselt werden, denn er ist im Transport Mode, der zur Absicherung von BGP verwendet werden soll, zur Datenvermittlung des Pakets notwendig.[Ken05b]
Der AH deckt dieselben Schutzziele wie die im vorherigen Abschnitt betrachteten TCP-MD5 bzw.TCP-AO Erweiterung ab. Ebenso wie bei TCP-AO kann zwischen mehreren Integritätssicherungsalgorithmen gewählt werden. Zusätzlich leistet der AH von IPSec eine automatische Aushandlung und periodische Erneuerung von Sitzungsschlüsseln, was die Sicherheit der Integritätsprüfung (vgl. Abschnitt 3.3.1.6) weiter erhöht. Bei Einsatz von ESP kann zusätzlich noch die Vertraulichkeit (vgl. Abschnitt 3.3.1.5) und der Schutz vor (D)DoS

des BGP Dienstes sichergestellt werden (vgl. Abschnitt 3.3.2.3). Leider setzt sich IPSec bei der Verwendung mit BGP bisher nicht durch. Dies ist vor allem auf die hohe Komplexität im Vergleich zu TCP-MD5 zurückzuführen. Allerdings gibt es auch Kompatibilitätsprobleme zwischen verschiedenen Implementierungen von IPSec.

4.1.4 Absicherung vor Nutzdatenüberlast durch QoS

Die bisher betrachteten Lösungen bieten keine Möglichkeit ein Abbrechen einer BGP Session bei Überlast der Netzwerkverbindung durch zu viel Nutzdatenverkehr zu verhindern. Dazu bietet sich allerdings der Einsatz von Quality of Service (QoS) Techniken an. Es muss damit sichergestellt werden, dass im Falle einer Überlast der BGP- dem Nutzdatenverkehr vorgezogen wird. Dies wird durch eine Priorisierung von BGP-Paketen sowohl auf Hardware- als auch auf Softwareebene erreicht. Durch QoS kann bei einer Überlast an Nutzdaten die rechtzeitige Zustellung von BGP Nachrichten sichergestellt werden. Dadurch kann das Unbeabsichtigte Abbrechen von BGP Sessions vermieden werden.

4.2 Lösungen für Schwachstellen in BGP

Während die Lösungsvorschläge des vorhergehenden Abschnitts sich auf die Schwächen der untergeordneten Protokolle konzentrieren, werden jetzt Verbesserungen der Sicherheitsgefahren im BGP Protokoll besprochen. Diese Lösungen haben eine gemeinsame Voraussetzung: Über eine Public Key Infrastructure (PKI) muss die Zuteilung der Internetressourcen nachvollzogen werden können. Dies wird in Abschnitt 4.2.1 betrachtet. Anschließend werden vorgeschlagene Erweiterungen zum BGP Protokoll in chronologischer Reihenfolge ihrer Publikation vorgestellt.

4.2.1 Public Key Infrastructure

Unter einer Public Key Infrastructure (PKI) versteht man ein System, das digitale Zertifikate ausstellen, verteilen, prüfen und widerrufen kann [Wik11b]. Den Inhabern eines digitalen Zertifikats lassen sich damit bestimmte Eigenschaften bzw. Berechtigungen zuteilen. Die Authentizität und Integrität der Zertifikate kann dabei mit Hilfe von asymmetrischen kryptographischen Algorithmen überprüft werden. Hierfür ist ein öffentlicher Schlüssel im Zertifikat enthalten. Den dazugehörigen privaten Schlüssel besitzt der Zertifikatsinhaber. Diese Infrastruktur ist in X.509 standardisiert, der durch RFC 5280 spezifiziert ist. Über eine solche Infrastruktur lassen sich genau die Erfordernisse erreichen, die nötig sind, damit eine Instanz zweifelsfrei nachweisen kann der berechtigte Besitzer von Internet Ressourcen zu sein. Genau diese Funktionalität fehlt im bestehenden BGP Protokoll, daher ist eine PKI nötig, um BGP besser absichern zu können.

4.2.1.1 PKI im Allgemeinen

Zum Betrieb einer PKI sind, wie in Abbildung 4.1 dargestellt ist, mehrere Komponenten nötig, die verschiedene Aufgaben übernehmen. Eine Zertifizierungsstelle (Certificate Authority, CA) signiert die Zertifikate und bestätigt somit deren Eigenschaften. Dazu sendet der Antragsteller Alice eine Zertifikatsanforderung, den sog. „Certificate Signing Request“

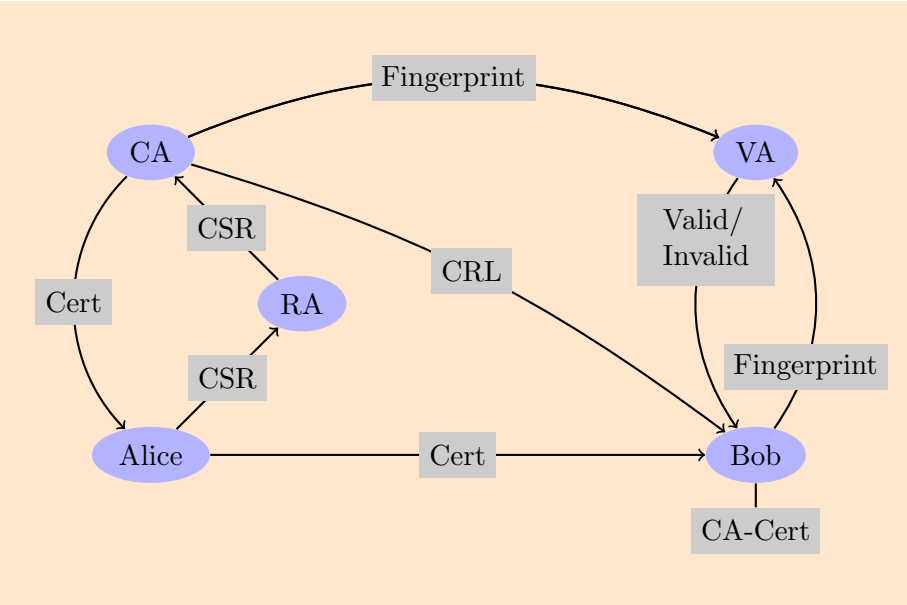

Abbildung 4.1: Aufbau einer PKI

(CSR) an die Registrierungsstelle (Registration Authority, RA). Diese überprüft die Eigenschaften, die ins Zertifikat aufgenommen werden sollen, und gibt diese an die CA weiter. Die CA signiert den CSR und erhält dadurch das Zertifikat (Certificate, CERT), welches an Alice weitergegeben wird. Zusätzlich gibt die CA den Fingerprint des Zertifikates an die Validierungsstelle (Validation Authority, VA) weiter, damit diese das Zertifikat von der Erstellung des Zertifikats Kenntnis hat.
Alice kann nun das Zertifikat veröffentlichen, damit andere Instanzen ihre zugewiesenen Eigenschaften überprüfen können. Im Gegensatz dazu muss Alice den privaten Zertifikatsschlüssel, der bei Erstellung des Antrags erzeugt wurde, möglichst sicher aufbewahren, da dieser nur für sie bestimmt ist. Dieser Schlüssel ist nötig, um weitere Dokumente zu signieren, die dann mittels Zertifikat auf seine Gültigkeit überprüft werden können.
Will Bob nun die Gültigkeit dieses Zertifikates überprüfen, benötigt er dazu mindestens ein Zertifikat einer CA (CA-Cert), dem er vertraut. Wurde das zu prüfende Zertifikat durch diese CA signiert, lässt sich dass anhand des CA-Zertifikats verifizieren. Für den Fall, dass eine CA den Antrag einer weiteren CA signiert, kann sich so eine sog. Zertifikatskette bilden. Die oberste CA dieser Kette entspricht dann der Wurzel-CA, deren Zertifikat vertraut werden muss, um die Validität dieser Kette zu bestätigen.
Um das Ausmaß des Schadens bei Abhandenkommen eines privaten Schlüssels zu minimieren, besitzen die Zertifikate eine beschränkte zeitliche Gültigkeit. Nach deren Ablauf muss ein Zertifikat erneuert werden. Ferner werden von den CAs sog. Widerrufslisten (CRL) erstellt, durch die Zertifikate vor Ablauf der Gültigkeit zurück gezogen werden können. Diese werden von den CAs veröffentlicht und müssen regelmäßig auf Aktualisierungen geprüft werden. Stellt der Betreiber einer CA mit einer VA den sog. „Online Certificate Status Protocol“ (OCSP) Dienst zur Verfügung, dann kann darüber überprüft werden, ob ein gewisses Zertifikat noch gültig ist oder von der CA widerrufen wurde.

4.2.1.2 RPKI

Eine spezialisierte Form einer PKI ist die Resource Public Key Infrastructure (RPKI). Sie wurde auf die Vergabe von Internet Ressourcen wie IP Adressen und AS Nummern angepasst. Dazu wurde eine X.509 Erweiterung geschaffen, die spezielle Felder für diese Ressourcen besitzt. Dem Zertifikatsinhaber können damit Zuständigkeiten für gewisse Bereiche der AS Nummern bzw. IP Adressen zugewiesen werden. Dabei kann jeder Zertifikatsinhaber als CA für die in seinem Zertifikat angegeben Ressourcen fungieren und somit diese weiter delegieren [LKS04].
In diesem System entspricht die IANA, der Wurzel-CA. Da durch die IANA die Ressourcen verwaltet werden (vgl. Abschnitt 2.3), enthält ihr Zertifikat alle Netzbereiche und AS Nummern. Die IANA signiert die Zertifikate der RIRs, die nur die Adress- bzw. Nummernbereiche enthalten, die vom jeweiligen RIR verwaltet werden. Diese stellen wiederum die Zertifikate für die LIRs aus. So wird die Zertifikatskette bis zum Endkunden fortgesetzt. Die Zertifikatskette für den Validierungspfad einer LIR ist in Abbildung 4.2 exemplarisch dargestellt [HE11].
Da sich das RPKI System gerade erst im Aufbau befindet, existiert noch kein IANA Zertifikat. Jedoch bieten alle RIRs bis auf ARIN ein entsprechendes System zur Zertifizierung an. Ebenso besteht bei diesen RIRs die Möglichkeit Zertifikate in ein öffentliches Repository aufzunehmen.
Durch das fehlende Wurzelzertifikat der IANA benötigt man zur Überprüfung von RPKI-Zertifikaten alle RIR-Zertifikate. Diese können inklusive dem Pfad zum öffentlichen Repository über einen sog. Trust Anchor (TA) angegeben werden.
Ein weiterer Nachteil am derzeitigen RPKI System ist, dass es nicht möglich ist, Ressourcen aus Legacy und PI Bereichen zu zertifizieren. Die RIRs haben jedoch angekündigt, auch für diese Bereiche zeitnah eine Zertifizierungsmöglichkeit zu schaffen [Ban11].

4.2.1.3 Verbindung RPKI und Routing

Die RPKI für sich ermöglicht nur Nachweis darüber zu führen, ob ein entsprechender Zertifikatsinhaber für gewisse Ressourcen berechtigt ist. Damit ein solcher Inhaber nun Routingbeziehungen verifizierbar angeben kann, muss er diese Informationen mit seinem Zertifikatsschlüssel signieren. Leider führt die Verbindung von RPKI und Routing zu einigen Nachteilen. Durch den hierarchischen Aufbau des Vergabe- und somit auch des Zertifizierungssystems, wird der IANA und den RIRs die Möglichkeit gegeben, Zertifikate von Internet Ressourcen zu widerrufen. Durch die enge Verflechtung von RPKI und Routing kann direkter Einfluss auf das Routing genommen werden. Diese Organisationen würden somit über eine Art Ausschaltknopf für bestimmte Netzbereiche oder im Fall der IANA sogar für das ganze Internet verfügen.
Obwohl sowohl die RIRs als auch die IANA faktisch unabhängige Organisationen sind, wäre dadurch die Neutralität und Unabhängigkeit des Internets gefährdet. Zum einen unterliegen die Organisationen der entsprechenden Gerichtsbarkeit des Staates, wo ihr Stammsitz liegt. Zum anderen kann bereits die Existenz einer Sperrinfrastruktur bei einigen Staaten Begehrlichkeiten hegen, diese zum Zwecke der Strafverfolgung oder gar der Zensur einzusetzen.
Mit ähnlichen Argumenten haben Gegner der Verknüpfung von RPKI und Routing sich auf

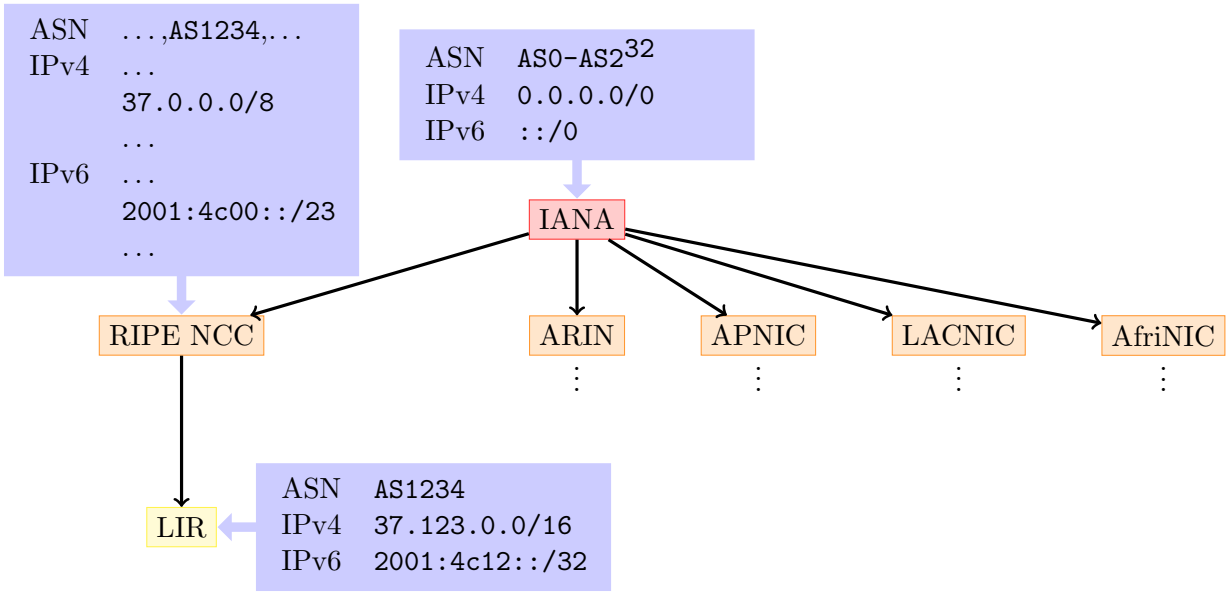

Abbildung 4.2: Zertifizierungspfad von Internet Ressourcen

der 63. RIPE Tagung im Herbst 2011 zu Wort gemeldet und einen Verzicht auf die Verbindung von Routing und RPKIs gefordert. Die Befürworter entgegneten, dass durch kritische Stellen die Sperrlisten genauer untersucht werden könnten und sollten und somit ein Missbrauch der RPKI aufgedeckt werden könnte. Schließlich sprach sich eine knappe Mehrheit auf dieser Tagung dafür aus RPKI und Routing weiterzuführen und damit das Routing abzusichern [Erm11].

4.2.2 sBGP

In der Entwicklungsabteilung von BBN Technologies wurden aufgrund der Sicherheitsgefahren von BGP bereits im Jahr 1997 mit der Entwicklung eines Absicherungsverfahren begonnen. Der hier vorgestellte Entwurf von sBGP wurde 2003 vorgestellt. Er führt eine Sicherheitsarchitektur ein, die RPKI, Attestations und IPSec vorsieht.
Die RPKI dient zur Verifizierung der Besitzverhältnisse von Ressourcen. Vorschläge aus diesem Draft wurden beim Aufbau der RPKI Infrastruktur der RIRs berücksichtigt (vgl. Abschnitt 4.2.1.2). Unter anderem baut das Format der Zertifikate in RFC 3779 auf diesem Draft auf. Die Attestations dienen zur Abbildung der Routingbeziehungen und werden im nachfolgenden Abschnitt gesondert behandelt. Die Verwendung von IPSec dient zur Vermeidung von Gefahren, die sich aus den Basisprotokollen ergeben (vgl. dazu Abschnitt 4.1.3).

4.2.2.1 Attestations

Bei den Attestations handelt es sich um Dokumente, die durch den entsprechenden Ressourceninhaber digital signiert sind. Es wird dabei zwischen zwei Arten unterschieden:

Address Attestations

Durch eine Address Attestation (AA) erteilt ein Inhaber eines Netzbereiches bestimmten ASen die Erlaubnis, einen Netzbereich bekanntzugeben. Solche AAs sind relativ statisch, da sie sich nur ändern, wenn ein Netzbereich durch ein weiteres AS bekanntgegeben werden soll. Aus den AAs wurden später die ROA-Dokumente abgeleitet. Diese entsprechen in Inhalt und Funktion den AA (vgl. Abschnitt 4.2.4.1).

Route Attestations

Eine Route Attestation (RA) wird von einem Router signiert. Dazu muss der jeweilige Router über den privaten Schlüssel eines Zertifikats verfügen, das den entsprechenden AS zugewiesen wurde. Die RA ist Teil jeder `UPDATE` Nachricht, die durch einen Router des ASes versendet wird. Diese RA wird in einem transitiven optionalen Pfadattribut übertragen (vgl. Abschnitt 3.2.3.2). Die RA sichert dabei den AS_PATH und den NLRI der `UPDATE` Nachricht ab. Wird eine bereits so signierte Nachricht von einem sBGP Router weitergeleitet, fügt er seine eigene AS Nummer an den AS_PATH an und signiert wiederum diese Nachricht. Die Signatur wird an die bereits bestehenden Signaturen im jeweiligen Pfadattribut angehängt. Diese zwiebelförmige Anordnung von Signaturen sichert den AS_PATH gegen ungerechtfertigte Manipulationen. Für die Beispieltopologie aus Abbildung 3.10 würden die `UPDATE` Nachrichten wie in Abbildung 4.3 dargestellt aussehen. Im Gegensatz zur AAs sind RAs sehr dynamisch, da sie bei jedem `UPDATE` generiert werden müssen.

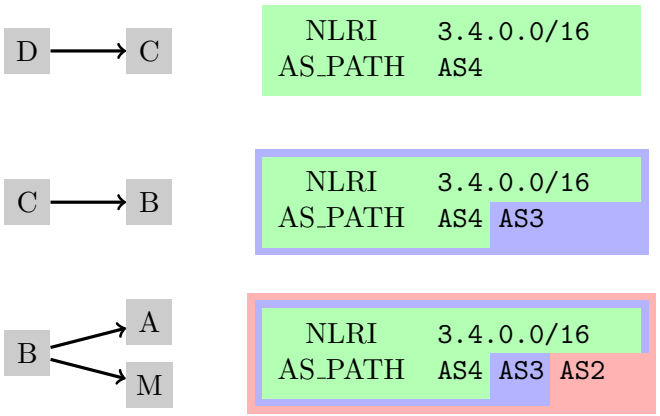

Abbildung 4.3: Mittels Route Attestations (RA) signierte `UPDATE` Nachrichten

4.2.2.2 Validierung von Updates

Um nun eine empfangene `UPDATE` Nachricht validieren zu können, benötigt ein sBGP-Router Zugriff auf einen sog. Validation Cache. Dieser speichert sämtliche Zertifikate, Widerrufslisten und AAs aus sBGP Repositories zwischen und validiert diese. Repositories werden verteilt von großen Internetknoten und Providern betrieben. Sie gleichen die Inhalte untereinander ab. Der Cache dient dazu die Anforderungen an CPU und Speicher auf den Routern zu minimieren.

