Border Gateway Protocol

Michael Rist

15. Januar 2008

Inhaltsverzeichnis

1 Einleitung **3**

2 Prinzipieller Aufbau des Internets **3**

3 Verwendungszweck für BGP und AS **3**

4 Besonderheiten des BGP: **4**
4.1 iBGP und eBGP 4
4.2 Routing Policy 4
4.3 Routing 4

5 Kommunikation zwischen 2 BGP Router: **5**
5.1 Message-Header 5
5.2 OPEN-Message 5
5.3 UPDATE-Message 6
5.4 KEEPALIVE-Message 7
5.5 NOTIFICATION-Message 8

6 Probleme des BGP und Internet Architektur **9**

7 Zusammenfassung **10**

1 Einleitung

Das Border Gateway Protocol ist eines der wichtigsten Protokolle des Internets, da es Verbindungen zwischen zwei Autonomen Systemen gewährleistet. Die aktuelle Version des Border Gateway Protokoll ist das BGP-4, im Folgenden BGP genannt. Ältere Versionen dieses Protokoll sind überholt und werden nicht mehr verwendet.

2 Prinzipieller Aufbau des Internets

Um die Funktionsweise des BGP zu verstehen braucht man zuerst Grundkenntnisse über den Aufbau des Internets. Das Internet besteht aus verschiedenen Routern, über die die Daten ausgetauscht werden. Diese Router gehören keiner zentralen Organisation, die diese verwaltet, sondern verschiedenen Gruppen (Internetdienstanbieter, kurz ISP, etc.). Die Router die zu einer solchen Gruppe gehören werden als autonome Systeme bezeichnet und werden von der jeweiligen Gruppe gewartet und verwaltet. Unter diese Verwaltung fällt auch das Routen von Daten. Somit hat das Internet unterschiedliche autonome Systeme (kurz AS) die unterschiedlich verwaltet werden und die Daten unterschiedlich weiterleiten. Diese AS sind jedoch noch nicht miteinander verbunden. Die Verbindung von verschiedenen AS und somit das weiterleiten von Routinginformationen ist Aufgaben des BGP. Das BGP dient dabei nicht zur Bestimmung der optimalen Route sondern gibt nur an über welchen Router bzw welche Netze man welche Adresse erreichen kann. Für die Verbreitung dieser Wege wird in jedem AS mindestens 1 Router benötigt, der das BGP beherrscht. Dieser Router sollte im Besten Falle am Rande des AS sein da er eine besondere Rolle im AS einnimmt. Wenn man nun von einem Rechner in einem AS auf einen anderen in einem anderen AS zugreifen will führt die Kommunikation über die spezielle Router in den jeweiligen AS, wobei beide Router das BGP beherrschen müssen. Das BGP nutzt dabei verschiedene Nachrichten um Informationen zwischen 2 BGP-Routern auszutauschen. Das BGP verbindet somit die AS miteinander und somit sind die Grundlagen für ein großes Netzwerk wie dem Internet gegeben.

3 Verwendungszweck für BGP und AS

Es gibt mehrere Gründe für die Verwendung von BGP und das Einteilen von Router zu AS. Zum einen gibt es im Internet zu viele Router als das man sie alle auf der selben Ebene verbinden könnte. Der entstehende Datenverkehr zum verbreiten von Routinginformationen wäre einfach zu groß und würde das ganze Netzwerk blockieren. Zum anderen gehören die AS, wie bereits erwähnt, zu verschiedenen Organisation und gehören zu verschiedenen Netzwerken. Da das Internet so groß ist und die Router verschiedenen Organisationen gehören, gibt es keine zentrale Organisation die alles verwaltet und einen einheitlichen Standard für das weiterleiten von Daten festlegt. Da man Router zwangsweise zusammenfassen muss um große Netzwerke zu ermöglichen ergeben sich neue Probleme. Wenn zu viele Router in einen AS, also einem Teilnetzwerk, zusammenfasst werden hat man das gleiche Problem wie im kompletten Netzwerk, nämlich das verbreiten von Routing Informationen. Dabei gibt es 2 Probleme, die als Delay und Overhead bezeichnet werden. Delay bezeichnet die Zeit die es braucht um die Routing Informationen im Netzwerk zu verbreiten. Bei N Routern im Netzwerk braucht man N Schritte um die Informationen zu verbreiten. Unter Overhead versteht man das Problem, das bei zu vielen Router die Länge der Nachrichten

zum verbreiten von Informationen immer länger wird, da jede mögliche Route im Netzwerk verbreitet werden muss. Somit muss die Größe der AS begrenzt sein (Richtlinie: etwa ein duzend Router in einem Wide Area Network und grob 5mal soviel in einem Local Area Network[2]).

