APT: A Practical Tunneling Architecture for Routing Scalability

Haoran Bai
(hrbai523@hotmail.com)

Seminar „Internet Routing“ ,
Technische Universität Berlin
SS 2009

1.Einleitung

Wegen des Anstiegs der Anzahl von Edge-Netzen wird die Größe der Routing-Tabelle
rasant angewachsen sowohl in Größen als auch in Dynamik, was zu zwei Begrenzungen der heutigen Netz-architektur zurückzuführen ist. Erstens, Der Existenz des Konfliktes zwischen der Anbieter- basierten Adressierung und dem Dedarf vom Multihoming des Edge-Netzes, zweitens, Flat Routing nicht in der lage ist, die Trennung der Edge-Dynamik zu erzielen.

Um die beiden Begrenzungen zu ausweichen bzw. die Skalierbarkeit des Netzes nicht dabei beeinflußt zu machen, wird eine neue Tunnel-Architektur APT (A Practical Tunneling), der eine Erweiterung des Internet Routing Systems unabhängig von der Erweiterung von Edge-netzen erlaubt, ins Visier genommen.

2.Was ist von APT

Da die Realisierung von APT die eingentliche Realisierung vom Map & Encap Schema ist, müssen wir zunächst mit den Funktionsideen vom APT beschäftigen.

2.1 Map & Encap

2.1.1 Die Idee vom Map & Encap

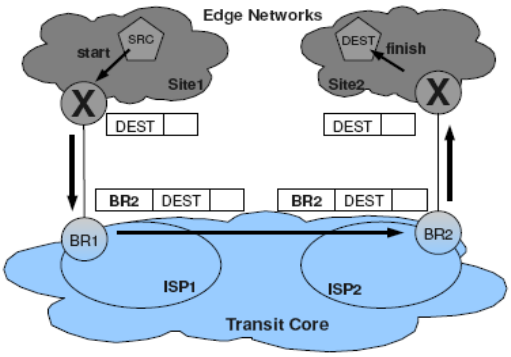

Abbildung1：Getrenntes Transit- und Edge-Network

Die Idee vom “Map & Encap” besteht darin, die stets sich vergrößernde Routingtabellen zu halten und verwalten, und gleichzeitig das Datenverkehr aufrecht zu halten, dem Dritten die Aufgaben übergeben, sich um die Routing-informationen zu kümmeren, um somit die ISP Routers zu erleichtern.

Wie Abbildung1, ein Host auf Seite 1 will nun ein Paket an Host auf Seite 2 übertragen. Das Paket wird zuerst beim Anbieter auf Seite 1 (ISP1) ankommen,

jedoch kann es nicht direkt an Seite 2 gesendet werden, weil Router auf Seite 1 keine Informationen über Edge-Prefix hat. Border Router des ISP1, der sogenannte BR1 bildet Mapping-Information der Ziel-adresse in Seite 2 ab (Map). Erst dann wird das Paket beim BR1(Ingress Tunnel Router ITR) komprimiert (Encap), danach durch dieses Transit-Kern an BR2 übermittelt. Nachdem das Paket beim BR2(Egress Tunnel Router ETR) angekommen und dekomprimiert ist, lässt es sich an den Host auf Seite 2 weiter senden.

2.1.2 Normen Vorstellungen

1). TR (Tunnel Router): TR sind die Routers, die sich auf der beiden Seiten einer erfolgreichen Daten- übertragung befinden und auch aus ITR, ETR bestehen.

- ITR (Ingress Tunnel Router):Der Eingangsrouter, der das Paket komprimeirt.

- ETR (Egress Tunnel Router): Der Ausgangsrouter, ein komprimiertes Paket dekompremiert.

2).Mapping-Information: Mapping-Information sind die Informationen, mit der die Kommunikation eines Edge-Präfix durch einen oder einigen ETR erzielt wird.

2.2Designsprinzipien

Bei der Realisierung vom Map & Encap hat APT die folgende Designprinzipien berücksichtigt.

1) Die bereitstehenden Internet-Dienste und derren Qualität nicht vernachlässigen. Die Skalierbarkeit erhöhen mit der gringsten Gefahr zum Verursachen von Störung zu den bereitstehenden Internet-Diensten.

2) Die Kosten (Investition) mit dem Profit ausrichten, indem Vor-Profit Fall (man profitiert von Instestionen der Anderen) ausgeschlossen ist.