Wird nun eine `UPDATE` Nachricht empfangen, wird zuerst überprüft, ob der Cache über eine valide AA verfügt, die dem Ursprungs-AS zum Veröffentlichen des Netzbereiches berechtigt. Anschließend werden die Signaturen der RAs überprüft. Für diese müssen im Cache passende und valide Zertifikate vorliegen. Anschließend wird der AS_PATH dahingehend überprüft, ob er mit den RA übereinstimmt. Werden alle diese Überprüfungen erfolgreich abgeschlossen, ist die Korrektheit der `UPDATE` Nachricht nachgewiesen. In der Router Konfiguration kann dann entschieden werden, in welcher Form sich das Ergebnis der Überprüfungen der RAs auf das Routing auswirkt. Eine `UPDATE` Nachricht kann daher drei Ergebnisse haben:

- Valid: Die Überprüfung der Attestations war erfolgreich
- Invalid: Mindestens eine Attestation war fehlerhaft
- Unkown: Es befinden sich Router auf den Pfad die sBGP nicht unterstützen

4.2.2.3 Fazit

sBGP kann die Umlenkung von Datenverkehr (vgl. Abschnitt 3.3.3.1) wirksam verhindern, denn sowohl Veränderungen am AS_PATH als auch unberechtigte Ankündigungen von Netzbereichen können dadurch erkannt werden.
Allerdings wird durch die Vielzahl an kryptographischen Operationen in sBGP mehr Hardwareressourcen in den Routern benötigt. Da trotz Einsatz eines Validation Caches, jede RA vom Router überprüft bzw. erstellt werden muss. Eine Untersuchung im Jahr 2000 kommt zu dem Schluss, dass in etwa die Leistung eines Desktop-PCs zusätzlich durch den Einsatz von sBGP benötigt wird [KLMS00]. Außerdem kam diese Arbeit zum Ergebnis, dass sich durch den Einsatz von sBGP die Konvergenzzeiten um mehr als das doppelte anwachsen könnten [BFMR10]. Dies war wohl der Grund, warum sich sBGP seinerzeit nicht durchsetzen konnte. Viele Ideen aus sBGP wurden jedoch für weitere Protokolle übernommen.

4.2.3 soBGP

Einen alternativen Mechanismus zur Absicherung der BGP Daten stellt Hardwareherstellers Cisco 2004 mit dem Entwurf von Secure Origin BGP (soBGP) vor. Er versucht durch spezielle Zertifikate den Overhead bei der Validierung zu minimieren. Des weiteren kann der Betreiber eines soBGP Router bestimmen, in welchem Maß die Routinginformationen abgesichert werden sollen. Dadurch nimmt er ebenfalls Einfluss darauf, wie viele Ressourcen dafür eingesetzt werden.
Eine zentrale Rolle bei der Validierung spielt die lokale Datenbank, die jeder Router bereithält. Anhand dieser werden eingehende `UPDATE` Nachrichten überprüft. Sie wird durch Validierung der eingehenden Zertifikate und anschließender Aufnahme der enthaltenen Informationen schrittweise aufgebaut.
Anstatt einer hierarchischen PKI, setzt soBGP auf ein Netz des Vertrauens, ein sog. „Web of Trust“ (WoT), zur Überprüfung der Echtheit von Zertifikaten [Whi05].

4.2.3.1 Web of Trust (WoT)

Der Ansatz des WoT stellt eine dezentrale Alternative zu einem PKI System dar. Dabei gibt jeder Teilnehmer dieses Netzes an, welchen Instanzen er in welchem Ausmaß vertraut. Dies geschieht durch Signatur des entsprechenden Zertifikats [Wik11c].

Zusätzlich hält jeder Teilnehmer eines WoT eine Anzahl an Entity Zertifikaten, denen er vertraut. Soll nun die Echtheit eines Zertifikats überprüft werden, wird versucht ausgehend von diesen vertrauten Zertifikaten eine Kette von Signaturen aufzubauen. Dadurch kann die Echtheit eines weiteren Zertifikats überprüft werden [Wei05].

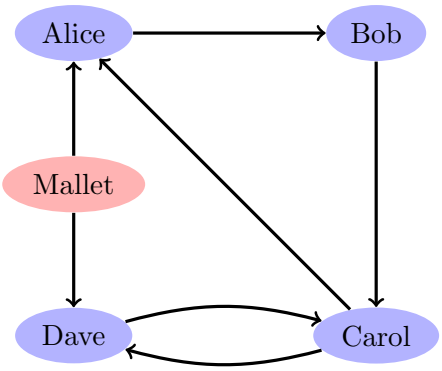

Abbildung 4.4: Vertrauensbeziehungen in einem Web of Trust

In Abbildung 4.4 ist ein solches WoT exemplarisch skizziert. Dabei drückt die Beziehung $X \rightarrow Y$ aus, dass X dem Zertifikat von Y vertraut. In diesem WoT akzeptieren Alice, Bob, Carol und Dave ihre Zertifikate gegenseitig. Da jedoch keiner dem Zertifikat von Mallet vertraut, wird dieses auf von keinem akzeptiert.

4.2.3.2 Zertifikatstypen bei soBGP

Bei soBGP kommen drei verschiedene Zertifikatstypen zum Einsatz: Entity, Auth und Policy Zertifikate. Zweck dieser Aufteilung ist, dass mit einem einzelnen Zertifikat nur eine beschränkte Information signiert werden soll. Dies soll den Verwaltungsaufwand sowohl bei Signierung als auch bei Validierung minimieren. Während es sich bei dem Entity Zertifikat um ein vollständiges Zertifikat mit Signatur und öffentlichen Schlüssel handelt, stellen Auth und Policy Zertifikat nur signierte Informationen bereit und beinhalten keinen öffentlichen Schlüssel.
Die Verteilung der Zertifikate ist aus Gründen der Interoperabilität auch innerhalb von BGP möglich. Dazu wird eine zusätzliche `SECURITY` Nachricht definiert. Damit der initiale Zertifikatsaustausch schneller vonstatten geht, empfiehlt es sich ähnlich wie bei sBGP öffentliche Repositories mit den gesamten Zertifikatsbestand anzulegen. Auth Zertifikate werden in der Regel nicht einzeln, sondern als Teil der Prefix Policy Zertifikate übertragen [Wei05].

Entity Zertifikat

Entity Zertifikate weisen eine Instanz im BGP Routing aus. Sie binden die AS Nummer an einen öffentlichen Schlüssel. Die Echtheit dieser Entity Zertifikate wird anhand des WoT Ansatzes (vgl. Abschnitt 4.2.3.1) überprüft. Dabei müssen die Organisationen, die AS Nummern zuweisen (RIRs, IANA), nicht zwingend den jeweiligen AS Zertifikaten vertrauen, denen sie AS Nummern zugewiesen haben. Es reicht aus, wenn vom Ausgangszertifikat über das WoT eine Vertrauenskette hergestellt werden kann. Über diese Entity Zertifikate können sich BGP Peers gegenseitig authentifizieren.

Auth Zertifikat

Über Auth Zertifikate werden einem Entity Zertifikat Adressblöcke zugewiesen. Sie entsprechen den Address Attestations von sBGP. Diese Auth Zertifikate werden anhand der Vergabehierarchie von IP Ressourcen signiert. Diese Zertifikate stellen somit den Bezug zwischen AS Nummer und zugewiesenen Netzbereichen her.

Policy Zertifikat

In den Policy Zertifikaten werden die Informationen festgelegt, die für den Aufbau der soBGP Topologie Datenbank nötig sind. Policy Zertifikate werden durch Entity Zertifikate signiert. Es wird zwischen Policy Zertifikaten mit Informationen zum AS und zu bestimmten Netzbereichen unterschieden:

- AS Policy Zertifikat
  - Festlegung der benachbarten ASe
  - Verbreitung der CRL des Entity Zertifikates
- Prefix Policy Zertifikat
  - Legt die zu verwendende Policy für einen bestimmten Netzbereich fest.
  - Aussteller Entity Zertifikat muss durch ein Auth Zertifikat den entsprechenden Netzbereich zugewiesen haben.

Bei der Prefix Policy kann bestimmt werden, ob und in welcher Form der dazugehörige Netzbereich validiert werden soll. Es ist möglich, dass der Netzbetreiber in diesem Fall auf eine Validierung durch andere BGP Router verzichtet oder lediglich die Ursprungsangabe überprüfen lässt.

4.2.3.3 Lokale Datenbank

Die lokale Datenbank, in die jeder Router die Informationen der validierten Zertifikate abspeichert, teilt sich in zwei Bereiche:
Zum einen gibt es die Autorisationsdatenbank, in diese werden die Zuordnungen von ASen zu Netzen und deren zugehörige Policy abgespeichert. Diese Informationen gehen aus den Prefix Policy Zertifikaten hervor.
Zum anderen wird aus den AS Policy Zertifikaten ein Verbindungsgraph, der sog. „Internetwork-Graph“ angelegt. Dieser zeigt die Verknüpfungen aller ASe untereinander, ausgehend vom eigenen AS [Wei05].
Im Vergleich zu den sehr dynamischen Routing Attestations von sBGP verhält sich die lokale Datenbank von soBGP statisch. Erst bei Änderung der Topologie oder der Netzbereiche, ändert sich auch die lokale Datenbank.

4.2.3.4 Validierung von Updates

Da die Validierung der Routinginformationen immer anhand der lokalen Datenbank stattfindet, ist diese erst dann möglich, wenn die Datenbank komplett aufgebaut ist. Bei Aufbau der Datenbank werden viele kryptographische Operationen benötigt, dagegen wird für die Überprüfung von Updates auf die Inhalte der Datenbank zurückgegriffen.

Soll nun eine eingehende `UPDATE` Nachricht überprüft werden, wird zuerst für jede Routing Information ein sog. „Security Preference“ Pfadattribut angelegt. Dieses Attribut wird ähnlich wie die Local Preference lediglich über iBGP Sessions ausgetauscht (vgl. Abschnitt 3.2.3.2).
Nun wird der Inhalt nach folgendem Schema validiert: Die Security Preference wird erhöht, wenn die Überprüfung erfolgreich ist und verringert, wenn diese negativ verläuft [Whi05].

- Überprüfung der Netzbereich - Ursprungs-AS Beziehung durch Abfrage an die Autorisationsdatenbank für jeden Netzbereich.
- Überprüfung des 2. AS im Pfad, ob eine Nachbarschaftsbeziehung im Topologiegraphen besteht.
- Überprüfung des Pfads, ob er nach dem Internetwork-Graph in beide Richtungen besteht, nur in eine Richtung besteht oder nicht möglich ist.

Mit Hilfe der Prefix Policy kann durch den Inhaber des Netzes gefordert werden, dass eine Route bei Fehlschlagen einer Prüfung sofort verworfen wird. Falls die Route dagegen bei den Überprüfungen keine Auffälligkeiten zeigt, wird die Security Preference während des Routeauswahlprozess berücksichtigt (vgl. Abschnitt 3.2.4).

4.2.3.5 Fazit

Im Vergleich zu sBGP geht soBGP ressourcensparender vor, da es durch Abspeicherung der Validierungsergebnisse die Anzahl der nötigen kryptographischen Operationen verringert. Außerdem muss nicht jede `UPDATE` Nachricht signiert werden, sondern es ist lediglich nötig bei Änderungen die entsprechenden Zertifikate neu zu erstellen.
Die Umlenkung von Datenverkehr (vgl. Abschnitt 3.3.3.1) wird dadurch wesentlich erschwert. Es ist nicht mehr möglich unberechtigte Adressbereiche anzukündigen. Allerdings findet keine vollständige Pfadvalidierung ähnlich wie bei sBGP statt, bei der jedes durchlaufene AS über eine Signatur verifiziert wird. Es wird lediglich überprüft, ob der Pfad im Hinblick auf die Topologie plausibel erscheint. Befindet sich ein AS mit bösen Absichten nahe im Umfeld des anzugreifenden AS, kann eine Manipulation des AS Pfades nicht gänzlich ausgeschlossen werden. Bemerkenswert sind die Möglichkeiten, die einem Inhaber des Netzbereichs zur Verfügung stehen, um die Validierung seines Netzes durch andere soBGP Instanzen zu beeinflussen.
Jedoch hat soBGP das Entwurfsstadium nicht verlassen und wurde nicht weiterverfolgt. Vermutlich waren die nicht vollständige Pfadvalidierung und das dezentrale WoT die Gründe hierfür.

4.2.4 Origin Validation mit dem RPKI/Router Protokoll

Im Jahr 2006 wurde bei der Internet Engineering Task Force (`IETF`), die die Weiterentwicklung der technischen Standards des Internets koordiniert, eine Arbeitsgruppe eingerichtet, die ein sicheres Inter-Domain Routing (SIDR) entwickeln soll. Diese Arbeitsgruppe steckte sich zwei Ziele: In einem ersten Schritt sollte der Ursprung von Routinginformationen gesichert werden. Darauf aufbauend soll die Validierung den gesamten Pfad einschließen.
Um den ersten Schritt zu erreichen schlägt diese Arbeitsgruppe das RPKI/Router Protokoll

vor. Es reduziert die nötigen Anpassungen von Routerplattformen auf geringfügige Softwareänderungen, da es die kryptographischen Operationen auf ein weiteres System, den Local Validation Cache auslagert. Dieser validiert durch Route Origin Attestations (ROAs) den Ursprung von `UPDATE` Nachrichten [LKK11].

4.2.4.1 Route Origin Attestations (ROAs)

Durch eine solche ROA wird einem AS die Erlaubnis erteilt, einen oder mehrere Netzbereiche bekanntzugeben. Sie entsprechen also den AAs bei sBGP und den Auth Zertifikaten bei soBGP. Zusätzlich zum Netzbereich kann mit dem Feld MaxLen angegeben werden, bis zu welcher Netzgröße der Netzbereich deaggregiert werden darf. Der Zertifikatsinhaber, der diese ROA signiert, muss über eine Berechtigung für alle angegebenen Netzbereiche verfügen. Eine ROA für das fiktive AS1234 ist in Abbildung 4.5 zu finden. Diese ROA würde dem AS erlauben die beiden Netze anzukündigen. Dabei darf vom IPv6 Netz kein angekündigtes Subnetz kleiner als `/64`. Beim IPv4 Netz ist die kleinstmögliche Subnetzgröße `/24` [LKK11].

version	`0`		
asID	`1234`		
	Family	Address	MaxLen
ipAddrBlocks	IPv4	`37.123.0.0/16`	`/24`
	IPv6	`2001:4c12::/32`	`/64`

Abbildung 4.5: Beispiel für ein ROA-Dokument

4.2.4.2 Architektur RPKI/Router Protokoll

Am RPKI/Router Protokoll sind wie in Abbildung 4.6 dargestellt drei Parteien beteiligt: Die globalen RPKI Server, Local Validation Caches und Router, die das RPKI/Router Protokoll aktiviert haben.
Auf den RPKI Servern stellen die RIRs Repositories zur Verfügung, die alle untergeordneten Zertifikate, ROAs und die aktuellen CRL enthalten. Diese stehen öffentlich zur Verfügung und jedes Dokument mit einer gültigen Signatur des entsprechenden RIRs bzw. nachfolgender Zertifikate kann dort abgelegt werden.
Die Local Validation Caches spiegeln die Repositories und prüfen diese periodisch auf Aktualisierung. Anschließend verifizieren sie die Gültigkeit aller Dokumente in den Repositories und erstellen daraus eine Datenbank von gültigen Beziehungen zwischen ASen und Netzbereichen.
Über das RPKI/Router Protokoll fragen Router nun diese Datenbank ab. Ein Router kann dabei mehrere Validation Caches verwenden. Bei der ersten Verbindung zum Cache wird eine Übertragung der gesamten Tabelle nötig. Anschließend können auch nur die Änderungen bezüglich einer bestimmten Version der Datenbank inkrementell übertragen werden. Anhand dieser Daten prüft nun der Router eingehende `UPDATE` Nachrichten auf die Gültigkeit des Ursprungs. Kann der Router mittels der Datenbank den Netzbereich dem Ursprung zuordnen, ist diese korrekt. Für den Fall, dass der Netzbereich nicht in der Datenbank gefunden

wird, kann der Router keine Aussage über die Korrektheit des Ursprungs treffen. Wird dagegen ein anderer Ursprung für den Netzbereich in der Datenbank gefunden, ist der Ursprung ungültig. Anhand dieses Ergebnisses wird dann der Routeauswahlprozess (vgl. Abschnitt 3.2.4) beeinflusst [BA11][Wä11].

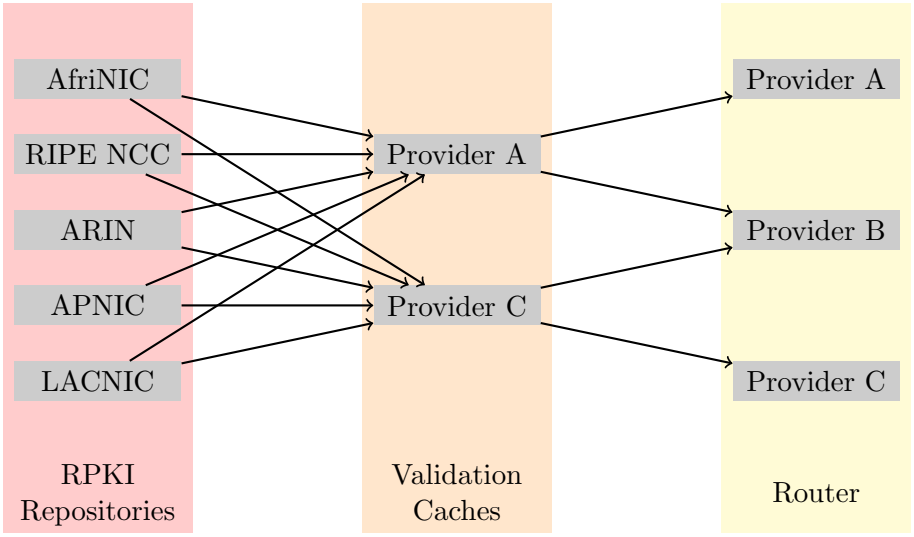

Abbildung 4.6: Architektur beim RPKI/Router Protokoll

4.2.4.3 Fazit

Das RPKI/Router Protokoll zielt lediglich auf die Überprüfung des Ursprungs von Routinginformationen. Es finden dabei keinerlei Veränderungen am BGP Protokoll statt, daher kann die Authentizität des Ursprungs nicht gesichert werden. Es kann jedoch die Korrektheit der Ursprungs-AS und Netzbereich Beziehung kontrolliert werden. Folglich kann das RPKI/Router Protokoll nicht vollständig gegen Manipulierung des Ursprungs schützen (vgl. Abschnitt 3.3.1.1). Somit ist eine Umlenkung des Traffic weiterhin möglich.
Allerdings lässt sich das RPKI/Router Protokoll mit im Vergleich zu sBGP und soBGP geringerem Aufwand einsetzen, da keine Anpassung des BGP Protokolls nötig ist. Bestehende Router könnten durch ein Softwareupdate nachgerüstet werden. So bietet Hardwarehersteller Cisco bereits eine Beta Version seines Betriebssystem IOS für einige Hardwareplattformen an. Auch das Aufsetzen eines Local Validation Cache ist dank bereits bestehender Implementierungen mit geringem Aufwand möglich (vgl. Abschnitt 7.4.1).
Weit wichtiger als die tatsächlichen Sicherheitseigenschaften des Protokolls könnte die Tatsache sein, dass eine Implementierung des RPKI/Router Protokoll für gängige BGP Router vorliegt und für weitere geplant ist. Dies könnte dazu führen, dass jene Netzverantwortliche ihre Netzbereiche zertifizieren lassen. Das RPKI/Router Protokoll könnte daher eine Übergangstechnologie zu einem abgesicherten BGP Protokoll werden. Das ist von der SIDR Arbeitsgruppe bewusst so angedacht und mittels der Erweiterung BGPSec, die im nachfolgenden Abschnitt behandelt wird, stellen sie dazu den Entwurf einer solchen BGP Erweiterung vor [BA11] [LK11].