Ein weiteres Problem liegt in dem Aufbau der AS, bzw. wenn ein Router nicht ein Exterior Gateway Protocol wie das BGP verwendet. Wenn z.B. 2 Router, jeweils die Gateway Router eines Lokalen Netzwerks sind, also der Router der das Lokale Netzwerk mit einem anderen Netzwerk verbinden, und ein dritter Router ebenfalls im Verbindungsnetzwerk angeschlossen ist kann es zu Problemen kommen wenn der 3. Router kein Exterior Gateway Protocol nutzt. Ein Router der nicht das selbe Protokoll nutzt, nutzt einen der Anderen Router als Standard Gateway und leitet seine Daten über diesen Router. Das kann dazu führen das Daten die z.B. Für Router 2 erst an Router 1 geschickt werden und dann nochmal über das selbe Netzwerk zu Router 2. Dies führt natürlich zu unnötigen Datenverkehr der sich vermeiden lässt indem alle Router ein einheitliches Exterior Gateway Protocol nutzen, im Falle des Internets das BGP. Dieses Phänomen wird als Extra Hop bezeichnet.

4 Besonderheiten des BGP:

Das BGP hat im Vergleich zu anderen Routingprotokollen verschiedene Besonderheiten.

4.1 iBGP und eBGP

Das BGP kann nicht nur dazu genutzt werden verschiedene AS zu verbinden sondern auch dazu innerhalb eines AS Gateway-Router zu verbinden, und so den Datenverkehr durch ein AS zu schleusen. Diese Router benutzen ebenfalls das BGP und sind mit anderen AS verbunden. Die Art des BGP das zur Verbindung von Router innerhalb eines AS benutzt wird, wird iBGP (internel Border Gateway Protocol) genannt, außerhalb wird es als eBPG (external Border Gateway Protocol) bezeichnet. Ein Administrator eines AS kann jedoch bestimmen ob eine solche iBGP Verbindung aufgebaut werden darf oder nicht. Solche Einstellungen fallen unter den Begriff „Routing Policy“.

4.2 Routing Policy

Beim BGP ist es möglich manuell bestimmt Regeln aufzustellen die das Routen von Daten steuern. Es ist möglich die Routen festzulegen die von einem bestimmten AS abgehen oder die Routen, die von einem AS, zu Ignorieren und somit dieses AS zu meiden. Dies hat Sinn wenn z.B. die Sicherheit in dem entsprechendem AS nicht gewährleistet ist oder das Routen innerhalb des Netzwerk nicht optimal ist, etc. Man kann auch sagen ob das eigene AS dazu genutzt werden darf Daten, die für ein anderes AS bestimmt sind, durch das eigene AS zu schleusen oder ob dies nicht erlaubt ist. Solche AS werden als transit bzw. nontrasit AS bezeichnet, also solche die ein Durchschleusen von Daten erlauben oder solche die es nicht erlauben.

4.3 Routing

Das BGP benutzt zum Übertragen von Daten kein Eigenes Protokoll welches die Verbindung aufbaut sondern nutzt TCP um einen sicheren Datenaustausch zu gewährleisten. Die

Routinginformationen die dabei vermittelt werden enthalten alle auf dem Pfad vorkommenden AS und nicht die genauen Netzwerkadressen der einzelnen Routern auf dem Weg. So sind die Informationen ausreichend um den Weg zurückverfolgen aber nicht unnötig groß. Des weiteren wird nur Pfad weiterverbreitet den das AS selbst wählen würde. Um die gesendeten Daten möglichst klein zu halten sammelt jeder BGP-Router alle erhaltenen Routen und es werden immer nur neue Routen weitergeleitet und nicht die kompletten Routingtabellen. BGP ermöglicht einen den Sender einer Information zu ermittelt und seit BGP-4 wird auch CIDR (Classless Inter Domain Routing) untersützt.

5 Kommunikation zwischen 2 BGP Router:

Die Kommunikation zwischen 2 oder mehreren BGP Routern erfolgt, wie schon zuvor erwähnt, über unterschiedliche Nachrichten. Dabei gibt es 5 verschiedene Typen von Nachrichten.

