Information Sciences 412–413 (2017) 101–115

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Scale the Internet routing table by generalized next hops of strict partial order

Qing Li[a,*], Mingwei Xu[c], Qi Li[b], Dan Wang[d], Yong Jiang[b], Shu-Tao Xia[b], Qingmin Liao[b]

[a] *Network Technology Laboratory, Huawei Technologies Co.Ltd, China*
[b] *Graduate School at Shenzhen, Tsinghua University, Shenzhen, Guangdong, China*
[c] *Tsinghua University, Beijing, China*
[d] *The Hong Kong Polytecnic University, Hong Kong*

ARTICLE INFO

Article history:
Received 8 December 2015
Revised 13 February 2017
Accepted 21 May 2017
Available online 25 May 2017

Keywords:
Routing scalability
Forwarding information base
FIB aggregation
SPO next hop

ABSTRACT

The Internet routing tables have been expanding at a dramatic and increasing rate. Although the latest high-performance routers provide enough capacities, Internet Service Providers (ISPs) cannot afford to upgrade their routers at the pace of routing table growth. Shrinking the routing table, especially the TCAM-based Forwarding Information Base (FIB), is more feasible. In this paper, we propose a scheme to aggregate the FIB based on generalized next hops of strict partial order (SPO). We first use generalized SPO next hops to construct the Nexthop-Selectable FIB (NSFIB), where each prefix has multiple next hops. Our NSFIB aggregation avoids the aggregation performance degrading with the network density increasing, which is one main defect of the traditional single-nexthop FIB aggregation. We then design different levels of aggregation algorithms to aggregate the NSFIB. Besides, we control the path stretch by setting an upper limit to filter bad next hops. We also introduce routing protection by the pre-computed SPO next hops. According to our simulation, our aggregation algorithms shrink the FIB to 5–15%, compared with 20–60% of single-nexthop FIB aggregation algorithms; our method works very well in controlling the path stretch; and SPO next hops protect 50–95% (topology-related) failure-affected packets.

1. Introduction

The report from the Internet Architecture Board (IAB) reveals that the current Internet routing system is facing severe scalability problem with the ever-increasing users [1]. The core Internet routing tables have been expanding at a dramatic rate for recent decades. Due to multi-homing, traffic engineering, policy routing, *etc.*, unaggregatable address fragments from edge networks are continuously pouring into the core Internet. By July 2012, the routing table size has reached 433,653 (from 15,100 of 1994) [2]. Although the latest high-performance routers provide enough memory capacities, the Internet Service Providers (ISPs) cannot afford to upgrade their routers at the pace of routing table growth [3,4]. Therefore, new solutions are required to control the routing table size.

One approach of handling this Internet routing scalability problem is to separate the edge network addresses from the backbone. In this way, only core network addresses (much fewer than the current global network addresses) will appear in

---

[*] Corresponding author.
E-mail address: li.qing@sz.tsinghua.edu.cn (Q. Li).

http://dx.doi.org/10.1016/j.ins.2017.05.033
0020-0255/

the core Internet. Numbers of such solutions have been proposed, including Shim6 [5], ILNP [6], LISP [7], APT [8], MILSA [9], SIX/ONE [10], *etc*. Although these solutions can control the routing table size, they change the current Internet routing protocols, thus a long time is required to deploy these solutions. Besides, there are also some non-IP solutions, such as compact routing [11,12] and geographic routing [13,14]. Nevertheless, these solutions do not target directly on the current Internet and they cannot immediately alleviate the pressure on the ISPs.

The most urgent requirement for ISPs is shrinking the Forwarding Information Base (FIB), which is the most expensive and scarce part of the router memory. Therefore, as an immediate local solution, FIB aggregation is proposed. In FIB aggregation, numerically matching prefixes with the same next hop can be aggregated into one. For example, if the prefixes 10* and 101* have the same next hop, they can be aggregated into one prefix of 10*. FIB aggregation can be quickly implemented in any one single router, which means it can alleviate the pressure of ISPs immediately. However, current single-nexthop FIB aggregation solutions [15–20] cannot provide a satisfactory shrinking performance in the high-density network: the performance degrades as the network density increases [21]. Besides, in [19,22–24], they propose the compressed encoding approaches to control the size of the software routing table, meanwhile guaranteeing the fast lookup. The scenario of this paper is the widely-used hardware FIB.