3) Flexibilitäten beim Ausgleich zwischen Leistung und Ressourcen schaffen.

2.3 Einsatz des APTs

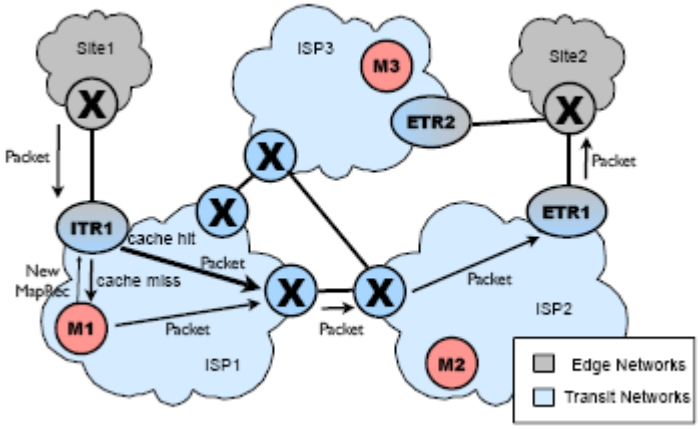

Abbildung2 Weiterleitung vom Daten-Paket aus [5]

Beim Einsatz von APT wird der Adress- Raum vom Internet in zwei Teilen aufgeteilt, einer für das Transit-Kern, der andere für das Edge-Netz.
APT positioniert TR(Tunnel-router)s auf der Anbieterseite der Verbindung zwischen Edge-netzwerken und dem Anbieter. Dazu gibt es zwei Hauptgründe, die unsereren Designprinzipien entsprechen.

Erstens, da Map & Encap zur Lösung der Erweiterungsproblems des routings und Erleichterung des ISP Routers dient, ist es selbstverständlich dass, die Anbieter die Invetitionen übernehmen.

Zweitens, ITR(Eingangs Tunnel-router) und ETR (Ausgangs Tunnel-router)können sich entweder in einem einzelnen ISP oder verschiedenen ISP befinden, das eine einseitige Anwendung des APTs von einem einzelnen Anbieter erlaubt.

Während des Einsatz von APT wird zuerst eine Mapping- Information vom ITR gebraucht, die dem ITR bei jeder Kommunikation zwischen Edge-Präfixen die Ortung vom richtigen ETR ermöglicht. Mapping-Informationen betreffen eine oder einigen Transit-Adressen für ein Edge-Präfix. Eine Mapping-Information aber gehört zu einem ETR in einem ISP und dient der bestimmten Edge-Netze, dazu sind zwei Normen P und W wichtig.

Kommt eine Mapping- Informationsanfrage von einem ITR, wird der ETR mit der höchsten Priorität ausgewählt. Sollte mehrere ETR-Adressen eine gleiche Priorität haben, werden sie ansicht Ihrer Werte proportional angewendet. Zur Generierung der Mapping-Information werden Prioritäten und Werte von Edge-Netzen mit Präfixen an ihren jeweiligen Provider gesendet. Dann baut ein DM jedes einzelnen Providers ein Mapset auf, z.B. ***MapSet (p, N)= {(d, w)*** | d ist eine ETR Adresse in N und d kann direkt mit p verbindet werden und w steht hier für die Priorität und den Wert für d.*}*

Im Cache vom ITR werden nur MapRecs, die abbildung von einem Edge-Präfix an die einzige ETR-Adresse, mit jeweiligem CIT (cache Idle timer) gespeichert. Der CIT wird nach jedem Nutzvorgang wiedergesetzt. Das Daten-Paket nach der

Komprimierung wird vom ITR am ETR direkt gesendet, beim positiven Suchergebnis von Mappinginformation im Cache des ITRs. Anderefalls beim fehlgeschlagenen Suchvorgang schickt der ITR das komprimierte Paket zu einem s.g. Default Mapper.

DM hier ist eine neue Anlage, spielt eine große Rolle beim APT. Darin wird eine vollständige Routing-Tabelle gespeichert und das ankommende Paket dekomprimiert, bei Anfrage vom ITR wird eine Suche nach der meisten anpassenden MapSets für Ziel-Adressen geführt,und ein MapRec mit CIT generiert. Der generierte MapRec wird zusammen mit seinem CIT an den ITR zurückgeschickt. Gleichzeitig wird das Paket mit der ausgewählten ETR-adresse wieder komprimiert ,und dazu getunnelt.

