Evolution of Internet Address Space Deaggregation: Myths and Reality

Luca Cittadini, Wolfgang Mühlbauer, Steve Uhlig

***Abstract*— Internet routing table size growth and BGP update churn are two prominent Internet scaling issues. There is widespread belief in a high and fast growing number of ASs that deaggregate prefixes, e.g., due to multi-homing and for the purpose of traffic engineering [1]. Moreover, researchers often blame specific classes of ASs for generating a disproportionate amount of BGP updates. Our primary objective is to challenge such widespread assumptions ("myths") and *not* solely to confirm previous findings [1]–[3]. Surprisingly, we find severe discrepancies between existing myths and reality. According to our results, there is no trend towards more aggressive prefix deaggregation or traffic engineering over time. With respect to update dynamics, we observe that deaggregated prefixes generally do not generate a disproportionate number of BGP updates, with respect to their share of the BGP routing table. On the other side, we observe much more widespread traffic engineering in the form of AS path prepending and scoped advertisements compared to previous studies [1]. Overall, our work gives a far more positive picture compared to the alarming discourses typically heard [1], [2], [4]: The impact of "bad guys" on routing table size growth and BGP churn has not changed for the worse in recent years. Rather, it increases at the same pace as the Internet itself.**

***Index Terms*— Routing table growth, update churn, address deaggregation, traffic engineering**

I. INTRODUCTION

The research community is now accustomed to alarmist discourse regarding the lack of *scalability* of the current routing system. By now, routing tables in the default-free zone of the Internet contain approximately $300,000$ prefixes and continue to grow super-linearly [3]. This goes along with a steady increase in the rate of changes to the routing and forwarding tables. Having to handle big tables which need frequent recomputations, routers waste more and more resources to maintain their routing tables rather than to forward packets. Considerable efforts have been made to quantify the evolution of routing table size and BGP update rates [2], [4], to characterize the BGP routing table growth [3] or to study the dynamics that can be observed in Internet routing [5], [6].

Address space *fragmentation*, i.e., the use of a large number of small, potentially overlapping prefixes in the routing system, is widely seen as one major driving force behind the inflation of routing tables and high BGP update rates. Fragmentation of the IPv4 address space can be caused by either *(i)* the Regional Internet Registries (RIR), who allocate small and unconnected blocks of IPv4 addresses to Internet Service Providers (ISP), or *(ii)* ISPs who use the Border Gateway Protocol (BGP) to inject multiple more specific prefixes instead of, or in addition to, their assigned prefix blocks.

For the latter case, the current routing system poses an inevitable conflict: While individual domains inject more specifics for what they perceive as legitimate reasons, e.g., to implement multi-homing or to better control how traffic enters their own network[1], other Autonomous Systems (ASs) pay a fee, e.g., in terms of increasing convergence times, for having to deal with an "inflated" routing table. We believe that a deeper understanding of the root causes of the Internet's routing scaling problems is urgently needed, in particular in the light of recent interest in re-designing the Internet routing architecture using "clean-slate" approaches [7].

We, in this paper, seek to shed light on the evolution of address space deaggregation and of update dynamics. For this purpose, we rely on BGP routing tables and update traces from RIPE [8] and RouteViews [9], as well as on the official allocation files as available at the RIRs [10]. An important contribution of our work is to extract information about traffic engineering practices from observed routing data and correlate this information with the deaggregation that we can observe for prefixes in BGP routing tables. Overall, our primary interest is *not* just to confirm previous findings [1]–[3]. Rather we challenge wide-spread assumptions that are frequently made whenever researchers discuss the root causes of the current scaling problems. Amongst others, such "myths" are:

(i) There exists a *large number of "bad guys"*, i.e., ASs that strongly contribute to routing table growth and high update churn via aggressive use of prefix deaggregation.
(ii) Recently, there has been a *strong shift towards a more widespread use of prefix deaggregation*. This can be explained by an increasing number of multi-homed ASs that apply traffic engineering.
(iii) With respect to BGP dynamics, there are *differences between various types of ASs*. For example, BGP update patterns differ depending on the business role of an AS, e.g., transit provider vs. access network or stub vs. tier-1.

Our results reveal discrepancies between such myths and reality. Surprisingly, there is no trend towards more aggressive prefix deaggregation or traffic engineering over time. Most of the prefixes advertised in BGP routing tables actually match the address block that has been allocated by the RIRs. Yet, there exists a roughly constant number of ("bad") guys, which (strongly) deaggregate their address space for possibly

Luca Cittadini is with the Department of Computer Science and Automation, Roma Tre University, Rome, Italy, Email: luca.cittadini@gmail.com. Wolfgang Mühlbauer is with ETH Zürich, Zurich, Switzerland, Email: wolfgang.muehlbauer@tik.ee.ethz.ch. Steve Uhlig is with Technische Universität Berlin/Deutsche Telekom Laboratories, Berlin, Germany. Email: steve@net.t-labs.tu-berlin.de.

[1] Note that there exist more reasons why ASs may deaggregate their address space. For example, ASs have announced sub-prefixes (e.g., /24) of their assigned prefix block (e.g., /16) to protect themselves against prefix hijacking.

legitimate reasons, e.g. insufficient support for traffic engineering [1] in the current Internet. With respect to BGP updates, we cannot claim that a specific class of prefixes or origin ASs misbehave more than others. Similar to routing table size, we cannot identify any trends, e.g., towards deaggregated prefixes causing a lot of churn and thus frequent recomputations of routing tables. All in all, our analysis shows that the impact of "bad guys" on routing table size growth and BGP churn has not changed for the worse in recent years: rather, it increases at the same pace as the Internet growth itself.