4.2.5 BGPSec

Auf den Grundlagen, die das RPKI/Router Protokoll schafft, baut die SIDR Arbeitsgruppe der `IETF` mit der BGP Erweiterung BGPSec auf. Sowohl Ursprung als auch der AS_PATH soll damit kryptographisch gesichert werden. BGPSec basiert auf dem RPKI System und setzt zur Bestimmung von gültigen Ursprungs-ASen auf ROAs (vgl. dazu Abschnitte 4.2.1.2 und 4.2.4.1). Zur Absicherung der TCP Verbindung selbst empfiehlt BGPSec den Einsatz von IPSec (vgl. Abschnitt 4.1.3) oder TCP-AO (vgl. Abschnitt 4.1.2) [BBW11].
Zu Beginn einer BGP Session überprüft ein BGPSec Router die Fähigkeiten seines Peers. Falls beide BGPSec unterstützen, wird es für diese Session aktiviert. Ein Grundsatz von BGPSec ist, dass nur korrekt signiert empfangene Routinginformationen auch beim Weitergeben wieder signiert werden. `UPDATE` Nachrichten mit falscher Signatur werden verworfen, während unsignierte Nachrichten auch nach der Weiterleitung unsigniert bleiben. Dadurch bilden sich Inseln, in denen Routinginformationen gesichert übertragen werden können (vgl. Abbildung 4.7). Diese Inseln können zusammen mit den entsprechender BGPSec-Verbreitung wachsen. Eine Signierung wird in einem nicht-transitiven Attribut übertragen (vgl. dazu Abschnitt 3.2.3.2), daher kann keine Signatur eine dieser BGPSec-Inseln verlassen. Auf Abbildung 4.7 kann deswegen die Insel F-G keine gesicherten Routinginformationen mit der Insel A-C-E austauschen. Obwohl Router H BGPSec unterstützt kann dieser keine gesicherte BGP Session aufbauen, da kein benachbarter Router BGPSec beherrscht. Eine signierte `UPDATE` Nachricht trägt immer nur ein einzelnes Prefix im NLRI. Außerdem werden verworfene Routen nicht signiert.

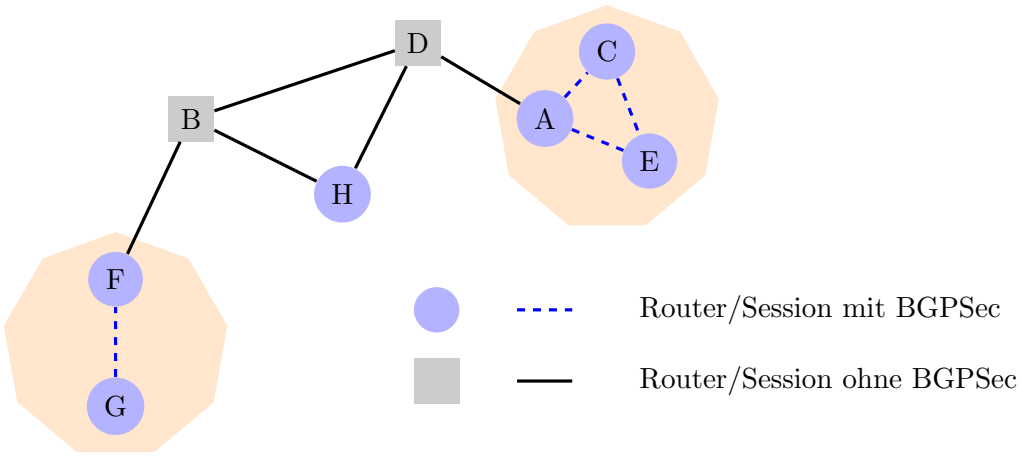

Abbildung 4.7: Mögliche Topologie bei BGPSec

Die Router können dabei entweder das Zertifikat des entsprechenden ASes verwenden oder es wird je Router ein eigenes End Entity Zertifikat (EE Zertifikat) ausgestellt. Die eingesetzte Variante sollte je nach Größe des ASes bei einem Vergleich zwischen zusätzlichen Aufwand beim Betreiben einer Zertifizierungsstelle gegenüber Sicherheitsüberlegungen abgewägt werden. [BBW11]

4.2.5.1 Signierung von Updates

Bei Erstellung der Signatur muss zwischen weitergeleiteten und neu erstellten `UPDATE` Nachrichten unterschieden werden. In Abbildung 4.8 sind diese zwei Arten von `UPDATE` Nachrichten abgebildet. Sie beziehen sich auf die Beispieltopologie aus Abbildung 3.10.

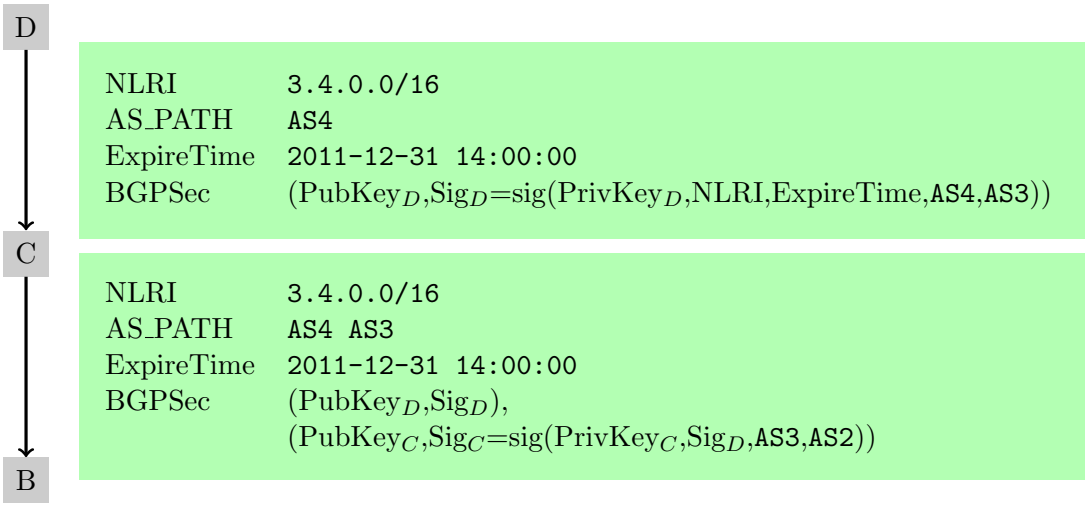

Abbildung 4.8: Signierte BGPSec `UPDATE` Nachrichten

Wenn eine korrekt signierte Routinginformation weitergereicht wird, muss die Signatur des vorhergehenden ASes, die eigene AS Nummer und die AS Nummer des Nachrichtenempfängers wiederum signiert werden. Die dadurch entstehende Signatur wird an sämtliche empfangene Signaturen angehängt und als Pfadattribut in der `UPDATE` Nachricht übertragen. Die ist in Abbildung 4.8 (C→B) auch klar zu erkennen. Die Expire Time einer empfangen BGPSec Nachricht wird dabei unverändert weitergereicht.
Soll dagegen ein neues Netz für das lokale AS angekündigt werdenm, muss zuerst eine Expire Time angegeben werden. Diese gibt das Ende der Gültigkeit der Bekanntmachung an. Genauer beschrieben wird der Zweck dieses Feldes in Abschnitt 4.2.5.3. Anschließend wird die Nachricht ähnlich wie zuvor signiert. Anstatt der vorherigen Signatur, die in dem Fall nicht existiert, wird die Expire Time und der NLRI verwendet. Dieser Fall ist in Abbildung 4.8 (D→C) dargestellt. [Lep11]

4.2.5.2 Validierung von Updates

Geht auf einem BGPSec Router eine signierte `UPDATE` Nachricht ein, wird die Korrektheit dieser durch mehrere Schritte sichergestellt. Dazu ist es nötig, dass ein BGPSec Router Zugriff auf alle gültigen Zertifikate als auch auf ROA-Dokumente besitzt. Dafür sollte auch hier ein Validation Cache eingesetzt werden. Dieser holt und prüft die Zertifikate und ROAs regelmäßig und stellt diese den Routern zur Verfügung. Damit alle nötigen Informationen übertragen werden können, muss das RPKI/Router Protokoll dafür noch angepasst werden (vgl. Abschnitt 4.2.4.2).
Zu Beginn prüft der Router das BGPSec Pfadattribut auf syntaktische Gültigkeit. Anschlie-

ßend werden die einzelnen Signaturen der Router geprüft.
Dazu wird mit der Signatur des Ursprungs AS begonnen. Diese Signatur muss durch ein gültiges Zertifikat des angegebenen ASes verifiziert werden können. Zusätzlich muss für den entsprechenden Netzbereich ein ROA-Dokument vorliegen, das dem AS die Bekanntmachung erlaubt. Außerdem darf die Expire Time nicht in der Vergangenheit liegen. Nun werden für jedes weitere AS im AS_PATH die signierten Angaben und die Gültigkeit der Signatur geprüft. Sind alle Prüfungen erfolgreich, wird die `UPDATE` Nachricht als gültig gekennzeichnet. Falls jedoch eine Prüfung fehlschlägt, wird die Route als fehlerhaft gekennzeichnet. Dieses Ergebnis kann dann im Routeauswahlprozess weiterverwendet werden (vgl. Abschnitt 3.2.4).[Lep11]

4.2.5.3 Unterschiede im Vergleich zu sBGP

Da BGPSec dem früheren Entwurf sBGP sehr ähnelt, wird in diesem Abschnitt nochmals speziell auf die Unterschiede zwischen beiden Sicherungsmechanismen eingegangen.
Die Verbreitung erfolgt bei sBGP über transitive Pfadattribute. Da auch nicht sBGP fähige Router diese weiterreichen, wirken sich die Folgen der Einführung auf das gesamte BGP Netz und daher auch auf das Internet aus. Eine Sicherung der Routinginformationen wird jedoch nur für sBGP Router möglich und auch nur auf Pfaden, auf denen alle Router sBGP einsetzen. Daher setzt BGPSec auf nicht-transitive Pfadattribute. Auf Grund dieser Verbreitungsweise bilden BGPSec Router Inseln im Internet in denen Routinginformationen gesichert ausgetauscht werden können. Benachbarte Router, die kein BGPSec unterstützen, werden deshalb von der Einführung von BGPSec nicht beeinflusst.
BGPSec setzt im Gegensatz zu sBGP auf das sog. „Beaconing“ von `UPDATE` Nachrichten. Dies bezeichnet das regelmäßige Neuankündigen von Netzbereichen, ohne dass sich die Routinginformationen geändert haben müssen. Daher tragen `UPDATE` Nachrichten eine Expire Time. Nach deren Ablauf wird die Nachricht ungültig. Dadurch kann die Zeitspanne verringert werden, in der zwischengespeicherte signierte `UPDATE` Nachrichten durch Replay erneut bekanntgegeben und validiert werden können. Dies erhöht zwar den Verkehr an `UPDATE` Nachrichten, kann jedoch das Ausnutzen von zwischengespeicherten Nachrichten auf einen Zeitraum beschränken. [Sri11]

4.2.5.4 Fazit

BGPSec stellt einen zuverlässigen Mechanismus dar, die Umlenkung von Datenverkehr zu verhindern (vgl. Abschnitt 3.3.3.1), da durch die vollständige Validierung des Pfades von Routinginformationen keine Manipulationen mehr möglich sind. Dies ist jedoch ebenfalls wie bei sBGP nur mit häufigem Einsatz von kryptographischen Operationen möglich. BGPSec Pfadattribute werden von nicht BGPSec fähigen Routern nicht weitergeleitet. Daher werden die nicht durch zusätzliches Datenaufkommen belastet. Durch eine Einführung von BGPSec bei wenigen großen Providern könnte die Routing Sicherheit für eine Vielzahl von daran angeschlossenen ASen verbessert werden. [Sri11]
Die erhöhten Hardwareanforderungen an Speicher und CPU Zeit für die kryptographischen Verfahren sind durch spezialisierte Hardwareanpassungen zu leisten. Da die Entwickler von mehreren Jahren bis zum ersten produktiven Einsatz von BGPSec ausgehen, sollte Hardware mit speziellen Kryptographie-Bausteinen und erhöhten Arbeitsspeicher bis dahin in der Lage sein, über die nötigen Ressourcen für einen globalen Einsatz von BGPSec im Internet zu

verfügen.
BGPSec könnte also die bestehenden Sicherheitsprobleme des BGP Protokolls lösen und für sicheres Inter-Domain Routing sorgen. [HB11]

4.3 Gegenüberstellung der Lösungsansätze

In diesem Teil der Arbeit werden die Ergebnisse der vorhergehenden Betrachtung noch einmal im Zusammenhang betrachtet. Dies verschafft einen besseren Überblick über die mögliche Maßnahmen zur Erhöhung der Sicherheit beim Betrieb von BGP. Auch hier wird die Unterscheidung zwischen Maßnahmen für die Basisprotokolle und für BGP selbst beibehalten.

4.3.1 Lösungen für TCP/IP

Bei den Sicherheitsverbesserungen für TCP/IP stellt sich die Situation ziemlich klar dar. IPSec (vgl. Abschnitt 4.1.3) bietet dabei die beste Lösung für die Sicherheitprobleme dieser Protokolle. Während es mit wählbaren kryptographischen Verfahrenen die Integrität und wahlweise auch die Vertraulichkeit der übertragenen Daten sichert, wird auch der verwendete Sitzungsschlüssel periodisch neu ausgehandelt. Somit kann IPSec, wie auch in Tabelle 4.1 zu erkennen ist, alle Sicherheitslücken von TCP/IP abdecken. Lediglich der (D)DoS durch Nutzdaten kann nicht verhindert werden. Es sollte daher, sofern es durch beide Peers technisch unterstützt wird, auf einer BGP Verbindung zur Absicherung des BGP Datenverkehrs immer IPSec eingesetzt werden. Beim Einsatz von ESP wird zusätzlich noch die Vertraulichkeit der Routinginformationen sichergestellt.
Falls es aus technischen oder administrativen Gründen nicht möglich ist IPSec einzusetzen, sollte auf TCP-MD5 bzw. TCP-AO zurückgriffen werden (vgl. Abschnitt 4.1.2). Diese Lösung kann ebenso die Integrität der TCP Pakete sichern und vor (D)DoS auf den BGP Dienst schützen. Da es kein automatisiertes Schlüsselmanagement gibt, sollten die Empfehlungen zur Beschaffenheit des Schlüssels aus RFC 3562 beachtet werden.
Der Einsatz von GTSM ist nicht zu empfehlen, da es lediglich den Kreis der möglichen Angreifer einschränken kann. Es kann damit weder die Integrität noch die Verfügbarkeit zuverlässig geschützt werden.
Als Schutz vor Überlast empfiehlt sich zusätzlich zum Einsatz von IPSec oder TCP-MD5/ TCP-AO ein QoS Mechanismus. Dieser muss dafür sorgen, dass immer eine ausreichende Bandbreite zum BGP Austausch zur Verfügung steht.

	GTSM	TCP-MD5 bzw. TCP-AO	IPSec	QoS
TCP-Reset / Session Hijacking	bedingt	ja	ja	nein
Vertraulichkeit	nein	nein	ja	nein
Dienst DDoS	bedingt	ja	ja	nein
Nutzdaten DDoS	nein	nein	nein	ja

Tabelle 4.1: Lösungen für Schwachstellen in TCP/IP

4.3.2 Lösungen für BGP

Die Lösungansatze bei BGP sind allesamt im Entwurfsstadium. Die Weiterentwicklung von soBGP und sBGP wurde zu Gunsten von BGPSec aufgegeben. Aktuell arbeitet die `IETF` die Entwürfe zu BGPSec und dem RPKI/Router Protokoll weiter aus, um diese als RFCs veröffentlichen zu können.
Für eine sichere Überprüfbarkeit des Besitzes von Internet Ressourcen ist eine RPKI obligatorisch (vgl. Abschnitt 4.2.1.2). Damit ist diese auch Grundlage der Absicherung von BGP. Die Eigenschaften dieser Erweiterungen sind in Tabelle 4.2 zusammengefasst.
Die Erweiterungen sBGP (vgl. Abschnitt 4.2.2) und BGPSec (vgl. Abschnitt 4.2.5) sind sich sehr ähnlich. Bei beiden wird durch den Einsatz von Signaturen bei jeder `UPDATE` Nachricht der Ursprung und sowie der Pfad verifizierbar. Dies ist möglich, da durch Zertifikate die Signatur als gültig für bestimmte AS Nummern erkannt werden kann. Der wesentliche Unterschied zwischen den beiden Protokollen liegt im Verhalten im Zusammenhang mit BGP Routern ohne Sicherheitserweiterung. Während bei sBGP die Signaturen auch über nicht sBGP fähige Router weitergeschickt werden, tauscht BGPSec lediglich Signaturen mit BGPSec fähigen Routern aus.
Bei soBGP steht der Ansatz des Web of Trust (WoT) im Mittelpunkt (vgl. Abschnitt 4.2.3. Dabei gibt jeder Teilnehmer dieses Netzes an, welchen anderen Zertifikaten er vertraut. Über diese Vertrauensverhältnisse wird versucht ein anderes Zertifikat zu überprüfen. Anhand von signierten Informationen über den Aufbau der Netze wird eine Topologiedatenbank von gültigen Beziehungen zwischen ASen aufgebaut. Jede `UPDATE` Nachricht wird nun gegenüber dieser Datenbank auf Plausibilität geprüft. Dabei kann jedoch die Korrektheit des Pfades nicht garantiert werden.
Das RPKI/Router Protokoll (vgl. Abschnitt 4.2.4) bietet lediglich eine Möglichkeit eingehende Routinginformationen auf Korrektheit der Netzbereich-AS Beziehung zu prüfen. Dabei kann die Umlenkung von Nutzdaten nicht wirksam verhindert werden, da weder Pfad noch Authentizität des Ursprungs AS überprüft werden. Die korrekten Beziehungen werden mittels ROAs ermittelt und durch die RPKI verifiziert.
Für den praktischen Betrieb werden also nur BGPSec und das RPKI/Router Protokoll relevant werden, da die anderen Protokolle nicht mehr weiterentwickelt werden. Dabei kann das RPKI/Router Protokoll unbeabsichtigte Konfigurationsfehler entdecken, jedoch nicht vor einem gezielten Angriff schützen. Dies wird erst einige Jahre später mit BGPSec möglich sein, da dafür eine neue Generation von Routern nötig sein wird. Diese werden dann über die entsprechenden Hardwareressourcen verfügen, die die Vielzahl an kryptographischen Operationen erfordert.

	sBGP	soBGP	RPKI/Router	BGPSec
Überprüfung von AS-Netz Beziehung	ja	ja	ja	ja
Pfadvalidierung	ja	eingeschränkt	nein	ja
Authentifizierung von benachbarten ASen	ja	ja	nein	ja

Tabelle 4.2: Lösungen für Schwachstellen in BGP

4.3.3 Tabellarischer Überblick über die Lösungsansätze

Die Ergebnisse der Betrachtung von Sicherheitslösungen im Zusammenhang mit BGP sind nun in Tabelle 4.3 aufgeführt. Diese Tabelle dient als Grundlage für die Erstellung eines Leitfadens für die Sicherheit des Inter-Domain Routing am LRZ in Kapitel 6.