Im DM sind die Auswahllogik bereits vorhanden mit Mapset's Priorität und Wert, dadurch wird der Treffen der Entscheidungen bei übertragung vom großen Menge von Daten sowie Regelungtreffen von ITR vermeidet. Weitere Pakete mit der gleichen Ziel-Addresse werden vom ITR so lange an den DM verschickt, bis DM ihm eine Rückmeldung mitteilt. DM leitet die dann weiter, unterdrückt jedoch die duplizierten Steuerrungsinformationen an den ITR, durch Einsatz von einem Deaf Timer. MapRec wird nur erneut übermittelt, wenn der CIT wieder verfallen ist.

Im Vergleich zu den TRs verwaltet DM eine größere Routing-Tabelle und dafür relativ geringe Datenverkehr, während die TRs eine kleine Menge von Routing-Tabellen in TRs aber viel Daten-Verkehren verwalten. Und somit erzielen wir eine Ausbalancierung der Verteilung von Lasten.

3. Vorteile und Herausforderungen von APT

3.1 Vorteile

Indem Präfixe der Edge-Seiten vom Inter-Domain-System isoliert werden, reduziert sich die Größe der globalen Routing Systeme und dadurch verbessert sich die Skalierbarkeit. Außerdem können Edge-Netze innerhalb der Routing-Infrastruktur nicht mehr direkt miteinander kommunizieren, um so bei jeder Veränderung der Edge-Netzen vom ISP Generieren- und Speichernsvorgang von neun Informationen zu vermeiden, und so die Sicherheit zu verbessern.

3.2 Herausforderungen

Dabei stehen wir bei der Nutzung vo APT aber auch noch vor Herausforderungen. Beim Verschoben von Mapping Informationen an ITR, müssen eventuelle Verlusten der daten und Verzögerungen, die von den zusätzlichen Schnitten vom Abrufs der Mapping Informationen verursacht sind, so weit wie möglich minimiert werden, um die Reduktion der Dienstqualität vom Internet zu vermeiden. Alle Mapping-informationen in allen ITR gespeichert sind, könnten die Verlusten und Verzögerung vermieden werden, das kann aber dazu führen dass dabei sich die Mapping-tabelle mehr Fach vergrößern. Anderer Falls, sind Mappinginformationen vom dritten zu

herabrufen, wodurch mehr Verlusten und Verzögerungen vorkommen könnten. Außerdem wegen der Löschung der Edge-Netzwerke von der Routing-Tabelle des Transit-Kerns aus muss eine neue Weise zur
Entdeckung der Fehler vom Map & Encap angeboten werden. Darüber hinaus bezieht sich auf Sicherheiten des Internet, dass Mapping einerseits anbietet, die Internet-sicherheit zu verbessern, andererseits in der Lage ist, z.B. die Verfälschung zu machen.

4. Lösungen der Fragen in APT

4.1. Erkennung der Fehler und Wiederherstellungen

1) Fehler vom Transit-Präfix: Wird sehen als Verloren vom Cache angesehen und das Paket wird nach dem lokalen Default Mapper,wie Abbildung3 ***M1***, gesendet, allerdings mit einem relativ kürzeren CIT in MapRec, was wieder an ITR1 zurückgesendet wird.

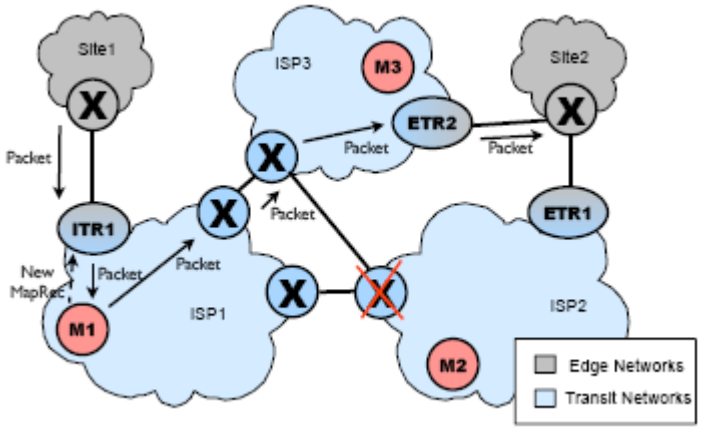

Abbildung3：Fehler vom Transit-Präfix aus[5]

2) Fehler von ETRs:Ist ein ETR nicht zu erreichen, nennt der DM dem ITR einen alternativen ETR und sendet das Paket weiter. Gleichzeitig stellt DM TBR (Time before Retry) Timer zu einer nicht-null Zahl, was nomalerweise in DM für jeden einzelnen ETR gespeicht und auf Null gestellt ist.

