Hochschule
Bonn-Rhein-Sieg
University of
Applied Sciences

Bachelor of Computer Science
Fachbereich Informatik

Seminararbeit zum Thema

IPv6-Anycast – Funktionsweise und Anwendung

von

Stefan Frohrath

im Rahmen des Bachelorseminars
Telekommunikation

WS 2012/13

Themenstellung: Prof. Dr. Martin Leischner
Verfasser: Stefan Frohrath
E-Mail: stefan.frohrath@smail.inf.h-brs.de
Eingereicht am: 19. Dezember 2012

Erklärung (Einzelarbeit)

Ich versichere, die von mir vorgelegte Arbeit selbständig verfasst zu haben. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich als entnommen kenntlich gemacht. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit benutzt habe, sind angegeben. Die Arbeit hat mit gleichem Inhalt bzw. in wesentlichen Teilen noch keiner anderen Prüfungsbehörde vorgelegen.

(Ort, Datum) (Unterschrift)

Inhaltsverzeichnis

IPv6-Anycast – Funktionsweise und Anwendung

Stefan Frohrath

Hochschule Bonn-Rhein-Sieg
stefan.frohrath@smail.inf.h-brs.de

Zusammenfassung

In dieser Arbeit wird der bei IPv6 neu eingeführte Adresstyp Anycast vorgestellt. Insbesondere wird die Funktionsweise von Anycast erklärt und darauf eingegangen, was bei der Anwendung von Anycast beachtet werden muss. Ein großer Teil dieser Arbeit befasst sich dabei mit dem Routing von Anycast-Adressen. Abgeschlossen wird die Arbeit durch die Erläuterung von konkreten Einsatzmöglichkeiten.

1 Einleitung

Anycast ist eine neue Kommunikationsmethode, bei der ein Paket an eine Gruppe von Hosts adressiert wird, aber nur einer der Hosts das Paket erhält. Dabei bieten die Hosts einer Gruppe für gewöhnlich den gleichen Dienst an. Um eine Gruppe zu adressieren, wird eine Anycast-IP-Adresse verwendet, die jedem Gruppenmitglied zugeordnet wird. Die grundlegende Idee hinter Anycast ist die Trennung von logischen Dienstkennungen von physikalischen Rechnern, d.h. die Anycast-Adresse wird auf Basis des Diensttyps zugewiesen, sodass der Netzwerkdienst als logischer Host agiert ([Web04, S. 1], [Doi04, S. 1]).

2 Grundlagen

Im Folgenden wird kurz auf die Historie von Anycast angegangen und die Adressierungsarten von IPv6 werden vorgestellt. Des Weiteren wird auf die nutzbaren Adressen von Anycast eingegangen und die Funktionsweise beider Anycast-Varianten erläutert. Zum Schluss dieses Kapitels wird noch die Lastverteilung mittels Anycast behandelt.

2.1 Historie

Im Jahr 1993 wurde Anycast schon für den Einsatz mit IPv4 als experimentelle Spezifikation in RFC 1524 definiert. Dabei wurde ein spezielles Präfix für Anycast vorgesehen, sodass sie von den Unicast-Adressen unterscheidbar sind. Dieses Verfahren wurde jedoch in der Praxis nicht implementiert. Stattdessen wurde eine Variante verwendet, die als Shared–Unicast-Address bekannt ist und bei der eine normale Unicast-Adresse mehreren Interfaces zugewiesen wird [Hag09, S. 63].

2.2 IPv6 Adressierungsarten

Das Internet Protocol version 6 (IPv6) unterscheidet drei Typen von IP-Adressen. Dies sind Unicast und Multicast wie in IPv4 sowie Anycast. Eine Unicast-Adresse ist eine eindeutige Kennung einer Netzwerkschnittstelle, wobei mehrere Schnittstellen nicht die

gleiche Unicast-Adresse erhalten dürfen. Pakete mit derselben Zieladresse erreichen denselben Netzwerkknoten. Im Gegensatz dazu wird eine Multicast-Adresse einer Gruppe von Knoten zugewiesen; alle Gruppenmitglieder erhalten die gleiche Multicast-Adresse und Pakete an diese Adresse werden an alle Mitglieder gleichzeitig ausgeliefert. Eine einzelne Anycast-Adresse wird wie eine Multicast-Adresse mehreren Knoten zugewiesen, jedoch kommuniziert immer nur einer der Knoten mit dem Absender. Die Tabelle 1 gibt einen Überblick über die drei Adressierungsarten [Doi04]:

	Unicast	Multicast	Anycast
Communication form	Point to point	Point to multipoint	Point to point
Target of address	Node	Group	Service type
Membership	Single	Multiple	Multiple
Roles in C/S model	Both	Client (listener)	Server

Tabelle 1: IPv6 Adressierungsarten [Doi04]

2.3 Ziele beim Einsatz von Anycast

Die Ziele beim Einsatz von Anycast sind [Abl06, S. 5f.]:

- Die grobe (unausgeglichene) Verteilung von Last über die Anycast-Nodes, um die Skalierbarkeit der Infrastruktur bei erhöhter Anzahl von Anfragen zu ermöglichen und vorrübergehenden Anfragespitzen entgegenwirken zu können.
- Abschwächung und Gebietseinschränkung von Denial of Service Attacken.
- Um im Fall eines Angriffes den Ort von Traffic verursachenden Quellen eingrenzen zu können, die gefälschte Quelladressen nutzen, da für die Wahl des Anycast-Knotens nicht die Quelladresse, sondern der Standpunkt des Quellknotens im Netzwerk entscheidend ist.
- Verbesserung von Anfrage-Antwort-Zeiten durch die Verringerung der Netzwerkentfernung zwischen Client und Server.
- Um mehrere Server über eine einzige Adresse nutzen zu können.
- Verbesserung der Dienstzuverlässigkeit durch die Bereitstellung einer automatischen Ausfallsicherung, d.h. Umschaltung auf einen anderen Anycast-Knoten [Abl06, S. 8].

