# On Routing Table Growth

Tian Bu [1], Lixin Gao[2], and Don Towsley[1]
[1] Department of Computer Science
University of Massachusetts Amherst
{tbu,towsley}@cs.umass.edu
[2] Department of Electrical and Computer Engineering
University of Massachusetts Amherst
lgao@ecs.umass.edu

*Abstract*— **The Internet has experienced explosive growth since its commercialization. The sizes of the routing tables have increased by an order of magnitude over the past six years. This dramatic growth of the routing table can decrease the packet forwarding speed and demand more router memory space. In this paper, we explore the extent that various factors contribute to the routing table growth and predict the future rate of growth of the routing table. We first perform measurement study to determine the extent that factors such as multi-homing, failure to aggregate, load balancing, and address fragmentation contribute to routing table size, and find that only $20-30\%$ of prefixes are due to multi-homing, $15-20\%$ of prefixes are due to failure to aggregate, $20-25\%$ of prefixes are due to load balancing, and more than $75\%$ of prefixes are due to address fragmentation. This leads us to group all prefixes that are not aggregated due to either failure to aggregate or address fragmentation. We find that the number of prefix clusters is no more than $20\%$ of the number of prefixes. We explore the extent that load balancing contributes to the number of prefix clusters. Furthermore, we predict the growth pattern of prefixes and prefix clusters by observing power-laws on prefixes and prefix clusters. The number of prefixes grows much faster than the number of prefix clusters does. To the best of our knowledge, this is the first study on the explosive growth of routing tables by systematically comparing factors that contribute to the growth and by observing routing table growth patterns.**

## I. Introduction

The Internet has experienced explosive growth since its commercialization. The Internet is divided into thousands of autonomous systems (ASes), each of which consists of networks of hosts or routers administrated by a single organization. Hosts and routers are identified with 32-bit IP addresses, which brings to a total of $2^{32}$ (more than 4 billion) possible IP addresses. To ensure the scalability of the Internet routing infrastructure, IP addresses are aggregated into contiguous blocks, called *prefixes*. Routers exchange reachability information for each prefix using the Border Gateway Protocol (BGP). As a consequence, each BGP routing table entry contains reachability information for a single prefix. The size of a BGP routing table is the number of prefixes contained in the routing table. The size of routing tables has risen from 10,000 to 100,000 over the past six years [1], [2]. This dramatic growth of the routing table can decrease the packet forwarding speed and demand more router memory space. Some experts have predicted that if the router memory is to keep pace with the growth of the routing tables, each router will require gigabits of memory within the next two years. In this paper, we explore the extent that various factors contribute to the routing table growth and predict the future rate of the growth of the routing table.

The authors were supported in part by NSF grant ANI-9977555, ANI-0085848, and NSF CAREER Award grant ANI-9875513. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

A prefix consists of a 32-bit IP address and a mask length (e.g., $1.2.3.0/24$ represents IP block 1.2.3.0-1.2.3.255). Since the introduction of Classless Interdomain Routing (CIDR) [3], [4], [5], [6], a prefix can be of any length. This enables more aggressive *route aggregation* in which a single prefix is used to announce the routes to multiple prefixes. For example, prefixes 1.2.3.0/24 and 1.2.2.0/24 can be aggregated as prefix 1.2.2.0/23, and prefixes 1.2.2.0/23 and 1.2.3.0/24 can be aggregated as prefix 1.2.2.0/23. Route aggregation, however, might not always be performed. First, an AS can aggregate its prefix with its provider's only when the AS is *single-homed*, i.e., the AS has only one provider. For a *multi-homed* AS, which has multiple providers, its prefix(es) cannot be aggregated by all of its providers. Second, an AS may choose not to aggregate prefixes originated by it. One reason that an AS originates several prefixes is that an AS *fails to aggregate* aggregatable prefixes originated by it. The second reason that an AS originates several prefixes is load balancing. An AS originates several prefixes so as to perform *load balancing* by reaching different prefixes via different AS paths. The third reason that an AS originates several prefixes is address fragmentation. *Address fragmentation* is caused by a set of prefixes originated by the same AS that cannot be summarized by one prefix.

We explore the extent that factors such as multi-homing, failure to aggregate, load balancing, and address fragmentation contribute to routing table size. We examine the BGP routing table from the Route Views server, present techniques to quantify and perform measurement study on these factors. We find that multi-homing introduces around $20-30\%$ extra prefixes. Next, we explore how load balancing can contribute to routing table size and show that load balancing introduces around $20-25\%$ extra prefixes. However, multi-homing and load balancing are necessary trends and cannot be eliminated. This leads us to consider how the failure to aggregate can affect the routing table size and find that failure to aggregate increases the routing table size by only $15-20\%$. Finally, we explore the extent that address fragmentation contributes to the routing table size and find that address fragmentation contributes to more than $75\%$ of routing table size. Clearly, address fragmentation contributes to the routing table size the most. This leads us to introduce the concept of the *prefix cluster*, the maximal set of prefixes originated by the same AS that are not aggregated due to either failure to aggregate or address fragmentation. In other words, a prefix cluster is a maximal set of prefixes among which no load balancing is performed, i.e., that are announced identically by any router. We show that the number of prefix clusters is no more than $20\%$ of the number of prefixes.

Clearly, the extent of load balancing directly affects the number of prefix clusters. We explore the extent of load balancing by determining the percentage of ASes that do not perform load balancing among its prefixes and the percentage of ASes that perform load balancing beyond the last hop. We observe that approximately half of the ASes do not perform load balancing and among those that do, more than half perform load balancing beyond the last hop. Moreover, the percentage is relatively constant over the last three years.

It is important to predict the future growth pattern of prefixes and prefix clusters. To do so we take advantage of an observation that both the number of prefixes and the number of prefix clusters originated by an AS can be approximated by power-laws. Using these power-law approximations, we estimate the number of prefixes and the number of prefix clusters given the number of ASes. We can predict the number of prefixes and prefix clusters as the number of ASes grows. We observe that the number of prefixes grows much faster than the number of prefix clusters does. To the best of our knowledge, this is the first study on the explosive growth of routing table by systematically comparing factors that contribute to the growth and by observing routing table growth patterns.