5.1 Message-Header

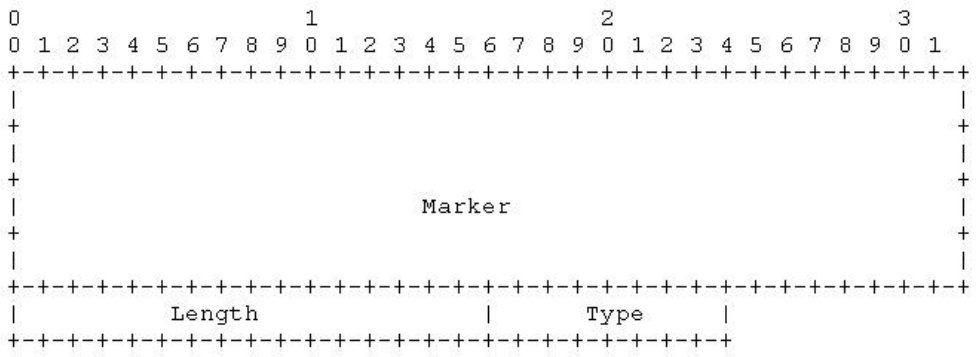

Abbildung 1: Massage-Header[3]

Der Mesage-Header ist bei jedem Nachrichtentyp gleich. Das Marker besteht nur aus Einsen, welches den Start einer neuen NAchricht markiert. Das anschließende Length Feld gibt die Gesamtlänge der Nachricht (zusammen mit dem Header) an. Die Länge der Nachricht ist mindestens 19 und maximal 4096. Das Auffüllen der Nachricht mit Daten ist nicht erlaubt und die angegebene Länge muss minimal sein. Das Feld Type gibt an welche Nachrichtenart übermittelt wird. Diese Nachrichten werden im Folgenden Abschnitt näher erläutert.

5.2 OPEN-Message

Nachdem eine TCP-Verbindung aufgebaut wurde wird von den einzelnen Routern eine Open-Message geschickt umd wenn die Open-Message akzeptiert wird, wird diese mit einer Keepalive-Message bestätigt. Mit einer OPEN-Message wird jede Verbindung zwischen 2 BGP Routern eröffnet.

Version

Gibt die verwendete BGP Version an. Die aktuelle Versionsnummer ist 4

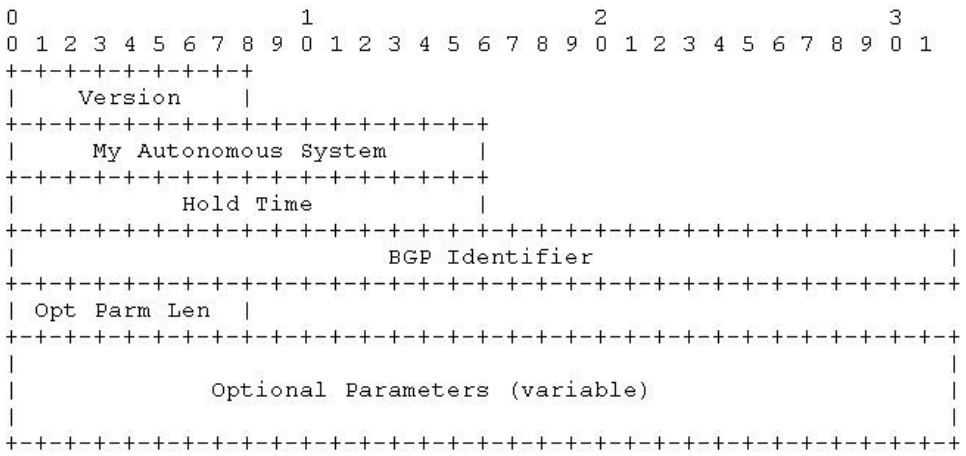

Abbildung 2: Open-Message[3]

My Authomonous System

Gibt die Nummer des AS des Senders an

Hold Time

Die Hold Time gibt an wie lange es dauern kann bis eine Keepalive oder Update Message, die vom Sender geschickt wurde, ankommt. Wenn keine Nachricht in dieser Zeit kommt wird der Sender als nichtmehr erreichbar angesehen und der Empfänger verschickt keine weiteren Daten.