The remainder of this paper is structured as follows: We start with a study of the allocations made by the RIRs in Section II. Then, Section III turns to deaggregation as observed only from static BGP data. In Section IV we compare allocated with observed IPv4 addresses, before we introduce in Section V a classification scheme to estimate the prevalence of traffic engineering and correlate this information with our results on prefix deaggregation from the preceding sections. Finally, Section VI explores whether specific classes of prefixes or ASs exhibit significant misbehavior in terms of BGP dynamics. This allows us to understand whether or not high updates rates are a direct consequence of an increasingly fragmented IPv4 address space. We summarize and conclude in Section VIII.

II. ALLOCATION OF ADDRESS SPACE

The current IP addressing architecture requires central coordination to ensure that different networks use unique non-overlapping IP prefixes. To this end, IP addresses are allocated by the Internet Assigned Numbers Authority (IANA) from pools of unused address space and delegated to the appropriate Regional Internet Registries (RIR). An Internet Service Provider (ISPs) that requests IP address space from these registries[2] is called Local Internet Registry (LIR) [11].

By the end of 2008 some 75% of the total usable IPv4 address space have been allocated. Current estimates for address space exhaustion are somewhere in 2011 for IANA address space and 2012 for RIRs [12]. The goal of this section is to investigate how the allocation policy adopted by RIRs has accommodated the topological growth of the Internet and to what degree it has promoted the inflation of routing tables.

The data sources we use for our analysis are the official allocation files [10] provided by the RIRs. These files summarize the current state of allocations and the assignment of Internet number resources. Figure 1 presents the evolution of the distribution of allocated blocks. The stacked area plot shows for each year between 1982 and 2008 the number of address blocks allocated up to and including that year.

Overall, allocations are dominated by /24 and /16 blocks, especially during the classful era of the Internet before 1994. /24 allocations still represent more than 40% of all allocated blocks today. Since 1994, the number of /24 and /16 blocks has been rather stable, while intermediate sized blocks (/19, /20 and /23) have been slowly increasing. This was made possible by the introduction of Classless Inter-Domain Routing (CIDR) [13] in 1993. Indeed, registries started to allocate CIDR blocks around 1994/1995 and, with CIDR, could allocate blocks that actually matched the expected use of the address space [14]: On the one hand, there was the fear that the class B space (/16) would soon run out, since it was consumed at a massive rate. On the other hand, some ISPs were afraid of too many small allocations (/24) that would tremendously inflate their routing tables and induce high costs for new router hardware. As a compromise, RIRs started to allocate intermediate size blocks, e.g., /19.

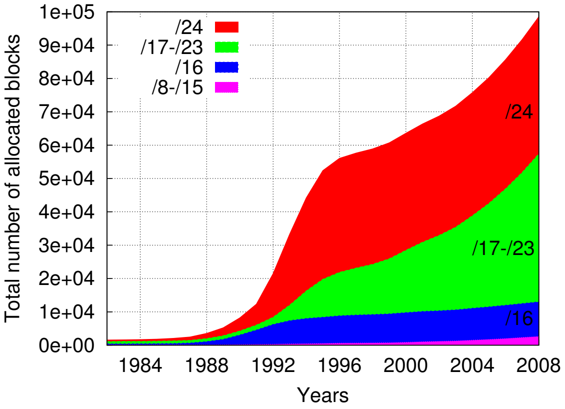


Fig. 1. Breakdown of prefix blocks allocated up to a certain year.

We point out that shorter prefix lengths correspond to significant fractions of the IPv4 address space although they only account for a negligible number of allocations: at the end of 2008, 50% of the total IPv4 address space has been allocated in blocks of length between /8 and /15.

In summary, the historical evolution of IPv4 address space allocation indicates a shift towards intermediate sized prefix blocks (/19-/23) since 1994. We conjecture that RIRs switched to a more parsimonious allocation policy to accommodate Internet's growth. Various allocation policies which are being discussed and adopted in the RIRs [15] envision much smaller sized blocks, e.g., /24 to /29 in an attempt to allow a long period of entry for newcomers despite the depletion of the main free pool of IPv4 address space. This is a change in the very conservative policies which attempted to minimize routing table impact, but are believed to have created barriers to entry.

III. OBSERVED ADDRESS SPACE FRAGMENTATION

We now turn to actually announced prefixes and observed deaggregation, an "avoidable" cause for the explosion of routing table entries. We investigate if patterns of deaggregation have changed recently (Section III-B) and if ASs of different business types follow different strategies with respect to deaggregation of assigned prefix blocks (Section III-C). To this end, we classify prefixes observed in BGP table dumps according to the inclusion relationships between prefixes, i.e., which one is more or less specific than another (Section III-A).

A. *Prefix Classification*

To study the advertised IPv4 address space, we rely on BGP table dumps from RIPE [8] and RouteViews [9]. Based on the

---

[2] In the APNIC and LACNIC region IP address allocations are mainly coordinated by National Internet Registries (NIRs)

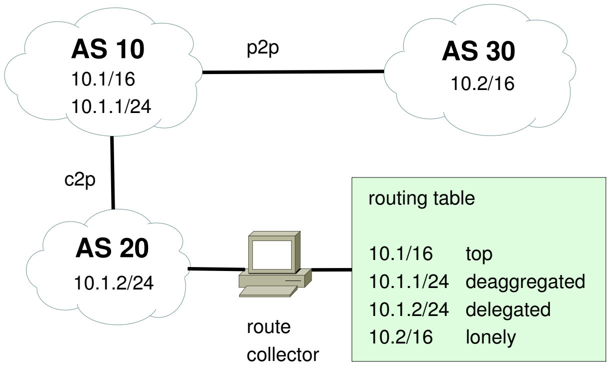


Fig. 2. BGP prefix classification.