The remainder of the paper is structured as follows. Section II presents the background on the Internet routing. In Section III, we explore how various factors contribute to the routing table growth and conclude that address fragmentation contribute to the routing table size the most. Section IV explore the effect of load balancing has on the number of prefix clusters. In Section V, we predict the growth patterns of the number of prefixes and the number of prefix clusters by observing power-laws. We conclude the paper in Section VI with a summary and further work.

## II. Internet Routing

In this section, we first describe the Internet architecture. We then present how IP addresses are allocated and route aggregations are performed to ensure the scalability of the Internet routing architecture. Finally, we describe the content of BGP routing tables.

### A. Internet Architecture

The Internet consists of a large collection of hosts interconnected by networks of links and routers. The Internet is divided into thousands of autonomous systems, each of which is administrated by a single organization. Each AS in the Internet is represented by a 16-bit AS number. An AS has its own routers and routing policies, and connects to other ASes to exchange traffic with remote hosts. ASes interconnect at dedicated point-to-point links or public Internet exchange points (IXPs) such as MAE-EAST or MAE-WEST. We model the connectivity between ASes in the Internet as an AS graph $G = (V, E)$, where the node set $V$ consists of ASes and the edge set $E$ consists of AS pairs that exchange traffic with each other. Two ASes that exchange traffic have either a customer-provider or peering arrangement. The relationships between ASes arise from contracts that define the pricing model and the exchange of traffic between two ASes [7], [8], [9]. In a customer-provider relationship, the customer is typically a smaller AS that pays a larger AS for access to the rest of the Internet. The provider may, in turn, be a customer of an even larger AS. In a peering relationship, the two peers are typically of comparable size and find it mutually advantageous to exchange traffic between their respective customers. Let $Provider(u)$ denote the set of AS $u$'s providers and $Peer(u)$ denote the set of AS $u$'s peers.

Throughout this paper, we use the AS relationships derived from the inference algorithm in [10]. [10] uses the routing policies implied by AS relationships to derive routing table entry patterns and to infer AS relationships based on these routing table entry patterns and on the heuristic that the size of an AS is typically proportional to its degree in the AS graph.

### B. IP Addresses

Each IP address is 32 bit long, and there are thus a total of $2^{32}$ (more than 4 billion) possible IP addresses. The scalability of the Internet routing infrastructure depends on the aggregation of IP addresses in contiguous blocks, called *prefixes*. We use $addr(p)$ and $len(p)$ to denote the *IP address* and the *mask length* of prefix $p$ respectively. Prior to the standardization of Classless Interdomain Routing (CIDR) in 1993, the mask length of a prefix could only be 8, 16 or 24 which represent Classes A, B or C respectively. With CIDR, a prefix can consist of any number of bits. This further improves the scalability of routing architecture. For example, an organization that supports 2,000 hosts needs to acquire at most $2^{11}$ IP addresses, which means that the prefix for the organization can have a mask length of 21 bits instead of 16 bits.

IP addresses are allocated in a hierarchical fashion. A host gets its IP address from the IP block of its organization's prefix. To acquire an IP address block, an organization typically contacts its ISP, which allocates addresses to the organization from its own block of addresses. An ISP acquires its address block from either its own provider or one of three regional routing registries: the American Registry for Internet Number (ARIN, which is responsible for North and South America), the Reseaux IP Europeans (RIPE, which covers Europe and nearby countries), and the Asia Pacific Network Information Center (APNIC, which covers Asia). Let $Prefix(u)$ denote all the prefixes originated by AS $u$.

### C. Route Aggregation

An AS employs an *inter-domain* routing protocol (Border Gateway Protocol or BGP) to advertise the reachability of these prefixes to neighboring ASes. The scalability of the Internet routing infrastructure depends on *route aggregation*. An AS performs route aggregation by using the minimum number of prefixes to summarize all of its IP addresses. Two prefixes are *aggregatable* if and only if the union of IP blocks represented by the two prefixes can be summarized by a prefix. For example, prefixes 1.2.3.0/24 and 1.2.2.0/24 can be summarized by prefix 1.2.2.0/23, and prefixes 1.2.2.0/23 and 1.2.3.0/24 can be summarized by prefix 1.2.2.0/23.

Route aggregation is typically performed in two situations. First, routes originated within the same AS can be aggregated if the aggregated IP addresses represented by their respective prefixes can be summarized by a single prefix. For example, suppose that AS 1 originates both 1.2.3.0/24 and 1.2.2.0/24. Instead of announcing both prefixes, AS 1 can announce 1.2.2.0/23 only. Second, routes originated by an AS can be aggregated with the route from the provider of the AS. When an AS announces a prefix that is contained in its provider's prefix, its provider can potentially aggregate the prefix and announce only the aggregation. For example, suppose AS 2 has only one provider, AS 1. AS 1 has IP address block 1.0.0.0/8 and allocates AS 2 IP block 1.2.0.0/15. Instead of announcing prefixes for AS 2 separately, AS 1 can announce address block 1.0.0.0/8 only. The longest prefix matching policy allows all traffic directed to AS 2 to go through AS 1 and in turn to be routed to AS 2.

