Leveraging Network Performances with IPv6 Multihoming and Multiple Provider-Dependent Aggregatable Prefixes

Cédric de Launois *, Bruno Quoitin **, and Olivier Bonaventure

Université catholique de Louvain
Department of Computing Science and Engineering
http://www.info.ucl.ac.be
{delaunois,quoitin,bonaventure}@info.ucl.ac.be

Abstract. Multihoming, the practice of connecting to multiple providers, is becoming highly popular. Due to the growth of the BGP routing tables in the Internet, the way to multihome in IPv6 is required to preserve the scalability of the interdomain routing system. A proposed method is to assign multiple provider-dependent aggregatable (PA) IPv6 prefixes to each site, instead of a single provider-independent (PI) prefix. This paper shows that the use of multiple PA prefixes per sites not only allows route aggregation but also can be used to reduce end-to-end delay by leveraging the Internet path diversity. We also quantify the gain in path diversity, and show that a dual-homed stub that uses multiple PA prefixes has already a better Internet path diversity than any multihomed stub that uses a single PI prefix, whatever its number of providers. We claim that the benefits provided by the use of IPv6 multihoming with multiple PA prefixes is an opportunity to develop the support for quality of service and traffic engineering.

Key words: BGP, IPv6 Multihoming, Path Diversity.

1 Introduction

Today, the Internet connects more than 17000 *Autonomous Systems* (AS) [1], operated by many different technical administrations. The large majority of ASes are *stub* ASes, i.e. autonomous systems that do not allow external domains to use their infrastructure. Only about 20% of autonomous systems provide transit services to other ASes [2]. They are called *transit* ASes. The Border Gateway Protocol (BGP) [3] is used to distribute routing announcements among routers that interconnect ASes.

* Supported by a grant from FRIA (Fonds pour la Formation à la recherche dans l'Industrie et dans l'Agriculture, rue d'Egmont 5 - 1000 Bruxelles, Belgium).

** Supported by the Walloon Government within the WIST TOTEM project. http://totem.info.ucl.ac.be

The size of the BGP routing tables in the Internet has been growing dramatically during the last years. The current size of those tables creates operational issues for some Internet Service Providers and several experts [4] are concerned about the increasing risk of instability of BGP. Part of the growth of the BGP routing tables [5] is due to the fact that, for economical and technical reasons, many ISPs and corporate networks wish to be connected via at least two providers to the Internet. For more and more companies, Internet connectivity takes on strategic importance. Nowadays, at least 60% of those domains are multihomed to two or more providers [1, 2]. Therefore, it can be expected that IPv6 sites will continue to be multihomed, primarily to enhance their reliability in the event of a failure in a provider network, but also to increase their network performances such as network latency. In order to preserve the scalability of the interdomain routing system, every IPv6 multihoming solution is required to allow route aggregation at the level of their providers [4]. Among the several IPv6 multihoming methods proposed at the IETF [6], a popular solution is to assign multiple provider-dependent aggregatable (PA) IPv6 prefixes to each site, instead of a single provider-independent (PI) prefix. Both IPv4 and IPv6 multihoming methods are described in section 2.

We show in this paper that the use of multiple PA prefixes introduces other benefits than simply allowing route aggregation. We explain in section 4 how stubs that use multiple PA prefixes can exploit paths that are otherwise unavailable. In other words, we explain how the use of PA prefixes increases the number of concurrent paths available. Next, we show that lower delays can often be found among the new paths. Our simulations suggest that a delay improvement is observed for approximately 60% of the stub-stub pairs, and that the delay improvement could be higher in the actual Internet.

In section 5, we quantify the gain in terms of Internet path diversity. We show that a dual-homed stub that uses multiple PA prefixes has already a better Internet path diversity than any multihomed stub that uses a single PI prefix, whatever its number of providers.

2 IPv4 and IPv6 Multihoming

This section provides some background on traditional IPv4 multihoming and IPv6 multihoming.

In the current IPv4 Internet, the traditional way to multihome is to announce, using BGP, a single prefix to each provider, see fig. 1 and 2. In fig. 1, AS 123 uses provider-aggregatable addresses. It announces prefix 10.0.123.0/24 to its providers AS 10 and AS 20. AS 10 aggregates this prefix with its 10.0.0.0/8 prefix and announces the aggregate to the Internet. In fig. 2, AS 123 announces a provider-independent prefix to its providers. This prefix is then propagated by BGP routers over the Internet. Throughout this paper, we will refer to this technique as *traditional IPv4 multihoming*, or simply *IPv4 multihoming*.