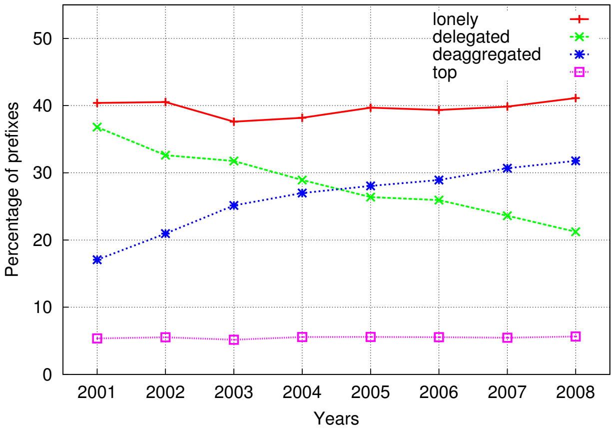


Fig. 3. BGP prefix classification distribution.

presence of overlapping prefix blocks, we assign every prefix in the routing table of our observation points to one of the following classes:

- *Lonely*: a prefix that does not overlap with any other prefix.
- *Top*: a prefix that covers one or more smaller prefix blocks, but is not itself covered by a less specific.
- *Deaggregated*: a prefix that is covered by a less specific prefix, and this less specific is originated by the same AS as the deaggregated prefix.
- *Delegated*: a prefix that is covered by a less specific, and this less specific is not originated by the same AS as the delegated prefix.

For illustration purposes, let us assume that the complete Internet consists of only three ASs, see Figure 2. There is a customer-to-provider (c2p) link between AS 10 and AS 20, i.e., AS 20 is a customer of AS 10. Moreover, there is a peer-to-peer (p2p) link between AS 10 and AS 30. Therefore, the following prefixes are observed at the observation point in AS 20: 10.1/16 (origin: AS 10), 10.1.1/24 (origin: AS 10), 10.1.2/24 (origin: AS 20) and 10.2/16 (origin: AS 30). Based on this information, we classify 10.2/16 as *lonely*, 10.1/16 as *top*, 10.1.1/24 as *deaggregated* and 10.1.2/24 as *delegated*. We point out that *delegated* corresponds to a provider aggregatable (PA) addressing model: A customer AS is reachable via its provider or, to put it differently, the AS path observed for a *delegated* prefix is going through the provider from which the address space is sub-allocated.

One important difference between our classification and the one presented in [1] is that we distinguish the case where the AS that originates a more specific prefix is not the same as the AS that originated the covering prefix. In that case, we talk about delegation. This difference is important if one wants to realistically quantify the portion of BGP routing tables that could be shrinked by aggregating more specific prefixes. In fact, we point out that delegated prefixes cannot be aggregated into their less specific without impairing the ability of some ASs to be reachable via multiple providers.

B. Prefix Deaggregation

We study multiple observation points for RIPE [8] and RouteViews [9] data, but only find trivial differences with respect to our prefix classification of Section III-A. In the following discussion, we rely on an observation point from RouteViews inside Level 3 (AS 3356). Since this router is part of the default-free zone (DFZ) of the Internet, it provides a representative picture of all globally-routable prefixes.

Figure 3 shows the evolution of these four prefix classes from 2001 to 2008. About 40% of all prefixes are of the *lonely* type. *Top* prefixes represent a stable 5% of the prefixes in the BGP routing tables. *Delegated* prefixes account for 35% of the prefixes in 2001, declining to a bit more than 20% in 2008. The decreasing fraction of *delegated* prefixes is matched by an increasing fraction of *deaggregated* prefixes, which increase from less than 20% in 2001 to more than 30% in 2008.

Clearly, it is possible to shrink routing tables by aggregating or filtering *deaggregated* and *delegated* prefixes. *Lonely* prefixes, on the other hand, are not aggregatable based only on the content of the BGP routing table. Consistent with other studies [2], [4], we find that, in theory, the current BGP routing tables could be reduced by a factor of slightly more than 2 if *delegated* and *deaggregated* were aggregated into *top* prefixes, leaving only non-overlapping prefixes in the routing tables. However, note that *delegated* prefixes announced by multihomed ASs cannot be aggregated without losing the ability to load balance incoming traffic among multiple providers.

From Figure 3 we learn that the combined fraction of *delegated* and *deaggregated* prefixes has remained approximately constant over the years. However, there is a shift from *delegated* to *deaggregated* prefixes. We speculate that this trend reflects the increasing popularity of provider-independent (PI) addresses, which avoid provider lock-in and renumbering issues when changing the provider.