2.4 Anycast-Adressen

Anycast-Adressen werden dem Unicast-Adressbereich entnommen, wobei jedes definierte Unicast-Adressformat genutzt werden kann. Dadurch sind Anycast-Adressen syntaktisch nicht von Unicast-Adressen unterscheidbar. Wenn eine Unicast-Adresse mehr als einer Netzwerkschnittstelle zugewiesen wird, so wird die Adresse zu einer Anycast-Adresse, die aber auch explizit als solche auf den entsprechenden Knoten deklariert sein muss [Hin06, S.12].

Es existieren auch reservierte Anycast-Adressen, auf die in RFC 2526 näher eingegangen wird [Joh99].

2.5 Anycast-Varianten

Bei Anycast können grundsätzlich zwei Varianten unterschieden werden, nämlich Off-Link-Anycast und On-Link-Anycast. Bei Off-Link-Anycast werden Routing-Protokolle und IP's Hop-by-Hop zielbasierendes Weiterleitungsparadigma genutzt, um Pakete zum „nächsten" Ziel zu leiten. Dies ist die herkömmliche Variante von Anycast, die hauptsächlich in dieser Seminararbeit behandelt wird und die mit IPv4 und IPv6 genutzt werden kann. On-Link-Anycast ist die formale Unterstützung von Anycast im Adressauflösungsprotokoll und ist nur für IPv6 standardisiert, mit der Einführung von gekennzeichneten Anycast-Adressen auf den Anycast-Hosts und dem Override-Flag in Neighbor Advertisements (NAs) des Neighbor Discovery Protokolls. Dafür gibt es in IPv4 keinen standardisierten Mechanismus [McP12, S. 3].

2.6 Funktionsweise von Off-Link-Anycast

In Abbildung 1 ist ein beispielhaftes Kommunikationsszenario dargestellt, an dem nun die Funktionsweise von Off-Link-Anycast erläutert wird. Die Anycast-Adresse A_{any} ist dabei drei Knoten zugewiesen. Wenn der Knoten mit der Unicast-Adresse S ein Paket an die Adresse A_{any} sendet, wird das Paket nur an einen der drei Anycast-Knoten ausgeliefert, in diesem Diagram an X_{uni}. Der Vorteil von Anycast ist somit, dass ein Knoten einen bestimmten Dienst nutzen kann, ohne die aktuellen Zustände von Dienstknoten und Netzwerken zu kennen. Wenn nämlich X_{uni} nicht mehr funktionsfähig ist, kann das Paket für A_{any} an einen anderen Host weitergereicht werden (Y_{uni} oder Z_{uni}) [Doi04].

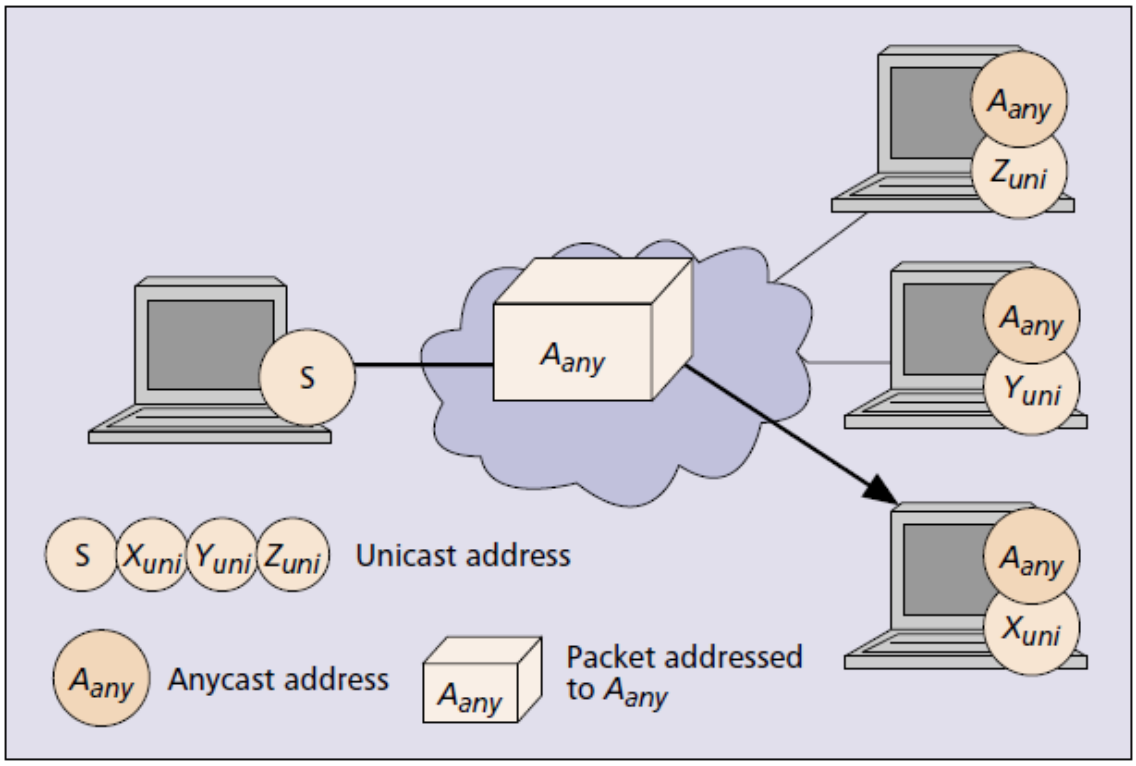

Abbildung 1: Anycast-Kommunikation [Doi04]

Das Routing-System entscheidet, an welchen Anycast-Netzknoten eine Anfrage geschickt wird, wobei dies vom Standort des anfragenden Systems abhängt. Für Anycast wird kein extra Routing-Protokoll benötigt. Stattdessen wird das gewöhnliche Unicast-Routing verwendet, um Anycast-Pakete auszuliefern, da die Anycast-Adressierung zum größten Teil der Adressierung von Unicast entspricht. Somit wird ein

Anycast-Paket immer an den „nächsten“ Anycast-Knoten ausgeliefert, der über die Metrik des eingesetzten Routing-Protokolls ermittelt wird ([Abl06, S. 5], [McP12, S. 3]).

