SFP: Toward Interdomain Routing for SDN Networks

Qiao Xiang	Chin Guok	Franck Le	John MacAuley	Harvey Newman	Y. Richard Yang
Tongji/Yale	LBNL	IBM	LBNL	Caltech	Tongji/Yale

ABSTRACT

Interdomain routing using BGP is widely deployed and well understood. The deployment of SDN in BGP domain networks, however, has not been systematically studied. In this paper, we first show that the *use-announcement inconsistency* is a fundamental mismatch in such a deployment, leading to serious issues including unnecessary blackholes, unnecessary reduced reachability, and permanent forwarding loops. We then design SFP, the first fine-grained interdomain routing protocol that extends BGP with fine-grained routing, eliminating the aforementioned mismatch. We develop two novel techniques, automatic receiver filtering and on-demand information dissemination, to address the scalability issue brought by fine-grained routing. Evaluating SFP using real network topologies and traces for intended settings, which are not global Internet but tens of collaborative domains, we show that SFP can reduce the amount of traffic affected by blackholes and loops by more than 50%, and that our proposed techniques can reduce the amount of signaling between ASes by 3 orders of magnitude compared with naive fine-grained routing.

CCS CONCEPTS

• **Networks** → **Network protocol design**; **Routing protocols**; **Programmable networks**;

KEYWORDS

Interdomain Routing, SDN, On-Demand Information Dissemination

ACM Reference Format:

Qiao Xiang, Chin Guok, Franck Le, John MacAuley, Harvey Newman, and Y. Richard Yang. 2018. SFP: Toward Interdomain Routing for SDN Networks. In *SIGCOMM Posters and Demos '18: ACM SIGCOMM 2018 Conference Posters and Demos, August 20–25, 2018, Budapest, Hungary*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3234200.3234207

SIGCOMM '18, August 20–25, 2018, Budapest, Hungary

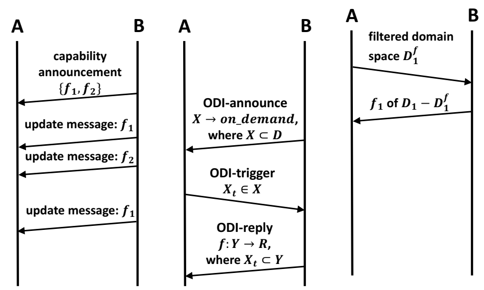
ACM ISBN 978-1-4503-5915-3/18/08…$15.00
https://doi.org/10.1145/3234200.3234207

1 INTRODUCTION

There are multiple important settings where multiple anonymous systems (ASes) interconnect to form collaborative networks (also called federations) to improve network performance of large scientific collaboration across member networks [2]. The *de facto* protocol to interconnect these ASes is the Border Gateway Protocol (BGP) [6]. Meanwhile, to support efficient usage of their network resources, members of federation networks commonly deploy software defined networking (SDN) [1, 3]. Though interdomain routing using BGP is well understood, the deployment of SDN in BGP-connected networks has not been systematically studied.

Specifically, we show that such a deployment reveals a fundamental mismatch between the fine-grained control by SDN and the coarse-grained routing by BGP, *i.e.*, the *use-announcement inconsistency*, which leads to serious issues. To illustrate these issues, consider two networks A, and B, connected using BGP. We focus on a prefix P. Assume that B drops all traffic sent to P with TCP destination port 22. If B still announces, through BGP, routes to P to its neighbor A, a subset of traffic (more specifically, traffic with destination port 22) may result in blackholes. Instead, if B does not announce any route to P to its neighbor A, such policy can result in reduced reachability for traffic to P with destination port other than 22. To further illustrate the issues, we assume that instead of dropping traffic with destination port 22, B decides to redirect this traffic to another network C and still announce the reachability of P to A via BGP. Such behavior can lead to permanent forwarding loops. Several solutions have been proposed to provide fine-grained routing. A particularly elegant one is SDX [5]. An issue of SDX, however, is that it requires a trusted third party to conduct integration.

Different from existing solution approaches, this paper investigates a simple, novel protocol named SFP (SDN Federation Protocol) that maintains the compatibility with BGP and at the same time provides *fine-grained routing*, where each network decides interdomain routes based on common packet header fields instead of destination IP only. We prove that SFP avoids all issues caused by use-announcement inconsistency. We develop two novel techniques, on-demand routing information dissemination and automatic receiver filtering, to address the messaging scalability issue brought by fine-grained routing. Evaluating SFP using real network topologies and traces, we show that by guaranteeing blackhole / loop-free routing, SFP can reduce the traffic affected



(a) Functional mapping dissemination. **(b)** On-demand information dissemination. **(c)** Automatic receiver filtering.

Figure 1: The message-time diagrams of SFP protocol.

by blackholes and forwarding loops by more than 50%, and the amount of signaling between ASes is reduced by 3 orders of magnitude compared with naive fine-grained routing.

2 SFP HIGH-LEVEL PROTOCOL DESIGN

Functional mapping information dissemination: SFP uses a functional mapping from the *general packet space* domain space – where each point is defined as a vector of common packet header fields (*e.g.*, the TCP/IP 5-tuple) and the address of an ingress, representing a packet entering the ingress of a network – to the *AS-path space*, where each point represents an AS-path, as the representation of the routing information base (RIB) of each network.