The way stubs multihome in IPv6 is expected to be quite different from the way it is done currently in IPv4. Most IPv6 multihoming mechanisms [6–8] pro-

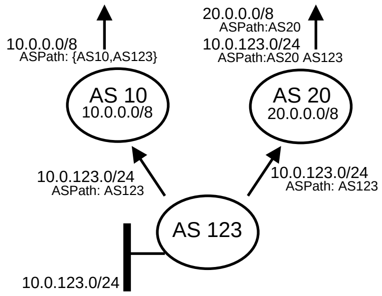

Fig. 1. IPv4 Multihoming using a provider-aggregatable prefix

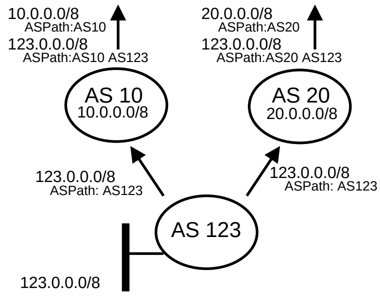

Fig. 2. IPv4 Multihoming using a provider-independent prefix

posed at the IETF rely on the utilization of several IPv6 provider-aggregatable prefixes per site. Figure 3 illustrates a standard IPv6 multihomed site.

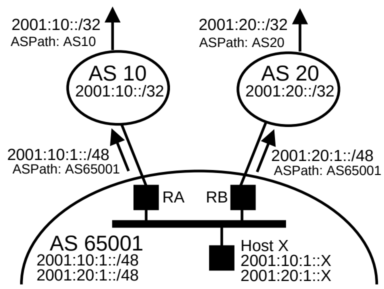

Fig. 3. IPv6 Multihoming

Suppose two Internet Service Providers, AS 10 and AS 20, provide connectivity to the multihomed site AS 123. Each provider assigns to AS 123 a site prefix, respectively 2001:10:1::/48 and 2001:20:1::/48. The two prefixes are advertised by the site exit routers RA and RB to every host inside AS 123. Finally, these prefixes are used to derive one IPv6 address per provider for each host interface. In this architecture, AS 123 advertises prefix 2001:10:1::/48 only to AS 10, and AS 10 only announces its own IPv6 aggregate 2001:10::/32 to the global Internet. This new solution is expected to be used only by stub ASes. Transit ASes are not concerned by these solutions since they will receive provider-independent IPv6 prefixes. Consequently, in this study, we focus only on stubs.

The use of multiple PA prefixes is natural in an IPv6 multihoming environment. However, it is not impossible to use the same multihoming technique in IPv4, i.e. to delegate two IPv4 prefixes to a site. Unfortunately, due to the current lack of IPv4 addresses, the need to delegate several IPv4 prefixes to a multihomed site makes this solution less attractive. Thus, throughout this document, we will call the new multihoming technique presented here for IPv6 simply as *IPv6 multihoming*; although the same concept could also be applied to IPv4 multihomed sites, and although other IPv6 multihoming techniques exist.

3 Simulation Setup

We describe here the tools and topologies used to conduct our simulations.

IPv6 multihoming with multiple provider-dependent aggregatable prefixes is currently not deployed. As a consequence, we set up simulations made on various Internet-like topologies instead of conducting measurement experiments on the actual IPv4 Internet. We use several topologies in order to delimit the impact of the topology on the results, and to explore possible evolution scenarios for the Internet.

In this study, we focus on the paths announced by BGP between each pair of stub ASes in a given topology. These paths depend on the topology but also on the commercial relationships between ASes together with their BGP routing policies. The commercial agreements between two ASes are usually classified as customer-provider or peer-to-peer relationships [9, 10]. These relationships are either inferred [9, 10], or directly provided by the topology description. We then compute, for each AS, the BGP configuration that corresponds to its commercial relationships with the other ASes. The BGP export policies basically define that an AS announces all the routes to its customers, but announces to its peers and providers only the internal routes and the routes of its customers. Moreover, the configuration defines that an AS prefers routes received from a customer, then routes received from a peer, and finally routes received from a provider [9, 10]. These filters ensure that an AS path will never contain a customer-to-provider or peer-to-peer edge after traversing a provider-to-customer or peer-to-peer edge. This property is known as the the *valley-free* property, and is defined in [9]. We announce one prefix per AS. The paths for a given topology are obtained by simulating the BGP route distribution over the whole topology. For this purpose, we use a dedicated BGP simulator, named C-BGP [11]. C-BGP supports import and export filters, and uses the full BGP decision process. The tie-breaking rule used by C-BGP to choose between two equivalent routes is to prefer the route learned from the router with the lowest router address. As soon as all the routes have been distributed and BGP has converged, we perform traceroute measurements on the simulated topology and deduce the paths.