BGP Identifier

Gibt eine IP-Adresse über die der Sender der Nachricht in dessen BGP identifiziert werden kann. Diese Adresse muss bei allen Verbindungen des Senders gleich sein.

Optional Parameter Length und Optional Parameter

Das Length Feld gibt die Länge des Parameterfeldes an, bei der Länge 0 gibt es kein Parameterfeld. Das Parameter Feld ist in 3 Teile unterteilt. Als erstes der Parameter Typ in dann die Länge des zusätzlichen Parameterfelds und dann die Parameterdaten. Diese zusätzlichen Daten werden z.B. dazu genutzt eine größere Anzahl von AS zu unterstützen.

5.3 UPDATE-Message

Die Update-Message dient dazu die Routinginformationen zu übertragen. Dabei kann eine neue Route übertragen werden und gleichzeitig mehrere Routen abgelegt werden. Diese Routen werden nicht mehr genutzt und werden dann auch vom Empfänger der Nachricht nicht mehr genutzt. Eine Route wird abgelegt, wenn z.B. eine Router auf dem Pfad seinen Routing Policy ändert und somit die Route dadurch nicht mehr gültig ist.

Withdrawn Routes

Das Withdrawn Routes Length gibt an wie lange das nachfolgende, variable Feld Withdrawn Routes ist. Die angegebene Zahl gibt dabei die Anzahl der Bytes an, die das Feld hat. Eine 0 in diesem Feld bedeutet das keine Routen abgelegt werden. Das Withdrawn Routes Feld gibt die Routen an die nicht mehr genutzt werden sollen. Eine Adresse besteht dabei aus einem 2er Tupel, das als erstes die Länge in Bits angibt und als 2. die Adresse. Dabei wird die eigentliche IP-Adresse verschlüsselt um klassenlose Adressierung zu unterstützen.

Total Path Atribute Length

Dieses 2 Byte großes Feld gibt die gesammte Länge der beiden nachfolgenden Felder an.

Path Attributes

Dieses Feld gibt verschiedene Eigentschaften der übertragenen Pfade, wenn in einer Update Nachricht keine neuen Pfade übertragen werden ist dieses Feld leer. Zum einen werden die Eigenschaften über den eigentlichen Pfad übermittelt, zum anderen Informationen über die Herkunft der Informationen und Information über den Sender dieser Informationen.

Network Layer Reachability

In diesem Feld stehen die neue Addresse die in der Update Nachricht verbreitet wird. Diese Adresse ist ebenfalls eine vom BGP verschlüsselte IP-Adresse um klassenlose Adressierung zu gewähreisten.

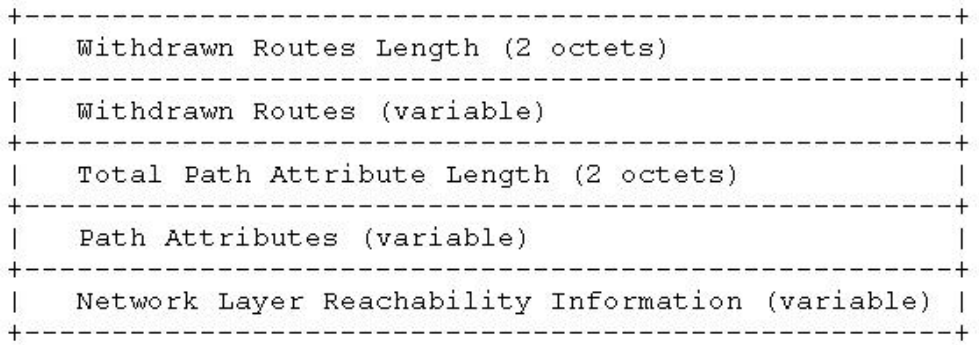

Abbildung 3: Update-Message[3]

5.4 KEEPALIVE-Message

Das BGP nutzt zwar TCP um eine Verbindung zwischen 2 Routern aufzubauen jedoch nicht die TCP eigenen Funktionen um eine Verbindung aufrecht zu halten. Daher wird eine eigene Funktion benötigt damit eine Verbindung erhalten bleibt. Ein maximaler Intervall zwischen 2 KEEPALIVE Nachrichten sollte ein Drittel der angegebenen Hold-Time sein. Dieser Nachrichtentyp besteht nur aus dem 19 Byte langen HEADER.