Lösungansatz	Ziel	Einsatz empfohlen
GTSM	TCP/IP	Nein
TCP-MD5 bzw. TCP-AO	TCP/IP	Ja, sofern Einsatz von IPSec nicht möglich ist
IPSec	TCP/IP	Ja
QoS	TCP/IP	Ja
sBGP	BGP	Nein
soBGP	BGP	Nein
RPKI/Router	BGP	Ja, als Übergangstechnologie bis BGPSec einsetzbar
BGPSec	BGP	Ja, sobald stabile Implementierungen verfügbar sind

Tabelle 4.3: Lösungen für Schwachstellen in BGP

5 Inter-Domain Routing am LRZ

In diesem Kapitel wird nun auf die lokalen Begebenheiten des Münchner Wissenschaftsnetzes (MWN) eingegangen. Dieses vom Leibniz-Rechenzentrum (LRZ) betriebene Netz verbindet die Universitäten und andere wissenschaftliche Einrichtungen im Großraum München sowohl untereinander als auch mit dem Internet.
In Abschnitt 5.1 wird genauer auf die Struktur des Netzes eingegangen und die verwendete Hardware angegeben. Anschließend erfolgt in Abschnitt 5.2 eine Betrachtung der BGP Instanzen im MWN.

5.1 Netzüberblick

Kernbereich des MWN ist der Backbone. Er besteht aus einer ringförmigen Verbindung der großen Areale im MWN. Der Ring verläuft dabei über diese Strecken: LRZ - Campus Garching - LMU Stammgelände - Campus Großhadern - TUM Stammgelände - LRZ. Eine zusätzliche Verbindung besteht zwischen dem Stammgelände der LMU und dem der TUM, wie auch aus Abbildung 5.1 hervorgeht. Durch diese redundanten Verbindungen, kann die Verfügbarkeit des Netzes erhöht werden.
Im Backbone kommen Router vom Typ Cisco Catalyst 6509 zum Einsatz. Diese Router verfügen über 10-Gigabit-Ethernet Glasfaser Verbindungen untereinander. Der Standort von einzelnen Routern aus dem Backbone ist in Tabelle 5.1 zu entnehmen.

Name	Standort	Besonderheiten
csr1-2wr	LRZ Garching	Anbindung an das DFN
vss1-2wr	LRZ Garching	Redundanter virtueller Router
csr1-kw5	Campus Garching	
csr1-0gz	Stammgelände LMU	Glasfaserstrecken und Transitverbindung M-net
csr1-0q1	Weihenstephan Campus	
csr1-kic	Großhadern Campus	
csr1-kb1	Stammgelände TUM	Glasfaserstrecken Telekom
csr2-kb1	Stammgelände TUM	
csr1-kra	FH München	

Tabelle 5.1: Router im Backbone des MWN

Die Hauptanbindung an das Internet erfolgt in Garching durch einen Anschluss an das Deutsche Forschungsnetz (DFN). Es werden dabei zwei getrennte physikalische 10-Gigabit Verbindungen genutzt, die am Router csr1-2wr gebündelt sind. Dadurch wird einerseits die

nutzbare Bandbreite andererseits die Ausfallsicherheit erhöht. Damit eine Internetanbindung trotz Ausfall des DFNs gewährleistet werden kann, steht am Router csr1-0gz eine Gigabitverbindung zum lokalen Provider M-net zur Verfügung. Somit handelt es sich beim MWN um ein Multihomed AS mit zwei Transitprovidern (vgl. Abschnitt 2.1.1).

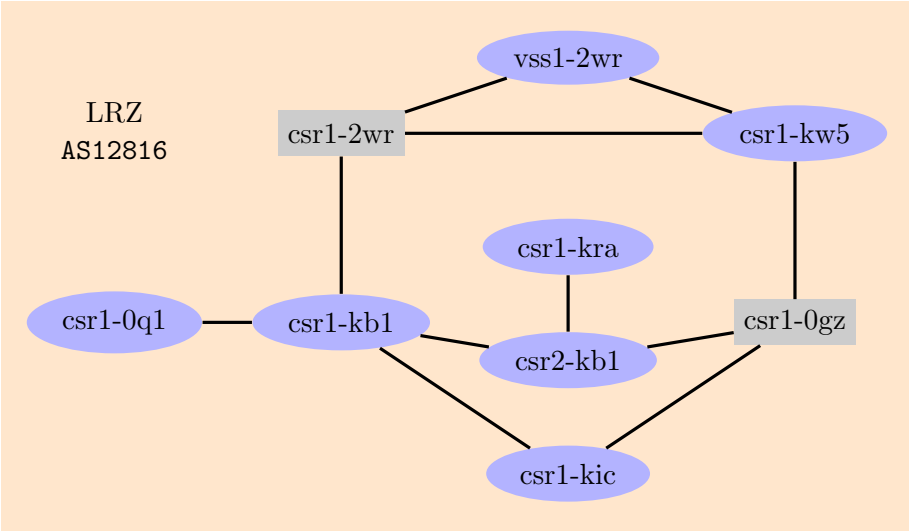

Abbildung 5.1: Topologie des MWN Backbone

An die Backbone-Router werden die zu versorgenden wissenschaftlichen Einrichtungen angeschlossen. Dies geschieht entweder anhand von eigenen Leitungswegen oder durch angemietete Leitungen. Während die angemieteten Glasfaserstrecken von M-net am Router csr1-0gz am LMU Stammgelände terminieren, sind die von der Deutschen Telekom bereitgestellten am Stammgelände der TUM an den Router csr1-kb1 bzw. csr2-kb2 angeschlossen. Zur Verbreitung von Routen innerhalb des MWN wird OSPF eingesetzt. Für die Anbindung an die Transitprovider DFN und M-net wird BGP verwendet. Detailliert wird das BGP Setup des MWNs in dem nachfolgenden Abschnitt behandelt.
Durch diese Maßnahmen werden also Redundanzen geschaffen, die dabei helfen sollen das MWN möglichst zuverlässig zu betreiben:

- Redundante Wege im Backbone durch Ringtopologie
- Internetanbindung an das DFN durch zwei physikalische Leitungen
- Backup-Internetanbindung an M-net
- BGP zur Umschaltung zwischen den Internetanbindungen
- OSPF zur Wegsuche im Backbone

5.2 BGP am LRZ

Damit das LRZ am globalen Inter-Domain Routing teilnehmen kann, ist es nötig, dass auf den Routern BGP eingesetzt wird. Im MWN werden auf den Routern csr1-2wr und csr1-0gz

jeweils über eBGP Routingdaten mit den Transitprovidern DFN bzw. M-net ausgetauscht. Die genauen Einzelheiten dazu sind in Abschnitt 5.2.1 zu finden. Die Weiterverbreitung im Netz des LRZ wird in Abschnitt 5.2.2 beschrieben.

Netzbereich	Verwendung durch/für
129.187.0.0/16	TUM/LRZ
131.159.0.0/16	Institut für Informatik der TUM
138.244.0.0/15	Medizin Grosshadern der LMU
138.246.0.0/16	LRZ
141.40.0.0/16	Hochschule Weihenstephan
141.84.0.0/16	LMU
192.54.42.0/24	TUM
192.55.197.0/24	Lehrstuhl für Integrierte Schaltungen
192.68.211.0/24	LRZ
192.68.212.0/22	Garchinger Technologie- und Gründerzentrum
2001:4ca0::/32	Gesamtes MWN
2002::/16	6to4

Tabelle 5.2: Vom LRZ bekanntgegebene Netzbereiche

5.2.1 Transitanbindung durch eBGP

Der Routinginformationsaustausch für die Hauptinternetanbindung an das DFN wird über den Router csr1-2wr abgewickelt. Als Backupverbindung dient die M-net Anbindung über Router csr1-0gz. Damit im Normalzustand alle Verbindungen über den DFN Uplink abgewickelt werden, wird der Local Preference Wert von Routen, die über M-net empfangen werden, verringert. Ebenso vermindert M-net die Local Preference der Routen, die direkt vom LRZ empfangen werden.
Die Integrität der ausgetauschten BGP Pakete wird über TCP-MD5 gesichert. Allerdings wird der dabei verwendete Schlüssel nicht nach den Empfehlungen von RFC 3562 gewechselt (vgl. Abschnitt 4.1.2). Denn ein solcher Schlüsselwechsel muss manuell erfolgen und verursacht zusätzlich eine Unterbrechung der BGP Session. Um bei einer Überlast von Nutzdatenverkehr den BGP Austausch zu gewährleisten wird durch QoS dieser bevorzugt behandelt. Dies kann die Wahrscheinlichkeit eines Abbruches bei einer solchen Situation bedeutend verringern (vgl. Abschnitt 4.1.4).
In den nachfolgenden Abschnitten wird auf die Details des eBGP Setups am LRZ bei IPv4 bzw. IPv6 eingegangen.

5.2.1.1 IPv4

Durch die beiden IPv4 eBGP Sessions auf den Routern csr1-2wr und csr1-0gz wird jeweils nur eine Default Route empfangen. Da mit der Default Route keine eigenständigen Routingentscheidungen getroffen werden können, wird dabei dem Routing der Transitprovider vertraut. Alle anderen eingehenden Netzankündigungen werden durch Filterregeln verworfen. Es werden keine empfangenen Routen weitergeleitet, sondern lediglich die eigenen Netze

angekündigt (siehe Tabelle 5.2).

5.2.1.2 IPv6

Bei IPv6 erhält das LRZ durch die Transitprovider jeweils eine vollständige Routingtabelle. Es kann dadurch vom LRZ Einfluss auf das Routing genommen werden. Angekündigt werden den Providern durch das LRZ die Netze aus Tabelle 5.2. Das Netz `2001:4ca0::/32` ist dabei ausschließlich dem LRZ zugewiesen. Der Netzbereich `2002::/16` wird nach dem Anycast Prinzip angekündigt. Er wird für den IPv6 Tunnelmechanismus 6to4 benötigt. Durch das Anycasting wird das Netz durch mehrere ASe angekündigt. Dies führt dazu, dass Pakete für dieses Zielnetz immer das nächste AS erreichen.

Aktion	Netzbereich	Netzgröße	Beschreibung
Verbiete	`2001:DB8::/32`	`32-128`	Beispielprefix
Erlaube	`2001:500::/30`	`40-48`	ARIN Microallocations für kritische Infrastrukturen
Erlaube	`2001:504::/30`	`48`	ARIN Microallocations für Internet Exchanges
Erlaube	`2620::/23`	`40-48`	ARIN PI Space
Erlaube	`2001:678::/29`	`40-48`	RIPE PI Space
Erlaube	`2001:7F8::/32`	`48`	RIPE Microallocations für Internet Exchanges
Erlaube	`2001:7FA::/32`	`48`	APNIC Microallocations für Internet Exchanges
Erlaube	`2001:C00::/23`	`40-48`	APNIC PI Space
Erlaube	`2001:43F8::/29`	`40-48`	AfriNIC PI Space
Erlaube	`2001:13C7:6000::/35`	`40-48`	LACNIC Microallocations für Internet Exchanges
Erlaube	`2801::/24`	`40-48`	LACNIC PI Space
Erlaube	`2002::/16`	`16`	6to4 Netzbereich
Verbiete	`2002::/16`	`16-128`	6to4 länger als /16
Erlaube	`2000::/3`	`12-36`	zugewiesene Netzbereiche an LIR
Verbiete	`::/0`	`0-128`	alle Netzbereiche

Tabelle 5.3: Filterregeln bei eingehenden IPv6 Routinginformationen

Die eingehende vollständige Routingtabelle wird durch die Regeln aus Tabelle 5.3 gefiltert. Dabei werden grundsätzlich alle eingehenden Netze, die nicht Teil von `2000::/3` sind und eine Netzgröße von `/12-36` haben, verworfen. Da die RIRs jedoch in einigen Netzbereichen für PI Space oder auch Internet Exchanges kleinere Netze erlauben, werden diese durch entsprechende Regeln akzeptiert. Es muss dabei beachtet werden, dass durch neue Vergaberichtlinien von RIRs auch die Regeln ggf. angepasst werden müssen.

5.2.2 Weiterverbreitung von Routen innerhalb des MWN

Zum Routingaustausch innerhalb des MWN wird zum einen iBGP benutzt. Auf den genauen Aufbau des Setups wird im nachfolgenden Abschnitt eingegangen. Des weiteren wird die über BGP empfangene IPv4 Default Route über OSPF weiterverbreitet.

5.2.2.1 iBGP

Im MWN unterscheidet sich die Topologie des iBGP Setups zwischen IPv4 und IPv6. Während bei IPv4 die beiden Border Router csr1-2wr und csr1-0gz jeweils eine iBGP Sessions zu den übrigen Routern halten und somit einen Route Reflector darstellen, wird bei IPv6 eine „fully meshed“ Topologie im Backbone verwendet (vgl. dazu auch Abschnitt 3.2.1).
Die Integrität der BGP Daten wird dabei nicht über TCP-MD5 sichergestellt. Jedoch wird der Backbone-Bereich durch VLANs vom restlichen Netz abgetrennt. Über iBGP werden von den entsprechenden Routern die zugehörigen öffentlichen IP Netze angekündigt. Diese Ankündigungen werden von den Border Routern empfangen und nach entsprechender Filterung ggf. aggregiert und über eBGP weitergegeben.

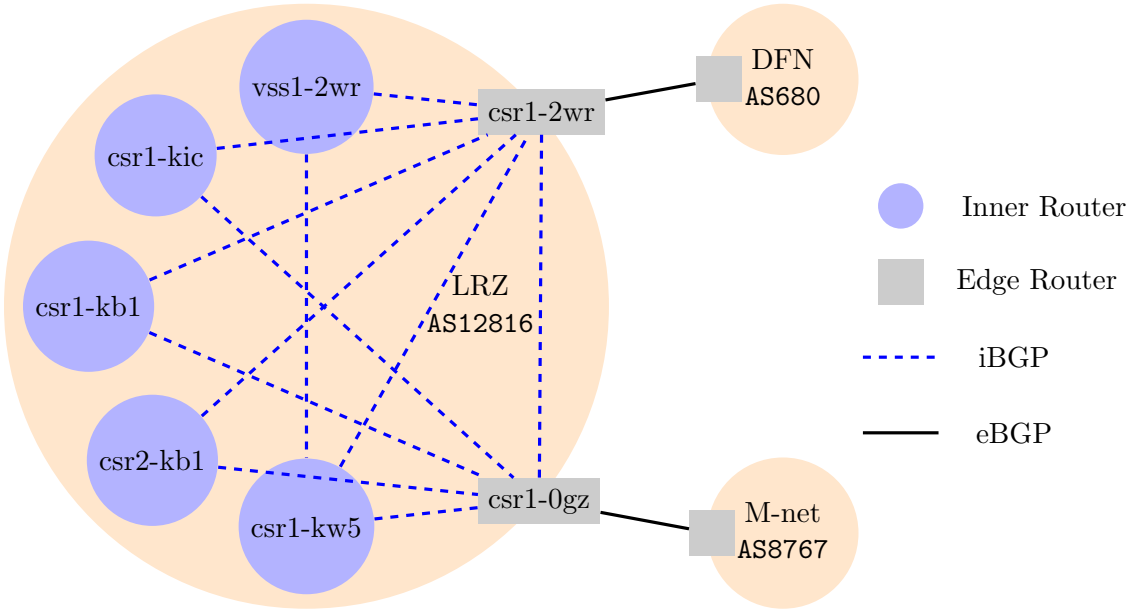

Abbildung 5.2: Übersicht von BGP Sessions bei IPv4 im MWN

5.2.2.2 OSPF

Durch Weiterverbreitung der über eBGP empfangenen IPv4 Default Routen über OSPF werden diese auf den Backbone Routern verteilt. Dadurch wird sichergestellt, dass der externe Datenverkehr auch bei Ausfall von Backbone Verbindungen zu den Border Router transportiert werden kann. Bei IPv6 ist dies nicht nötig, da dort die iBGP Verbindungen „fully meshed“ ausgeführt sind.

6 Leitfaden zur Absicherung von BGP am LRZ

In diesem Kapitel werden die Lösungsmöglichkeiten beschrieben, die für die Verbesserung des Inter-Domain Routing am LRZ empfohlen werden. Diese Vorschläge resultieren aus der Anwendung der Ergebnisse von Kapitel 4 auf die Ausgangslage am LRZ. Die Sicherheitsverbesserungen sind in der Reihenfolge aufgeführt in der sie implementiert werden sollten. Eine prototypische Implementierung wird in Kapitel 7 dargelegt.
In Abschnitt 6.1 werden die Netze des LRZ zertifiziert. Dadurch ist es möglich dass sich das LRZ als berechtigter Inhaber der Netzblöcke ausweisen kann. Diese Zertifizierung ist nötig, damit das LRZ, wie in Abschnitt 6.2 beschrieben, ROA-Dokumente herausgeben kann, anhand derer die Korrektheit der Ursprungsangabe bei Netzankündigen von Netzen des LRZ durch andere ISPs überprüft werden kann.
Im Abschnitt 6.3 wird mit IPSec ein Sicherheitsmechanismus für IP- und TCP-Schwachstellen empfohlen. IPSec sichert diese ab und verhindert dadurch eine Manipulation des Datenaustausches bei BGP.
Eine Möglichkeit die Netzbereiche des LRZ in den globalen Routingtabellen zu überwachen, wird in Abschnitt 6.4 vorgestellt. Der Webdienst `BGPmon.net` kontrolliert dabei an vielen verschiedenen Standorten, ob dort in der Routingtabelle die Netze des LRZ Unregelmäßigkeiten zeigen.
Mit dem RPKI/Router Protokoll wird in Abschnitt 6.5 die Einführung von Origin Validation auf den Routern des LRZ empfohlen. Zum Betrieb von wird ein entsprechender Validation Cache benötigt.

6.1 Zertifizierung der Internet Ressourcen des LRZ

Die Zertifizierung der Ressourcen muss nach der Verteilungshierarchie von Internet Ressourcen erfolgen. Somit ist für die Zertifizierung der Ressourcen des LRZ die RIPE zuständig. Diese bietet zwei verschiedene Varianten an: Local Certification Service und Hosted Certification Service. Die Varianten unterscheiden sich darin, wer die jeweils zu dem Ressource Zertifikat zugehörige CA betreibt. Während dies beim Hosted Certification Service von der RIPE übernommen wird, muss sich beim Local Certification Service der jeweilige Ressourceninhaber selbst um den Betrieb kümmern.