4 Improving Delays with Multiple Prefixes per Site

We show in section 4.1 that stubs that use multiple PA prefixes can exploit paths that are otherwise unavailable. Among the new paths, some of them offer lower delays. We try to evaluate here how often this improvement in network latency occurs. We detail in section 4.2 the topology used to perform this evaluation. We then present the results of this simulation in section 4.3.

4.1 Impact of PI and PA Prefixes on Available AS Paths

Fig. 4 shows an AS-level interdomain topology with shared-cost peerings and customer-provider relationships. An arrow labelled with "$" from AS x to AS y means that x is a customer of y. A link labelled with "=" means that the ASes have a shared-cost peering relationship [9]. In this figure, both S and D are dual-homed ASes.

An IPv4 multihomed stub D traditionaly announces a single provider-independent prefix to each of its providers. This PI prefix is then propagated by BGP routers all over the Internet. In particular, if AS S is single-homed, it will receive a single route from its provider to reach the dual-homed AS D. This route is the best route known by the provider to join AS D. If AS S is also dual-homed, as illustrated in figure 4, S will receive two routes $ECAD$ and $FCAD$ towards D one from each of its provider, as shown in fig. 5.

When stubs use IPv6 multihoming with multiple PA prefixes, additional routes exist.

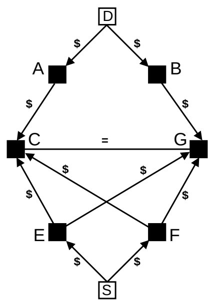

Fig. 4. Topology

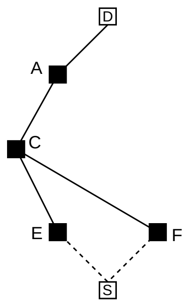

Fig. 5. IPv4 path tree

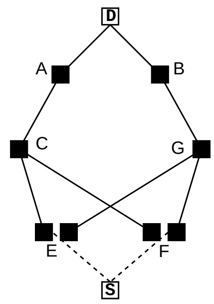

Fig. 6. IPv6 path tree

Suppose that both AS D and AS S use IPv6 multihoming with multiple PA prefixes. Each host in S has two IPv6 addresses. One is derived from the prefix allocated by the first provider E, while the other one is derived from the prefix allocated by the second provider F. Similarly, each host in D has two IPv6 addresses. When selecting the source address of a packet to be sent, the host in S could in theory pick any of its two addresses. However, for security reasons, most providers refuse to convey packets with source addresses outside their address range. For example, E refuses to forward a packet with a source address belonging to F. Therefore, depending on the source address selected, the upstream provider used to convey the packet is either E or F. Moreover, the path used to reach D also depends on which destination address of D is used. For instance, if we use the destination address derived from the prefix allocated by A, then the packet will flow through A. Thus, a host in S may have up to four paths to reach a host in D, depending on both the source and the destination addresses used.

Fig. 5 shows that two BGP routes towards D ($SECAD$ and $SFCAD$) are learned by S in the IPv4 scenario. In this example, we see that these routes join early at AS C. In an IPv6 scenario, since both S and D have two prefixes, S can reach D via A or B depending on which destination prefix is used, and via E or F depending on which source prefix is used. Thus, S has a total of four paths to reach D : $SECAD$, $SEGBD$, $SFCAD$ and $SFGBD$, see fig. 6.

4.2 A Two-Level Topology with Delays

In order to simulate delays along paths, we cannot rely on topologies provided by Brite [12], Inet [13], or GT-ITM [14] since they either do not model business relationships or do not provide delays along links. A topology that contains both delays and commercial relationships is available at [15]. In this topology, the interdomain links and the business relationships are given by a topology inferred from multiple routing tables [10]. For each peering relationship found between two domains in this topology, interdomain links are added.

