This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE INFOCOM 2010 proceedings

Empirical evaluation for the impact of core-edge separation on Internet routing scalability

Yangyang Wang, Jun Bi, Jianping Wu
Network Research Center
Department of Computer Science & Technology,
Tsinghua University, Beijing, China

Abstract— **BGP is the de facto standard protocol for the inter-domain routing. Due to multi-homing and traffic engineering, the BGP routing table size of default free zone (DFZ) is growing rapidly. Inter-domain routing is facing the scaling challenge. Such routing scaling issue has being concerned by research community and IETF, and many solutions have been proposed. Among them, the core-edge separation scheme gets more attentions than others due to its practical advantages. The typical core-edge separation solutions include LISP, eFIT, Ivip, etc. They prevent edge network specific prefixes from entering into transit core, and reduce the DFZ BGP routing table size. However, there has no evaluation on how much scalability can be improved from core-edge separation. In this paper, we use real BGP routing data of Routeviews Project to empirically analyze the possible impacts of core-edge separation. We find that it can reduce 43% routing table size and prevent more than half of announcement and withdraw messages.**

Keywords-BGP routing table; routing scalability; inter-domain routing;Internet routing

I. INTRODUCTION

BGP Protocol is the de facto standard protocol for the Internet inter-domain routing. In recent years, the fast growth of DFZ BGP routing table size indicates a potential scaling challenge in the migration of IPv6 deployment. The routing scaling issue has been got more attentions from research community and IETF.

A workshop report [1] of the Internet Architecture Board (IAB) summarized the Internet routing scaling challenge. Due to multi-homing and traffic engineering at the Internet edge, more specific prefixes are announced into transit core networks. In addition, edge networks prefer provider-independent (PI) address space to provider-allocated (PA) address space for avoiding renumbering on changing upstream ISPs (Internet Service Providers). PI address cannot be aggregated in upstream provider networks, and then increase the number of prefixes of global routing table.

A number of solutions for the routing scaling issue have been proposed. Some of them are clean-slate designs that need to change the addressing format, routing scheme and the whole routing architecture. However, most of the solutions focus on how to practically improve the Internet routing scalability in a short term future. Some of them, called core-edge separation scheme, get more attentions. The typical solutions of core-edge separation scheme include LISP[2], eFIT[3], Ivip[4], etc.

The core-edge separation scheme separate the customer networks at the Internet edge from the provider networks at the transit core. Therefore, the non-aggregateable specific prefixes announced by the edge networks cannot enter into the transit core networks, and the global routing table size of the transit core will be reduced. However, in this situation, only the address space of transit core is globally routable, while the edge network address space cannot be seen at the global routing system and is not globally routable. When a host of an edge network communicates with a remote host in another edge network across the transit core, the source and destination IP addresses of edge networks need to be mapped to the corresponding transit IP addresses. As shown in figure 1, R1~R4 routers are termed as ingress/egress tunneling routers (ITR/ETRs) in some proposals[2], which separate edge networks ASX and ASY from transit core. When a packet from host A to host B reaches R1, R1 will look up mapping system, and encapsulate the packet in a IP-in-IP tunnel using the R1 and R4 globally routable IP addresses as the tunneling source and destination addresses. After receiving the packet, R4 will de-encapsulate the tunnel and forward the original packet to host B. During this process, mapping and tunneling are two important steps. Therefore, core-edge separation is also called Map-Encap scheme.

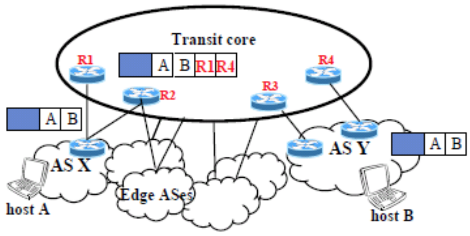

Fig. 1. Core-edge separation scheme

The gains of core-edge separation include reducing routing table size, and no changes to legacy infrastructures and hosts. But, it also has many costs. It relies on an extra mapping system and bilateral deployment of ITR/ETR devices. It also causes the map-encap delay during packets delivery. More discussions about the advantages and challenges of core-edge separation are described in [5].

Although there are many work [6][7] on the design and implementation of core-edge separation solutions, but less evaluation on core-edge separation scheme has been done, and it is not clear how much impacts of the core-edge separation on routing scalability. This paper takes a first step to quantifying

978-1-4244-6739-6/10/$26.00 ©2010 IEEE

the impacts of the core-edge separation on real Internet inter-domain routing.

II. METHODOLOGY FOR EVALUATION

A. Methodology

The basic metrics to be examined is routing table size reduction. We can count how many edge prefixes are prevented from entering into the global routing table. The network reachability information for separated prefixes is provided in the mapping tables of mapping system. In other words, the burden of global routing table growth is transformed as the mapping table growth. Therefore, the mapping table size is also an important metrics. The core-edge separation also prevents the routing dynamics of the edge prefixes from propagating into the transit core, and this benefit can be estimated via statistics on update messages.