6.1.1 Local Certification Service

Zum Betrieb einer lokalen Zertifizierungsstelle hat die RIPE eine Java Applikation Local Certificate Authority LCA entwickelt. Durch diese wird die Erstellung von nachgelagerten Ressource Zertifikaten ermöglicht, die über ein öffentliches Repository zur Verfügung gestellt werden. Diese Aktionen können über ein Webinterface abgewickelt werden. Leider bietet diese Software noch nicht den vollständigen Umfang an Funktionen, die für einen produktiven

Betrieb nötig wären. Es müssen viele Vorgänge, die automatisiert ablaufen könnten, aktuell noch händisch durchgeführt werden. Dazu gehört z.B. die Erneuerung von Zertifikaten nach Ablaufdatum, so wie die Ausstellung von ROA-Dokumenten.
Eigentlich wäre der Betrieb dieser CA lokal am LRZ wünschenswert, da dann nur dort die privaten Schlüssel der zugehörigen Zertifikate vorliegen würden. Weil die Software noch nicht vollständig implementiert ist, stellt dies derzeit keine Option dar. Außerdem ist es für das LRZ nicht nötig Zertifikate für Ressourcen nachgelagerter Institutionen (wie z.B. LMU,TU,...) auszustellen, da diese selbst kein Inter-Domain Routing betreiben. Sobald die Software über eine vollständige Funktionalität verfügt, wäre es erstrebenswert die LCA Software am LRZ einzusetzen. Daher ist die Installation und Erstkonfiguration in Abschnitt 7.1 beschrieben.

6.1.2 Hosted Certification Service

Eine Alternative zur LCA stellt der Hosted Certification Service durch die RIPE dar. Dabei wird sowohl das Zertifikatsrepository als auch die zu den Ressourcen zugehörige CA durch die RIPE betrieben. Dies führt zu dem Nachteil, dass der private Schlüssel der CA bei der RIPE vorliegt.
Die Verwaltung dieser Zertifizierung kann über das LIR-Portal durchgeführt werden. Die Erstellung der Ressource Zertifikate ist in Abschnitt 7.2.1 beschrieben.

6.1.3 Zertifizierbare Ressourcen

Zu den Internet Ressourcen des LRZ zählen neben der AS Nummer `AS12816` auch die Netze aus Tabelle 5.2. Da alle IPv4 Adressbereiche zugeteilt worden waren bevor die heutige Ressourcenvergabehierarchie existierte, handelt es sich um Netze aus dem Legacy Space.
Später wurde dieser Legacy Bereich mit Netzgrößen von `/8` durch die IANA einzelnen RIRs zugeteilt (vgl. dazu [IAN11]). Die Netze aus `141.0.0.0/8` werden demnach von der RIPE verwaltet, während für alle weiteren IPv4 Netze des LRZ die ARIN zuständig ist. Da die RPKI der ARIN jedoch noch nicht im Produktivbetrieb zur Verfügung steht, können diese Ressourcen nicht zertifiziert werden. Die Zuteilung des IPv6 Netzbereiches erfolgte bereits nach den heutigen Strukturen. Obwohl die beiden Netze `141.40.0.0/16` und `141.84.0.0/16` durch die RIPE verwaltet werden, können diese derzeit nicht durch das LRZ zertifiziert werden, da sie dem LIR Kundenkonto des LRZ nicht zugeordnet sind. Daher kann aktuell nur für das bereits nach der Vergabehierarchie zugeteilte IPv6 Netz `2001:4ca0::/32` ein Zertifikat ausgestellt werden.

6.2 ROA für Netze ausstellen

Nachdem die Ressourcen des LRZ nun zertifiziert sind, kann mittels eines ROA-Dokuments Einfluss auf das Routing genommen werden. Die Erstellung eines solchen Dokumentes anhand des LIR-Portals ist in Abschnitt 7.2.2 beschrieben.
Für das LRZ ist lediglich ein ROA-Dokument nötig. Dieses wird in Abbildung 6.1 dargestellt. Darin werden der AS Nummer des LRZ erlaubt, die zertifizierten Netze bekanntzugeben. Da die Netze vom LRZ in voller Größe angekündigt werden, entspricht die maximale Prefixlänge im ROA der Prefixlänge des jeweiligen Netzes. Dies verhindert die Deaggregation des Netzbereiches.

version	0		
asID	12816		
ipAddrBlocks	Family	Address	MaxLen
	IPv6	2001:4ca0::/32	/32

Abbildung 6.1: ROA-Dokument des LRZ

6.3 Absicherung von TCP/IP Gefahren

Das LRZ setzt auf den eBGP Sessions zur Absicherung von Gefahren, die sich aus den Basisprotokollen TCP/IP ergeben, TCP-MD5 ein. Durch die Schwächen von TCP-MD5 (vgl. Abschnitt 4.1.2) sollte stattdessen IPSec (vgl. Abschnitt 4.1.3) eingesetzt werden. Dies erfordert zum einen entsprechende Unterstützung der jeweiligen Gegenseiten beim DFN bzw. bei M-net Zum anderen sind die aktuell dafür eingesetzten Router Cisco Catalyst 6509 ohne entsprechende Hardwareunterstützung ausgestattet. Dies würde eine erhöhte CPU-Last nach sich ziehen, die die Zuverlässigkeit der BGP Session gefährdet. Aus diesem Grund ist eine Verwendung von IPSec in dieser Konstellation nicht empfohlen. Zu Verbesserung der Sicherheit bis eine neue Routergeneration eingesetzt wird, könnte der Schlüssel von TCP-MD5 nach den Kriterien von RFC 3562 erneuert werden. Dies würde eine Erneuerung nach höchstens 90 Tagen erforderlich machen. Allerdings muss der Schlüsselwechsel manuell auf beiden Routern durchgeführt werden. Da der Schlüsselwechsel einerseits mit dem entsprechenden Provider der Gegenseite koordiniert werden muss und andererseits die bestehende BGP Session unterbricht, stellt dies keine praktikable Lösung dar. Daher wird, obwohl es für die Sicherheit beim Betrieb von BGP wünschenswert wäre, von einem regelmäßigen Schlüsselwechsel abgeraten.
Auf den iBGP Sessions wird derzeit kein Absicherungsmechanismus betrieben. Dort wird ebenfalls der Einsatz von IPSec nach Anschaffung neuer Routerhardware empfohlen. Bis dahin sollte dort TCP-MD5 verwendet werden.

6.4 Monitoring

Aktuell werden die globalen Routingtabellen nicht dahingehend überwacht, ob die Netzbereiche des LRZ darin vorkommen bzw. ob sie durch Dritte missbräuchlich verwendet werden. Daher sollte der Dienst von `BGPmon.net` in Anspruch genommen werden. Die nötigen Schritte zur Konfiguration des Dienstes sind in Abschnitt 7.3 erläutert.
`BGPmon.net` überprüft die Sicht auf die globale Routingtabelle von mehreren Standorten aus. Dabei werden die Daten aus dem RIS Projekt der RIPE verwendet (Standorte siehe [RIP11]). Über das Webinterface des Dienstes können die zu überwachenden Netzbereiche bzw. ASe angegeben werden. Diese werden auf Unregelmäßigkeiten überwacht und sofern welche entdeckt werden wird man unverzüglich per E-Mail darauf hingewiesen. Der kostenlose Dienst kann folgende Unregelmäßigkeiten (vgl. Abschnitt 3.3.3.1) erkennen:

- AS kündigt einen weiteren Netzbereich an
- Netzbereich mit falschem Ursprung

- Überprüfung der Korrektheit des Pfades über regulären Ausdruck
- Netzankündigung wurde verworfen
- Spezifischer Netzbereich als konfiguriert werden angekündigt

Durch diesen Dienst können Unregelmäßigkeiten der eigenen Ressourcen in den BGP Daten erkannt werden und ggf. manuelle Gegenmaßnahmen eingeleitet werden.

6.5 Origin Validation mit dem RPKI/Router Protokoll

Eine wirksame Abwehr gegenüber `UPDATE` Nachrichten mit gefälschtem Ursprung stellt das RPKI/Router Protokoll dar (vgl. Abschnitt 4.2.4.2). Für dessen Betrieb wird ein Validation Cache benötigt, welcher in periodischen Abständen die ROAs der RPKI auswertet. Dieser Cache stellt dann entsprechend ausgestatteten Routern die Informationen der ROAs zur Verfügung. Die Installation und Konfiguration eines solchen Caches ist in Abschnitt 7.4.1 beschrieben.
Damit ein Router diese Caches über das RPKI/Router Protokoll abfragen kann, ist es nötig, dass dessen Software das unterstützt. Für die beim LRZ eingesetzten Router Cisco Catalyst 6509 existiert noch keine entsprechende Software. Sobald diese verfügbar ist, kann das RPKI/Router Protokoll anhand der Beschreibung in Abschnitt 7.4.6 konfiguriert werden. Beim LRZ sollten dann die Border Routern csr1-0gz und csr1-2wr das RPKI/Router Protokoll einsetzen. Eine Absicherung ist dabei nur für die IPv6 Tabelle möglich, da das LRZ bei IPv4 lediglich eine Default Route erhält.
Solange noch keine aktualisierte Software zur Verfügung steht, kann durch Erstellung von statischen Filtern versucht werden unerlaubte `UPDATE` Nachrichten zu unterbinden. Diese werden mit Hilfe eines Skriptes aus den Informationen der ROA-Dokumente generiert und anschließend auf den Border Routern händisch eingespielt. Dazu wurde ein Prototyp dieses Skriptes als Ergänzung zum Leitfaden entwickelt. Da das LRZ lediglich bei IPv6 eine vollständige Routingtabelle hält, hat auch nur dort der Einsatz dieser Filter Sinn. Das Erzeugen der Filterregeln ist in Abschnitt 7.4.4 erläutert. Die Filter prüfen mittels der maximalen Prefixlänge im ROA-Dokument, ob eingehende `UPDATE` Nachrichten spezifischer sind. Sofern das der Fall ist, werden entsprechende Nachrichten verworfen.

7 Prototypische Implementierung des Leitfadens

In diesem Kapitel wird das Vorgehen bei der prototypischen Implementierung der Empfehlungen des Leitfadens aus Kapitel 6 beschrieben. Die durchzuführenden Schritte werden dabei erläutert und deren Auswirkungen dokumentiert.

7.1 Local Certification Service

In diesem Abschnitt wird beschrieben, wie die Software Local Certification Authority (LCA) der RIPE installiert und grundlegend konfiguriert wird. Diese Software dient dazu eine lokale Ressourcen CA zu betreiben. Da die Software sich z.Z. noch in der Entwicklung befindet, deckt sie noch nicht alle Anforderungen des LRZ ab (vgl. Abschnitt 6.1.1). Zu einem späteren Zeitpunkt soll diese Software jedoch produktiv eingesetzt werden.

7.1.1 Installation

Für die Installation wird ein SUSE Linux Enterprise Server in der Basis Installation vorausgesetzt. Das Programm Rsync wird zum Betreiben des Repositories benötigt. Außerdem benötigt die LCA Software eine Java Virtual Machine von Oracle. Die nötigen Schritte sind hier aufgeführt:

```
# Rsync installieren
zypper install rsync
# Oracle JAVA installieren
wget http://javadl.sun.com/webapps/download/AutoDL?BundleId=56691
    -O java.bin
chmod +x java.bin
./java.bin
ln -s /usr/java/jre1.6.0_29/bin/java /usr/bin


## Verzeichnisse erstellen
# Programmverzeichnis
mkdir -p /usr/local/rpki/
# Oeffentliches Repository
mkdir -p /var/local/rpki/repository
# Lokale Datenbank
mkdir -p /var/local/rpki/data/db

# Herunterladen und Entpacken der Software
```

```
cd /usr/local/rpki
wget "https://certification.ripe.net/content/public-repo/releases/
    net/ripe/lca/rpki-lca/1.0.2/rpki-lca-1.0.2-bin.zip"
unzip rpki-lca-1.0.2-bin.zip
ln -s rpki-lca-1.0.2 rpki-lca

# Kopieren der Standard Konfigurationen
cd rpki-lca
cp config/env_vars.example config/env_vars
cp config/lca.properties.example config/lca.properties
```

Listing 7.1: Installation des Local Certification Services

Anschließend muss in der Datei `/usr/local/rpki/rpki-lca/config/env_vars` das Verzeichnis und der Port für das öffentliche Repostiory gesetzt werden:

```
#!/bin/bash
CERTREPO_DIR=/var/local/rpki/repository
RSYNCD_PORT=873
```

Listing 7.2: /usr/local/rpki/rpki-lca/config/env_vars

In der Datei `/usr/local/rpki/rpki-lca/config/lca.properties` muss das Datenverzeichnis korrekt gesetzt werden:

```
# Specify the path to a directory and database name
# Make sure that this dir exists!
embedded.db.storage.dir=/var/local/rpki/data/db
```

Listing 7.3: /usr/local/rpki/rpki-lca/config/lca.properties

Außerdem wird die Beschreibung des Repositories (comment) in `/usr/local/rpki/rpki-lca/config/rsyncd.conf.template` noch angepasst:

```
uid = nobody
gid = nobody
use chroot = no
timeout = 600
dont compress = *
pid file = @RUNTIME_DIR@/rsyncd.pid
lock file = @RUNTIME_DIR@/rsyncd.lock

[lca]
comment = TESTING: LRZ RPKI Repository
path = @CERTREPO_DIR@/published
read only = true
```

Listing 7.4: /usr/local/rpki/rpki-lca/config/rsyncd.conf.template

Anschließend wird zum Starten der Dienste eine Datei `/etc/init.d/lca` mit diesem Inhalt erstellt:

```
#!/bin/sh
### BEGIN INIT INFO
# Provides: lca
# Required-Start: $network $syslog
# Required-Stop: $network $syslog
# Default-Start: 3 5
# Default-Stop: 0 1 2 6
# Description: RIPE Local Resource Certificate
### END INIT INFO

# Starten der Dienste
/usr/local/rpki/rpki-lca/bin/rsyncd_ctl.sh $1
/usr/local/rpki/rpki-lca/bin/lca_ctl.sh $1
```

Listing 7.5: /etc/init.d/lca

Mit diesen Befehlen werden die Dienste schließlich aktiviert:

```
# Startskript ausfuehrbar machen
chmod +x /etc/init.d/lca

# Startskript beim Systemstart ausfuehren
chkconfig --add

# Dienste jetzt starten
/etc/init.d/lca start
```

Listing 7.6: Start des Local Certification Services

7.1.2 Konfiguration und Zertifizierung der Ressourcen

Nun kann die Erstkonfiguration bequem über das Webinterface durchgeführt werden. Dazu sind drei Schritte nötig. Das Webinterface ist unter der URL `http://localhost:8082/lca` zu erreichen. Am besten man verbindet sich zu dieser Seite durch einen SSH Tunnel:

```
ssh -L 8082:localhost:8082 username@rpkitest.srv.lrz.de
```

Im Schritt eins werden Informationen zum Repository abgefragt. Dieser ist in Abbildung 7.1 dargestellt. Diese Informationen müssen eingetragen werden:

- Hostname: rpkitest.srv.lrz.de
  Dieser Domainname sollte sowohl in eine IPv4 als auch in eine IPv6 Adresse aufgelöst werden können
- Port: 873
  Standard Port des Rsync Dienstes
- Module: lca
  Standard Name des Repositories

- Base directory: /var/local/rpki/repository
  Verzeichnis des Repositories

Durch einen Klick auf `Save Configuration` gelangt man zu Schritt zwei (vgl. Abbildung 7.2). Dort muss bei **1.** die Datei identity.xml heruntergeladen werden. Diese enthält Informationen sowohl zur Authentifizierung der RIPE gegenüber dem LCA als auch den Hostnamen und Port des Repositories.
Nun muss der Link zum LIR-Portal der RIPE bei **2.** geöffnet werden. Nachdem man sich dort erfolgreich angemeldet hat, muss die identity.xml aus Schritt eins hochgeladen werden. Anschließend werden Daten zur Authentifizierung des LCA gegenüber der RIPE erstellt und können in der Datei issuer-identity.xml heruntergeladen werden. Jetzt kann man sich am LIR-Portal wieder abmelden und gelangt nun durch einem Klick auf `Next` zu Schritt drei.
Dort muss die Datei issuer-identity.xml hochgeladen werden (vgl. Abbildung 7.3). Mit einem Klick auf `Upload and request resource certificate` wird die nun ein Ressourcen Zertifikat bei der RIPE beantragt und im LCA eingerichtet.

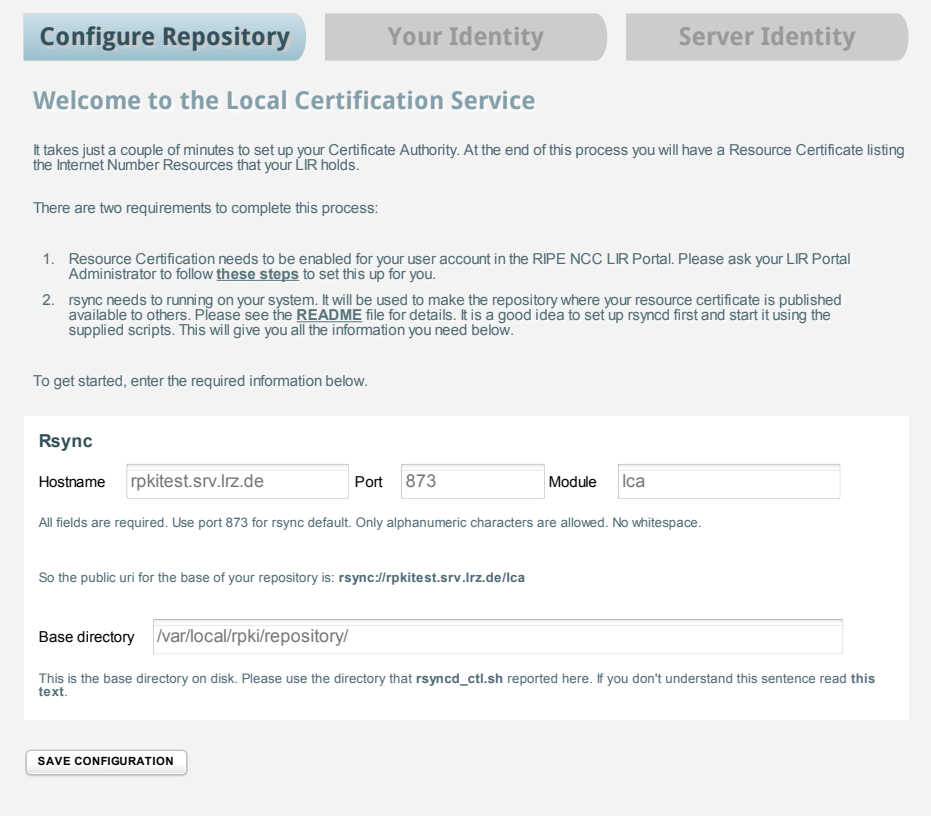

Abbildung 7.1: Konfiguration des Repositories

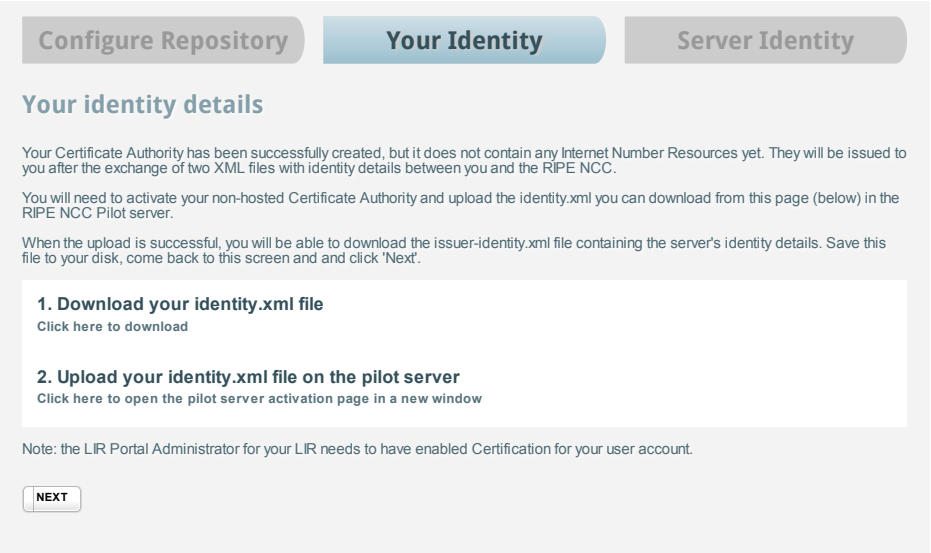

Abbildung 7.2: Authentifizierungsdaten für die RIPE

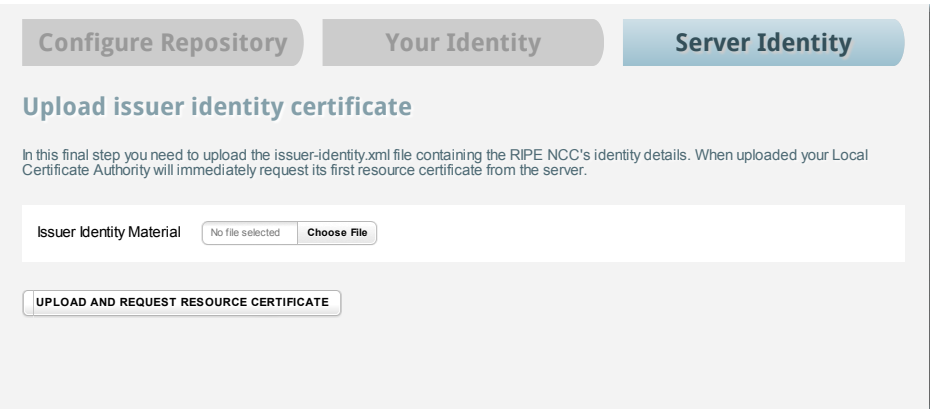

Abbildung 7.3: Authentifizierungsdaten für die LCA

7.2 Hosted Certification Services

Dieser Abschnitt erläutert, wie man bei der RIPE über den Hosted Certification Service bestehende Internet Ressourcen zertifiziert. Außerdem wird beschrieben wie ein zugehöriges ROA-Dokument erzeugt wird.