C. Business Type of Originating ASs

According to [16] not all parts of the Internet grow at the same rate. It is widely agreed that the Internet has a tiered structure [17] which reflects AS business relationships, e.g., [17]–[19]. A few tier-1 ISPs form the *core*. A larger number of transit providers buy service from other providers, including tier-1 providers, and provide connectivity to other ASs. Stub ASs get their connectivity from transit or tier-1

AS class	Prefix type	# Prefixes	% of prefixes in class
Enterprise Customers (EC)	lonely	57,567	44.7%
	delegated	35,819	27.8%
	deaggregated	29,459	22.9%
	top	5,893	4.6%
Content/Access/-Hosting Provider (CAHPs)	lonely	5,884	54.2%
	delegated	1,260	11.6%
	deaggregated	2,700	24.9%
	top	1,008	9.3%
Small Transit Provider (STP)	lonely	28,260	38.3%
	delegated	13,206	17.9%
	deaggregated	27,131	36.8%
	top	5,147	7%
Large Transit Provider (LTP)	lonely	5,456	50.2%
	delegated	1,695	15.6%
	deaggregated	2,505	23.1%
	top	1,211	11.1%

TABLE I

BREAKDOWN OF PREFIX CLASSES WITHIN AS TYPES AS DEFINED IN [16].

providers and form the *edge*. Most of the growth of the AS-level topology is due to networks at the edge [16]. Accordingly, one might expect that deaggregation is also mostly due to edge networks. To verify this, we now look closely at the origin ASs where prefixes of different classes are injected into the routing system.

We rely on the classification of ASs according to their business type taken from [16]. Based on manual classification, an initial training set and on machine-learning techniques, Dhamdhere et al. distinguish between four different types of ASs: Enterprise Customers (EC), Small Transit Providers (STP), Large Transit Providers (LTP) and Content/Access/Hosting Providers (CAHPS). Most of the ASs (92%) are found to be EC ASs. For more details about this classification refer to Section 4 of [16].

For each classified prefix (see Section III-A), we check the business type of its originating ASs according to the categories of Dhamdhere et al. [16]. Table I shows the breakdown of our four types of prefixes into the four business classes of ASs. The first column gives the AS class as defined in [16], the second the prefix class as defined above, the third column the number of prefixes originated by that class among those originated by the ASs of the given type.

In Table I, we first notice that the majority of prefixes are originated by EC and STP networks. Content/Access/Hosting providers and large transit providers do not advertise a large number of prefixes. Thus, most prefixes are advertised by networks that are not in the "core."

Irrespective of the business type of ASs, there is always a large fraction of *lonely* prefixes, then *delegated* and *deaggregated* follow, and finally *top* prefixes represent the smallest fraction of prefixes. Still there are differences. For example, STPs show a bit more deaggregation than other AS types; EC networks have slightly more delegated prefixes than the others. CAHPs and LTPs do not have so many deaggregated prefixes, but this is only due to the fact that these AS types advertise less prefixes, not because they have a smaller percentage of deaggregated prefixes. However, overall the proportions between our four prefix types are similar for every AS type.

This finding is important as it shows that there are no fundamental differences in how ASs of different business types announce prefixes in the Internet. In particular, it disproves the widespread belief that edge networks, such as EC or CAHP domains, tend to deaggregate more than other ASs. In this respect, the predicted growth of the Internet at the edge [16] does not worsen address space deaggregation as is widely thought.

IV. ALLOCATED VS. OBSERVED PREFIXES

In the preceding two sections we have studied the fragmentation of allocated address space and the deaggregation of announced prefixes as separate phenomena. In this section we aim to understand the combined effect of these two aspects on the routing tables in today's Internet. Consequently, we now investigate to what degree individual ASs deaggregate the allocations they obtained from the registries, possibly announcing multiple prefixes for each allocated address block. Section IV-A defines categories that characterize the use of allocated address blocks while Section IV-B presents the results. Then, Section IV-C studies the ratio between the number of announced prefixes and the number of allocated address blocks for each AS,

A. *Use of allocated blocks – classification*

We again rely on the official allocation files [10], already used in Section II. In addition, we obtain routing table snapshots for AS 3356 for every November 1st between the year 2001 and 2008[3]. For each allocated address block, we determine all the prefixes from the routing tables which overlap with that allocated block. If the routing table contains at least one such prefix, we classify the allocated block into one of the following five categories:

- *Only root*: The complete allocated address block (called "root prefix") is announced and nothing else.
- *root/MS-complete*: The root prefix and at least two sub-prefixes are announced. The set of all sub-prefixes spans the whole root prefix.
- *root/MS-incomplete*: The root prefix and at least one sub-prefix is announced. Together, the set of announced sub-prefixes does not cover the root prefix.
- *no root/MS-complete*: The root prefix is not announced. However, there are at least two sub-prefixes which together cover the complete root prefix.
- *no root/MS-incomplete*: The root prefix is not announced. There is at least one sub-prefix. Taking all sub-prefixes together, they do not cover the complete root prefix.

B. *Is Deaggregation Popularity Increasing?*

To study the use of allocated blocks we consider the time between November of a specific year and November of the previous year: for each address block that has been allocated during that time, we compare against BGP routing tables

---

[3] This is the same data source as used in Section III, but we only rely on the observation point in AS 3356.

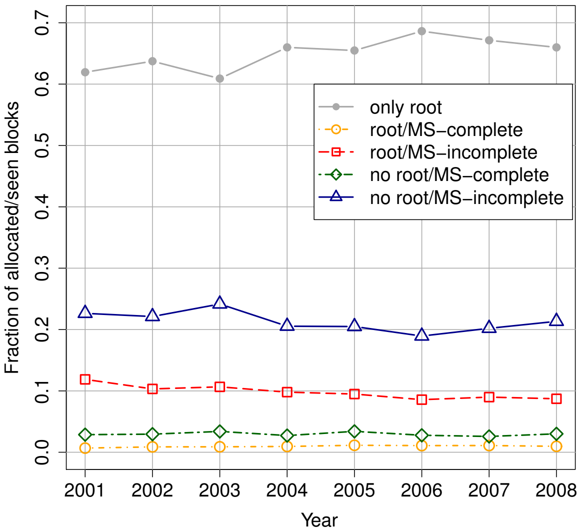


Fig. 4. Use of address blocks allocated by year.

and classify the allocated block according to the categories described in Section IV-A. Figure 4 shows the breakdown of newly allocated blocks in fractions.