The different points of presence of each domain are geographically determined by relying on a database that maps blocks of IP addresses and locations worldwide. The intradomain topology is generated by first grouping routers that are close to each other in clusters, and next by interconnecting these clusters with backbone links.

The delays along the links is the propagation delay computed from the distance between the routers. The IGP weights used are the delays for links shorter than 1000 km, twice the delay for links longer than 1000 km but shorter than 5000 km and 5 times the delay for links longer than 5000 km. This is used to penalize the long intradomain links and favor hot-potato routing.

The resulting topology is described in more details at [15]. It contains about 40,000 routers, 100,000 links and requires about 400,000 BGP sessions. Next, C-BGP [11] is used to simulate the BGP protocol in order to obtain the router-router paths and their delays between multihomed stubs. For computation time reason, we conduct the simulation for 2086 multihomed stubs randomly chosen among the 8026 multihomed stubs.

In this topology, 55% of the delays along the BGP route are comprised between 10 and 50ms. About 20% of the delays are below 10ms and 25% sit between 50 and 100ms. These delays are considered as minimal bounds for the real delays, since only the propagation delay is taken into account. Factors that increase delays like limited bandwiths or congestion delays are not considered here. Although the simulated delays are inferior bounds to delays observed in the global Internet, their order of magnitude is preserved.

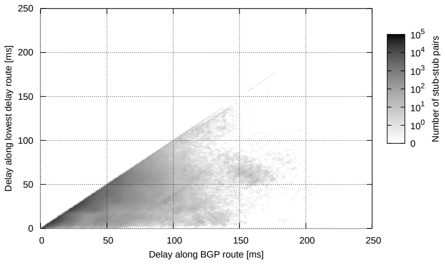

Fig. 7. Delay along the BGP route versus delay along the lowest delay route

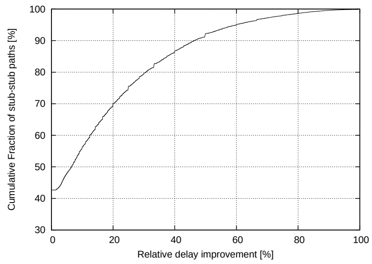

Fig. 8. Distribution of the relative delay improvement

4.3 Simulation Results

Figure 7 plots the best delay obtained when stubs use traditional IPv4 multihoming (x-axis), against the best delay obtained when stubs use IPv6 multihoming with multiple PA prefixes (y-axis). The gray-scale indicates the number of stub-stub paths, on a logarithmic scale. The diagonal line that appears represents stub-stub pairs for which both multihoming techniques yield to the same delay.

As explained in section 4.1, the use of multiple PA prefixes provides additional paths, beside traditional paths that are still available. As a consequence, delays can only improve, and no dot can appear above the diagonal line. A dot under this diagonal line indicates that the use of multiple PA prefixes introduces a new path with a delay lower than the delay along the best BGP path obtained when a single PI prefix is used. We can see that a lot of dots are located under this line. Sometimes, the improvement can even reach 150ms in this topology.

Figure 8 shows the cumulative distribution of the relative delay improvement. It shows that no improvement is observed for approximately 40% of stub-stub pairs. However, the relative improvement is more than 20% for 30% of stub-stub paths. Delays are cut by half for about 8% of stub-stub pairs.

As said in section 4.2, the delays observed in this topology are expected to be minimal bounds to those seen in the real Internet. Thus, we can reasonably assume that the absolute delay improvements presented in figure 7 will not be lower in the actual Internet.

These simulation results show that improving delays is a benefit of IPv6 multihoming with multiple PA prefixes, without increasing the BGP routing tables.

5 Leveraging Internet Path Diversity with Multiple Prefixes

We have seen in section 4.1 that stubs that use multiple PA prefixes can exploit paths that are otherwise unavailable. In other words, the use of multiple PA prefixes increases the number of paths available, i.e. the Internet path diversity. We have shown that a benefit is that better delays can often be found among the new paths. An increase in path diversity also yields other benefits, like better possibilities of load balancing. In this section, we use simulations to measure the Internet path diversity that exists when a multihomed stub uses either multiple PA prefixes or a single PI prefix.

We first detail the inferred and generated topologies that are used in the simulations. Next, we present and discuss the results of simulations made on an inferred AS-level Internet topology. Finally, we evaluate the impact of the topology on the path diversity.

