

Diss. ETH No. 22897
TIK-Schriftenreihe Nr. 158

Centralizing Routing Control Across Domains: Architectural Approach and Prominent Use Cases

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

VASILEIOS KOTRONIS

Diplomate Engineer, Electrical and Computer Engineering
National Technical University of Athens, Greece

born on March 10, 1989

citizen of Greece

accepted on the recommendation of
Prof. Dr. Bernhard Plattner, examiner
Prof. Dr. Nick Feamster, co-examiner
Asst Prof. Dr. Xenofontas Dimitropoulos, co-examiner
Dr. Bernhard Ager, co-examiner

2015

Abstract

The coordination and management of inter-domain routing in the Internet are crucial tasks that happen “behind the scenes”, while we use the Internet for checking our emails, surfing the web, making Skype video calls, watching Netflix episodes or engaging in social networking. The basic protocol that governs inter-domain routing, i.e., routing between the diverse administrative network domains, is the [Border Gateway Protocol \(BGP\)](#). Researchers have pinpointed several issues with the design and operation of BGP over the years, such as slow convergence times or lack of end-to-end service guarantees, and have proposed various solutions and improvements. Despite this large body of proposals, BGP is extremely resistant to change, albeit the only core protocol that virtually every [Internet Service Provider \(ISP\)](#) uses today to route packets between domains. Therefore, we need to find ways to address the requirements of new Internet services and evolve routing while maintaining compatibility with BGP. We thus propose a new architectural approach which can be incrementally deployed in today’s Internet and which can gradually tackle BGP-related challenges while its adoption gains momentum. Within the scope of this work, we address the challenges related to the suboptimal convergence behavior, suboptimal path computations and lack of inter-domain service guarantees. The approach is general enough to be used for optimizing further aspects of inter-domain routing beyond the scope of this work.

As our basic approach, we make the case for gradual routing centralization across domains based on [Software Defined Networking \(SDN\)](#) principles and outsourcing mechanisms. The envisioned setting is composed of contractors offering [Routing as a Service \(RaaS\)](#), serving—contiguous or disjoint—clusters (i.e., groups) of [Autonomous Systems \(AS\)](#), which are their clients. In general, contractors may operate on abstract views of their client networks. We show how the proposed architecture may be used as the basis for such improvements, by investigating two concrete and prominent use cases.

As a first use case, the RaaS contractors talk to external ASes via BGP, while providing inter-domain optimizations and new routing services within

their respective client clusters via SDN mechanisms. These optimizations, such as more stable routing and better convergence, may also benefit non-client ASes. In fact, we show that routing convergence improves w.r.t. time and load of routing updates, as the sphere of influence of RaaS contractors expands over hybrid BGP-SDN multi-domain networks. We have thus developed an emulation framework to run hybrid routing experiments.

As a second use case, we investigate RaaS contractors that work independently from BGP by performing overlay stitching of partial paths crossing ISP domains. The goal of this process is to form end-to-end routes with certain properties, such as guaranteed bandwidth, latency or availability. The stitching can be implemented, e.g., using well placed and connected programmable switching points at ISP interconnects. The setup on which we simulate our proposal is the rich overlay fabric that ISPs and their peering interconnections compose, utilized in a novel way for the provisioning of guaranteed end-to-end services across domains. This setup is aligned to the status quo in the ISP ecosystem, in terms of inter-domain topologies and peering relationships. In fact, our research is based on [Internet eXchange Points \(IXPs\)](#); these are physical infrastructures where multiple ISPs exchange data with each other.

To investigate the potential of the aforementioned peering fabric in terms of AS-level path diversity and associated metrics, we further develop a graph transformation algorithm. This algorithm can be used for the computation of min-cuts on arbitrary network graphs governed by policies. One of the application contexts of this computation is an inter-domain setting governed, e.g., by valley-free routing policies. Metrics that rely on policy-compliant min-cuts can be calculated centrally by a multi-domain RaaS contractor to evaluate routing availability, multipath throughput and reliability against link failures.

We investigate the benefits, implications and outlook of the general approach and its associated use cases, supporting the following hypothesis: *Inter-domain routing centralization, performed in a staged and consistent manner, can help to evolve and improve Internet routing on the AS level.* In the context of the studied use cases, the routing algorithms, protocols and services are evolving on top of the RaaS contractor's platform. This platform is the logically centralized routing control plane which operates on an abstract global view of the (inter-)network.