Wie Abbildung4, aufgrund des Fehlers vom ETR1 wird das Paket nach ***M2*** geschickt und durch ETR2 an Seite2 angekommen. Auch die Fehler-Nachrichten werden vom M2 an M1 zurückgegeben. Indem TBR für ETR1 auf nicht-Null gestellt ist, wird ETR1 als nicht erreichbar gesehen.

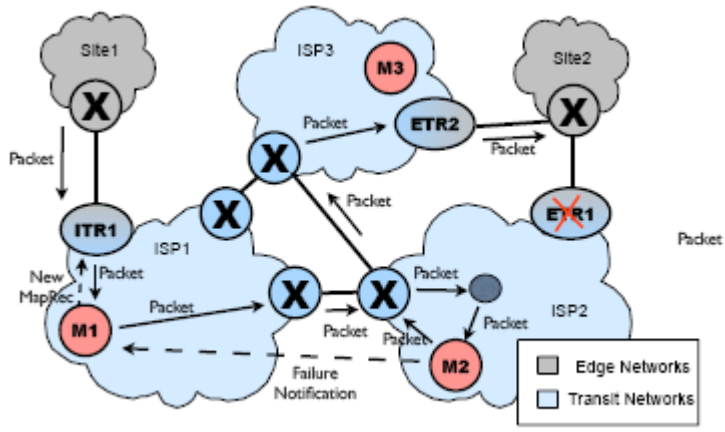

Abbildung4： Fehler von ETRs aus[5]

3) Fehler von der Edge-Netz-Erreichbarkeit Dieser Fehler weist sich auf sowohl fehlende Linke zwishen ETR und seinem Edge-Netz als auch Versagen des Edge-Netzes. Es wird genau so behandelt wie 2), nur die Nicht-erreichbare Information von Edge-Netzen wird zum DM des ITRs zurückgesendet wie die Abbildung 5.

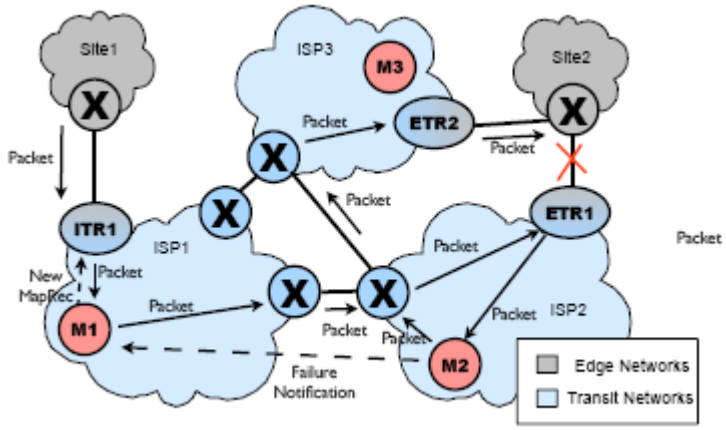

Abbildung5： Fehler von der Edge-Netz-Erreichbarkeit aus[5]

4.2. MDP

Eine MDP Information (Mapping Distribution Protocol) wird auch vom DM generiert, die aus der MapSet (p,X), DMany_ext und Header-Information besteht. Header- Information d.h. Kontrollinformationen enthalten AS Nummer, Sequenz Nummer, die für die Identität der MDP Information gültig sind. Durch DMany_ext, der mit anderem DMany_ext in benachbartem AS verknünpft, wird die Information MapSet (p,X) von allen DM in Transit-Netzwerken erhalten. Diese Verbindung der DM bildet ein DM-Netz ab. Wenn ein MDP seine Lebensdauer erreicht, der

generierende DM aktualisiert sie und meldet alle anderen DM mit Hilfe des DM-Netz.