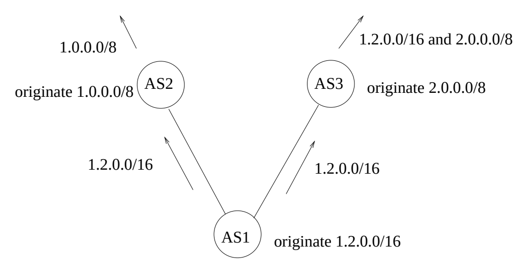



Fig. 1. Announcing a prefix for a multihomed AS

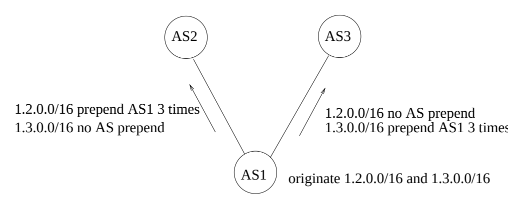



Fig. 2. Load balancing for multihomed AS

Route aggregation, however, cannot be performed all the time. First, an AS may not be able to aggregate its prefixes with its provider's. One reason that an AS does not aggregate with its provider is multi-homing. An AS is *multi-homed* if it has multiple providers to ensure connectivity even under the failure of some providers. That is, AS $u$ is *multi-homed* if and only if $|Provider(u)| > 1$. Note that we do not classify AS that is multi-homed to a single provider as a multi-homed AS in this paper. A multi-homed AS gets its address blocks from either some or all of its providers, or the Routing Registry directly. In any case, a multi-homed AS cannot aggregate its prefix(es) with those of its providers'. For example, in Figure 1, AS1 has two providers: AS2 and AS3. AS1 originates prefix 1.2.0.0/16, AS2 originates prefix 1.0.0.0/8, and AS3 originates prefix 2.0.0.0/8. To ensure that AS 1 can be reached via both providers, AS 1 announces 1.2.0.0/16 to both of its providers. AS 2 can announce 1.0.0.0/8 only since it contains 1.2.0.0/16. However, AS 3 has to announce both 2.0.0.0/8 and 1.2.0.0/16. Therefore, at least one of the providers of a multi-homed AS has to announce the prefixes originated by the AS. Otherwise, there is no redundancy provided to the multi-homed connections. Second, prefixes originated by the same AS might not be aggregated. One reason that an AS may not aggregate its prefixes is due to the desire to perform load balancing. For example, in Figure 2, AS1 announces 1.2.0.0/16 to AS2 prepending AS1 3 times and 1.3.0.0/16 to AS2 without AS prepend. AS1 announces 1.3.0.0/16 to AS3 prepending AS1 3 times and 1.2.0.0/16 to AS3 without AS prepend. This ensures that most of ASes reach 1.2.0.0/16 via AS3 and 1.3.0.0/16 via AS2. Therefore, the two provider links of AS1 can share the traffic to AS1. Another reason that an AS may not aggregate its prefixes is that an AS may fail to aggregate its prefixes even if they are aggregatable and no load balancing is performed among them. An AS may fail to aggregate due to the artifact of pre-CIDR practice, where only 8, 16 or 24 bit prefixes are announced. For example, an AS might originate prefixes 1.2.0.0/16 and 1.3.0.0/16 and announce them identically to others. Another reason that an AS may not aggregate its prefixes is address fragmentation. For example, an ISP might expand to have more customers and thus have an insufficient number of IP addresses. The ISP has to request additional IP address blocks which might not be aggregatable with its previously acquired IP address block.

### D. Routing Tables

Each BGP speaking router maintains a *BGP routing table*, which stores routes received from its neighbors. There is one entry for each destination prefix, which contains a set of candidate routes to reach the prefix. Each route contains a set of route attributes that includes an AS path and next hop. Formally, let $RouteEntry_u(p)$ denote the set of routes for prefix $p$ announced to AS $u$. An AS path is an ordered sequence of ASes that must be traversed in order to reach the destination prefix. The *last hop* of an AS path is the last edge (in the AS graph) of the AS path. Formally, the last hop of AS path $u_1, \ldots, u_n$ is $u_{n-1}, u_n$.

The size of the routing table is the number of entries in the table. Since each routing table entry represents a single prefix, the routing table size is the number of prefixes appeared in the routing table. Therefore, the extent that route aggregations are performed directly affects the routing table size. Next, we explore how various factors contribute to routing table growth. These factors include multi-homing, failure to aggregate, load balancing, and address fragmentation.

## III. Contribution to Routing Table Size by Various Factors

In Section II, we described various factors that contribute to the size of routing table. In this section, we quantify the extent that these factors contribute to the routing table size. In particular, we investigate to what extent that the routing table has been inflated due to multi-homing, failure to aggregate, load balancing, and address fragmentation. To this end, we choose to use BGP routing tables from the Route Views router in Oregon [11], which has the most complete view currently available. The Router Views router establishes 41 BGP peering sessions with ISPs such as tier-1 US providers and European providers. For a detailed description of the Route Views server, see [11]. The Route Views server collects the BGP routing table once every night [12] and has done so since November 1997. Figure 3 plots the routing table size of the Route Views router [11] between 1997 and 2000.

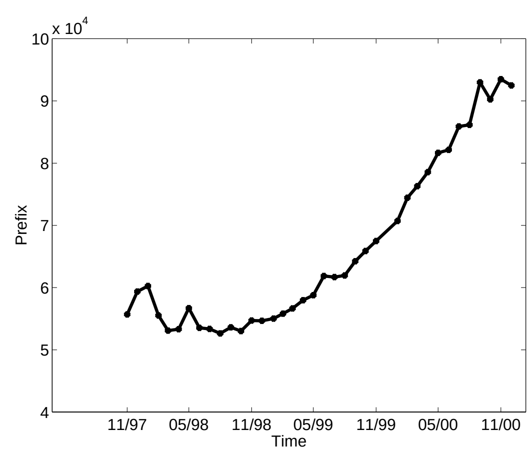



Fig. 3. Routing table size over three years

It is clear that the sizes of BGP routing tables might differ for different routers or ISPs. Although the Route

Views server is a special passive router, our measurement data show that it can capture the size of a typical BGP routing table. Figure 4 shows the number of prefixes announced by Sprint, Cable&Wireless, and RIPE(Reseaux IP Europeans) NCC to the Route Views server. We see that the number of prefixes from Sprint, Cable&Wireless, and RIP NCC are all quite close to the routing table size of the Route Views. In fact, the actual routing table sizes of these three peers might be larger than the numbers of prefixes announced by them, since they might not announce all the prefixes in their table to the Route Views server. Therefore, we use Route Views server's routing table size to indicate a typical routing table size throughout this paper. Furthermore, all of measurement data in this paper are derived from Route Views server's BGP routing tables from November 1997 to December 2000.