5.1 Internet Topology

In section 4, we used a large router-level Internet topology that models delays. In this section, we use AS-level topologies instead, for two reasons. A first reason is the computation time. The topology used in section 4 is unnecessarily complex for an AS-level simulation since it models routers and delays. A second reason is that we want to consider other topologies in order to estimate the variability of our results with respect to the topology.

We first use an AS-level Internet topology inferred from several BGP routing tables using the method developed by Subramanian et al. [10].

We next generate several AS-level Internet-like topologies, using the GHITLE topology generator [16]. A topology is generated level by level, from the dense core to the customer level. Four levels are usually created : a fully-meshed dense core, a level of large transit ASes, a level of local transit ASes and a level of stubs. Additional levels can be created if needed. We developed GHITLE for this analysis because no existing generator could produce a topology that provides details about customer-provider and peer-to-peer relationships, and where the number of Internet hierarchy levels and nodes in each level could be specified. In particular, Inet [13] does not provide commercial relationships between ASes. Brite [12] and GT-ITM [14] do not produce a hierarchical topology with more levels than just transit and stub levels.

5.2 Path Diversity Metric

In order to measure the path diversity for a given destination AS, we first build the tree of paths from all source AS towards the destination AS. As explained in section 4.1, this path tree depends on the multihoming technique used. Next we use a new, fine-grain, path diversity metric to evaluate the diversity of this tree. This metric takes into account the lengths of the paths and how much they overlap. We define this new path diversity metric, from a source AS S to a destination AS D, as follows.

Let $P_1, P_2, ..., P_n$ be the n providers of S. We first build the tree of all paths starting from providers P_i of S to destination D, for $i = 1, ..., n$. This tree represents all the BGP paths for D that are advertised by the providers P_i to S. Our path diversity metric is computed recursively link by link, from the leaves to the root. It returns a number between 0 and 1. We first assign an initial diversity of 0.5 to each link in the tree. This number is chosen in order to best distribute the values of the path diversity metric in the range $[0, 1]$. At each computation step, we consider two cases, to which all other cases can be reduced. Either two links are in sequence, or the links join in parallel at the same node.

Alg. 1. Computing Diversity Metric

```
Diversity(root)
{
    d = 0 ;
    if ( Children(root) == ∅ )
        return 1 ;

    foreach child ∈ Children(root) {
        d_child = 0.5 · Diversity(child) ;
        d = d + d_child − d · d_child ;
    }

    return d ;
}
```

In the first case, two links with diversity d_1 and d_2 in sequence can be merged into a single link with a combined diversity $d_{1,2} = d_1 \cdot d_2$. The combined diversity $d_{1,2}$ is a number in $[0, 1]$ lower than both d_1 and d_2, so that the metric favors short paths over longer one. This computation step also implicitely gives higher importance to the path diversity that exists near the root. This behaviour ensures that the metric prefers trees where paths join lately near the destination node over trees where paths merge near the source node.

In the second case, when a link with a diversity d_1 and another link with a diversity d_2 join in parallel, we merge the two links into a single link with a combined diversity $d_{1,2}$, computed as $d_{1,2} = d_1 + d_2 - d_1 \cdot d_2$. This gives a number greater than both d_1 and d_2, which corresponds adequately to an improvement in terms of diversity.

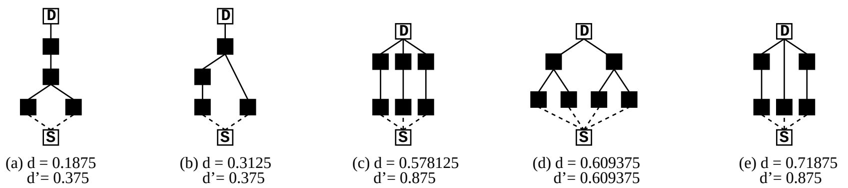

Fig. 9. Path diversity metric examples

A recursive algorithm to compute this metric is presented in Alg. 1. A detailed description of this metric together with a comparison with other path diversity metrics is available at [17]. Examples of values for this metric are shown in figure 9.