Kurzfassung

Die Koordination und die Verwaltung des Inter-Domain-Routings sind unverzichtbare Aufgaben im heutigen Internet, welche weitgehend unbemerkt “hinter den Kulissen” stattfinden, während wir das Internet nutzen um E-Mails zu lesen, auf Webseiten zu surfen, zu skypen, NetFlix-Serien anzuschauen oder Zeit in sozialen Netzwerken zu verbringen. Das Border-Gateway-Protocol (BGP) ist das Standardprotokoll, welches das Routing zwischen verschiedenartigen administrativen Netzwerkdomänen ermöglicht. Zahlreiche Forscher haben in den letzten Jahren verschiedene Schwachstellen der Gestaltung und Implementierung des BGP-Protokolls aufgezeigt, wie zum Beispiel hohe Konvergenzzeiten oder das Fehlen von Ende-to-Ende Service Garantien. In der Folge haben sie verschiedene Änderungen am BGP-Protokoll vorgeschlagen, welche die aufgezeigten Probleme beheben können. Trotz der zahlreichen Änderungsvorschläge ist das BGP-Protokoll sehr schwer zu ändern. Es ist zwar das einzige Protokoll, das heutzutage von praktisch jedem Internet-Service-Provider (ISP) unterstützt wird, um Pakete zwischen verschiedenen Domänen zu leiten. Daher benötigen wir einen neuartigen Ansatz für die Entwicklung des Inter-Domain-Routings, welcher auf der einen Seite die gestiegenen Anforderungen heutiger Internet-Dienste berücksichtigt, und auf der anderen Seite kompatibel zum BGP-Protokoll ist. Deshalb schlagen wir einen neuartigen, inkrementellen Ansatz vor, welcher schrittweise im heutigen Internet eingesetzt werden kann, um die erwähnten Herausforderungen nach und nach zu bewältigen. Im Rahmen dieser Arbeit befassen wir uns mit den Problemen, die aufgrund des suboptimalen Konvergenzverhaltens, der suboptimalen Pfadberechnungen und des Fehlens von Dienstgütegarantien im Betrieb des Inter-Domain-Routings zu lösen sind. Unser generischer Ansatz kann auch angewendet werden, um weitere Aspekte des Inter-Domain-Routings zu optimieren, die in dieser Arbeit nicht behandelt werden. Mit unserem Ansatz werben wir für eine stufenweise Zentralisierung des Inter-Domain-Routings, basierend auf den Prinzipien des Software-Defined-Networkings (SDN) und mittels Outsourcing-Mechanismen. In dem von uns betrachteten Szenario bieten Dienstleister Routing-as-a-Service (RaaS) an, welche von zusammenhängenden oder disjunkten Gruppen von Autonomen Systemen (AS) erworben werden können. Die Dienstleister (RaaS-Provider) operieren dabei mit einer abstrakten Sicht auf ihre Kunden-Netzwerke. In dieser Doktorarbeit untersuchen wir unseren Ansatz in zwei konkreten und prominenten Anwendungsfällen.

In unserem ersten Anwendungsfall kommunizieren die RaaS-Provider mit den AS ihrer Kunden über SDN-Mechanismen und mit externen AS über das BPG-Protokoll. Sie bieten dabei ihren Kunden sowohl Optimierungen des Inter-Domain-Routings als auch neue Routing-Dienstleistungen innerhalb der Kundengruppen an. Hierbei können auch AS, die keine Kunden sind, von den angebotenen Optimierungen, wie ein stabileres Routing oder ein besseres Konvergenzverhalten, profitieren. In der Tat zeigen wir, dass sich das Konvergenzverhalten des Routings in Bezug auf die Zeit und die Anzahl an Routing-Updates verbessert, wenn sich der Einflussbereich der RaaS-Provider über hybride BGP-SDN Multi-Domain Netzwerke ausweitet. Dazu haben wir ein Emulations-Framework entwickelt, mit dem wir hybride Routing-Experimente durchführen können.

In unserem zweiten Anwendungsfall betrachten wir RaaS-Provider, die virtuelle Routen—sog. Pathlets—die sich über verschiedene ISP-Domänen hinweg erstrecken, zusammenknüpfen, ohne dabei das BGP-Protokoll zu verwenden. Hierbei ist es das Ziel, Ende-to-Ende-Verbindungen aufzubauen, die bestimmte Eigenschaften erfüllen, wie zum Beispiel eine garantierte Übertragungskapazität, Latenz oder Verfügbarkeit. Pathlets verschiedener Domänen können über programmierbare Vermittlungsstellen miteinander verknüpft werden. Diese Vermittlungsstellen sind so ausgewählt, dass sie als leistungsfähige Schnittstellen zwischen verschiedenen ISPs dienen können. In Frage dafür kommen vor allem sog. Internet eXchange Points (IXP); dies sind physische Infrastrukturen, an welchen verschiedene ISPs untereinander Daten austauschen. Dabei wird das virtuelle Netzwerk so aufgebaut, dass Ende-zu-Ende Dienste mit garantierter Dienstgüte zwischen verschiedenen Domänen angeboten werden können. Unsere Simulationsumgebung berücksichtigt den gegenwärtigen Stand des ISP-Ökosystems, in Bezug auf die Inter-Domain-Topologien und Peering-Beziehungen.

Darüber hinaus haben wir einen Algorithmus für die Transformation der Inter-Domain-Graphen entwickelt, so dass wir das erwähnte virtuelle Netzwerk in Bezug auf die Diversität von Pfaden zwischen Autonomen Systemen sowie weitere assoziierte Metriken untersuchen können. Dieser Algorithmus kann dazu verwendet werden, minimale Schnitte in beliebigen Netzwerkgraphen zu berechnen. Diese Netzwerkgrahpe werden durch Richtlinien (Policies) gesteuert, wie zum Beispiel eine Valley-free Routing-Policy bei einem Inter-Domain-Setup. Metriken, welche auf Policy-konforme, minimale Schnitte angewiesen sind, können von einem Multi-Domain RaaS-Provider zentral berechnet werden. Hierbei ist es das Ziel, das Routing für einzelne Attribute wie die Verfügbarkeit, den Multi-Path-Durchsatz und die Ausfallsicherheit gegenüber Verbindungsfehler zu bewerten.

Wir untersuchen den Nutzen, die Auswirkungen und den Ausblick unseres generischen Ansatzes und der assoziierten Anwendungsfälle, um damit die folgende Hypothese zu stützen: *Eine stufenweise und konsistente Zentralisierung des Inter-Domain-Routings kann das Routing zwischen Autonomen Systemen verbessern und weiterentwickeln.* In den von uns untersuchten Anwendungsbeispielen werden die Routing-Algorithmen, die Routing-Protokolle und die Routing-Services auf der Plattform jedes RaaS-Providers weiterentwickelt. Diese Plattform stellt eine logisch zentralisierte Steuerung für das Routing dar und verwendet eine abstrakte, globale Sicht auf die beteiligten Netzwerke und die dazugehörigen netzwerkübergreifenden Links.