In [21], Nexthop-Selectable FIB (NSFIB) aggregation is proposed, which achieves the better performance in FIB shrinking. In NSFIB aggregation, multiple selectable next hops are constructed for each IP prefix. Two numerically matching prefixes can be aggregated into one only if they share one common next hop. For example, if the prefixes 10* and 101* have the different next hops *a* and *b*, they cannot be aggregated. However, if we can construct another next hop *c* for the two prefixes and guarantee no forwarding error, there two prefixes can still be aggregated into 10* with the next hop of *c*. Although NSFIB aggregation provides another feasible path towards a scalable Internet, there are still some challenges to address, *e.g.*, **an effective NSFIB construction algorithm, the compatibility with other aggregation approaches** [15–18] and **a feasible method to control the path stretch.**

In this paper, to solve the above challenges, we (1) propose the generalized next hops of strict partial order to construct the NSFIB, (2) prove NSFIB's compatibility by applying different single-nexthop FIB aggregation approaches on the constructed NSFIB, and (3) provide the effective approaches to control the path stretch caused by NSFIB aggregation.

We first propose the generalized next hop of strict partial order (generalized SPO next hop) to construct the NSFIB. As an extended concept of next hop, the generalized next hop can be any possible forwarding actions, such as a physical interface, an IP tunnel or an MPLS tunnel. In this way, the possibility of two prefixes having one common next hops increases sharply. Thus, more prefixes can be aggregated to shrink the FIB. The SPO next hops for each prefix is computed according to the corresponding egress BGP router of that prefix. Based on the incremental shortest path first approach, we design the effective algorithms of *SPO-Nexthop* and *Ger-SPO-Nexthop* to compute all the SPO next hops for each destination. We control the complexity of *SPO-Nexthop* as the same as Shortest Path First (SPF). It means that the NSFIB can be re-constructed effectively when the topology changes.

Then, given the constructed NSFIB where each prefix have multiple SPO next hops, we prove the compatibility by applying different single-nexthop FIB aggregation approaches [15–18] on the NSFIB. For the four levels of FIB aggregation in [16], we design the corresponding NSFIB-based aggregation algorithms by the bottom-up dynamic method. Both optimality and stability are taken into account in these algorithms. Besides, we design the NSFIB-ORTC algorithm to compute the minimal aggregated FIB for a given NSFIB, as ORTC [15] does in the traditional single-nexthop FIB. The space complexity of NSFIB-ORTC is controlled by using the path-compressed trie. All the corresponding NSFIB aggregation algorithms are more efficient in FIB compression, as more aggregation can be triggered by multiple next hops for each prefix.

As NSFIB aggregation may use the non-optimal next hops, path stretch is inevitable. In the implementation of NSFIB aggregation, based on the definition of one-step increase amplitude, we propose the approach of setting an upper limit to filter the next hops that may cause the higher path stretch. In this way, we provide an efficient method for a router to make a trade-off between the aggregation performance and the path stretch. Different routers can set different upper limits according to the memory capacity. Besides, as multiple SPO next hops are computed during NSFIB construction, we use them as backup next hops to protect packets during network failures, which can decrease the packet dropping ratio during network failures.

We demonstrate the performance of our algorithms by comprehensive simulations on China Education and Research Network (CERNET) [25], Rocketfuel [26] topologies and BRITE-generated topologies [27].

- The results show that NSFIB-based aggregation algorithms perform better than the corresponding single-nexthop FIB aggregation algorithms. NSFIB Level1, NSFIB Level2 and NSFIB-ORTC can respectively shrink the FIB to about 11%, 9% and 6%.
- The path stretch ratios of NSFIB aggregation are less than 5% in all the six Rocketfuel topologies. While in denser BRITE topologies, the path stretch ratios are more than 30%. Our method of controlling path stretch works well in these denser topologies.
- The SPO next hops can avoid about 50–95% packet loss during the failure. In denser networks, the protection ratios are generally higher.

The above results are based on SPO next hops. If the generalized next hops are employed to construct the NSFIB, the FIB can be shrunk to 2.7–6.8% of the original size. Consequentially, the path stretch increases by 20–40%.