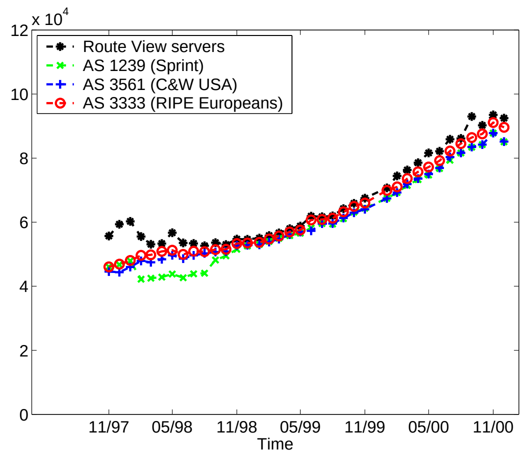



Fig. 4. Routing table size from different views

## A. Multi-homing

Many ASes connect to more than one provider for the purpose of fault tolerance. Multi-homing may create "holes" in the routing table. A hole is an address block that is contained in another announced address block but is announced separately. If a multi-homed AS originates a prefix, $p$, that is contained in a prefix announced from one of its providers, then $p$ has to be announced to the Internet by one of the multi-homed AS' providers for the purpose of fault tolerance as explained in Section II-C. On the other hand, if an AS is single-homed, it is not necessary that the AS announces the prefix beyond its providers. Therefore, we can evaluate the extent that multi-homing contributes to the routing table size by identifying *multi-homed prefixes*, i.e., prefixes that are originated by a multi-homed AS and contained in the prefixes originated by one of its providers. Formally, prefix $p_1$ contains prefix $p_2$ if and only if $len(p_2) > len(p_1)$ and $addr(p_2)/2^{32-len(p_1)} = addr(p_1)/2^{32-len(p_1)}$. Prefix $p$ is a multi-homed prefix if and only if $p \in Prefix(u)$, $u$ is a multi-homed AS, and $\exists q, v$ such that $q \in Prefix(v)$ and $v \in Provider(u)$ and $q$ contains $p$. Figure 5 plots the total number of prefixes and the number of prefixes that are not multi-homed prefixes over the last three years on the Route Views router. The number of multi-homed prefixes is on the rise and multi-homing introduces approximately $20-30\%$ more prefixes.

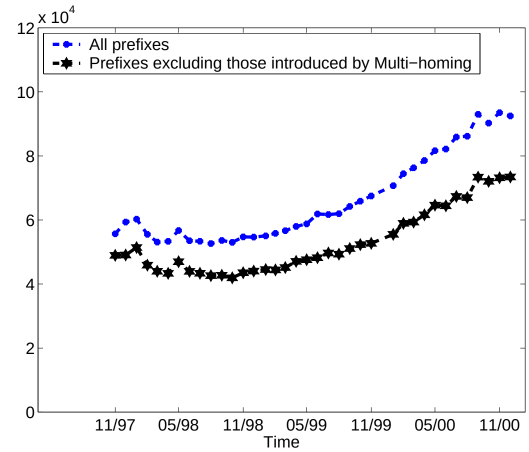



Fig. 5. Contribution of multi-homing to routing table size

Although multi-homing increases the routing table size significantly, it cannot be eliminated. Multi-homing is necessary to ensure the reliability of the Internet connection. Failure to aggregate, on the other hand, can be avoided by making aggregation mandatory. We address the extent of failure to aggregate in the next subsection.

## B. Failure to Aggregate

Some AS may fail to aggregate its aggregatable prefixes even though no load balancing is performed among those prefixes. In order to understand to what extent that failure to aggregate contributes to the routing table size, we aggregate all aggregatable prefixes that are originated by the same AS and are announced identically. First, we classify prefixes into prefix clusters, in each of which prefixes are announced identically. Formally, a *prefix cluster* is a maximal set of prefixes whose routing table entries are the same in the Route Views server's routing table. That is, two prefixes, $p_1$ and $p_2$, belong to the same prefix cluster if and only if $RouteEntry_v(p_1) = RouteEntry_v(p_2)$ for Route View server $v$. Note that although the Route View server has a limited view of the Internet, it does have a good sample of routes since it peers with many tier-1 ISPs.

Second, we perform aggregation for prefixes from the same prefix cluster iteratively as follows. Initially, we remove all prefixes that are contained in another prefix. That is, all prefixes contained in a prefix, $p$, are aggregated by prefix $p$. In each iteration, we first sort all prefixes in an increasing order of their addresses. We then aggregate each pair of consecutive prefixes that is aggregatable. A pair of consecutive prefixes, $p_1$ and $p_2$, are aggregatable if and only if $len(p_1) = len(p_2)$, $addr(p_1)/2^{32-len(p_1)} + 1 = addr(p_2)/2^{32-len(p_2)}$, and $addr(p_1)\%2^{33-len(p_1)} = 0$. The aggregated prefix has the address of $p_1$ and the length of $p_1$ minus 1. We repeat the iteration until no aggregation can be performed. The total number of prefixes after the aggregation is the number of prefixes excluding those that are introduced by failure to aggregate. Figure 6 plots the number of prefixes and the number of prefixes excluding those that are introduced by failure to aggregate. We observe from the figure that approximately $15-20\%$ prefixes could be aggregated beyond what network operators have done.

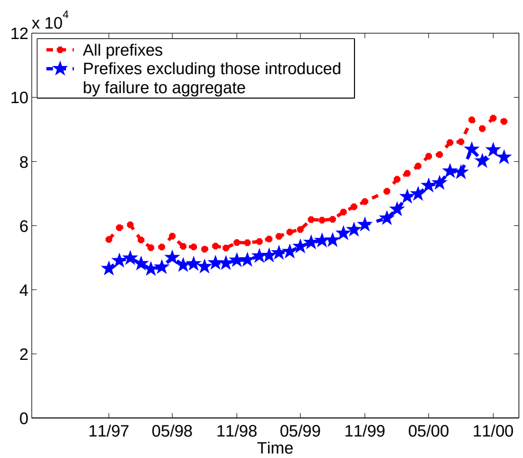



Fig. 6. Contribution of failure-to-aggregate to routing table size

## C. Load Balancing

Failure to aggregate introduces more prefixes since an AS does not aggregate its aggregatable prefixes even though those prefixes are announced identically. Another reason that route aggregation cannot be performed for prefixes originated by the same AS is load balancing. Two aggregatable prefixes might not be aggregated since they are announced differently. To quantify the effect of load balancing on the routing table size, we first compute the number of prefixes resulting from aggregating all aggregatable prefixes originated by the same AS independent of whether those prefixes are announced identically or not. That is, we perform aggregation for prefixes from the same AS iteratively as shown in Section III-B. We compare the total number of prefixes after the aggregation with the number of prefixes excluding those introduced by failure to aggregate. The difference between the two numbers quantifies that load balancing contributes to routing table size. We plot the two numbers and the number of prefixes in Figure 7. It is observed from the figure that the load balancing introducing an additional $20-25\%$ more prefixes.