5.5 NOTIFICATION-Message

Die Notification Nachricht dient zur Fehlererkennung und -meldung. Wenn ein Fehler gemeldet wird, wird auch die TCP Verbindung unterbrochen. Dies wird auch dazu genutzt eine Verbindung regulär zu beenden. Die Nachricht besteht aus 3 Feldern. Die ersten 2 Felder sind Codefelder mit deren Hilfe man erkennen kann bei welchen Nachrichtentyp welcher Fehler aufgetreten ist.

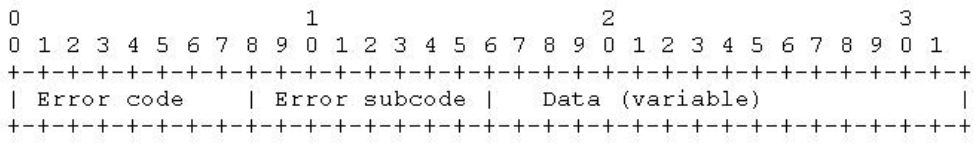

Abbildung 4: NOTIFICATION-Message[3]

Tabelle 1: Error Codes[3]

Error Code	Symbolic Name
1	Message Header Error
2	OPEN Message Error
3	UPDATE Message Error
4	Hold Timer Expired
5	Finite State Machine Error
6	Cease

Message Header Error

Ein Message Header Error hat 3 mögliche Ursachen, die im Error Subcode ablesbar sind. Wenn ein Fehler beim lesen des Header entsteht wird der Error Code 1 Übertragen. Ein 2. Fehler kann auftregten wenn das Length-Feld einen falschen wert enthält, der z.B. kleiner als 19 ist, also kleiner als die minimal erforderlichen Bytes. Ein solcher Fehler hat den Code 2. Der 3. Fehler wird gemeldet wenn das Type-Feld des Headers einen flaschen Wert hat.

OPEN Message Error

Ein OPEN Message Error hat 5 Subcodes die den Fehler näher beschreiben. Zum einen wenn eine andere BPG Version verwendet wurde (Subcode 1), oder die Adresse des Empfängers (Subcode 2) oder des AS des Empängers (Subcode 3) falsch sind. Zum anderen gibt es einen Fehler wenn einer der Optionalen Parameter der OPEN Nachricht falsch sind (Subcode 4) oder der angegebene Hold Timer vom empfänger nicht akzeptiert wird (Subcode 6). Frühere Varianten des BGP hatten noch einen Subcode 5 der angibt ob bei der Authentifizierung ein Fehler auftrat.

UPDATE Message Error

Der UPDATE Message Error hat 11 Subcodes. Der 1. Subcode wird gesendet wenn die angegebene Länge der Attribute oder die angegebene Länge der abgelegten Pfade falsch ist. Der 2. Subcode bezieht sich auf ein falsches Attribute

das gesetzt wurde und vom Empfänger nicht unterstützt wird. Der 3. Subcode wird gesendet wenn ein bestimmtes Attribut fehlt, also vom Sender nicht genutzt wird obwohl es der Empfänger braucht. Wenn ein Attribut gesetzt ist aber die dazugehörigen Daten falsch sind wir der 4. Subcode verwendet. Der 5. Subcode bezieht sich auf eine falsch angegebene Länge des Attributsfelds. Wenn die Herkunft der Information einen falschen, bzw. nicht erwünschten Wert hat wird der 6. Subcode gesendet. Der 7. Subcode wird nicht mehr verwendet und verweiste auf einen Loop beim Routen. Dies wird beim BGP durch die Routing Policies vermieden. Subcode 8 tritt auf wenn die angegebene Next-Hop Adresse falsch ist. Subcode 9 wird gesendet wenn ein Fehler bei den Optionalen Attributen auftritt. Wenn das Network Layer Reachability Information Feld einen Fehler hat wird der Subcode 10 gesendet, Subcode 11 bezieht sich auf einen Fehlerhaften AS-Pfad in der Update Nachricht

Bei allen Subcodes die durch einen Fehler in den Attributen ausgelöst werden, wird das Attribut das den Fehler verursacht hat im Data-Feld der Notification Nachricht übermittelt um den Sender über das Attribut das den Fehler versucht hat zu unterrichten.

HOLD Message Error

Wenn in die angegebene Hold Time abgelaufen ist und keine Nachricht angekommen ist,wird eine Notification Nachricht, mit dem entsprechenden Fehldercode, gesendet und die Verbindung unterbrochen.