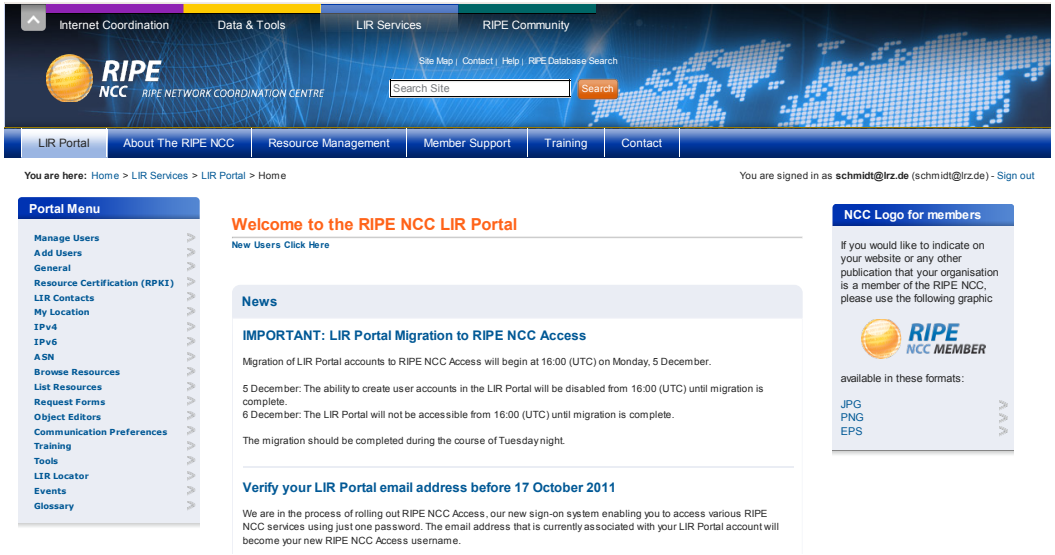

Abbildung 7.4: LIR-Portal der RIPE

7.2.1 Erstellung der Zertifikate

Die Hosted Certification Services lassen sich über das LIR-Portal der RIPE nutzen. Das Portal, das auch in Abbildung 7.4 dargestellt ist, ist unter `https://lirportal.ripe.net` zu erreichen. Zuerst muss über den Admin-Login ein entsprechend berechtigter Benutzer der Gruppe Certification hinzugefügt werden. Sobald dieser Benutzer sich im LIR-Portal anmeldet, erscheint nun der Menüpunkt „Resource Certification (RPKI)“.

Certificate Authority Name	CN=de.lrz
Certified Resources	2001:4ca0::/32

Abbildung 7.5: Zertifizierte Ressourcen

Wird dieser Menüpunkt nun aufgerufen, muss zunächst den Nutzungsbedingungen des Hosted Certification Services zugestimmt werden. Anschließend werden automatisch die entsprechenden Ressourcen zertifiziert. Unter „My Certified Resources“ können diese eingesehen werden (vgl Abbildung 7.5)

7.2.2 ROA-Dokument erstellen

Jetzt kann für das zertifizierte Netz ein ROA-Dokument erstellt werden. Dazu muss aus dem Menü die Seite „My ROA Specifications“ aufgerufen werden. Mit einen Klick auf „Add ROA Specification“ gelangt man zu dem ROA Formular, das auch in Abbildung 7.6 abgebildet ist. Dort muss entsprechend der Abbildung als Ursprungs AS das `AS12816` eingetragen werden. Nun müssen zugehörige Netze aus „My certified resources“ mit der Maus in das Formular gezogen werden. Anschließend muss noch die maximale Prefixlänge eingetragen werden, die der Länge der Bekanntmachung, nämlich `32` entspricht. Ein Klick auf „Add ROA“ erstellt diesen.

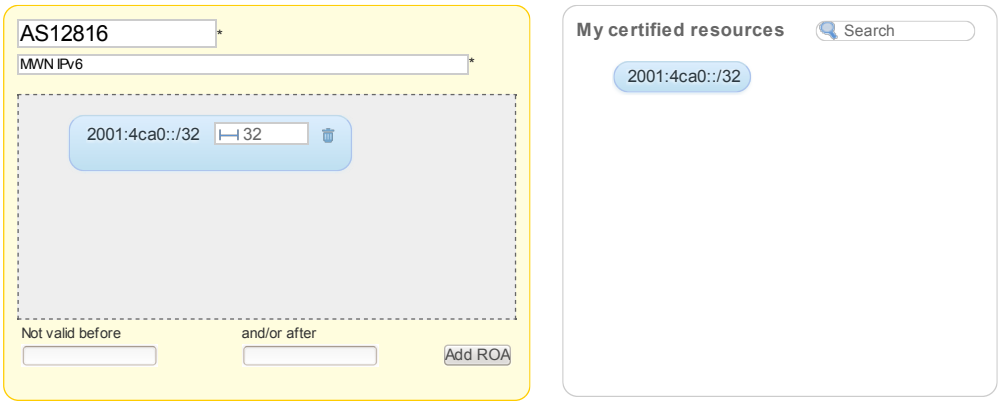

Abbildung 7.6: ROA Record wird erstellt

Unter „My ROA Specifications“ sollte man nun im Bereich „Current BGP announcements“ das Feld „Route Validity“ betrachten. Dieses sollte nun, sofern die ROA Erstellung erfolgreich war, für die entsprechenden Netzbereiche `VALID` wie in Abbildung 7.7 anzeigen.

Origin AS	Prefix	Route Validity
AS12816	2001:4ca0::/32	VALID

Abbildung 7.7: Status von BGP Ankündigungen

7.3 Monitoring mit BGPmon.net

Dieser Abschnitt beschreibt die Einrichtung eines Monitorings, dass die Netzbereiche des LRZ in mehreren Sichten auf die globalen Routingtabellen überwacht. Sobald Unregelmäßigkeiten entdeckt werden, wird man von dem kostenlosen Dienst über automatisierte E-Mail Benachrichtigungen auf diese hingewiesen. Zuerst muss ein Konto auf der Internetseite (`http://www.bgpmon.net`) des Dienstes erstellt werden. An die bei dieser Registrierung hinterlegte E-Mail Adresse werden die Benachrichtigungen versandt. Diese Adresse wird bevor das Konto erstellt wird, durch eine Bestätigungsnachricht getestet.
Nach dem ersten Anmelden müssen die Netze des LRZ hinzugefügt werden. Dazu reicht es aus die AS Nummer des LRZ, wie in Abbildung 7.8 dargestellt, anzugeben. Nach einem Klick auf „Auto Detect from BGP data“ werden die Netze des LRZ aus den BGP Daten bestimmt.

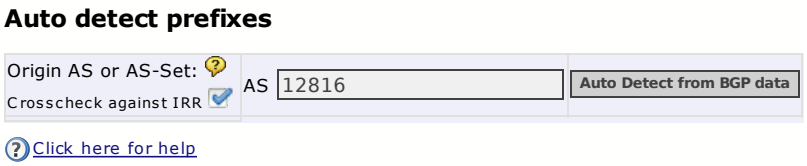

Abbildung 7.8: BGPMon: „Auto detection“

Die entdeckten Netzbereiche werden wie in Abbildung 7.9 angezeigt. Diese Netzbereiche sollten mit der Tabelle 5.2 übereinstimmen. Nach einem Klick auf „Bulk submit“ werden diese ins Monitoring übernommen.

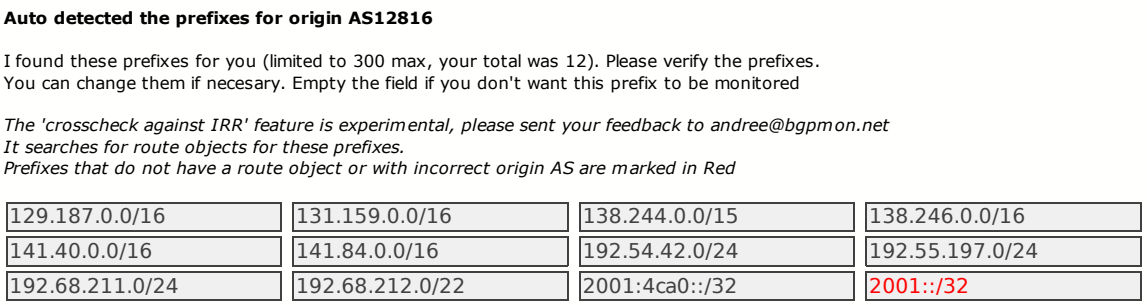

Abbildung 7.9: BGPMon: Resultate der „Auto detection“

Die gerade importierten Netzbereiche, werden bereits auf Ankündigungen von spezifischeren Netzbereichen bzw. falschen Ursprungsangaben überwacht. Damit zusätzlich überprüft werden kann, dass das erste Transit-ASe im AS-Pfad stimmt, müssen diese bei jedem Prefix hinzugefügt werden. Nach einem Klick auf „edit“ muss auf der folgenden Seite das Feld „Upstream AS“ mit den AS Nummern der beiden Transitprovider ausgefüllt werden: `680 8767`. Nun sollte die Übersicht auf der Seite „My Prefixes“ der Abbildung 7.10 entsprechen. Damit ist eine Überwachung aller LRZ Netze in den BGP Tabellen an einer Vielzahl von Internetknoten eingerichtet. Die entsprechenden Alarmmeldungen, die an die bei der Registrierung angegebene Adresse versandt werden, sollten an die zuständigen Stellen am LRZ

weitergeleitet werden.

My Prefixes

Remove	Edit	Prefix	Ignore More Specifics	Origin AS	Upstream AS	AS path Regex	Email alert setting	notify on withdraw	Minimum peer Threshold	Must Match
☐	Edit	129.187.0.0/16	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	131.159.0.0/16	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	138.244.0.0/15	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	138.246.0.0/16	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	141.40.0.0/16	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	141.84.0.0/16	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	192.54.42.0/24	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	192.55.197.0/24	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	192.68.211.0/24	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	192.68.212.0/22	No	AS12816	680 8767		Inherit	No	1	
☐	Edit	2001:4ca0::/32	No	AS12816	680 8767		Inherit	No	1	
Delete selected										

Abbildung 7.10: BGPMon: Fertig konfigurierte „My Prefixes“

7.4 Einrichten von Origin Validation

7.4.1 Validation Cache

Es wird nun die Installation und Konfiguration eines Validation Cache beschrieben. Es kommt dazu die Software rcynic zum Einsatz, welche Teil der Programmsammlung Rpkid ist, die von Rob Austein entwickelt wurde. Damit der Cache Validator bei einer möglichen Kompromittierung das restliche System nicht gefährden kann, wird dieser in einer Chroot Umgebung unter einem nicht privilegierten Benutzer ausgeführt.

7.4.1.1 Installation von Rpkid

Auch für diese Installation wird ein SUSE Linux Enterprise Server in der Basis Installation vorausgesetzt. Zur Installation des Caches sind die nachfolgenden Schritte nötig:

```
# Installation der Vorraussetzungen
zypper install busybox python-devel subversion

# Programm zur Absicherung des System vor dem Validator (chroot)
# kompilieren und installieren
mkdir -p /usr/src/rpki/
cd /usr/src/rpki
wget http://ftp.de.debian.org/debian/pool/main/c/chrootuid/
   chrootuid_1.3.orig.tar.gz
tar xvfz chrootuid_1.3.orig.tar.gz
cd chrootuid-1.3
make clean
make
make install

## Rpkid herunterladen
cd /usr/src/rpki
svn co http://subvert-rpki.hactrn.net/trunk/ validator
cd validator

# Kompilieren
./configure --disable-django
make

# Rpkid Installieren
useradd -g rcynic -d /var/rcynic --system rcynic
make install
```

Listing 7.7: Installation von Rpkid

7.4.1.2 Einrichten der Chroot Umgebung

Mit diesen Befehlen wird die Chroot Umgebung angepasst, so dass alle nötigen Programme und Bibliotheken vorhanden sind:

```
# Kopieren von notwendigen Programmen und Bibliotheken
cp /lib64/libpam.so.0 /var/rcynic/lib64
cp /lib64/libpam_misc.so.0 /var/rcynic/lib64
cp /lib64/libreadline.so.5 /var/rcynic/lib64
cp /lib64/libncurses.so.5 /var/rcynic/lib64
cp /lib64/libresolv.so.2 /var/rcynic/lib64
cp /lib64/libselinux.so.1 /var/rcynic/lib64/
cp /lib64/libaudit.so.0 /var/rcynic/lib64/
cp /lib64/librt.so.1 /var/rcynic/lib64/
cp /lib64/libacl.so.1 /var/rcynic/lib64/
cp /lib64/libpthread.so.0 /var/rcynic/lib64/
cp /usr/bin/id /var/rcynic/bin/
cp /bin/ping /var/rcynic/bin/
cp /bin/sh /var/rcynic/bin/
cp /bin/ls /var/rcynic/bin/

# Teste Wechsel in das Chroot als Benutzer rcynic
/usr/local/bin/chrootuid /var/rcynic rcynic /bin/sh
```

Listing 7.8: Einrichtung und Test der Chroot Umgebung

Die Chroot Umgebung lässt sich nun mit dem Befehl aus Zeile 18 von Listing 7.8 testen. Der Befehl darf keine Fehler ausgeben und muss einen neuen Kommandozeilen-Prompt zeigen. Wird dort der Befehl `ls -l /` ausgeführt, sollte der Inhalt von Verzeichnis `/var/rcynic` gezeigt werden. Der Befehl `id` sollte diese Ausgabe ergeben:

```
uid=445(rcynic) gid=1000(rcynic) groups=1000(rcynic)
```

Wenn bei dieser Ausgabe weder die `uid` noch die `gid` 0 ist, war das Erstellen der Chroot Umgebung erfolgreich.

7.4.2 Konfiguration von Rcynic

Nun muss die Konfigurationsdatei von Rcynic editiert werden. Es wird dort durch Trust Anchors (TAs) definiert, welchen Stellen der Cache vertraut. Jeder TA stellt dabei eine Wurzel eines Zertifikatbaumes im RPKI System dar. Daher sollten darin nur die produktiven TAs der RIRs eingetragen werden. Außerdem werden in der Datei noch die für das Programm nötigen Pfade festgelegt. Es wird folgender Inhalt für die /var/rcynic/etc/rcynic.conf empfohlen:

```
[rcynic]
rsync-program           = /bin/rsync
authenticated           = /data/authenticated
old-authenticated       = /data/authenticated.old
unauthenticated         = /data/unauthenticated
lockfile                = /data/lock
jitter                  = 600
use-syslog              = true
log-level               = log_usage_err
```

```
10 xml-summary                  = /data/stats.xml

12 # Vertraute RPKI Repositories
   trust-anchor-locator.1   = /etc/trust-anchors/afrinic.tal
14 trust-anchor-locator.2   = /etc/trust-anchors/apnic.tal
   trust-anchor-locator.3   = /etc/trust-anchors/arin.tal
16 trust-anchor-locator.4   = /etc/trust-anchors/lacnic.tal
   trust-anchor-locator.5   = /etc/trust-anchors/ripe-ncc-root.tal
```

Listing 7.9: /var/rcynic/etc/rcynic.conf

7.4.3 Konfiguration des RPKI/Router Protocol Servers

Die Router können bisher auf die Daten des Validation Caches noch nicht zugreifen. Deshalb wird nun ein RPKI/Router Protocol Server installiert und eingerichtet. Dieser Server wird sowohl über TCP als auch über SSH erreichbar gemacht. Als Server wird das Python Skript Rtr-origin aus dem Rpkid Archiv eingesetzt. Die Konfiguration läuft so ab:

```
1 # Rtr-origin Datenverzeichnis erstellen
  mkdir -p /var/local/rpki/rtr-origin
3 # User erstellen und Rechte fuer Datenverzeichnis anpassen
  useradd -r -g nogroup -s /bin/bash -d /var/local/rpki/rtr-origin/
      rtr-origin
5 chown rtr-origin:nogroup -R /var/local/rpki/rtr-origin

7 # Installation von Xinet.d
  zypper install xinetd
```

Listing 7.10: /var/rcynic/etc/rcynic.conf

Damit der Server über SSH erreichbar ist, muss diese Zeile in der Datei `/etc/ssh/sshd_config` eingefügt werden:

```
  # Rtr-origin Subsystem
2 Subsystem        rpki-rtr        /usr/local/bin/rtr-origin
```

Listing 7.11: /etc/ssh/sshd_config

Anschließend muss für jeden Router noch der Public Key in der Datei `/var/local/rpki/rtr-origin/.ssh/authorized_keys` hinzugefügt werden. Dabei werden alle weiteren SSH Dienste deaktivert und es wird automatisch der RPKI/Router Protocol Server gestartet:

```
command="/usr/local/bin/rtr-origin --server /var/local/rpki/rtr-
    origin",no-port-forwarding,no-X11-forwarding,no-agent-
    forwarding,no-pty ###PubKey Router1###
```

Listing 7.12: /var/local/rpki/rtr-origin/.ssh/authorized_keys

Anschließend wird der Server Dienst zusätzlich über TCP von außen erreichbar gemacht. Dazu wird Port 42420 verwendet und Xinetd entsprechend konfiguriert:

```
# default: off
# description: An RPKI/Router Protocol server.

service rpki-rtr
{
        socket_type     = stream
        protocol        = tcp
        port            = 42420
        user            = rtr-origin
        wait            = no
        server          = /usr/local/bin/rtr-origin
        server_args     = --server /var/local/rpki/rtr-origin
        FLAGS           = IPv6 IPv4
}
```

Listing 7.13: /etc/xinetd.d/rpki-rtr

In der Datei `/etc/services` muss der Dienst durch die folgende Zeile noch ergänzt werden:

```
rpki-rtr          42420/tcp          # RPKI/Router Protocol
```

Listing 7.14: /etc/services

Schließlich muss noch durch Einträge in der `/etc/hosts.allow` sichergestellt werden, dass alle Routern Quelladressen sich mit den beiden Diensten verbinden können.