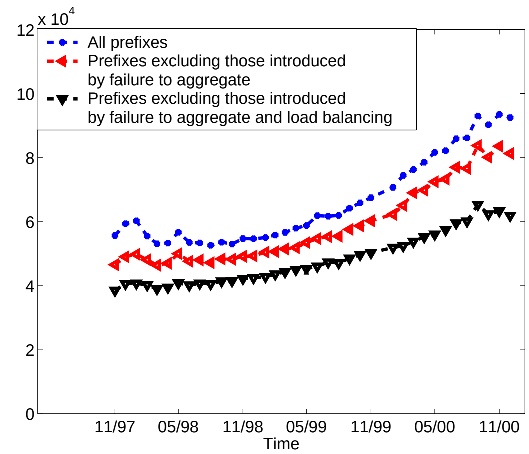



Fig. 7. Contribution of load balancing to routing table size

## D. Address Fragmentation

Both multi-homing and load balancing are necessary trend of the Internet. Although it is possible to eliminate the prefixes that are due to failure to aggregate, the reduction on the routing table size is not significant. We conjecture that this is due to the address fragmentation. Since all of the prefixes within the same prefix cluster are announced identically, a single routing table entry would be sufficient for them if these prefixes could be represented by one prefix cluster. However, the Internet addresses covered by these prefixes may not be summarized by one prefix due to either failure to aggregation or address fragmentation. In this section, we investigate the effect of address fragmentation by comparing the number of prefixes excluding those contributed by failure to aggregate with the number of prefix clusters.

We plot the number of prefix clusters in Figure 8. The number of prefix clusters is only about $1/5$ of the size of current routing table. The contribution of the address fragmentation to the routing table size is the gap between the number of prefixes excluding those introduced by failure to aggregate and the number of prefix clusters. It is suggested by the plot that address fragmentation contributes to more than 75% of the routing table size and is the most significant factor.

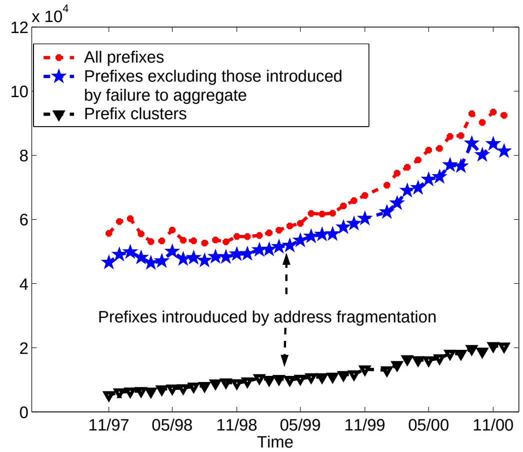



Fig. 8. Contribution of address fragmentation to routing table size

## IV. Effect of Load Balancing on the Number of Prefix Clusters

We have learned from Section III that the number of prefix clusters is significantly less than the size of routing table. In this section, we first show the extent that load balancing contributes to the number of prefix clusters. We then investigate the extent that load balancing is performed on the AS level.

### A. *Contribution of Load Balancing to the Number of Prefix Clusters*

To quantify the effect of load balancing, we first determine the number of prefix clusters required under the assumption that no load balancing is performed. As we described in Section III-B, a prefix cluster includes all prefixes that are announced identically. Prefixes originated by different ASes belong to different prefix clusters since they are announced with different AS paths. That is, there is at least one prefix cluster from each AS. AS performs If no AS perform load balancing, there are at most one prefix cluster from each AS. Therefore, the number of prefix clusters is the number of ASes in the BGP routing table under the assumption that no load balancing is performed. We plot both the number of prefix clusters and the number of prefix clusters if no load balancing is performed in Figure 9. The effect of load balancing on the number of prefix clusters is characterized by the difference between the two plots. It is observed from the figure that load balancing contributes more than 70% of the prefix clusters.

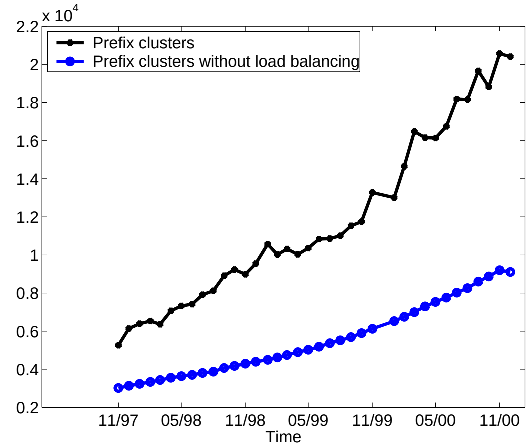



Fig. 9. Contribution of load balancing to the number of prefix clusters

We have observed that load balancing significantly contributes to the total number of prefix clusters. In the next subsection, we examine the extent that load balancing is performed at the AS level.

### B. *Extent of Load Balancing*

To understand the extent that load balancing is performed at the AS level, we first show the fraction of ASes that do not perform load balancing. The number of ASes that do not perform load balancing is the number of ASes that originate one prefix cluster. Figure 10 plots the fraction of ASes that did not perform load balancing over the last three years. We observe from the figure that approximately 50% of ASes do not perform any load balancing and that this fraction has remained relatively constant over time.