We observe that the majority (more than 60%) of address blocks that have been newly allocated between 2001 and 2008 fall into the "only root" class. Then the "no root/MS-incomplete" class follows. All the curves remain roughly constant over time. Apparently, there are no significant changes in how newly allocated blocks are used by the people requesting them. Moreover, Figure 4 even reveals that, in recent years, we observe a slightly decreasing popularity of address space deaggregation. In fact, the percentage of "only root" allocations is slightly increasing since 2001. Here we point out that "only root" address blocks are almost exclusively announced as "lonely prefixes" when comparing the classification of allocated blocks (Section IV-A) with the classification of prefixes (Section III-A). Overall, from a general perspective, we cannot confirm a trend towards increasing fragmentation, with more and more sub-prefixes being announced for individual allocated blocks.

C. "Bad" Guys

Results in the previous section show that most allocations are announced exactly as received by RIRs. This must not mislead us into thinking that address block deaggregation is not a problem: even if the percentage of deaggregated blocks is small, the corresponding number of prefixes might be relevant. In order to take this into account, we define the *deaggregation factor* of an AS to be the ratio between the number of announced prefixes and the number of allocated address blocks. We investigate if there exist ASs where the number of announced prefixes significantly exceeds the number of address blocks that have been allocated to it. Although we call such ASs "bad guys", it is not our objective to blame such ASs for the scaling problems of the routing system. After all, there may be legitimate reasons for them to deaggregate their available address space (see Section V).

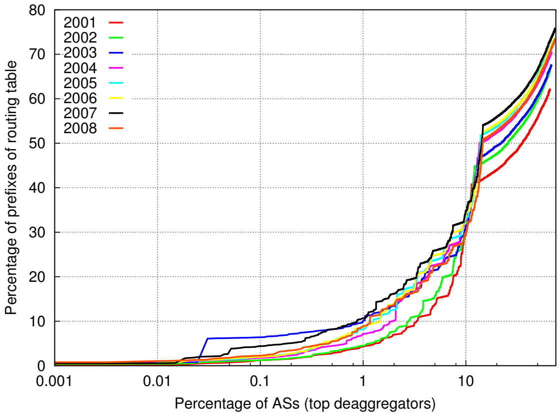


Fig. 6. Percentage of the IP prefixes injected by heavy deaggregators. A data point (x,y) means that the x% top deaggregator ASs are responsible for injecting y% of the prefixes in the routing table.

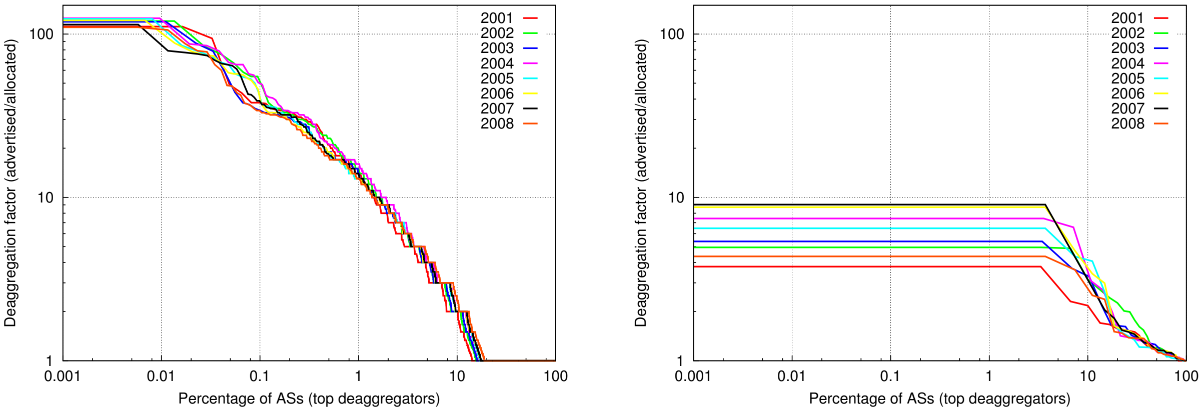
Rather, our goal is to quantify how much "bad guys" contribute to the routing load in today's Internet and to check if the number and impact of "bad guys" has gone for the worse or not over time.

Figure 5 shows the deaggregation factor for Enterprise Customers (EC) and Large Transit Providers (LTP), respectively. For a point (x,y) on the curve, the x value on the logarithmic x-axis indicates the percentage of ASs that have a deaggregation factor of more than y. As shown in Figure 5(a), some EC domains do deaggregate their allocated space by a factor of up to one hundred. About 1% of EC ASs split each allocation into more than 10 more specifics on average. More than 10% of all ASs deaggregate their allocations (i.e., they have deaggregation factor greater than 1), with no notable trends since 2001.

Figure 5(b) shows the completely different behavior of LTP ASs: most of them advertise more than one prefix per allocation on average, but their deaggregation factor never exceeds 10. We run the same analysis on Small Transit Providers (STPs) and on Content, Access and Hosting Providers (CAHPs). We find that STPs behave similarly to LTPs, while CAHPs behave more like ECs.