The placement of ITR/ETRs impacts the effectiveness of core-edge separation. Placing ITR/ETRs close to the tier-1 networks will improve the reduction of the area of the transit core and its routing table size. There is no specific borderline for the separation in the Internet. In our primary evaluation, we choose the place between stub ASes (Autonomous Systems) and their direct upstream providers as the separation place where ITR/ETRs are deployed. Thus, all of stub ASes form into the edge, and the non-stub ASes forms the transit core.

B. Preliminary results

Routeviews project [8] provides BGP routing data. We collect the IPv4 routing tables for several different of days of every month from Jan 2004 to the Nov 2009. To facilitate routing dynamics analysis, we collect three months of IPv4 routing updates from the "route-views2.oregon-ix.net" collector of Routeviews from September to November of 2009. The primary results are presented in the following.

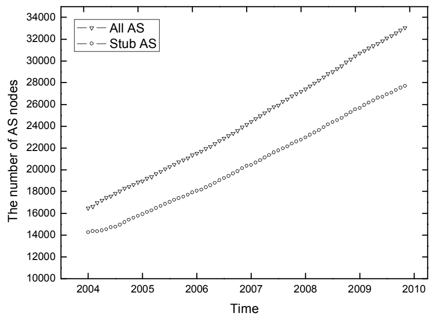

Fig. 2. The growth of ASes

We first examine the growth of the number of ASes. As shown in figure 2, the total number of ASes is 33050 in Nov. 2009,, among which there are 27701 stub ASes and 5349 non-stub ASes. The sub ASes is more than 83%. The total number of ASes and non-stub ASes are doubled over the past six years.

The figure 3 shows the number of blocked prefixes of stub ASes. We can see that the stub ASes (83% ASes) announce about 43% prefixes. It shows that blocking all the prefixes of the large amount of stub ASes does not imply a significant reduction in routing table size. The mapping table size is nearly the same as the global routing table size. After the year 2007, the mapping table size gets larger than the routing table size. This may be caused by the increasing multi-homing at the edge.

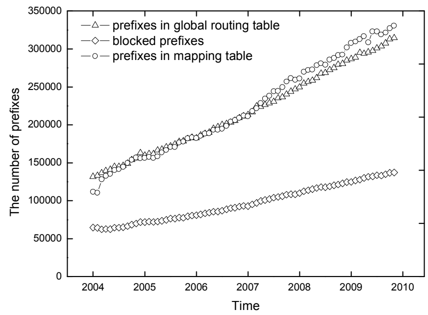

Fig. 3. Reduction of routing table size

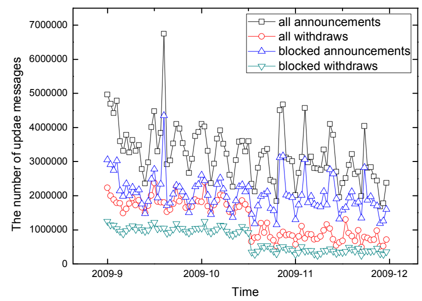

Fig. 4. Reduction of update messages

As shown in figure 4, more than half (63% announcements and 52% withdraws) of BGP update messages can be blocked entering into Internet transit core.

III. CONCLUSION

We use real Internet routing data to quantify the impact of core-edge separation on improving the Internet routing scalability. The empirical result shows that the reduction is not significant. Separating all of stub ASes (83%) produces 43% reduction in routing table size and more than half of reduction in routing dynamics. Further detailed study is ongoing.

REFERENCES

[1] "Report from the IAB Workshop on Routing and Addressing",RFC4984

[2] D. Farinacci, V. Fuller, D. Meyer, and D. Levis. Locator/ID Separation Protocol (LISP). draft-farinacci-lisp-12, Sep. 2009

[3] The eFIT project, http://netlab.cs.memphis.edu/efit/index.php

[4] Ivip, http://www.firstpr.com.au/ip/ivip/

[5] D. Jen, M. Meisel, H. Yan, D. Massey, L. Wang, B. Zhang, L. Zhang, Towards A New Internet Routing Architecture: Arguments for Separating Edges from Transit Core, in: Proc. HotNets-VII, 2008.

[6] Luigi Iannone and Olivier Bonaventure, "On the cost of caching locator/ID mappings", ACM CoNEXT 2007

[7] Hong Zhang, Maoke Chen, Yuncheng Zhu: Evaluating the Performance on ID/Loc Mapping. GLOBECOM 2008: 2276-2280

[8] Routeviews project: http://www.routeviews.org/