7.4.4 Filterregeln erzeugen

Als Erweiterung des Leitfadens dieser Arbeit wurde ein Skript entwickelt, welches aus den validierten ROA-Dokumenten des Caches statische Prefixfilterlisten für Cisco Router erzeugen kann. Durch diese Filterlisten werden Routen, die spezifischer sind als das ROA-Dokument erlaubt, ausgefiltert. Das Skript nutzt dabei die Python-Module von Rpkid. Zur Installation müssen die folgenden Schritte durchgeführt werden:

```
# Installiere Abhaengigkeiten
zypper install git python-setuptools python-lxml

# Installiere notwendiges Python Modul
easy_install ipaddr

# Installiere Prefixfilter-Skript + weitere Python Module
git clone git://github.com/simonswine/PyRPKI.git /usr/local/pyrpki
```

Listing 7.15: Installation des Prefixfilter-Skriptes

Durch diesen Befehl kann die Prefixfilterliste angezeigt werden:

```
# Generiere Filterliste
python /usr/local/pyrpki/source/build_prefixlist_from_rpki.py

# Ausgabe
```

```
! DATA from RPKI validated at lrz.de
! LAST UPDATE: 2011-12-10 15:56:16.568592
! Prefix 217.8.160.0/19 bis /19 Origins: AS6714
ip prefix-list ROA_FILTER seq 1005 permit 217.8.160.0/19 ge 20
! Prefix 89.187.96.0/19 bis /19 Origins: AS21371
ip prefix-list ROA_FILTER seq 1010 permit 89.187.96.0/19 ge 20
! Prefix 80.220.0.0/14 bis /24 Origins: AS1759
...
```

Listing 7.16: Generieren der Filterliste

Wird das Skript aus Abschnitt 7.4.5 erstellt und aktiviert, wird die Filterliste bei jeder Aktualisierung des Caches erneut generiert. Sie ist dann über die URL `http://rpkitest.srv.lrz.de/rpki_filter_more_specific.txt` abrufbar und muss händisch in die Router Konfiguration eingespielt werden. Anschließend kann man die Filterliste durch Anwenden auf die aktuelle BGP Tabelle auf den Routern testen:

```
cisco1#show bgp ipv6 unicast prefix-list ROA_FILTER
   Network                Next Hop            Path
*  2001:6D8::/35          2001:4CA0::5        12816 680 20965 8501 8267
   i
*  I2A03:1180::/33        2001:4CA0::5        12816 680 3356 12956 5610
   20884 23456 198002 i
*  I2A03:1180:8000::/33   2001:4CA0::5        12816 680 3356 12956 5610
   20884 23456 198002 i
```

Listing 7.17: Testen der Filterliste

7.4.5 Periodische Aktualisierung einrichten

Damit der Datenbestand aus der RPKI periodisch aktualisiert wird, muss unter dem Pfad `/usr/local/sbin/update_rpki.sh` folgendes Skript erstellt werden:

```
#!/bin/bash

# Daten rtr-orgin
RTR_DATA=/var/local/rpki/rtr-origin

# Setze Pfadvariablen
export PATH=$PATH:/usr/bin:/usr/local/bin:/usr/local/sbin #$

# Starte Rcynic chrooted, aktualisiere RPKI Repositories
chrootuid /var/rcynic rcynic /bin/rcynic -c /etc/rcynic.conf

# Erneuere Datenverzeichniss des RPKI/Router Protokoll Server
OWD=$(pwd)
cd $RTR_DATA
su rtr-origin -s /bin/bash -c 'PATH=$PATH:/usr/local/bin rtr-
   origin --cronjob /var/rcynic/data/authenticated/'
```

```
cd $OWD

# Erzeuge Statistiken von Rcynic Durchlauf
/usr/bin/python /usr/src/rpki/validator/rcynic/rcynic.py /var/
    rcynic/data/stats.xml > /srv/www/htdocs/stats.html

# Erzeuge prefix-filter aus RPKI ROAs
/usr/bin/python /usr/local/pyrpki/source/
    build_prefixlist_from_rpki.py > /srv/www/htdocs/
    rpki_filter_more_specific.txt
```

Listing 7.18: /usr/local/sbin/update_rpki.sh

Dieses Skript ruft zuerst Rcynic auf, wobei die Repositories der RIRs mit den lokalen Daten abgeglichen und anschließend validiert werden. Darauf wird mit einem Aufruf von Rtr-origin der Datenbestand des RPKI/Router Protocol Servers erneuert. Dann erzeugt das Skript noch Statistiken der Validierung von Rcynic. Diese sind in der Datei `/srv/www/htdocs/stats.html` einzusehen. Zuletzt werden noch die Filterregeln aus Abschnitt 7.4.4 in der Datei `/srv/www/htdocs/rpki_filter_more_specific.txt` aktualisiert. Beide Dateien sind auch im Browser über die URLs `http://rpkitest.srv.lrz.de/stats.html` bzw. `http://rpkitest.srv.lrz.de/rpki_filter_more_specific.txt` aufrufbar.
Mit den folgenden Eintrag in die Crontab des Benutzers root wird das Skript alle zwei Stunden gestartet:

```
0 */2 * * *     /usr/local/sbin/update_rpki.sh
```

Listing 7.19: Crontab für Benutzer Root

7.4.6 Konfiguration des RPKI/Router Protokolls am Router

Im nächsten Schritt werden nun die Router mit dem Validation Cache verbunden. Da die aktuelle Firmware Version der Cisco Catalyst 6509 Router des LRZ dies nicht unterstützt, wurde das mit Hilfe eines virtualisierten Routers mit Cisco IOS 15.2 getestet. Ein Validation Cache über TCP kann wie folgt im Konfigurationsmodus hinzugefügt werden:

```
router bgp 12816
 bgp rpki server tcp <IP des Caches> port 42420 refresh 120
```

Listing 7.20: Router: Konfiguration des Validation Caches

Dabei gibt der Refresh Parameter die Zeitdauern in Sekunden an, nach welcher der Router nach Aktualisierungen abfragen soll. Eine Konfiguration von SSH war mit dieser IOS Version (noch) nicht möglich. Der Router validiert in den Routingtabellen die Beziehung zwischen Ursprungs-AS und Netzbereich. Um das Ergebnis der Validierung anzusehen, können folgenden Befehle genutzt werden:

```
# Zeige Validierungsstatus der IPv6 Routingtabelle
router>show bgp ipv6 unicast summary | include RPKI
Path RPKI states: 210 valid, 5767 not found, 9 invalid

```

```
# Zeige alle ungueltigen IPv6 Eintraege
router>show bgp ipv6 unicast | include I
*  I2001:6D8::/35     2001:4CA0::5  0 12816 680 20965 8501 8267 i
...

# Zeige alle ungueltigen IPv6 Eintraege
router>show bgp ipv6 unicast | include V
*> V2001:468::/32     2001:4CA0::5  0 12816 680 20965 11537 i
...
```

Listing 7.21: Router: Diagnose der Verbindung zwischen Router & Cache

Im Testsystem, das über die Routingtabelle des LRZ verfügt, können also 210 Routen anhand von ROAs validiert werden. Es wurden 9 falsche Routen entdeckt. Die fehlerhaften Routen werden, wie die obigen Ausgaben zeigen, nicht für den Routeauswahlprozess herangezogen.

8 Fazit & Ausblick

Durch den Leitfaden aus Kapitel 6 kann das Inter-Domain Routing am LRZ nach den derzeitigen Rahmenbedingungen bestmöglich abgesichert werden. Dritte können nun die Netzankündigungen des LRZ anhand der veröffentlichten ROA-Dokumente überprüfen. Für einige Netzbereiche aus den IPv4 Netzen des MWNs ist dies aufgrund von fehlender Zertifizierung von Legacy Netzen nicht möglich. Dies sollte jedoch, sobald die Möglichkeit dazu besteht, nachgeholt werden. Durch gezieltes Monitoring werden die Netzbereiche des LRZ auf Unregelmäßigkeiten, die durch andere Teilnehmer am globalen BGP verursacht wurden, überwacht.
Eingehende IPv6 Routing Informationen können nach Freigabe einer entsprechenden Firmware durch Cisco von den Routern des LRZ auf deren korrekte Ursprungs-Angabe geprüft werden. Bei IPv4 ist dies aktuell nicht möglich, da dort nur Default Routen von den Providern empfangen werden.
Diese Maßnahmen, die in den Leitfaden aufgenommen wurden, sind diejenigen die kurz- bis mittelfristig umsetzbar sind. Daher können sie keine absolute Sicherheit bei der Korrektheit der Routinginformationen garantieren. Diese wird erst mit der BGP Erweiterung BGPSec ermöglicht, die eine vollständige Absicherung der Routinginformationen erlaubt. BGPSec kann dabei sowohl den Ursprung als auch den Pfad einer Routinginformation absichern. Dies wird durch den Einsatz von Signaturen für jede Zwischenstation einer Netzankündigung erreicht. Über diese Signaturen kann durch Zertifikate die Authentizität der entsprechenden Router überprüft werden, da den Zertifikaten über die RPKI entsprechende Internet-Ressourcen zugeordnet werden können. Das Verifizieren der Routinginformationen geht daher mit einer Vielzahl an kryptographischen Operationen einher. Da diese Operationen, wenn sie von Softwareimplementierungen ausgeführt werden, sehr viele Ressourcen beanspruchen, wird eine neue Hardwaregeneration von Routern benötigt. Diese verfügen dann über Krypto-Hardwarebausteine, die es ermöglichen, dass die kryptographische Operationen effizient und schnell ausgeführt werden können. Da aber BGPSec sich derzeit noch in der Entwurfsphase befindet, gibt es noch keine Implementierungen der Erweiterung. Daher wird es noch mehrere Jahre dauern bis BGPSec-Router mit der erforderlichen Qualität zur Verfügung stehen. Selbst wenn die Technik dann zur Verfügung steht, wird es nach meiner persönlichen Einschätzung noch einige Zeit lang dauern bis diese von den Provider eingesetzt wird. Diese Erfahrung konnte zumindest bei der Einführung von IPv6 gemacht werden. Obwohl die entsprechenden Produkte das neue Protokoll schon einige Zeit unterstützten, zögerten die Provider bei der letztendlichen Einführung.
Sobald diese Router dann zur Verfügung stehen, hat der Einsatz von BGPSec nur Sinn, wenn einer der Transitprovider ebenfalls diese Technik einsetzt. Daher sollte dann durch weitere Arbeiten in Zusammenarbeit mit dem DFN untersucht werden, inwiefern im DFN der Einsatz von BGPSec sinnvoll ist. Denn dann könnte das DFN eine Vorreiterrolle bei der Absicherung von Routinginformationen einnehmen. Zudem könnte dadurch die Zuverlässigkeit des Datenaustausches zwischen den angeschlossenen Forschungseinrichtungen verbessert werden.

Glossar

AA : Eine Address Attestation gibt bei sBGP bestimmten ASen das Recht einen entsprechenden Netzbereich anzukündigen. 42, 43, 47

`ACK` TCP Flag, das beim Datenaustausch zur Bestätigung von Paketen verwendet wird. 12, 13

`Adj-RIB-In` : Die Routingtabelle `Adj-RIB-In` enthält alle Routen, die von einem BGP Peer eingehen. 24

`Adj-RIB-In` : Die Routingtabelle `Adj-RIB-Out` enthält alle Routen, die an einem BGP Peer gesandt werden. 24, 25

AFI : Der Address Family Identifier identifiziert das verwendete Protokoll bei MBGP. 26, 27

AH : Der Authentication Header stellt Authentizität und Integrität von IP Paketen bei IPSec sicher. 37

ARPANET : Aus dem ARPANET entwickelte sich das heutige Internet. 17

AS Autonomes System: Ansammlung von Netzbereichen die unter einheitlicher Verwaltung stehen und Routinginformationen intern über eines oder mehrere IGPs austauschen. 7–9, 11, 12, 14, 17, 18, 25, 27, 28, 31–33, 42–51, 53, 58, 62, 63, 72, 79

Backbone bezeichnet den Kernbereich eines Netzwerkes. Dieser zeichnet sich i.A. durch hohe Bandbreiten und mehrfache Redundanzen aus. 17, 55, 56

Beaconing : Als Beaconing wird das regelmäßige Versenden von Informationen bezeichnet, ohne dass sich diese geändert haben müssen. 51

BGP Border Gateway Protokoll: EGP Routingprotokoll, das im Internets eingesetzt wird. 2, 4, 5, 8, 15, 17–19, 21–31, 33, 35–38, 41, 49, 51–54, 56, 57, 61

BGP Identifier : Der BGP Identifier ermöglicht eine eindeutige Identifikation eines Router anhand von einer IP Adresse des Systems. 18, 21

BGP Peer : Ein BGP Peer ist ein Router, der das BGP Protokoll einsetzt. 21, 26, 28, 29, 33

BGP Session : Eine BGP Session bezeichnet eine BGP Sitzung zwischen zwei Peers. 18, 24, 28–30, 35, 36, 38, 49, 57, 63

BGPSec : Die BGPSec Sicherheitserweiterungen wird von der SIDR Arbeitsgruppe der IETF entwickelt und soll BGP absichern. 49–54, 81

Blackholing bezeichnet das Verwerfen von Paketen. 27

Bogon : Als Bogons werden Netzbereiche bezeichnet, deren Verwendung im Internet (noch) nicht vorgesehen ist. 33

CA : Eine Certificate Authority stellt die Zertifikate in einer PKI aus. 38–40, 61, 62, 65

CIDR : Classless Inter-Domain Routing beschreibt ein Verfahren zur effizienten Aufteilung des Adressbereiches bei IP Netzen. 10, 12, 18

CRL : In der Certificate Revocation List werden widerrufene Zertifikate aufgeführt. 39, 45, 47

CSR : Mit einem Certificate Signing Request kann ein Zertifikat angefordert werden. 39

(D)DoS (Distributed) Denial of Service bezeichnet das beabsichtige Herbeiführen einer Überlast in Rechnernetzen. 28, 35–37, 52

Default Route : Eine Default Route gilt für den gesamten Netzbereich und wird somit immer dann verwendet, wenn keine spezifischere Route zur Verfügung steht. 4, 7, 8, 81

DFN : Das Deutsche Forschungsnetz betreibt ein Rechnernetz, das wissenschaftliche Einrichtungen in Deutschland untereinander und mit dem Internet verbindet. 55–57, 81

Diffie-Hellman : Beim Diffie-Hellman-Verfahren kann über einen unsicheren Kanal, ein Schlüssel sicher getauscht werden. 37

Eavesdropping bezeichnet das Mitscheinen von Pakten. 27

eBGP bezeichnet eine BGP Session zwischen verschiedenen ASen. 18, 24, 25, 57, 59

Edge Router verfügen über Datenverbindungen zu anderen ASen (auch Border Router genannt). 7, 17, 18

EE Zertifikat : Ein End-Entity Zertifikat wird an eine Instanz vergeben, die keine weiteren Zertifikate mehr vergeben muss. 49

EGP : Exterior Gateway Protokolle sind für den Routingdatenaustausch im Inter-AS Bereich zuständig. 14, 17, 23

ESP : Die Encapsulating Security Payload stellt Vertraulichkeit, Authentizität und Integrität von IP Paketen bei IPSec sicher. 37, 52

Expire Time : Die Expire Time bezeichnet den Zeitpunkt an dem die Gültigkeit der Information ausläuft. 50, 51

`FIN` TCP Flag, das beim Verbindungsabbau verwendet wird. 13, 29

Flapping : Unter Flapping versteht man den wiederholten schnellen Wechsel zwischen zwei Zuständen. 28

Flooding bezeichnet das „Überfluten“ eines Netzwerkes durch massenhaften versandt von Paketen. 14

GGP Gateway to Gateway Protocol: Historisches EGP Routingprotokoll nach dem Distanzvektorverfahren. 17

GTSM Generalised TTL security mechanism: Mechanismus, um Pakete von Hosts, die mehr als einen Hop entfernt sind, zu filtern. 35, 52, 54

Hoplimit ist das zur TTL äquivalente Feld bei IPv6. 35

Hops : Ein Hop bezeichnet den Weg von einem Netzwerkknoten zum nächsten. 15

IANA Internet Assigned Numbers Authority: Abteilung der ICANN, die sich um die Verwaltung der Internet-Ressourcen kümmert. 15, 27, 40, 44, 62

iBGP bezeichnet eine BGP Session innerhalb eines ASes. 18, 23–25, 46, 59

ICANN : Die Internet Corporation for Assigned Names and Numbers kümmert sich um die Verwaltung des Internets. 7, 15

ICMP : Das Internet Control Message Protocol dient zum Austausch weiterer Informationen beim Betrieb von IP. 29

`IETF` : Die Internet Engineering Task Force befasst sich mit der Weiterentwicklung der Techniken des Internets und stellt dazu eine Plattform bereit, um diese Standards zu definieren. 46, 49, 53

IGP : Interior Gateway Protokolle sind für den Routingdatenaustausch im Intra-AS Bereich zuständig. 14, 17, 23

IKE Internet Key Exchange ist ein Protokoll zum Schlüsselaustausch bei IPSec. 37

Inner Router verfügen ausschließlich über Datenverbindungen zu anderen Router des eigenen ASes. 7

Inter-AS bezeichnet Datenaustausch zwischen mehreren ASen. 7, 14, 17

Intra-AS bezeichnet Datenaustausch innerhalb eines AS. 7, 14, 17

IPSec Internet Protocol Security ist eine Sicherheitserweiterung für das IP Protokoll. 36–38, 41, 49, 52, 54, 61, 63

ISP Internet Service Provider: Dienstleister, der seinen Kunden durch technische Lösungen einen Zugang zum Internet verschafft. 7, 61

`KEEPALIVE` Mit der BGP `KEEPALIVE` Nachricht wird getestet, ob eine BGP Session immer noch besteht. 20, 30

LCA : Die Local Certification Authority ist eine Software zum lokalen Betrieb einer CA im RPKI System. 61, 62, 65, 68

Legacy Space bezeichnet Netzbereiche, die zugeteilt wurden bevor die heutige Vergabehierarchie für Internet-Ressourcen bestand. 16, 40, 62, 81

LIR Bei Local Internet Registries handelt es sich meist um ISPs, die als Mitglied einer RIR, Internet-Ressourcen beanspruchen. 16, 40, 58, 62

`Loc-RIB` : Die Routingtabelle `Loc-RIB` enthält die Routen, die vom Routeauswahlprozess ausgewählt wurden. 24, 25

Longest Prefix Match Router wählen bei Routingentscheidungen immer das spezifischste (=längste) Netz. 3, 32

LRZ : Das Leibniz-Rechenzentrum betreibt das zentrale Rechenzentrum der wissenschaftlichen Einrichtungen im Großraum München. 4, 55

MAC : Über einen Message Authentication Code kann die Integrität und Authentizität einer Nachricht sichergestellt werden. 37

MBGP bezeichnet BGP mit Multiprotocol Erweiterungen. 27

MITM Kann ein Angreifer die Kommunikation zwischen zwei Systemen mitlesen, verändern und unterdrücken, so ist dieser in der Lage sog. Man-in-the-Middle Angriffe durchzuführen. 29