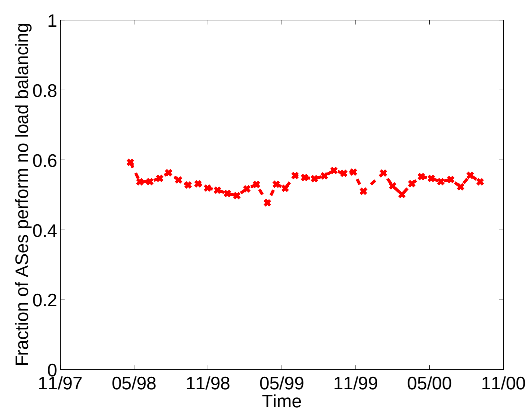



Fig. 10. The fraction of ASes that do not perform load balancing

We then show the extent of load balancing for those

ASes that do perform load balancing. To this end, we compute the fraction of aggressive-LB ASes among those ASes that perform load balancing. *An aggressive-LB AS* is an AS on which load balancing is performed beyond the last hop, i.e., two prefix clusters originated by the AS use the same last hop but use different AS-level hop before the last hop. Since the last hop AS is the provider or peer of the AS, we classify the an AS as an aggressive-LB AS if the total number of providers and peers of the AS is less than the number of prefix clusters originated by the AS. That is, an AS $u$ is an aggressive-LB AS if $|Provider(u)| + |Peer(u)| <$ the number of prefix clusters originated by $u$. The fraction of aggressive-LB ASes among the ASes that perform load balancing gives a lower bound on the fraction of ASes on which load balancing is performed beyond the last hop. We plot the fraction of aggressive-LB ASes in Figure 11. The figure suggests that there are around $60\%$ ASes on which load balancing is performed beyond the last hop and and this fraction has been relatively constant over time. We

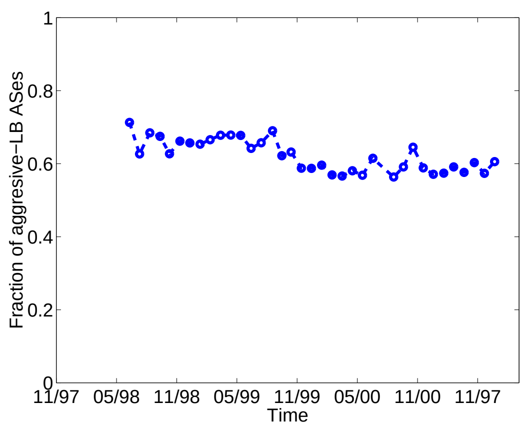



Fig. 11. The fraction of aggresive-LB ASes

see that both the fraction of ASes that do not perform load balancing and of ASes that perform load balancing beyond the last hop is relatively constant over time. This leads us to conjecture that the number of prefix clusters is a simple function of the number ASes over time. To predict the growth of the number of prefixes and the number of prefix clusters, we explore the use of power law approximations to predict the number of prefixes and the number of prefix clusters in the next section. We will see that the number of prefix clusters grow much slower than the number of prefixes.

## V. Predicting routing table size and the number of prefix cluster

Power laws have been observed in the Internet topology [13]. In this section, we first demonstrate power law approximations on the number of prefixes and the number of prefix clusters. That is, both the number of ASes containing at least $d$ prefixes and the number of ASes containing at least $d$ prefix clusters can be approximated by power laws. We then use the power laws to estimate the total number prefixes and the total number of prefix clusters given the number of ASes.

Let $F^{prefix}(d)$ and $F^{aggregate}(d)$ denote the number of ASes containing at least $d$ prefixes and prefix clusters respectively. We plot $F^{prefix}(d)$ in log-log scale in Figure 12 and $F^{aggregate}(d)$ in log-log scale in Figure 13. In order to present the power laws, we use linear regression to fit a line in a set of two-dimensional points. The technique is based on the least-square errors method. The validity of the power law is indicated by the correlation coefficient which is a number between $-1.0$ and $1.0$. A correlation coefficient of $1.0$ or $-1.0$ indicates the perfect linear correlation, i.e., the data points form a line. The dashed lines in the graphs are the results of the linear regression and the correlation coefficient for each linear regression are shown in the graph. We observe that all plots are approximately linear except for a small number of outliers on both the left side and the right side of each figure.

This leads to the following power law approximations.

**Power-Law on the number of prefixes**

*The number of ASes that contains at least $d$ prefixes is approximately proportional to $d$ to the power of a constant, $\alpha$:*

$$F^{prefix}(d) \propto d^{\alpha} \quad (1)$$

*where the prefix exponent $\alpha$ is the slope of the plot of $F^{prefix}(d)$ versus $d$ in log-log scale.*

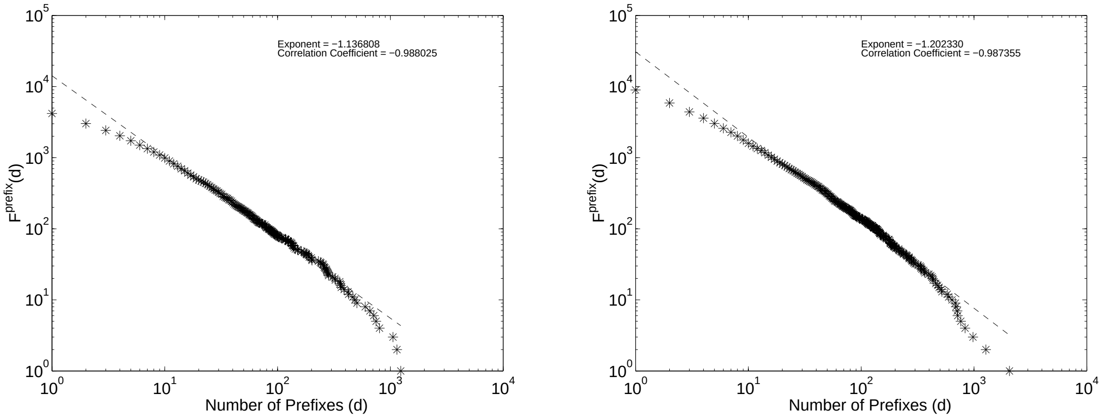
**Power-Law on the number of prefix clusters**