According to our findings, there exist only a few "bad guys" amongst EC and CAHP edge networks which deaggregate their allocated address space a lot. Yet, we point out that this situation has been the case since 2001, and probably even earlier. A limited fraction of "bad guys" has existed since quite some time and their total fraction has not significantly increased despite the natural growth of the Internet at the edge [16]. Overall, the vast majority of edge ASs refrain from deaggregation, with more than 80% of EC domains in the Internet advertising their allocations exactly as received by routing registries.

Figure 5 helps to identify heavy deaggregators, namely our "bad guys". But what is their impact on routing table size? To this end, Figure 6 illustrates the contribution of the top deaggregators in terms of the prefixes they contribute to the routing tables in today's Internet. We sort all ASs by



(a) Enterprise Customers (b) Large Transit Providers

Fig. 5. Deaggregation factor of ASs, differentiated by business type.

decreasing deaggregation factor (x-axis) and display on the y-axis the cumulative fraction of prefixes for which the top $x\%$ deaggregator ASs are responsible[4].

We observe a skewed distribution: The 1% top deaggregators inject almost 10% of the prefixes in the Internet. ASs that deaggregate their address space, which account for more than 10% of ASs, are responsible for almost half of the routing table size. Overall, such top deaggregators make up a considerable fraction of the routing table entries. This justifies the alarmist discourses about prefix deaggregation. However, we point out that the fraction of prefixes announced by heavy deaggregators has not strongly increased since 2001.

Our analysis contradicts the widespread belief that huge prefix deaggregation is more dominant today than it has been in the past: over time, address block deaggregation is not becoming more popular (Figures 4 and 5) and heavy deaggregators are not inflating routing tables at a growing pace (Figure 6). We conjecture that the growth of the Internet topology is actually the main factor that is the cause of the widespread belief that the routing system is going worse than in the past.

It may be considered unacceptable that a limited fraction of all ASs inflates the routing tables, damaging the whole Internet. If we are to mitigate the impact of such ASs, we believe that the Internet community should discuss the reasons why these ASs are relying on deaggregation, and find alternative ways for those ASs to do what they want without such an impact on the routing system.

V. DEAGGREGATION AND TRAFFIC ENGINEERING

So far, we have been mainly concerned with understanding the extent to which prefix deaggregation is applied and whether deaggregation has become more popular or more intense recently. Many researchers consider traffic engineering as the main driver that causes network operators to split available address blocks into multiple prefixes. For example, in order to influence where traffic enters its network, an ISP can announce different sets of more specific prefixes to different upstreams.

In this section, we explore how prevalent such traffic engineering actually is. An important contribution of our work is to infer information about traffic engineering practices from observed routing data and to correlate this information with observed deaggregation, as discussed in the preceding sections. Contrary to previous sections, we also consider AS path information that is announced together with prefixes. In Section V-A we propose a classification scheme that we use in Section V-B to estimate the prevalence of traffic engineering in the current Internet.

A. BGP Traffic Engineering – Classification

While traffic engineering in its general sense is about how to optimize the performance of a network, this section focuses on mechanisms that BGP provides to control route selection and thus traffic flow to certain prefixes. In this regard, BGP routing data allow us to identify two different types of traffic engineering activities: we refer to them as *AS-path prepending* and *scoped advertisements*.

AS-path prepending adds consecutive repetitions of the same AS number in the AS-path attribute. The goal is to make BGP announcements less appealing for the BGP decision process, which prefers routes with shorter AS-path length. AS path prepending provides some degree of control over inbound traffic, since an AS A can induce a peering AS B to prefer other shorter routes over the prepended one that is announced by A to B. We only consider prepending performed by the last two distinct AS hops on the AS path. Many origin ASs either perform prepending themselves, or ask their upstream providers to do it on their behalf through BGP communities [20]. Prepending made by other ASs on the path is unlikely to be related to traffic engineering activities at the origin AS.

Another traffic engineering technique are *scoped advertisements*. Here ASs selectively announce distinct prefixes to different upstream providers. For example, some prefixes are not advertised to selected upstreams, such that traffic destined to those prefixes cannot be received via those upstreams. For more details about BGP-based traffic engineering, see [21].

---

[4] The y-axis does not span the 100% of prefixes since the deaggregation factor metric does not capture *delegated* prefixes.

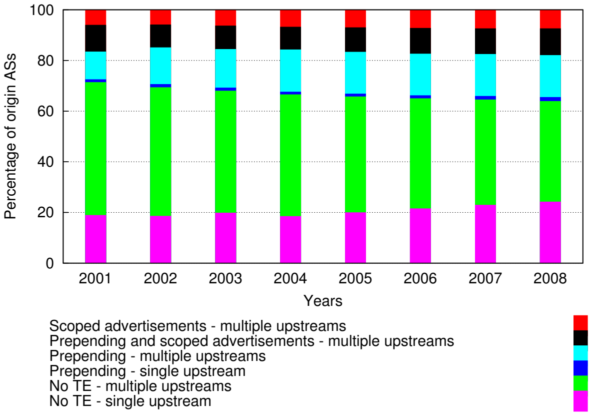
We are interested in understanding which ASs exhibit a behavior that can be related to some form of BGP traffic engineering. To this end, we investigate on a per-prefix basis to what degree the two traffic engineering techniques described above are applied in the current Internet. The data source we rely on are the same BGP routing tables as used in Section IV, but we consider all RouteViews RIBs, not only the view from AS 3356. Thus, our analysis is not restricted to only two vantage points as in [1]. Yet, visibility provided by observed BGP data is inherently limited [22]. To cope with this, we rely on a threshold that considers that some paths may not be visible from the data. Limited visibility is especially relevant for *scoped advertisements*, so our threshold requires that we see scoping across at least 80% of the prefixes of an AS before concluding that scoping is actually used. Overall, our analysis provides a lower bound on the amount of actual traffic engineering performed in the Internet.