Damit das Routing funktioniert, muss jedoch jeder Anycast-Knoten selbst eine Route für die Anycast-Adresse bekanntgeben; siehe hierzu auch Kapitel 4.1 [Abl06, S. 3, 9].

2.7 Funktionsweise von On-Link-Anycast

Für die Auflösung von IPv6 Anycast-Adressen in Link Layer-Adressen, wird wie bei IPv6 Unicast-Adressen auch, das Neighbor Discovery-Protokoll verwendet [Hag09, S. 99, 121].

Wenn ein Knoten eine Link Layer-Adresse eines lokalen Empfängers ermitteln möchte, so schickt er eine Neighbor Solicitation-Nachricht an die Solicited-Node Multicast-Adresse des Empfängers [Hag09, S. 121].

Bei der Solicited-Node Multicast-Adresse handelt es sich um eine Multicast-Adresse, für die sich jeder Knoten für jede seiner Anycast- und Unicast-Adressen registrieren muss. Eine Solicited-Node Multicast-Gruppe für eine Anycast-Adresse besteht somit aus mehreren Mitgliedern und für eine Unicast-Adresse nur aus einem Mitglied. Gebildet wird die Solicited-Node Multicast-Adresse aus einem festgelegten Präfix und einem Teil der Interface ID der IPv6-Empfängeradresse [Hag09, S. 70f.].

Vor dem Versand der Neighbor Solicitation-Nachricht wird dann noch aus der Solicited-Node Multicast-Adresse nach festgelegten Regeln die entsprechende MAC Multicast-Adresse erstellt; für die Auflösung von IPv6 Multicast-Adressen in Link Layer-Adressen (MAC-Adressen) wird das Neighbor Discovery-Protokoll also nicht benötigt [Hag09, S. 71, 121].

Die Empfänger der Neighbor Solicitation-Nachricht schicken anschließend eine Neighbor Advertisement-Nachricht mit ihrer Link Layer-Adresse als Antwort zurück. Bei der Auflösung von Anycast-Adressen gibt es aber nun zwei Unterschiede im Vergleich zur Unicast-Adressauflösung: Die Versendung der Neighbor Advertisement-Nachricht sollte durch eine Zufallszeit zwischen 0 und 1 Sekunde verzögern werden, um die Wahrscheinlichkeit einer Netzwerküberlastung zu verringern. Außerdem sollte das Override-Flag in der Neighbor Advertisement auf 0 gesetzt werden, sodass das erste empfangene Advertisement genutzt wird, statt dem zuletzt erhaltenen Advertisement ([Hag09, S. 107-109], [Nar07, S. 67, 78]).

Durch das Override-Flag wird dem Empfänger signalisiert, dass er bestehende Einträge im Neighbor Cache mit den Informationen im Advertisement überschreiben und die Link Layer-Adresse im Cache aktualisieren soll. Ist das Override-Flag jedoch nicht gesetzt, wie dies bei Anycast-Adressen der Fall sein sollte, so werden im Cache bestehende Link Layer-Adressen nicht aktualisiert; existiert jedoch keine Link Layer-Adresse für einen Neighbor Cache-Eintrag, so wird die Link Layer-Adresse aus dem Advertisement eingetragen [Hag09, S. 108].

2.8 Lastverteilung

Die Lastverteilung zwischen Anycast-Knoten ist im Allgemeinen nicht ausgeglichen, da bei der Auswahl des Anycast-Knotens nur die Metrik des Routing-Protokolls eine Rolle spielt, die für gewöhnlich unabhängig von der Last des Zielhosts ist. Jedoch kann die

grobe Verteilung der Last zwischen Knoten zum Zwecke der Verbesserung der Zuverlässigkeit und der Skalierbarkeit genutzt werden [Abl06, S. 5].

3 Routing-Systeme

In diesem Kapitel wird auf die Bedeutung der Routing-Stabilität und Transaktionsdauer beim Einsatz von Anycast eingegangen. Zudem werden die Charakteristiken der Routing-Systeme in IGP-Netzwerken und dem Internet aufgezeigt.

3.1 Routing-Stabilität und Transaktionsdauer

Wenn ein Dienst per Anycast bereitgestellt wird, entscheidet das Routing-System, welcher Anycast-Knoten für einen Client ausgewählt wird. Da für gewöhnlich eine Client-Server-Interaktion zwischen einem Client und demselben Server für die Dauer der Transaktion ausgeführt werden muss, folgt daraus, dass sich die Knoten-Auswahl des Routing-Systems mindestens für die erwartete Transaktionsdauer nicht mehr ändern sollte [Abl06, S. 6].

Manche Dienste, wie z.B. DNS über UDP, haben sehr kurze Transaktionszeiten und bestehen sogar nur aus einem einzelnen Anfragepaket und einem einzelnen Antwortpaket. Transaktionen anderer Dienste können jedoch auch viel langlebiger sein, wie z.B. große Dateidownloads oder Videostreaming. Dies gilt auch analog zur Stabilität von Routing-Systemen. Manche Routing-Systeme sind sehr vorhersehbar und bleiben für eine lange Zeitspanne stabil, wie z.B. innerhalb eines gut gemanagten und einfach aufgebauten IGP-Systems, wo die Knotenauswahl nur bei Ausfall von Netzknoten geändert wird. Auf der anderen Seite gibt es jedoch auch Routing-Systeme die viel weniger vorhersehbar agieren [Abl06, S. 6f.].

Bei der Entscheidung ob ein Dienst für die Verteilung mit Anycast geeignet ist, sollte die Stabilität des Routing-Systems zusammen mit der Transaktionszeit des Dienstes genau verglichen werden.
Für neue Protokolle kann es in einigen Fällen sinnvoll sein, große Transaktionen in eine Initialisierungsphase aufzuteilen, die von Anycast Servern gehandhabt wird und in eine ununterbrochenen Phase, die zwischen nicht Anycast-Servern stattfindet und vielleicht während der Initialisierungsphase gewählt wurden [Abl06, S. 7].