5.3 Simulation Results

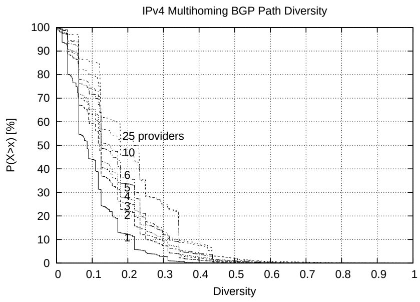

Fig. 10. AS-level path diversity for the inferred Internet topology, using traditional IPv4 multihoming

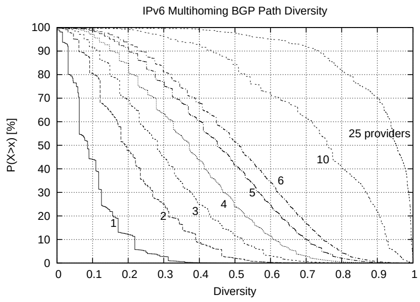

Fig. 11. AS-level path diversity for the inferred Internet topology, using IPv6 multihoming

Figure 10 presents the path diversity available to stub ASes that use traditional IPv4 multihoming in the inferred AS-level Internet topology. Figure 11 shows the path diversity when all stubs use IPv6 multihoming with multiple PA prefixes, in the same inferred topology.

The figures show $p(x)$: the percentage of couples *(source AS, destination AS)* having a path diversity greater than x. The results are classified according to the number of providers of the destination stub. The number of providers is indicated beside each curve. In fig. 10, we see for example that only 12% of single-homed stubs using traditional IPv4 multihoming have a diversity better than 0.2. This percentage raises to more than 20% for dual-homed stubs. In fig. 11, about 50% dual-homed IPv6 stubs have a quality better than 0.2, compared to less than 15% single-homed IPv6 stubs.

We can observe that the diversity remains the same when considering only single-homed destinations. Indeed, only one prefix is announced by a single-homed stub, using either IPv4 or IPv6 multihoming technique. The use of IPv6 multihoming does not introduce any benefit in this case.

When comparing fig. 10 and 11, it appears that the AS-level path diversity is much better when stubs use IPv6 multihoming with multiple PA prefixes than when stubs use traditional IPv4 multihoming. For example, when considering dual-homed IPv6 stubs, we see in fig. 11 that the path diversity observed is already as good as the path diversity of a 25-homed stub that uses traditional IPv4 multihoming. The path diversity obtained by a 3-homed stub AS that uses IPv6 multihoming completely surpasses the diversity of even a 25-homed stub that uses traditional IPv4 multihoming.

5.4 Impact of Topology on Path Diversity

The way Internet will evolve in the future remains essentially unknown. In order to delimit the range of variation for our results, we perform simulations with three distinct topologies.

The first is a topology that tries to resemble the current Internet.

The second is a small-diameter Internet topology, consisting of stubs directly connected to a fully meshed dense core. This simulates a scenario where ASes in the core and large transit ASes concentrates for commercial reasons. At the extreme, the Internet could consist in a small core of large transit providers, together with a large number of stubs directly connected to the transit core. This could lead to an Internet topology with a small diameter.

The third is a large-diameter topology, generated using eight levels of ASes. This topology simulates a scenario where the Internet continues to grow, with more and more core, continental, national and metropolitan transit providers. In this case, the Internet might evolve towards a network with a large diameter.