Multicast bezeichnet die Datenkommunikation mit einer Quelle und mehreren Empfängern. 26

Multihoming : Beim Multihoming ist ein AS durch mehrere Anbindungen an das Internet angeschlossen. 8, 56

MWN : Das Münchner Wissenschaftsnetz verbindet wissenschaftliche Einrichtungen im Großraum München untereinander und mit dem Internet. 55, 56, 59

NLRI : Die Network Layer Reachability Information enthält die Netzbereiche in einer BGP `UPDATE` Nachricht. 22, 23, 26, 31, 42, 49, 50

`NOTIFICATION` Die BGP `NOTIFICATION` Nachricht wird im Fehlerfall versandt. Sie löst den Abbau einer BGP Session aus. 20, 21

NSFNET : Das NSFNET löste das ARPANET ab und war somit eine weitere Entwicklungsstufe zum heutigen Internet. 17

OCSP : Über das Online Certificate Status Protocol kann die Gültigkeit von Zertifikaten bei einer VA abgefragt werden. 39

`OPEN` Die BGP `OPEN` Nachricht dient zum Aufbau einer BGP Session. 19–21, 24, 26

OSPF : Open Shortest Path First ist ein Link State IGP Routingprotokoll. 14, 56, 59

PA : Provider Aggregated Netze sind Netze, die ein Endkunde aus dem größeren Bereich seiner LIR erhält. 16

Payload : Bei der Payload handelt es sich um die Nutzdaten eines Paketes, die frei von Protokollinformationen sind. 10, 12, 13

Peering beschreibt eine AS Beziehung in denen zwei AS als gleichberechtigte Partner Datenverkehr i.d.R. kostenfrei austauschen. 7–9

PI : Provider Independent Netze sind Netze, die ein Endkunde von einer RIR erhält und unabhängig von seinem Provider verwenden kann. 16, 40, 58

PKI : Eine Public Key Infrastructure stellt ein System dar, dass Zertifikate ausstellen, prüfen, widerrufen und verteilen kann. 37, 38, 40, 43

PoP : Ein Point of Presence stellt einen Knotenpunkt dar, an dem mehrere ASe Daten austauschen. 9

PSK : Ein Pre Shared Key bezeichnet einen zuvor über einen anderen Kanal getauschten Schlüssel. 36, 37

QoS : Als Quality of Service wird in Rechnernetzen, die Priorisierung von bestimmten Datenverkehr bezeichnet. 38, 52, 54, 57

RA : Eine Registration Authority prüft in einer PKI, ob ein CSR berechtigt ist und reicht diesen ggf. an die CA weiter. 39, 43

RA : Über Route Attestation werden bei sBGP die Routinginformationen in der `UPDATE` Nachricht abgesichert. 42, 43

RFC Request for Comments: Technische und organisatorische Dokumente zum Betrieb des Internets. 5, 17

RIB : Eine Route Information Base bezeichnet eine Routingtabelle. 24

RIP : Das Routing Information Protocol ist Distanzvektorprotokoll, das als IGP eingesetzt wird. 15

RIR Regional Internet Registry: Verwalten die Internet-Ressourcen auf regionaler Ebene, die sie von der IANA zugeteilt bekommen. 16, 40, 41, 44, 47, 58, 62, 75, 79

ROA : Ein Route Origin Authorisation Dokument gibt einem AS das Recht einen entsprechenden Netzbereich anzukündigen. 42, 47, 49–51, 53, 61, 62, 64, 70, 80, 81

Route Reflector ist ein Router, der als Alternative zur „fully meshed“ Topologie, alle internen Router mit den Routinginformationen aus den eBGP Session versorgt. 18

RPKI : Die Resource Public Key Infrastructure ist eine Spezialisierung der PKI zur Zertifizierung von Internet Ressourcen. 40, 41, 49, 53, 62, 64, 75, 81

`RST` TCP Flag, das zum Zurücksetzen der Verbindung verwendet wird. 13, 29

SA : Eine Security Association beschreibt bei IPSec in welcher Form Pakete zu einem bestimmten Ziel abgesichert werden. 37

SAFI : Der Subsequent Address Family Identifier wird zur Unterscheidung von Uni-/Multicast bei MBGP verwendet. 26, 27

sBGP Secure BGP ist eine Sicherheiterweiterungen für BGP. 41–48, 51, 53, 54

`SECURITY` : soBGP führt diese neue BGP Nachricht ein. Sie dient zur Absicherung der Routinginformationen. 44

SIDR : Die Secure Inter-Domain Routing Arbeitsgruppe der IETF will Standards definieren, die das Inter-Domain Routing im Internet sicherer machen sollen. 46, 48, 49

soBGP Secure Origin BGP ist eine Sicherheitserweiterung für BGP. 43–48, 53, 54

Spoofing bezeichnet das Manipulieren von Pakten, um eine andere Identität zu erlangen. 29

`SYN` TCP Flag, das beim Verbindungsaufbau zur Synchronisation verwendet wird. 12, 13

TA : Über einen Trust Anchor wird eine CA festgelegt, der vertraut wird. 40, 75

TCP Transmission Control Protocol: TCP ist das zuverlässige und verbindungsorientierte Transportprotokoll im Internet. 10, 12, 18

TCP-AO : Die TCP Authentication Option ermöglicht die Absicherung von TCP Datenverkehr über den Hashwert einer kryptographischen Hashfunktion. 36, 37, 49, 52, 54

TCP-MD5 ermöglicht die Absicherung von TCP Datenverkehr über einen MD5 Hashwert. 18, 36–38, 52, 54, 57, 59, 63

Tier-1 Provider haben Peerings zu alle anderen Tier-1 Providern. 9

Tier-2 Provider haben ein Transitabkommen mit mindestens einem Tier-1 Provider. 9

Tier-3 Provider haben lediglich Transitabkommen mit einem Tier-2 oder Tier-3 Provider. 9

Transit beschreibt eine AS Beziehung bei der ein Provider-AS einem Kunden-AS ermöglicht sämtlichen Datenverkehr über sein Netz abzuwickeln. 2, 7–9

TTL : Die Time-to-Live gibt an nach wie vielen durchlaufenen Routern ein Paket verworfen werden soll. Dazu wird diese an jedem Hop um den Wert eins verringert. 11, 12, 35

Unicast bezeichnet die Datenkommunikation mit einer Quelle und einem Empfängern. 26

`UPDATE` Über BGP `UPDATE` Nachrichten werden die Routinginformationen übertragen. 20, 21, 23–28, 31–33, 42, 43, 46, 47, 49–51, 53, 64

VA : Eine Validation Authority prüft in einer PKI Zertifikate. 39

VLAN : Über ein virtuelles LAN kann ein physisches Netz in mehrere logische unterteilt werden. 59

VPN Virtuelles privates Netzwerk: Mehrere Teile eines Netzwerkes werden über ein weiteres verbunden. 1

WAN Wide Area Network: Rechnernetz, das sich über ein geographisch sehr großes Gebiet erstreckt. 1

WoT : Ein Web-of-Trust stellt den dezentralisierten Ansatz einer PKI dar. Dabei gibt jeder Teilnehmer diejenigen Teilnehmer an, denen er vertraut. 43, 44, 46, 53

Abbildungsverzeichnis

Literaturverzeichnis

[BA11] Bush, R. und R. Austein: *The RPKI/Router Protocol*, Oktober 2011. http://tools.ietf.org/html/draft-ietf-sidr-rpki-rtr-19.

[Ban11] Band, Alex: *Resource Certification*, November 2011. http://ripe63.ripe.net/presentations/32-RIPE63-RPKI-Session.pdf.

[BBW11] Bellovin, S., R. Bush und D. Ward: *Security Requirements for BGP Path Validation*, April 2011. http://tools.ietf.org/html/draft-ietf-sidr-bgpsec-reqs-01.

[BCC06] Bates, T., E. Chen und R. Chandra: *BGP Route Reflection: An Alternative to Full Mesh Internal BGP (IBGP)*. RFC 4456 (Draft Standard), April 2006.

[BCKR07] Bates, T., R. Chandra, D. Katz und Y. Rekhter: *Multiprotocol Extensions for BGP-4*. RFC 4760 (Draft Standard), Januar 2007.

[Bei02] Beijnum, Iljitsch van: *BGP*. O'Reilly Media, 1. Auflage, 2002.

[BFMR10] Butler, Kevin, Toni R. Farley, Patrick McDaniel und Jennifer Rexford: *A Survey of BGP Security Issues and Solutions*. Proceedings of the IEEE, Seiten 100–122, Januar 2010. http://ix.cs.uoregon.edu/~butler/pubs/bgpsurvey.pdf.

[Cym11] Cymru, Research NFP: *The Bogon Reference*, August 2011. http://www.team-cymru.org/Services/Bogons/.

[DH98] Deering, S. und R. Hinden: *Internet Protocol, Version 6 (IPv6) Specification*. RFC 2460 (Draft Standard), Dezember 1998. Updated by RFCs 5095, 5722, 5871.

[Dom02] Domhan, Gregor: *Link-State-Routing mit OSPF*, Mai 2002. http://www.domhan.de/OSPF.pdf.

[Erm11] Ermert, Monika: *RIPE: IP-Routen werden weiterhin gesichert*, November 2011. http://www.heise.de/netze/meldung/RIPE-IP-Routen-werden-weiterhin-gesichert-1371882.html.

[FLYV93] Fuller, V., T. Li, J. Yu und K. Varadhan: *Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy*. RFC 1519 (Proposed Standard), September 1993. Obsoleted by RFC 4632.

[FW11] Forschungsgruppe Wahlen, e.V.: *Internet-Strukturdaten - Repräsentative Umfrage - II. Quartal 2011*, Juli 2011. http://www.forschungsgruppe.de/Aktuelles/Internet-Strukturdaten/web_II_11.pdf.

[GHM+07] Gill, V., J. Heasley, D. Meyer, P. Savola und C. Pignataro: *The Generalized TTL Security Mechanism (GTSM)*. RFC 5082 (Proposed Standard), Oktober 2007.

[HB96] Hawkinson, J. und T. Bates: *Guidelines for creation, selection, and registration of an Autonomous System (AS)*. RFC 1930 (Best Current Practice), März 1996.

[HB11] Huston, Geoff und Randy Bush: *Securing BGP with BGPsec*, Juli 2011. `http://www.potaroo.net/ispcol/2011-07/bgpsec.html`.

[HE11] Hurricane Electric, Inc.: *About Resource Public Key Infrastructure (RPKI)*, 2011. `http://rpki.he.net/about_rpki.html`.

[Hef98] Heffernan, A.: *Protection of BGP Sessions via the TCP MD5 Signature Option*. RFC 2385 (Proposed Standard), August 1998. Obsoleted by RFC 5925.

[Hol02] Holtkamp, Heiko: *Einführung in TCP/IP*, Februar 2002. `http://www.rvs.uni-bielefeld.de/~heiko/tcpip/tcpip.pdf`.

[Hor09] Horn, Christian: *Understanding IP Prefix Hijacking and its Detection*, Juni 2009. `http://www.net.t-labs.tu-berlin.de/teaching/ss09/IR_seminar/talks/prefix_hijacking_horn.handout.pdf`.

[HS82] Hinden, R.M. und A. Sheltzer: *DARPA Internet gateway*. RFC 823 (Historic), September 1982.

[Hus11] Huston, Geoff: *IPv4 Address Report*, September 2011. `http://www.potaroo.net/tools/ipv4/index.html`.

[IAN11] IANA: *IANA IPv4 Address Space Registry*, Februar 2011. `http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xml`.

[IBM95] IBM, Corporation: *TCP/IP Tutorial and Technical Overview*, Juni 1995. `http://www.pms.ifi.lmu.de/mitarbeiter/ohlbach/multimedia/IT/IBMtutorial/3376c32.html#arparte`.

[ICA11] ICANN: *Bylaws for Internet Corporation for Assigned Names and Numbers*, Juni 2011. `http://www.icann.org/en/general/bylaws.htm`.

[IWS11] Internet World Stats, Miniwatts Marketing Group: *Internet Usage Statistics*, März 2011. `http://www.internetworldstats.com/stats.htm`.

[Ken05a] Kent, S.: *IP Authentication Header*. RFC 4302 (Proposed Standard), Dezember 2005.

[Ken05b] Kent, S.: *IP Encapsulating Security Payload (ESP)*. RFC 4303 (Proposed Standard), Dezember 2005.

[KLMS00] Kent, Stephen, Charles Lynn, Joanne Mikkelson und Karen Seo: *Secure Border Gateway Protocol (S-BGP) - Real World Performance and Deployment Issues*. BBN Technologies, Februar 2000. `http://www.ece.cmu.edu/~adrian/731-sp04/readings/KLMS-SBGP.pdf`.

[Koz05] Koziero, Charles M.: *The TCP/IP Guide: BGP Overview, History, Standards and Versions*, September 2005. http://www.tcpipguide.com/free/t_BGPOverviewHistoryStandardsandVersions-2.htm.

[KSM07] Kuhn, Rick, Kotikalapudi Sriram und Doug Montgomery: *Border Gateway Protocol Security*, Juli 2007. http://csrc.nist.gov/publications/nistpubs/800-54/SP800-54.pdf.

[Lee03] Leech, M.: *Key Management Considerations for the TCP MD5 Signature Option*. RFC 3562 (Informational), Juli 2003.

[Lep11] Lepinski, M.: *BGPSEC Protocol Specification*, Oktober 2011. http://tools.ietf.org/html/draft-ietf-sidr-bgpsec-protocol-01.

[LK11] Lepinski, M. und S. Kent: *An Infrastructure to Support Secure Internet Routing*, Mai 2011. http://tools.ietf.org/html/draft-ietf-sidr-arch-13.

[LKK11] Lepinski, M., S. Kent und D. Kong: *A Profile for Route Origin Authorizations (ROAs)*, Mai 2011. http://tools.ietf.org/html/draft-ietf-sidr-roa-format-12.

[LKS04] Lynn, C., S. Kent und K. Seo: *X.509 Extensions for IP Addresses and AS Identifiers*. RFC 3779 (Proposed Standard), Juni 2004.

[LR89] Lougheed, K. und Y. Rekhter: *Border Gateway Protocol (BGP)*. RFC 1105 (Experimental), Juni 1989. Obsoleted by RFC 1163.

[LR90] Lougheed, K. und Y. Rekhter: *Border Gateway Protocol (BGP)*. RFC 1163 (Historic), Juni 1990. Obsoleted by RFC 1267.

[LR91] Lougheed, K. und Y. Rekhter: *Border Gateway Protocol 3 (BGP-3)*. RFC 1267 (Historic), Oktober 1991.

[LRZ09] LRZ: *WiN-Nutzung bayerischer Hochschulen*, März 2009. http://www.lrz.de/services/netz/statistik/diagbay/.

[LRZ11] LRZ: *Überblick über das Münchner Wissenschaftsnetz*, 2011. http://www.lrz.de/services/netz/mwn-ueberblick/mwn-karte.jpg.

[Man07] Manral, V.: *Cryptographic Algorithm Implementation Requirements for Encapsulating Security Payload (ESP) and Authentication Header (AH)*. RFC 4835 (Proposed Standard), April 2007.

[Mil84] Mills, D.L.: *Exterior Gateway Protocol formal specification*. RFC 904 (Historic), April 1984.

[Mur06] Murphy, S.: *BGP Security Vulnerabilities Analysis*. RFC 4272 (Informational), Januar 2006.

[Mü10] Müller, Yves: *Why the Internet Sucks: A Core Perspective*, Januar 2010. http://cst.mi.fu-berlin.de/teaching/WS0910/19510b-PS-TI/mueller10why_slides.pdf.

[Pos81a] Postel, J.: *Internet Protocol*. RFC 791 (Standard), September 1981. Updated by RFC 1349.

[Pos81b] Postel, J.: *Transmission Control Protocol*. RFC 793 (Standard), September 1981. Updated by RFCs 1122, 3168, 6093.

[Rei11] Reiser, Helmut: *Netzsicherheit - Schicht 3: Network Layer - IPSec*, Januar 2011. http://www.nm.ifi.lmu.de/teaching/Vorlesungen/2010ws/itsec/_skript/itsec-k11-v6.0.pdf.

[RIP08] RIPE, NCC: *YouTube Hijacking: A RIPE NCC RIS case study*, März 2008. http://www.ripe.net/internet-coordination/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study.

[RIP11] RIPE, NCC: *RIS Raw Data*, Februar 2011. http://www.ripe.net/data-tools/stats/ris/ris-raw-data.

[RLH06] Rekhter, Y., T. Li und S. Hares: *A Border Gateway Protocol 4 (BGP-4)*. RFC 4271 (Draft Standard), Januar 2006. Updated by RFC 6286.

[RP94] Reynolds, J. und J. Postel: *Assigned Numbers*. RFC 1700 (Historic), Oktober 1994. Obsoleted by RFC 3232.

[Sri11] Sriram, K.: *BGPSEC Design Choices and Summary of Supporting Discussions*, Juli 2011. http://tools.ietf.org/html/draft-sriram-bgpsec-design-choices-00.

[Tan03] Tanenbaum, Andrew S.: *Computernetzwerke*. Pearson Studium, 4. überarb. Auflage, 2003.

[Tel11] Telekom, Deutsche: *CSS INTERACTIVE NETWORK MAP*, 2011. http://www.deutschetelekom-icss.com/dtag/cms/content/ICSS/en/interactivemap.

[TMB10] Touch, J., A. Mankin und R. Bonica: *The TCP Authentication Option*. RFC 5925 (Proposed Standard), Juni 2010.

[Too11] Toonk, Andree bgpmon.net: *BGPMon.net – Blog / BGP Hijacks*, Mai 2011. http://bgpmon.net/blog/?cat=9.

[vdB08] Berg, Rudolph van der: *How the 'Net works: an introduction to peering and transit*, September 2008. http://arstechnica.com/old/content/2008/09/peering-and-transit.ars.

[WDMC07] Wang, Mei, Larry Dunn, Wei Mao und Tao Chen: *Reduce IP Address Fragmentation through Allocation*, Februar 2007. http://140.116.82.38/members/html/ms07/yslin/papers/Reduce%20IP%20Address%20Fragmentation%20through%20Allocation.pdf.

[Wei05] Weis, Brian: *Secure Origin BGP (soBGP) Certificates*. Internet Engineering Task Force, Juli 2005. ftp://ftp-eng.cisco.com/sobgp/drafts/draft-weis-sobgp-certificates-02.txt.

[Whi05] White, R.: *Architecture and Deployment Considerations for Secure Origin BGP (soBGP)*. Internet Engineering Task Force, November 2005. `ftp://ftp-eng.cisco.com/sobgp/drafts/draft-white-sobgp-architecture-01.txt`.

[Wik11a] Wikipedia: *Denial of Service — Wikipedia*, 2011. `http://de.wikipedia.org/w/index.php?title=Denial_of_Service&oldid=95600473`.

[Wik11b] Wikipedia: *Public-Key-Infrastruktur — Wikipedia*, 2011. `http://de.wikipedia.org/w/index.php?title=Public-Key-Infrastruktur&oldid=94349189`.

[Wik11c] Wikipedia: *Web of Trust — Wikipedia*, 2011. `http://de.wikipedia.org/w/index.php?title=Web_of_Trust&oldid=94036967`.

[Wä11] Wählisch, Matthias: *Beta Version of the RPKI RTR Client C Library Released*, September 2011. `https://labs.ripe.net/Members/waehlisch/beta-version-of-the-rpki-rtr-client-c-library-released`.