3.2 Anycast innerhalb eines IGP-Netzwerkes

Wenn ein Dienst innerhalb eines IGP-Netzwerkes mit Hilfe von Anycast verteilt wird, steht das Routing-System normalerweise unter derselben Leitung der Organisation, die auch den Dienst zur Verfügung stellt. Dadurch ist es sehr wahrscheinlich, dass die Charakteristiken der Diensttransaktionen und der Netzwerkstabilität gut verstanden werden. Folglich ist diese Einsatzart für eine größere Anzahl von Anwendungen passend, als bei der Internet-weiten Anycast-Dienstverteilung (siehe Kapitel 3.1) [Abl06, S. 8].

Bei einem IGP-System gibt es im Allgemeinen für zu importierende Routen keine Längeneinschränkung des Präfixes. Aus diesem Grund muss kein Covering-Präfix für Dienst- bzw. Anycast-Adressen erstellt werden. Stattdessen können für die Dienstadressen Host-Routen in das Routing-System eingefügt werden. Weitere

Informationen über die Erfordernis von Covering-Präfixen finden sich in Kapitel 5.2 [Abl06, S. 8].

IGP-Systeme weisen oft nur wenig oder keine Aggregation von Routen auf, welches teilweise an der algorithmischen Komplexität zur Unterstützung dieser Methode liegt. Es gibt auch nur wenig Motivation für die Aggregation in vielen IGP-Netzwerken, da sich die zu verwaltenden Routing-Informationen in Grenzen halten. Skalierungsprobleme treten bei den Routern also nicht auf [Abl06, S. 8].

Die Aggregationsrisiken bei anderen Routing-Systemen werden in Kapitel 5.6 behandelt.

3.3 Anycast innerhalb des Internets

Anycast kann auch genutzt werden, um einen Dienst im Internet zu verteilen. Die Unterschiede zwischen dieser Einsatzart und der Verteilung im IGP-Bereich sind Folgende [Abl06, S. 9]:

- Andere Leute kontrollieren im Allgemeinen das Routing-System und
- das zuständige Routing-Protokoll (BGP) und übliche Praktiken in seinem Einsatz bringen einige zusätzliche Beschränkungen mit sich, auf die nachfolgend eingegangen wird.

4 Routing-Besonderheiten

Dieses Kapitel behandelt mehrere Besonderheiten beim Routing von Anycast-Adressen.

4.1 Signalisierung der Dienstverfügbarkeit

Wenn in einem Routing-System eine Route zu einem bestimmten Anycast-Knoten eingetragen ist, werden an die entsprechende Anycast-Adresse gerichtete Pakete an diesen Knoten weitergeleitet. Da der Knoten bereit sein muss Anfragen zu akzeptieren, bevor Pakete an ihn weitergeleitet werden, ist eine Kopplung zwischen der Route und der Verfügbarkeit eines Dienstes auf einem bestimmten Knoten wünschenswert [Abl06, S. 9].

Eine solche Kopplung kann erreicht werden, indem vereinbart wird, dass für jede einzelne Dienst- bzw. Anycast-Adresse eine eigene Route verwaltet wird. Die Verfügbarkeit eines Dienstes würde dann eine Routen-Bekanntgabe auslösen und die Nichtverfügbarkeit eine Routen-Zurücknahme. Dies kann erreicht werden, indem auf den Anycast-Knoten zusätzlich Routing-Software eingesetzt wird, die entsprechend mit der Dienst-Software verknüpft werden muss. Diese Technik kommt auch beim F-Root-Name-Server zum Einsatz ([Abl06, S. 9], [Abl04, S. 1]).

4.2 Covering Präfix

In manchen Routing-Systemen, wie z.B. dem BGP-basierenden Routing-System des Internets, kann man sich nicht darauf verlassen, dass eine propagierte Host-Route auch im ganzen Netzwerk verteilt wird. Dies ist aber keine Protokolleinschränkung, sondern eine Konsequenz von Betriebsrichtlinien [Abl06, S. 10].

In solchen Fällen müssen Routen propagiert werden, die die Dienstadresse abdecken und ein ausreichend kurzes Präfix haben, welches nicht durch häufig eingesetzte Importrichtlinien vom Routing-System verworfen wird [Abl06, S. 10].

4.3 Kostengleiche Netzwerkpfade (Equal-cost paths)

Kostengleiche Pfade zum selben Ziel werden von einigen Routing-Systemen unterstützt. In Fällen wo mehrere kostengleiche Pfade existieren und zu unterschiedlichen Anycast-Knoten führen, besteht das Risiko, dass unterschiedliche Anfrage-Pakete einer Transaktion zu mehr als einem Knoten ausgeliefert werden [Abl06, S. 10].

Für Dienste die mit Hilfe von BGP im Internet verteilt werden, stellen kostengleiche Pfade normalerweise kein Problem dar, weil der verwendete Algorithmus von BGP meist einen einzelnen gleich bleibenden Endknoten für ein bestimmtes Ziel auswählt, unabhängig davon ob mehrere mögliche Pfade vorhanden sind. Es gibt jedoch auch BGP-Implementationen die eine Multi-Pfadauswahl ermöglichen [Abl06, S. 10f.].

IGPs unterstützen häufig kostengleiche Pfade. Eine Multi-Knotenauswahl für eine einzige Transaktion kann aber in den meisten Fällen vermieden werden durch die sorgfältige Wahl von IGP-Link-Metriken oder durch die Anwendung von equal-cost-multi-path (ECMP) Auswahlalgorithmen, durch die nur ein einzelner Knoten für eine Multi-Pakettransaktion ausgewählt wird [Abl06, S. 11].

Es gibt aber auch ECMP-Algorithmen die die zu verwendende Route für jedes Paket einzeln bestimmen, statt pro Transaktion. Deshalb sagt man auch, dass diese Art von Algorithmen eine „Per Packet Load Balancing“ (PPLB) ausführen. Die Benutzung von PPLB kann aber in manchen Fällen dazu führen, dass Pakete einer einzelnen Multi-Pakettransaktion, die durch einen Client verschickt wurden, unterschiedlichen Anycast-Knoten ausgeliefert werden und somit den Anycast-Dienst praktisch nicht verfügbar machen. Weitere Details hierzu können in RFC4786 nachgelesen werden [Abl06, S. 11].