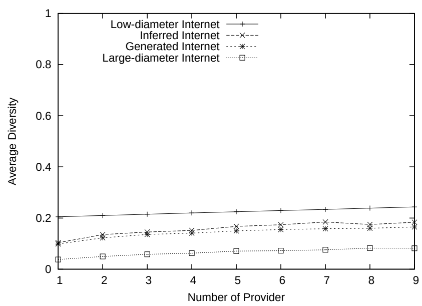

Fig. 12. Average path diversity using traditional IPv4 Multihoming

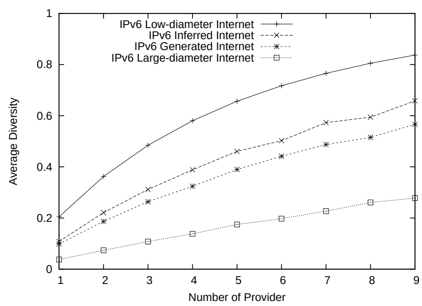

Fig. 13. Average path diversity using IPv6 multihoming

Figures 12 and 13 show the average path diversity in function of the number of providers for all topologies. For each stub AS x, we compute the mean of the path diversities for all paths towards x. We then group the stub ASes according to their number of providers, and compute their mean. In fig. 12 and 13, we can first observe that the average diversity of the inferred Internet is included between the average diversities of the two extreme cases. In fig. 12, we see that the average path diversity using traditional IPv4 multihoming does not rise much in function of the number of providers, whatever the topology. Fig. 10 and 12 suggest that it is nearly impossible that a stub achieves a good path diversity using traditional IPv4 multihoming, whatever the number of providers. In contrast, the path diversity that is obtained using IPv6 multihoming with multiple PA prefixes is much better, see fig. 13. When comparing fig. 12 and 13, we can observe that a dual-homed stub using IPv6 multihoming already gets a better diversity than any multihomed stub that uses traditional IPv4 multihoming, whatever its number of provider and for all considered topologies. In a small-diameter Internet, this diversity rises fast in function of the number of providers, but also shows a marginal gain that diminishes quickly. In a large-diameter Internet, the diversity rises more slowly.

So far, we have analyzed the AS-level path diversity considering one router per AS. However, a factor that can impact the path from a source to a destination is the intradomain routing policy used inside transit ASes. We have also evaluated the path diversity that exists when ISP routing policies in the Internet conform to hot-potato routing. In hot-potato routing, an ISP hands off traffic to a downstream ISP as quickly as possible. Results show that hot-potato routing has no significant impact on the AS-level path diversity. Due to a lack of space, this simulation is not detailed in this document.

5.5 Impact of BGP on Path Diversity

We now discuss in this section how this path diversity is affected by the BGP protocol.

Multihoming is assumed to increase the number of alternative paths. However, the AS-level path diversity offered by multihoming depends on how much the interdomain routes, as distributed by BGP, overlap.

The results presented in the previous section suggest that BGP heavily reduces the path diversity, at the level of autonomous systems. Two factors explain why the diversity is so much reduced.

The primary factor is that BGP only announces one route, its best one. This behavior heavily reduces the number of paths. Unfortunately, BGP is designed as a single path routing protocol. It is thus difficult to do better with BGP.

A second factor exists that further reduces the path diversity. The tie-breaking rule used by BGP to decide between two equivalent routes often prefers the same next-hops. Let us consider a BGP router that receives two routes from its provider towards a destination D. According to the BGP decision process, the shortest AS path is selected. However the diameter of the current Internet is small, more or less 4 hops [1]. As a consequence, paths are often of the same length, and do not suffice to select the best path. It has been shown that between 40% and 50% of routes in core and large transit ASes are selected using tie-breaking rules of the BGP decision process [18]. In our model with one router per AS, the tie-breaking rule used in this case is to prefer routes learned from the router with the lowest router address. Unfortunately this rule yields to always prefer the same next-hop, a practice that worsen path diversity.

The first factor suppresses paths, while the second factor increases the probability that paths overlap. An IPv6 multiaddress multihoming solution circumvents the first factor by using multiple prefixes. The IPv6 multihoming solution has no impact on the second factor, since it does not modify BGP and its decision process in particular.

6 Related Work

A work about IPv4 multihoming path diversity appears in [19], where the authors define two path diversity metrics to quantify the reliability benefits of multihoming for high-volume Internet servers and receivers. They notice however that their metrics have an undesirable bias in favor of long paths. Their study draws empirical observations from measurement data sets collected at servers and monitoring nodes, whereas our work is based on inferred and generated global-scale AS-level topologies.

A comparison of Overlay Routing and Multihoming Route Control appears in [20]. In this study, the authors demonstrate that an intelligent control of BGP routes, coupled with ISP multihoming, can provide competitive end-to-end performance and reliability compared to overlay routing. Our results agree with this finding. In addition, our work explicits the impact of the path diversity on performances, and shows that IPv6 multihoming with multiple PA prefixes is able to provide these benefits.

It is well known that the use of provider-dependent aggregatable prefixes preserves the scalability of the interdomain routing system [21]. To our knowledge, this is the first study that shows that the use of multiple PA prefixes also increases network performances by leveraging the Internet path diversity, compared to the use of traditional multihoming with a single PI prefix.