Property 1. *SFP with general packet space announcements has the same space partition structure as SDN, and hence can avoid unnecessary blackholes, reachability, or loops caused by use-announcement inconsistency.*

A major concern of fine-grained routing is its messaging scalability, as the RIB can become extremely large due to the multiple packet header fields, and the ensuing cross-product when composing fine-grained routes. To achieve scalability, we develop two novel techniques in SFP.

On-demand information dissemination (ODI): This technique allows a neighbor AS B to send incomplete functional mappings to AS A, and later send updated mappings to A based on A's demand triggers. In Figure 1b), B first sends an *ODI-announce* message $X \to on\_demand$ to A to indicate that the mapping of the subspace X is on-demand to A. When A needs the missing information, it sends an on-demand trigger X_t to B, where $X_t \subset X$, to ask for the information of the domain subspace X_t. Upon receiving the trigger, B looks in its local RIB for the mapping of X_t, invokes its SDN program to compute the mapping if it is not in RIB, sends other triggers to query other ASes if its SDN program does not provide the mapping, and returns an *ODI-reply* message $Y \to R$, the mapping information of a subspace $Y \supset X_t$. We prove that ODI does not introduce any new convergence issue. Given a single point in the general packet space, under the Gao-Rexford conditions [4], ODI always converges.

Automatic receiver filtering: Consider a packet *pkt*. If A can already make the decision for *pkt* without the information from a neighbor B, then the mapping sent from B to A does not need to include *pkt*. With this observation, SFP allows the receiver AS A to provide *receiver filtering* to notify the neighbor B which part of a mapping is of no use to A. In Figure 1c, A sends $\mathcal{D}_1'$, a filter on the domain space $\mathcal{D}_1$ of mapping f_1, to B. When B sends f_1 to A, it only needs to send the mapping from a smaller space $\mathcal{D}_1 - \mathcal{D}_1'$, reducing the amount of exchanged routing information.

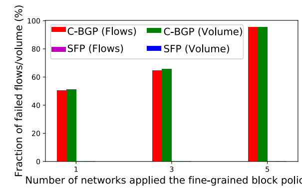
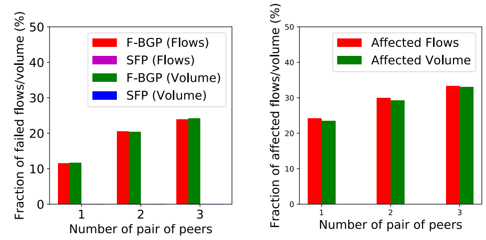


Figure 2: Loss when blocking one unused destination TCP port.



(a) Failed traffic **(b)** Traffic affected by loops

Figure 3: Loss when deflect traffic between neighboring peers.

3 PERFORMANCE EVALUATION

We evaluate SFP on the topology of LHCONE using real trace from the CMS experiment [2], a main source of traffic in LHCONE. We compare SFP with two variations of BGP: C-BGP, where a route to a destination IP will not be announced by an AS if the AS only uses this route for a subset of traffic for this IP, and F-BGP, where such a route will still be announced. Figure 2 shows that with 1 transit network deploying one fine-grained policy blocking traffic to one unused destination transport port, 50% of the flows, and 51% of the traffic volume can be dropped in C-BGP. In contrast, SFP ensures all 100% of flows to successfully reach their destinations. Figure 3a shows that with 1 pair of providers deploying one fine-grained policy to deflect large data transfers sent to their customer, 11% of the flows result in loops in F-BGP. In addition, flows traversing the affected links – although not resulting in loops – may still suffer high packet losses. Figure 3b show that 23% of the flows can get affected in F-BGP, whereas no flow is affected in SFP.

ACKNOWLEDGMENTS

The authors thank Jingxuan Zhang, Xin Wang and Yang Liu (Tongji University) for their help during the preparation of the paper. This research is supported in part by NSFC grants #61672385, #61472213 and #61702373; China Postdoctoral Science Foundation #2017-M611618; NSF grant CC-IIE #1440745; NSF award #1246133, ANSE; NSF award #1341024, CHOPIN; NSF award #1120138, US CMS Tier2; NSF award #1659403, SANDIE; DOE award #DE-AC02-07CH11359, SENSE; DOE/ASCR project #000219898; Google Research Award, and the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001.

REFERENCES

[1] Bosshart, P., Daly, D., Gibb, G., Izzard, M., McKeown, N., Rexford, J., Schlesinger, C., Talayco, D., Vahdat, A., Varghese, G., and Walker, D. P4: Programming protocol-independent packet processors. *SIGCOMM Comput. Commun. Rev. 44*, 3 (July 2014), 87–95.

[2] Collaboration, T. C. The CMS experiment at the CERN LHC. *Journal of Instrumentation 3*, 08 (2008).

[3] Foundation, O. N. Openflow switch specification 1.4.0. Open Networking Foundation (on-line), Oct. 2013.

[4] Gao, L., and Rexford, J. Stable internet routing without global coordination. *IEEE/ACM Transactions on Networking (TON)* (2001).

[5] Gupta, A., Vanbever, L., Shahbaz, M., Schlinker, S. P. D. B., Feamster, N., Rexford, J., Shenker, S., Clark, R., and Katz-Bassett, E. SDX: A Software Defined Internet Exchange. In *SIGCOMM'14*.

[6] Rekhter, Y., Hares, S., and Li, D. T. A Border Gateway Protocol 4 (BGP-4). RFC 4271, Jan. 2006.