4.4 Routen-Dämpfung (Route Dampening)

Häufige Bekanntgaben und Zurückziehungen von einzelnen Präfixen in BGP sind bekannt als Flaps. Schnelles Flapping kann zu hohen CPU-Auslastungen auf Routern führen, die von der Quelle der Instabilität weit entfernt sind und aus diesem Grund werden schnelle Routen-Schwankungen oftmals gedämpft [Abl06, S. 12].

Ein gedämpfter Pfad wird von Routern für eine Zeitspanne unterdrückt und nicht propagiert, wobei sich die Zeitspanne entsprechend der Frequenz der beobachteten Schwankung erhöht. Ein einzelner Router kann also die Propagierung eines flapping Präfixes zum Rest eines autonomen Systems verhindern und schützt so andere Router im Netzwerk vor Instabilität [Abl06, S. 12].

Um die Dämpfung eines Pfades und somit auch die Nichtverfügbarmachung eines Anycast-Knotens zu verhindern, sollten die Anycast-Knoten so konfiguriert werden, dass schnelle Schwankungen im Routing-System verhindert werden. Dies kann z.B. durch die Einführung einer minimalen Verzögerung nach einer Routen-Zurückziehung erreicht werden, bevor die Route wieder bekanntgegeben werden kann. Weitere

Details und Möglichkeiten zur Minderung des Problems befinden sich in RFC4786 [Abl06, S. 6, 12].

4.5 Reverse Path Forwarding Checks

Reverse Path Forwarding (RPF) Checks werden häufig als Teil von Ingress-Schnittstellenpaketfiltern im Internet auf Routern eingesetzt, um Pakete abzulehnen, deren Quelladresse gefälscht ist [Abl06, S. 13].

Bei einigen Betriebsarten von RPF ist es aber möglich, dass nicht gefälschte Pakete abgelehnt werden, wenn sie von multi-homed Sites stammen. Eine multi-homed Site bezeichnet hierbei einen Standort, der über mehrere globale IP-Adressen mit dem Internet verbunden ist ([Abl06, S. 13], [Hag09]).

Da für ein Routing-System eine Sammlung von Anycast-Knoten, die im Internet eingesetzt werden, größtenteils nicht unterscheidbar von einer verteilten Multi-homed Site ist, existiert das Risiko der Paketablehnung auch für Anycast-Knoten, selbst wenn einzelne Knoten gar nicht multi-homed sind. Für die Problematik mit multi-homed Sites bietet RFC 3704 jedoch brauchbare Lösungen an, weshalb sichergestellt werden sollte, dass jeder Anycast-Knoten als multi-homed Netzwerk behandelt wird und entsprechende Empfehlungen in RFC 3704 bezüglich RPF-Checks beachtet werden ([Abl06, S. 13], [Bak04]).

4.6 Aggregationsrisiken

Die Propagierung einer einzelnen Route für jeden Anycast-Dienst skaliert nicht gut für Routing-Systeme, bei denen die zu transportierende Last an Routing-Informationen von Bedeutung ist und wo es viele Dienste gibt, die verteilt werden müssen [Abl06, S. 14].

Die weit verbreitete Praktik minimale Präfixlängenfilter in Import-Richtlinien im Internet einzusetzen, bedeutet, dass für eine Covering-Route einer Dienstadresse die Präfixlänge wesentlich kleiner sein muss als jene für die Propagierung einer Host-Route. Durch die weitverbreitete Bekanntgabe von kurzen Präfixen für einzelne Dienste werden aber auch IP-Adressen verschwendet, da von den abgedeckten IP-Adressen der Covering-Route nur eine einzige Adresse genutzt wird [Abl06, S. 14f.].

Beide Probleme können einigermaßen durch die Benutzung eines einzelnen Covering-Präfixes für mehrere Anycast-Adressen gemildert werden, wenn mehrere Dienste auf der gleichen Gruppe von Anycast-Knoten betrieben werden. D.h. es werden Host-Routen von mehreren Anycast-Adressen zu einer Covering-Route aggregiert. Dadurch gibt es jedoch keinen Zusammenhang mehr zwischen der Routen-Bekanntgabe und der Verfügbarkeit einzelner Dienste (siehe Kapitel 5.1); dies kann wiederrum die Stabilität der Dienste gefährden. Weitere Informationen hierzu liefert Kapitel 4.8 von RFC4786 [Abl06, S. 15, 17].

Das hier beschriebene Skalierungsproblem verhindert im Allgemeinen, dass Anycast als nützlicher, genereller Ansatz für die Dienstverteilung im Internet genutzt werden kann. Anycast ist jedoch eine nützliche Technik für die Verteilung einer begrenzten Anzahl von Internet-kritischen Diensten sowie in kleineren Netzwerken, wo die Routen-Aggregation nicht benötigt wird ([Abl06, S. 8, 15], [Doi04, S. 3]).

5 Datensynchronisation

Es gibt ein paar Dienste, welche in lokalisierter Form eingesetzt werden, sodass Clients abhängig von ihrer Lage im Netzwerk einen anderen Inhalt präsentiert bekommen. Viele Dienste haben jedoch die Eigenschaft, dass Antworten auf Client-Anfragen gleich bleiben sollen, unabhängig vom Standort des Clients. Für einen mit Anycast verteilten Dienst bedeutet dies, dass alle Anycast-Knoten in einer einheitlichen Art und Weise betrieben und ggf. auch relevante Daten zwischen den Knoten synchronisiert werden müssen [Abl06, S. 15f.].

Der Synchronisationsmechanismus hängt dabei von der Art des Dienstes ab; Beispiele sind Zone-Transfers für autoritative DNS-Server und Rsync für FTP-Archive. Wird die Datensynchronisation über öffentliche Netzwerke ausgeführt, so sollten geeignete Authentifizierung- und Verschlüsselungstechniken eingesetzt werden [Abl06, S. 16].