7 Conclusion

A proposed way to improve network performances of the interdomain is to enhance BGP. We revealed that another way is to use multiple provider-dependent aggregatable (PA) prefixes per sites, in an IPv6 Internet in particular. In this paper, we have shown that stubs that use multiple PA prefixes can exploit paths that are otherwise unavailable. Thus, the use of multiple PA prefixes increases the number of paths available, i.e. the Internet path diversity. Next, our simulations have shown that 60% of the paths can benefit from a lower delay. Finally, we used simulations on various topologies to quantify the path diversity benefits offered by the use of multiple PA prefixes. We have shown that a dual-homed stub that uses multiple PA prefixes has already a better Internet path diversity than any multihomed stub that uses a single provider-independent (PI) prefix, whatever its number of providers.

Our observations show that, from a performance point of view, IPv6 multihomed stubs get benefits from the use of multiple PA prefixes and should use them instead of a single PI prefix as in IPv4 today. This study thus strongly encourages the IETF to pursue the development of IPv6 multihoming solutions relying on the use of multiple PA prefixes. The use of such prefixes reduces the size of the BGP routing tables, but also provides lower delays, more diverse Internet paths, which in turn yields to better possiblities to balance the traffic load and to support quality of service.

Acknowledgments

We thank Steve Uhlig and Marc Lobelle for their useful comments and support. We also thank the authors of [10] for providing the inferred Internet topology.

References

1. Huston, G.: BGP Routing Table Analysis Reports. `http://bgp.potaroo.net` (2004)
2. Agarwal, S., Chuah, C.N., Katz, R.H.: OPCA: Robust interdomain policy routing and traffic control. In: Proceedings OPENARCH. (2003)
3. Stewart, J.W.: BGP4: Inter-Domain Routing in the Internet. Addison-Wesley (1999)
4. Atkinson, R., Floyd, S.: IAB concerns & recommendations regarding internet research & evolution. Internet Draft, IAB (2004) <draft-iab-research-funding-03.txt>, work in progress.
5. Bu, T., Gao, L., Towsley, D.: On routing table growth. In: Prceedings IEEE Global Internet Symposium. (2002)
6. Huitema, C., Draves, R., Bagnulo, M.: Host-centric IPv6 multihoming. Internet Draft (2004) <draft-huitema-multi6-hosts-03.txt>, work in progress.
7. Terakoa, F., Ishiyama, M., Kunishi: LIN6: A solution to mobility and multi-homing in IPv6. Internet Draft (2003) <draft-teraoka-multi6-lin6-00.txt>, work in progress.
8. de Launois, C., Bonaventure, O., Lobelle, M.: The NAROS approach for IPv6 multihoming with traffic engineering. In: Proceedings QoFIS 2003, LNCS 2811, pp. 112-121. (2003)
9. Gao, L.: On inferring autonomous system relationships in the internet. IEEE/ACM Transactions on Networking **vol. 9, no. 6** (2001)
10. Subramanian, L., Agarwal, S., Rexford, J., Katz, R.H.: Characterizing the internet hierarchy from multiple vantage points. In: Proceedings IEEE Infocom. (2002)
11. Quoitin, B.: C-BGP - An efficient BGP simulator. `http://cbgp.info.ucl.ac.be/` (March 2004)
12. Medina, A., Lakhina, A., Matta, I., Byers, J.: BRITE: An Approach to Universal Topology Generation. In: Proceedings MASCOTS '01. (2001)
13. Jin, C., Chen, Q., Jamin, S.: Inet: Internet topology generator. Technical Report CSE-TR-433-00 (2000)
14. Calvert, K., Doar, M., Zegura, E.: Modeling internet topology. IEEE Communications Magazine (1997)
15. Quoitin, B.: Towards a POP-level Internet topology. `http://cbgp.info.ucl.ac.be/itopo/` (2004)
16. de Launois, C.: GHITLE. `http://openresources.info.ucl.ac.be/ghitle/` (April 2004)
17. de Launois, C.: Leveraging internet path diversity and network performances with IPv6 multihoming. `http://www.info.ucl.ac.be/people/delaunoi/diversity/` (2004)
18. Bonaventure, O., et al.: Internet traffic engineering. In: COST 263 Final Report, LNCS 2856, pp.118-179. (2003)
19. Akella, A., et al.: A measurement-based analysis of multihoming. In: Proceedings ACM SIGCOMM'03. (2003)

20. Akella, A., et al.: A comparison of overlay routing and multihoming route control. In: Proceedings ACM SIGCOMM'04. (2004)
21. Fuller, V., Li, T., Yu, J., Varadhan, K.: Classless inter-domain routing (CIDR): an address assignment and aggregation strategy. RFC 1519, IETF (1993)