4.3. Kryptografische Schutz

Zur Sicherheiten der ganzen Routing Systeme fügt man den kryptographischen Schutz für alle Kontrolle Informationen hinzu. Es ist vorausgesetzt, dass jedes Transit-Netz über sein eigene öffentlich-private Schlüsselpaar verfügt. Folgend wird die Frage diskutiert, wie die verteilung der Schlüssel in APT ohne PKI(Public Key Infrastructure) funktioniert.

4.3.1 Verteilung der Schlüssel

Beim APT haben wir durch Einsatz vom DM enorme Vorteile Betreff Sicherheit im Vergleich zu den herkömmlichen Internetdiensten.

In APT lassen sich die öffentlichen Schlüssel jederes Transit-Netzes durch DM-Netze übertragen. Damit die Verfälschungen der öffentlichen Schlüssel von Angreifer verhindert werden können, ist jedes Netzwerk von seiner Nachbaren zu überprüfen und dessen eigenen Schlüssel zu unterzeichnen. Nach einem begrenzten Zeitdauer werden Schlüssel ungültig, auch regelmäßig gewechselt.

4.3.2 Erkennung der Angriffe

Obwohl Edge-Netze nicht an Verteilung der Mapping Informationen teilnehmen, kann sie durch Einrichtung eines MDP-Überwachungs mit ihrem Anbieter schnell Fehler entdecken, bzw. sich von externen Angriffen beschützen, indem die fehlhaften Präfixe wieder durch neu generierten richtigen ersetzt werden.

5. Einsatzbare Möglichkeiten

Trotz der vielen Vorzüge von APT für jetztige Netzwerke, brauchen alle Netzwerke aber nicht zur gleichen Zeit APT anzuwenden. APT bietet die Abwärtskapatibilitäten für die APT langsam nutzende Edge-Seite an, indem die Mapping-Informationen zu BGP-Routen umgewechselt werden können.

Es muss klar sein, dass die Anwendung von APT auf gegenwärtige Internet nicht vernachlässigt. Die einzige zusätzliche neue Aufgabe eines Edge-Netzwerks bezieht sich darauf, die Verkehrsvorliebe der Informationen ihrem Provider abzugeben. Im ganzen Transit AS werden bei vollem Einsatz von APT die herkömmlichen BGP unverändert benutzt um die Transit-Präfixe zu bekommen. Die Border Routers werden in TRs umgewandelt und die Edge-Präfixe von derren BGP-Routing-Tabelle entfernt.

Nachdem man APT-Technik einsetzt, können APT Netzwerke und Nicht-APT-Netzwerke miteinander gut kooperieren. Dabei können AS mit APT mit AS komplett oder nur teilweise mit APT sowie AS ohne APT, kommunizieren.

6.Zusammenfassung

Heute wegen der schnellen Entwicklung des Internets kann die originale Architektur langsam nicht mehr den funktionalen Forderungen von Internet decken. Dafür ist eine neue Routing-Architektur notwendig. Dennoch bestehen einige Herausforderungen, etwa wie die Frage der Zieldivergierung von verschiedenen Parteien sowie die Frage der daraus entstandenen sich wiedersprechenden Anforderungen gelöst werden sollen. Danach stehen noch die Verbesserung der Zuverlässigkeiten und Leistungen der Edge- und Transit-Netzwerken und Aufteilung und Nutzung der Ressourcen in Frage. Außerdem muss sich auf die Frage der Ausrichtung der Investition mit dem Profitierung hinweisen.

7. Literaturverzeichnis

[1]S. Deering. The Map & Encap Scheme for Scalable IPv4 Routing with Portable Site Prefixes. Presentation, Xerox PARC, March 1996

[2]R. Hinden. New Scheme for Internet Routing and Addressing (ENCAPS) for IPNG. RFC 1955, 1996.

[3] X. Meng, Z. Xu, B. Zhang, G. Huston, S. Lu, and L. Zhang. IPv4 Address Allocation and BGP Routing Table Evolution. In ACM SIGCOMM CCR, Janurary 2005.

[4] D. Massey, L. Wang, B. Zhang, and L. Zhang. A scalable routing system design for future Internet. In Proc. of the ACM SIGCOMM Workshop on IPv6 and the Future of the Internet, Aug. 2007.

[5] Dan Jen, Michael Meisel, Daniel Massey, Lan Wang, Beichuan Zhang, Lixia Zhang. APT: A Practical Tunneling Architecture for Routing Scalability_