The classification scheme that we propose is along three dimensions. In addition to *scoped advertisements* and *AS-path prepending*, we check if an AS has a *single upstream* or *multiple upstream* providers according to the AS path information from our BGP routing information. Altogether, this results in the following six categories:

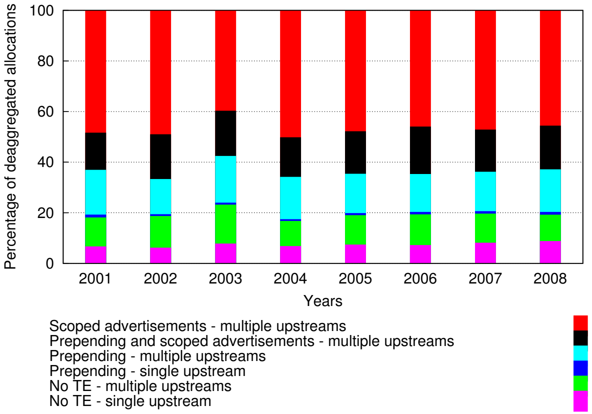
(i) *No TE - single upstream*: no AS-path prepending nor scoped advertisements are identified and the AS has a single visible upstream.
(ii) *No TE - multiple upstreams*: no AS-path prepending nor scoped advertisements are identified, and the AS has multiple visible upstreams.
(iii) *Prepending - single upstream*: some of the prefixes originated by the AS are prepended, no scoped advertisements are identified, and the AS has a single visible upstream.
(iv) *Prepending - multiple upstreams*: some of the prefixes originated by the AS are prepended, no scoped advertisements are identified, and the AS has multiple visible upstreams.
(v) *Prepending and scoped advertisements - multiple upstreams*: some of the prefixes originated by the AS are prepended and scoped advertisements are identified.
(vi) *Scoped advertisements - multiple upstreams*: no AS-path prepending is identified and scoped advertisements are identified.

B. Prevalence of BGP Traffic Engineering

Figure 7 quantifies the extent of traffic engineering practices we can identify from our BGP data. Each origin AS is assigned to one of the six traffic engineering classes from above. While Figure 7(a) considers the evolution over time for *all* origin ASs, Figure 7(b) only takes into account those ASs that advertise *deaggregated* prefixes (according to the classification in Section III-A).



(a) All origin ASs



(b) Deaggregating origin ASs

Fig. 7. Quantifying BGP traffic engineering in observed routing data.

As shown by Figure 7(a), approximately 20% of origin ASs appear to be single homed. Despite some 80% of multi-homed origin ASs, 50% of all origin ASs are multi-homed and do not show evidence of applying any of the BGP traffic engineering techniques introduced above. On the contrary, 30% of all ASs fall into the categories related to traffic engineering, suggesting that BGP-based traffic engineering plays a non-negligible role in the Internet.

Contrary to [1] we take into account the evolution since 2001. Here, we observe that the proportion of origin ASs that apply BGP deaggregation has only slightly increased from 11.3% in 2001 to 14.6% in 2008. Besides, we observe more ASs with single upstreams and without traffic engineering, possibly reflecting the growth of the Internet at the edge [16]. Overall, our findings contradict the widespread belief that the importance of BGP traffic engineering has massively increased and threatens to collapse the current routing system.

Given the results of the previous sections, we investigate whether there is any correlation between prefix deaggregation and traffic engineering. For example, is it true that *deaggregated* prefixes are deaggregated for reasons of traffic engineering, e.g., to influence where traffic enters the network? The approach we take is to study how much traffic engineering is related to the different classes of prefixes defined in Section III-A. For this purpose, we generate in Figure 7(b) the same plot as in Figure 7(a), but only consider those prefixes that are classified as *deaggregated* in Section III-A.

Comparing the two plots of Figure 7, it becomes evident that there is a correlation between prefix deaggregation and traffic engineering. Around 80% of the ASs for which we see *deaggregated* prefixes in our routing tables fall into the traffic engineering categories. This percentage is less than 40% for

2008 if we consider *all* ASs irrespective of the type of prefixes that they originate, see Figure 7(a).

Our analysis is not limited to *deaggregated* prefixes only. The 40% of *lonely* prefixes (see Section III-A) are apparently not affected by BGP traffic engineering as defined in this section. As expected, we do not observe scoped advertisements for delegated prefixes, only prepending. This supports our classification as scoped advertisements are expected to be strongly related to deaggregation.

To sum up, given the differences in the two plots of Figure 7, we conclude that traffic engineering is a significant driver for address space deaggregation. However, based on the evolution since 2001, we cannot identify any sudden shift towards a more aggressive and widespread use of prefix deaggregation in order to achieve traffic engineering. This suggests that it is not an increasing demand for traffic engineering that inflates routing tables: rather, inflation is a consequence of the topological growth of the Internet combined with the lack of support for traffic engineering in the current routing system.