6 Transportprotokolle

Zustandsorientierte Transportprotokolle (z.B. TCP) sind im Allgemeinen nicht für die Verwendung mit Anycast-Adressen geeignet, da sie die Eigenschaften von Anycast nicht kennen und somit beim Vorhandensein normaler Routing-Dynamik fehlschlagen. Wenn nämlich Pakete, die mit einer aktiven Transaktion verbunden sind, zu einem neuen Anycast-Endsystem geroutet werden, fehlen diesem Endsystem die Zustandsdaten der aktiven Transaktion, sodass die Sitzung rückgesetzt wird und wieder initiiert werden muss. UDP ist daher heutzutage die „Verkehrssprache“ für Anycast [McP12, S. 8, 10].

Trotzdem ist es möglich, auch TCP-Anwendungen mit Anycast zu nutzen, falls das Routing-System relativ stabil ist; siehe hierzu auch Kapitel 3.1 [Abl06, S. 6f.].

7 Sicherer Einsatz von Anycast-Adressen

Im Folgenden wird beschrieben, welche Voraussetzungen erfüllt sein müssen, damit Anycast-Adressen ohne Probleme als Ziel- oder Quelladresse genutzt werden können.

7.1 Anycast-Adressen als Ziel

Anycast-Adressen können „sicher“ als Zieladresse für eine Anwendung genutzt werden, wenn alle der folgenden Punkte erfüllt sind [McP12, S. 8f.]:

- Eine Anfragenachricht oder „einmalige“ Nachricht passt komplett in ein einziges Transportpaket.
- Ein zustandsloses Transportprotokoll, wie z.B. UDP, wird zum Verschicken der Nachricht verwendet.
- Bei Antworten wird immer eine Unicast-Adresse als Zieladresse genutzt. Dadurch ist gewährleistet, dass eine Antwort nicht fälschlicherweise von einem Anycast-Knoten empfangen wird, der diese Antwort gar nicht angefordert hat. Außerdem können Antworten somit auch aus mehreren Paketen bestehen, da das Zielsystem immer gleich bleibt.

- Die Server-Seite der Anwendungssoftware hält keinen Zustand über verschiedene Anfragen hinweg.
- Die wiederholte Versendung von Nachrichten ist idempotent.

7.2 Anycast-Adressen als Quelle

Wenn alle der folgenden Punkte erfüllt sind, können Anycast-Adressen „sicher“ als Quelladresse für eine Anwendung genutzt werden [McP12, S. 9]:

- Es wird keine Antwortnachricht durch den Empfänger erstellt, welche als Zieladresse die Anycast-Quelladresse der Anfragenachricht beinhaltet, außer die Anwendung verfügt über eine Zustandssynchronisation zwischen den Anycast-Knoten, sodass die Antwortnachricht auch korrekt von einem anderen Anycast-Knoten bearbeitet werden kann.
- Falls Unicast Reverse Path Forwarding (RPF) Checks auf Routern eingesetzt werden, so muss die als Quelle verwendete Anycast-Adresse für die Router erreichbar sein, d.h. für die Quelladresse muss eine Route vorhanden sein oder die Anycast-Adresse muss von den RPF-Checks ausgeschlossen werden [Kae07, S. 31f.]. Zusätzliche Informationen zu RPF-Checks liefert Kapitel 4.5.

8 Forschung

Es existieren einige Ansätze, welche das erwähnte Skalierungsproblem (siehe Kapitel 4.6) lösen und die Unterstützung von Anycast durch zustandsorientierte Transportprotokolle verbessern sollen. Diese Ansätze wurden aber bisher weder standardisiert noch haben sie weite Verbreitung erlangt. Dazu existieren auch keine RFCs oder aktive Internet-Drafts der IETF, außer RFC 1546, der die Modifikation von TCP vorschlug, um die Unterstützung von Anycast zu verbessern ([McP12, S. 10f.], [Wan09], [Wan07], [Szy06], [Has06], [Bal05], [Doi04], [Web04], [Aus07], [Kan07], [Doi05], [Mat05], [Doi04-2]).

9 Anwendungen

In diesem Kapitel werden ein paar Einsatzmöglichkeiten für Anycast vorgestellt, die auch in Dokumenten der IETF Erwähnung finden.

9.1 Domain Name System (DNS)

Bei DNS-Servern wird Anycast gerne genutzt um Redundanz hinzuzufügen, die die eigene Redundanz der DNS-Architektur ergänzen soll [Abl06, S. 3].

Im Jahr 2009 haben mindestens 10 von 13 Root-Name-Server Anycast benutzt. Anycast wird nun schon seit über einem Jahrzehnt für den DNS-Dienst eingesetzt und wird aktuell von einigen DNS Top-Level-Domain (TLD) Betreibern verwendet [McP12, S. 3].

Der Einsatz von Anycast bei DNS bietet auch zusätzlich den Vorteil, dass damit die Beschränkung der physikalischen Anzahl von autoritativen Nameservern überwunden werden kann [Lee07].

DNS kann gut mit Anycast eingesetzt werden, da in den meisten Fällen für die Namensauflösung eine zustandslose Verbindung verwendet wird und nur ein UDP-Paket für eine DNS-Anfrage und ein weiteres für eine DNS-Antwort benötigt wird. Autoritative Nameserver benötigen jedoch TCP für den Zone-Transfer, weshalb für diesen Prozess keine Anycast-Adresse als Zieladresse genutzt werden sollte [Lee07].

Ein Beispiel für den Einsatz des DNS-Dienstes mit Anycast wird im Artikel „IPv6 Anycast-based DNS Service Model“ erläutert [Lee07].

In RFC 4339 [Jeo06] wird der Einsatz von Anycast auch als ein möglicher Ansatz zur DNS-Konfiguration von IPv6 Hosts vorgestellt. Genauer gesagt sollen hierbei Hosts mit Adressen von rekursiven DNS-Servern, d.h. Servern die einen rekursiven DNS-Auflösungsdienst zur Verfügung stellen, konfiguriert werden. Diese Server werden von Hosts benötigt, um auf Internetdienste zugreifen zu können, die durch einen DNS-Namen identifiziert werden (z.B. http-Dienste) [Jeo06, S. 3].