*The number of ASes that contains at least $d$ prefix clusters is approximately proportional to $d$ to the power of a constant, $\beta$:*

$$F^{aggregate}(d) \propto d^{\beta} \quad (2)$$

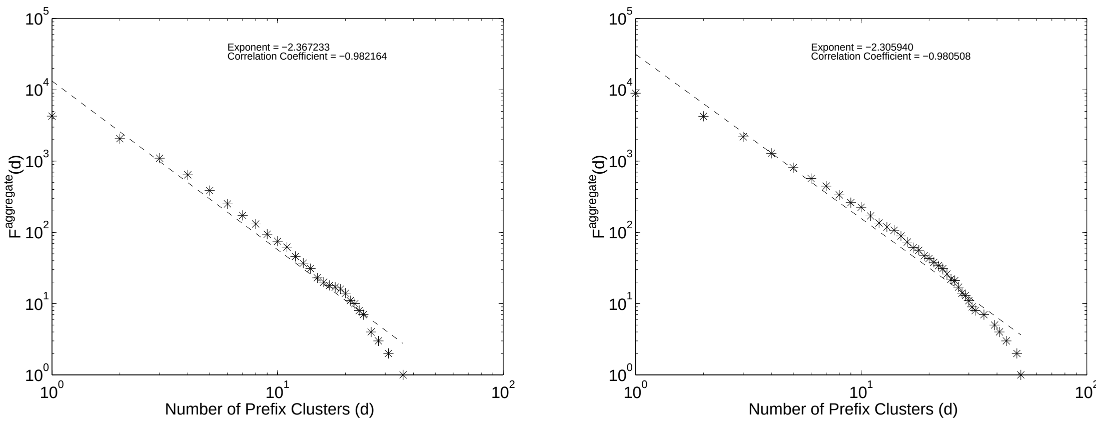
*where the prefix cluster exponent $\beta$ is the slope of the plot of $F^{aggregate}(d)$ versus $d$ in log-log scale.*

Note that the power law is only one of the possible approximations to the curves in Figure 12 and Figure 13. We are investigating better approximations that will account for the outliers and that will allow us provide more




Nov. 1998 Nov. 2000

Fig. 12. $F^{prefix}(d)$ vs. $d$ in loglog plot




Nov. 1998 Nov. 2000

Fig. 13. $F^{aggregate}(d)$ vs. $d$ in loglog plot

accurate prediction on the number of prefix and prefix clusters

We plot the exponents of prefix power laws and the exponent of prefix cluster power laws over the past three years in Figures 14 and 15 respectively. We observe that the values of the exponents have changed very slowly over this time period and have remained constant over the past two years. Note we exclude a small number of outliers ASes in each routing table when we compute power law exponents. Specifically, let $Frequency(d)$ be the number of ASes that contain $d$ prefixes (or prefix clusters). We only include those ASes that contain $d$ prefix(prefix cluster) where $d$ starts from 1 to the minimum value of $n$ such that $Frequency(n) = 1$. This results in the exclusion of fewer than 1% ASes.

If we require that only one AS originates the maximum number of prefixes or prefix clusters (this is observed from the routing table of the Route Views server), we can approximate the number of prefixes (prefix clusters) given the number of ASes and the power law exponents.

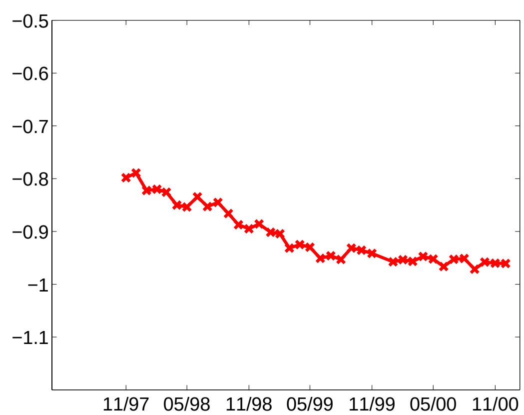



Fig. 14. Prefix exponents

*Approximation 1:* The total number of prefixes, $P$, within a routing table can be estimated as a function of the total number of ASes, $N$, and the prefix exponent, $\alpha$,

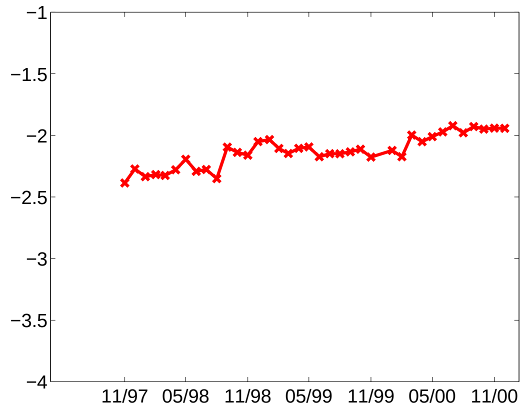



Fig. 15. prefix cluster exponents

as follows:

$$P = N \sum_{d=1}^{p_{max}} d^{\alpha} \tag{3}$$

where $p_{max} = N^{-1/\alpha}$.

*Approximation 2:* The total number of prefix clusters, $S$, within a routing table can be estimated as a function of the total number of ASes, $N$, and the prefix exponent, $\beta$, as follows:

$$S = N \sum_{d=1}^{s_{max}} d^{\beta} \tag{4}$$

where $s_{max} = N^{-1/\beta}$.

We now show how to obtain Approximation 1; Approximation 2 can be obtained similarly.

According to the power law, we have

$$F^{prefix}(d) = C_{prefix} d^{\alpha}. \tag{5}$$

We know that $F^{prefix}(1) = N$. That is, $C_{prefix} = N$. Therefore,

$$F^{prefix}(d) = N d^{\alpha}. \tag{6}$$

Let $p_{max}$ denote the maximum number of prefixes that an AS originates. To derive $p_{max}$, we use the fact that there is only one AS that originates the maximum number of prefixes. We have $F^{prefix}(p_{max}) = 1$ and obtain

$$p_{max} = N^{-1/\alpha} \tag{7}$$

The total number of prefixes is

$$P = \sum_{d=1}^{p_{max}} d(F^{prefix}(d) - F^{prefix}(d+1)) = N \sum_{d=1}^{p_{max}} d^{\alpha} \tag{8}$$

We compare the number of prefixes (or prefix clusters) derived from Approximation 1 (or 2) with the actual number of prefixes (or prefix cluster). The estimated number of prefixes differs by $5\%$ to $15\%$ from the actual number of prefixes and the estimated number of prefix clusters differs by $8\%$ to $20\%$ from the actual number of prefix clusters.