Finite state machine Error

Dieser Fehler entsteht wenn ein BGP Router auf ein unerwartetes Ereignis trifft. Z.B. wenn ein Router sich schon in einer Verbindung mit einem Router steht und ein 3. ebenfalls mit dem Router eine Verbindung öffnen will.

Cease

Eine Notification Nachricht mit dem Cease Code wird gesendet um eine Verbindung zu beenden.

6 Probleme des BGP und Internet Architektur

Wenn man das BGP bei einem sehr großen Netzwerk wie dem Internet anwendet ergeben sich zusätzliche Probleme. Bei 2 großen Netzwerken ist sehr gut möglich das mehrere Verbindungen zwischen den beiden Netzwerken vorhanden sind. Das BGP nutzt aber immer nur eine Verbindung obwohl pysikalisch gesehen mehrere Verbindungen gibt. Wenn nun das Netzwerk einfach nur die vom BGP auomatisch gewählten Routen nehmen würde, würde alle Daten zwischen den beiden Netzwerken durch nur eine Verbindung geleitet werden obwohl mehrere vorhanden sind. Um von mehreren Routern im eigenen Netz ein anderes zu erreichen müssen die einzenen Router manuell eingestellt werden, da sonst das BGP immer nur die für ihn optimale Route wählt. Ein weiteres Problem besteht darin, dass das BGP zwar die länge des Pfades weiß, und danach zwar die Route mit den wenigsten AS auf dem Weg auswählt, diese Route jedoch nicht die optimale sein muss. Wenn z.B. 2 Routen in ein bestimmtes AS führen und Route A 1 AS durch 1 AS führt und Route B durch 2 AS wählt das BGP in der Regel die Route A da sie nur durch ein AS geleitet wird.

Da das BGP nicht die internen Routingvorgänge kennt kann es sein das Route B schneller ist obwohl sie durch 2 AS führt. Bei Route A kann es nämlich vorkommen das die Daten durch etliche Interne Router geleitet werden bevor sie wieder durch ein BGP-Router in das Ziel-AS geleitet werden. Bei Route B hingegen werden die Daten jeweils direkt von einem BGP-Router im AS zum anderen BGP gesentet der die Daten dann wieder zum nächsten AS schickt. Ein Weiteres Problem besteht darin das das Internet nicht statisch ist sonder sich immer wieder ändert, neue Router werden ins Netzwerk eingebunden und alte Router verschwinden. Dies erfordert eine immer wieder neue Routen zu einem bestimmten Ziel. Des weiteren müssen Routen auch zum richtigen Ziel führen. Um dies zu gewährleisten wurde Routing Registries eingeführt. Diese Route Registries werden eingesetzt um festzulegen welche Adressblöcke zu welchen AS gehören. Dadurch ergibt sich ein neues Problem, nämlich ist es nicht möglich festzustellen ob ein Eintrag in einer solchen Route Registry richtig ist oder nicht. Dadurch ist es möglich das eine bestimmte Adresse aus dem gesammten Internet kurzzeitig nicht mehr zu erreichen ist. Auch ist zu bedenken wie weit das Internet noch wachsen kann, bzw. wird. Je mehr Router sich in einem Netzwerk befinden desto mehr Routinginformationen sind nötig und um so schwieriger wird es einen BGP Router zu warten.

7 Zusammenfassung

Das BGP, oder besser gesagt die Idee dahinter, ist eines der wichtigsten Grundbausteine für das Internet. Ohne die Einteilung von Routern und die besondere Art des Routens wäre es kaum möglich ein weltweites Netzwerk zu betreiben. Durch den Informationsaustausch zwischen den BGP-Routern wird gewährleistet das neue Routen verbreitet werden und alte Routen nicht mehr genutz werden. Obwohl, wie im vorherigen Kapitel beschrieben, einige Probleme auftreten ist das BGP das Grundprotokol auf dem das Internet basiert.

Literatur

[1] Uyless D. Black. *IP Routing Protocols*. Prentice Hall PTR, 2000.

[2] Douglas E. Cormer. *Internetworking with TCP/IP*. Prentice Hall International, 5. edition, 2005.

[3] Y. Rekhter, T. Li, and S. Hares. RFC 4271, 2006.

Abbildungsverzeichnis

1 Massage-Header[3] 5
2 Open-Message[3] 6
3 Update-Message[3] 7
4 NOTIFICATION-Message[3] 8