Bei dem erwähnten Ansatz werden den rekursiven DNS-Servern allgemein bekannte Anycast-Adressen zugewiesen, welche dann in die Resolver-Konfigurationsdateien der Clients von Anfang an als Voreinstellung eingetragen werden. Dadurch müssen die DNS-Serveradressen nicht nachträglich z.B. per DHCP verteilt werden. Weitere Details, insbesondere zu Vor- und Nachteilen dieses und weiteren Konfigurationsansätzen sowie mögliche Einsatzszenarien, können dem erwähnten RFC entnommen werden [Jeo06, S. 1, 9].

9.2 Umlenkungsmechanismus für das Internet Key Exchange Protocol Version 2 (IKEv2)

Das IKEv2 ist ein Protokoll um Virtual Private Network (VPN) Tunnel von einem entfernten Standort zu einem Gateway aufzubauen, sodass der VPN-Client auf Netzwerkdienste hinter dem Gateway zugreifen kann. RFC 5685 [Dev09] definiert eine Erweiterung von IKEv2, die es einem VPN-Gateway erlaubt, einen VPN-Client zu einem anderen Gateway umzuleiten. Dies ist z.B. sinnvoll, wenn das VPN-Gateway überlastet ist [Dev09, S. 1].

Um VPN-Clients die IP-Adresse des VPN-Gateways mitzuteilen, wird oft DNS eingesetzt. Statt aber im DNS-System eine Liste von Unicast-IP-Adressen der VPN-Gateways zu hinterlegen, kann auch eine einzige Anycast-Adressen, die allen Gateways zugewiesen werden muss, eingetragen werden [Dev09, S. 2f.].

Der Ablauf der Umlenkung sieht folgendermaßen aus [Dev09, S. 4-6]:

1. Der VPN-Client führt einen DNS-Lookup für das VPN-Gateway aus.
2. Durch die Antwort des DNS-Servers erhält der Client die Anycast-Adresse des VPN-Gateways.
3. Eine IKE_SA_INIT-Anfrage wird dann vom VPN-Client an die Anycast-Adresse gesendet, welche den REDIRECT_SUPPORTED-Payload beinhaltet. IKE_SA_INIT ist hierbei eine Nachricht des IKEv2-Protokolls, welche als erste zum Aufbau einer Security Association verschickt werden muss [Kau05, S. 1, 4].
4. Anschließend wird die IKE_SA_INIT-Anfrage zu einem der VPN-Gateways, das Mitglied der Anycast-Gruppe ist, geroutet.

5. Das VPN-Gateway, das die Anfrage erhalten hat, antwortet darauf mit einer IKE_SA_INIT-Antwort von der Anycast-Adresse, wobei der REDIRECT-Payload enthalten sein muss, um den VPN-Client zu einer Unicast-Adresse der VPN-Gateways umzuleiten. Falls der angefragte VPN-Gateway selbst nicht überladen ist, so kann er den Client auch zu seiner eigenen Unicast-Adresse umleiten.
6. Als Nächstes verschickt der Client eine neue IKE_SA_INIT-Nachricht an den VPN-Gateway, der in der REDIRECT-Payload angegeben wurde. Dabei muss der Client der Nachricht den REDIRECTED_FROM-Payload mit der IP-Adresse des ursprünglichen Gateways anhängen.
7. Nun wird die normale IKEv2-Kommunikation zwischen zwei Unicast-Adressen fortgeführt.

Weitere Details können den referenzierten RFCs entnommen werden.

9.3 Anycast Rendezvous Point (RP) Mechanismus für Protocol Independent Multicast (PIM)

PIM-SM ist ein Multicast Routing Protokoll, das einen unidirektionalen Verteilungsbaum für jede Multicast-Gruppe erstellt. Ein Router der so konfiguriert wurde, dass er als Wurzel eines Verteilungsbaums fungiert, wird als Rendezvous Point (RP) bezeichnet. Beitrittsnachrichten von Empfängern für eine Gruppe werden zu dem RP gesendet und Daten von Sendern, sodass die Empfänger herausfinden können wer die Sender sind und mit dem Empfang von Traffic für die Gruppe beginnen können [Fen06, S. 1-6].

Bei PIM-SM darf aber nur ein einziger RP pro Gruppe aktiv sein, was mehrere negative Auswirkungen hat wie eine langsame Konvergenz wenn ein aktiver RP ausfällt und eine schlechte Skalierbarkeit. RFC 3446 schlägt daher einen Mechanismus vor, der eine beliebige Anzahl von RPs pro Gruppe in einer PIM-SM Domäne erlaubt. Dies wird realisiert, indem allen RPs einer Gruppe eine identische Anycast-Adresse zugewiesen wird. Die Anycast-Adresse darf jedoch nicht immer verwendet werden, da dies zu Problemen führen kann; Einzelheiten hierzu befinden sich im entsprechenden RFC [Kim03].

Für den in RFC 3446 beschriebenen Anycast-RP Mechanismus wird das Multicast Source Discovery Protocol (MSDP) benötigt. Anycast-RP kann jedoch auch ohne MSDP genutzt werden; diese Möglichkeit wird in RFC 4610 beschrieben, bei der das PIM-Protokoll erweitert wird ([Kim03], [Far06]).

9.4 Weitere Anwendungen

In RFC2526 wurde außerdem eine Anycast-Adresse für Mobile IPv6 Home-Agents reserviert und RFC3068 legt eine Anycast-Adresse für 6to4 Relay Routers fest ([Joh99], [Hui01]).

Des Weiteren definiert RFC 4291 eine Subnet-Router Anycast-Adresse, welche zur Adressierung aller Router in einem Subnetz gedacht ist [Hin06, S. 12f.].

Anycast kann aber auch zum Aufbau von Server-Clustern genutzt werden sowie für Content Distribution Networks ([Abl06, S. 9], [McP12, S. 13]).