We can use these approximations to predict the trend of the routing table growth as well as the growth of the number of prefix clusters. Before providing any details, we mention that the prefix growth and the prefix cluster growth are due to many factors, it is very difficult to predict their growth very accurately by only knowing two metrics. However, we believe that our approximations are sufficiently accurate to allow us to understand the growth trend of prefixes and prefix clusters.

It is observed that the exponents of power laws have been practically constant for the past two years. More precisely, the median of the exponents of the power laws on prefixes is $-0.95$ during past two years and all the prefix exponents are within $8\%$ of the median. The median of the exponents of the power law on prefix clusters is $-2.06$ during past two years and all the prefix cluster exponents are within $10\%$ of the median. Since the exponents have not changed much over the past two years, it is reasonable to assume that the exponents will not change significantly in the near future unless there are significant changes on the Internet architecture. Assume that the prefix exponent is $-0.95$ and the route aggregate exponent is $-2.06$. The number of ASes is bounded by $2^{16}$ since AS number is only 16 bit long. We estimate the total number of prefixes and prefix clusters when the number of ASes is between $1000$ and $64,000$ in Figure 16. Note that if $N = 10,000$, then our approximations conclude that $P/N = 13$ and $S/N = 1.5838$. If $N = 64000$, then our approximations conclude that $P/N = 16.38$ and $S/N = 1.5889$. In general, $S/N$ is close to a constant when $N$ is sufficient large whereas $P/N$ grows larger as $N$ grows. This suggests that the routing table grows faster than the number of prefix clusters.

## VI. Conclusions and Future Work

In this paper, we first investigated the extent that various factors contribute to the routing table size. Among multi-homing, failure to aggregate, load balancing, and address fragmentation, address fragmentation contributes the most to the routing table size. This led us to introduce the concept of prefix cluster, a set of prefixes that are announced identically, and present the extent that load balancing contributes to the number of prefix clusters. We

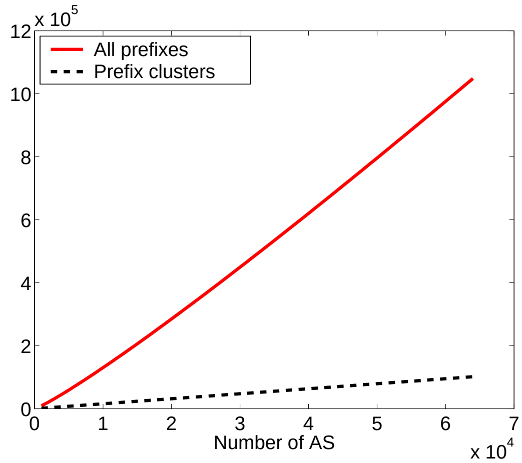



Fig. 16. Estimation of the number of prefixes and prefix clusters

then study the growth pattern of the routing table. By observing power laws on the number of prefixes and the number of prefix clusters, we predict the number of prefixes and the number of prefix clusters given the number of ASes. The prediction suggests that the number of prefix clusters grows more slowly than the number of prefixes.

Several interesting topics remain to be studied in the future. First, this work is done based on the routing table collected by the Route Views server. We would like to see whether our conclusions are valid from other routers' view. Second, the large discrepancy between the number of prefixes and the number of prefix clusters suggests that it might be beneficial to propose an extension of the current routing protocol. The key idea is to use prefix clusters instead of prefixes to identify the route whenever possible. By pushing a prefix cluster identifier to an IP packet by an ingress domain and adding a prefix cluster identifier to a route announcement, we can effectively use the prefix cluster as an index for packet forwarding. Since matching a prefix cluster identifier does not require longest prefix matching, we can increase the packet forwarding speed. Similar to pushing an Multiple Protocol Label Switching (MPLS) [14] header, we can push a prefix cluster identifier to an IP packet. This "end-to-end MPLS" might "solve" some of the route-aggregation problems we discovered in this paper. The proposal on extending the current routing protocol is preliminary. As a part of our future work, we will investigate how MPLS could be exploited to implement prefix clusters, design the extended protocol and build a prototype to evaluate the performance.

## References


[1] `http://www.telstra.net/ops/bgptable.html/`.

[2] G. Huston, "Analyzing the internet bgp routing table," in *Internet Protocol Journal*, March 2001.

[3] R. Hinden, "Applicability statement for the implementation of classless inter-domain routing (CIDR)." Request for Comments 1517, September 1993.

[4] Y. Rekhter and T. Li, "An architecture for IP address allocation with CIDR." Request for Comments 1518, September 1993.

[5] V. Fuller, T. Li, J. Yu, and K. Varadhan, "Classless inter-domain routing (CIDR):an address assignment and aggregation strategy." Request for Comments 1519, September 1993.

[6] Y. Rekhter and C. Topolcic, "Exchanging routing information across provider boundaries in the CIDR environment." Request for Comments 1520, September 1993.

[7] G. Huston, "Interconnection, peering and settlements–Part I," in *Internet Protocol Journal*, March 1999.

[8] G. Huston, "Interconnection, peering and settlements–Part II," in *Internet Protocol Journal*, June 1999.

[9] G. Huston, "Interconnection, peering, and settlements," in *Proc. INET*, June 1999.

[10] L. Gao, "On inferring autonomous system relationships in the Internet," in *Proc. IEEE GLOBE INTERNET*, November 2000.

[11] `http://www.antc.uoregon.edu/route-views/`.

[12] National Laboratory for Applied Network Research, `http://moat.nlanr.net/AS/`.

[13] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proc. ACM SIGCOMM*, August 1999.

[14] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol label switching architecture." Request for Comments 3031, January 2001.