10 Zusammenfassung

Anycast hat einige Vorteile wie z.B. die Verbesserung der Verfügbarkeit und wird in der Praxis eingesetzt. Es eignet sich aber nur bedingt für TCP-Anwendungen und die Anzahl möglicher Internet-weiter Anycast-Dienste ist beschränkt. Außerdem müssen beim Routing von Anycast-Adressen einige Dinge beachtet werden. Durch neue Forschungsansätze können jedoch vorhandene Probleme mit Anycast gelöst und somit die Einsatzmöglichkeiten erweitert werden ([Abl06, S. 5 - 15], [McP12, S. 3-11], [Wan09], [Wan07], [Szy06], [Has06], [Bal05], [Doi04], [Web04], [Aus07], [Kan07], [Doi05], [Mat05], [Doi04-2]).

11 Literatur

[Abl04] Abley J.: A Software Approach to Distributing Requests for DNS Service using GNU Zebra, ISC BIND 9 and FreeBSD, ISC-TN-2004-1, 2004.

[Abl06] Abley, J.; Lindqvist, K.: Operation of Anycast Services, RFC 4786, 2006.

[Aus07] Aus, M. S.; Borhanuddin, M. A.; Sabira, K.; Gopakumar, K.: An Enhanced IPv6 Anycast Routing Protocol Using Protocol Independent Multicast-Sparse Mode (PIM-SM). In: IEEE (Hrsg.), Proceedings of the 2007 IEEE International Conference on Telecommunications and Malaysia International Conference on Communications, Penang, Malaysia, S. 588 – 593, 2007.

[Bak04] Baker, F.; Savola, P.: Ingress Filtering for Multihomed Networks, RFC 3704, 2004.

[Bal05] Ballani, H; Francis, P.: Towards a Global IP Anycast Service. In: ACM (Hrsg.), SIGCOMM'05, Philadelphia, Pennsylvania, USA, S. 301 – 312, 2005.

[Dev09] Devarapalli, V.; Weniger, K.: Redirect Mechanism for the Internet Key Exchange Protocol Version 2 (IKEv2), RFC 5685, 2009.

[Doi04] Doi, S.; Ata, S.; Kitamura, H.; Murata, M.: IPv6 Anycast for Simple and Effective Service-oriented Communications. In: IEEE (Hrsg.), IEEE Communications Magazine, S. 163 – 171, 2004.

[Doi04-2] Doi, S.: Design, Implementation and Evaluation of Routing Protocols for IPv6 Anycast Communication, Master's Thesis, Osaka University, 2004.

[Doi05] Doi, S.; Ata, S.; Kitamura, H.; Murata, M.: Design, Implementation and Evaluation of Routing Protocols for IPv6 Anycast Communication. In: IEEE (Hrsg.), Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05), S. 1 – 6, 2005.

[Far06] Farinacci, D.; Cai, Y.: Anycast-RP Using Protocol Independent Multicast (PIM), RFC 4610, 2006.

[Fen06] Fenner, B.; Handley, M.; Holbrook, H.; Kouvelas, I.: Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised), RFC 4601, 2006.

[Hag09] Hagen, S.: IPv6 Grundlagen – Funktionalität – Integration, 2. Auflage, Sunny Edition, 2009.

[Has06] Hashimoto, M.; Kitamura, H.; Ata, S.; Murata, M.: Making practical use of IPv6 anycasting: Mobile IPv6 based approach. In: IEEE (Hrsg.), Proceedings of the 2005 Symposium on Applications and the Internet (SAINT'06), 2006.

[Hin06] Hinden, R.; Deering S.: IP Version 6 Addressing Architecture, RFC 4291, 2006.

[Hui01] Huitema, C.: An Anycast Prefix for 6to4 Relay Routers, RFC 3068, 2001.

[Jeo06] Jeong, J.: IPv6 Host Configuration of DNS Server Information Approaches, RFC 4339, 2006.

[Joh99] Johnson, D.; Deering, S.: Reserved IPv6 Subnet Anycast Addresses, RFC 2526, 1999.

[Kae07] Kaeo, M.: Current Operational Security Practices in Internet Service Provider Environments, RFC 4778, 2007.

[Kan07] Kang, Y.-H.; Jung, B.-G.: IPv6 Anycast Routing aware of a Service Flow. In: IEEE (Hrsg.), IEEE International Symposium, 2007

[Kau05] Kaufman, C.: Internet Key Exchange (IKEv2) Protocol, RFC 4306, 2005.

[Kim03] Kim, D.; Meyer, D.; Kilmer, H.; Farinacci, D.: Anycast Rendevous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP), RFC 3446, 2003.

[Lee07] Lee, S.; Suh, J.; Song, K.: IPv6 Anycast-based DNS Service Model. In: IEEE (Hrsg.), Proceedings of Asia-Pacific Conference on Communications, S. 247 - 250, 2007.

[Mat05] Matsunaga, S.; Ata, S.; Kitamura, H.; Murata, M.: Design and Implementation of IPv6 Anycast Routing Protocol: PIA-SM. In: IEEE (Hrsg.), Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05), S. 1 – 6, 2005.

[McP12] McPherson, D.; Oran, D.; Thaler D.; Osterweil E.: Architectural Considerations of IP Anycast, Internet draft, draft-iab-anycast-arch-implications-05, 2012 (expires April 2013).

[Nar07] Narten, T.; Nordmark, E.; Simpson, W.; Soliman, H.: Neighbor Discovery for IP version 6 (IPv6), RFC 4861, 2007.

[Szy06] Szymaniak, M.; Pierre, G.; van Steen, M.: Versatile Anycasting with Mobile IPv6. In: ACM (Hrsg.), AAA-Idea'06, Pisa, Italy, 2006.

[Wan09] Wang, X.; Qian, H.: Design and implementation of Anycast communication model in IPv6. In: John Wiley & Sons Ltd. (Hrsg.), INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT, S. 175 – 182, 2009.

[Wan07] Wang, X.; Qian, H.: Analysis and discussion of Anycast scalability in IPv6. In: John Wiley & Sons Ltd. (Hrsg.), INTERNATIONAL JOURNAL OF NETWORK MANAGEMENT, S. 321 – 328, 2007.

[Web04] Weber, S.; Cheng, L.: A Survey of Anycast in IPv6 Networks. In: IEEE (Hrsg.), IEEE Communications Magazine, S. 127 - 132, 2004.