VI. BGP DYNAMICS

Large routing tables in the Internet are not only dangerous *per se*, but also because they increase the likelihood that the best routes frequently need to be recomputed. For this reason, we now investigate the dynamics of routing information. Past work has claimed the existence of highly unstable edge networks [6], [23]. Therefore, we aim at understanding whether or not there are specific classes of ASs or prefixes that generate a disproportionate number of BGP messages. To this end, we start in Section VI-A by studying the relationship of the business class of originating ASs and update dynamics. In a similar way, we explore whether specific prefix classes (see Section III) can be held responsible for a significant fraction of BGP update churn (Section VI-B). Finally, in Section VI-C we focus our analysis on subsets of prefixes that are responsible for a high number of observed BGP updates and summarize our findings in Section VI-D.

A. BGP update rate by business type of originating AS

Based on the classification of ASs by business type such as LTP, STP, CAHP, and EC (see Section III and [16]), we compute the number of updates seen for all prefixes of each class at a given observation point. The data we use are taken from RIPE collectors [8] *rrc00*, *rrc01* and *rrc03* and contain routing updates recorded at 180 observation points in more than 142 ASs. We analyze BGP updates for the first week in November of the years 2001 to 2008, but only present the results for 2008, as the findings for other years are very similar.

Figure 8 is generated as follows: We sort the observation points by the number of prefixes they receive along the *x* axis. For each observation point, we classify BGP prefix updates into four categories (EC, STP, CAHP and LTP) according to the business type of the AS which originates the BGP update. The *y*-value represents the total number of updates collected at an observation point.

Figure 8 reveals that a large fraction of the updates (for most observation points, more than 60%) are due to prefixes originated by enterprise networks (EC). This number is lower for STP (around 30%). CAHP and LTP networks are responsible for a small share, around 5% each. Unfortunately, not all AS types are equally represented in publicly available data sets. Most observation points are inside CAHP ASs, only 3 observation points are inside enterprise networks. However, further analysis confirms that, irrespective of the location of the observation point, most observed BGP updates are originated by enterprise networks and small transit providers.

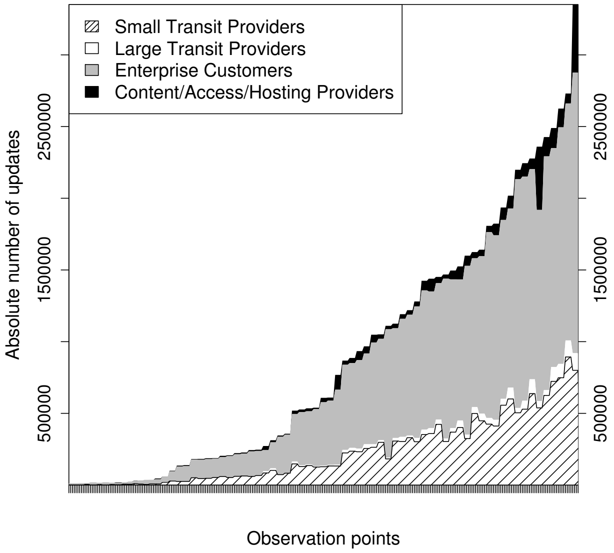


Fig. 8. BGP updates by the four AS classes of [16].

Since Internet domains that are either EC or CAHP belong to the edge of the Internet, people may argue that the edge of the Internet should be blamed for the scaling issues with respect to BGP dynamics. However, we believe that such statements over-simplify the root cause of the high BGP update rates. After all, EC and CAHP networks together make up 95% of the total number of ASs in today's Internet and this number is continuously increasing. Obviously, prefixes originated by these two types of ASs contribute considerably to BGP update churn due to their sheer number, but not necessarily due to the fact that such prefixes are more likely to experience changes in routing policies or are subject to more aggressive traffic engineering practices. With respect to the topological growth in terms of ASs at the edge, core-edge separation approaches [24] for the future Internet, if they come to pass, can be hoped to contain the increase in update rates.

B. BGP update rates by prefix class

We now rely on the same BGP update trace as in Section VI-A, but study the breakdown of the number of BGP updates according to the four prefix classes introduced in Section III: *lonely*, *delegated*, *deaggregated* and *top*. Our goal is to understand to what extent specific prefix classes, e.g., *delegated* prefixes used by stub networks, are more frequently involved in update churn than others.

Figure 9 illustrates the breakdown according to our four prefix types in a manner similar to Figure 8: The observation

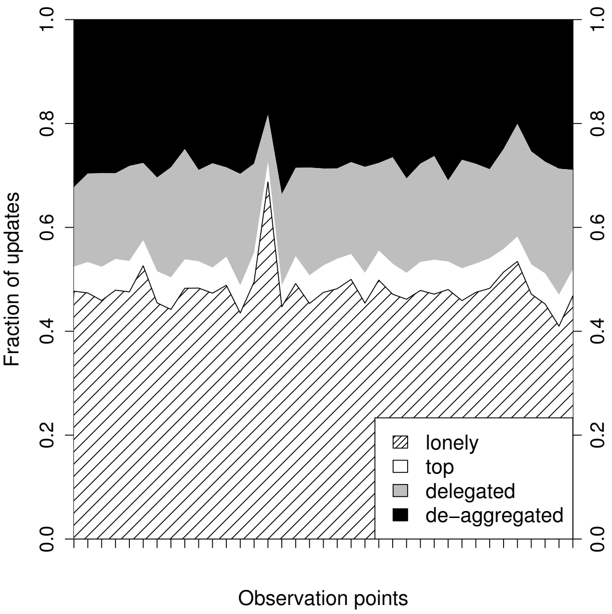


Fig. 9. Fraction of BGP updates by prefix type.

points are aligned along the x axis, while the y-value represents updates collected at an observation point. In order to directly compare BGP updates with the data shown in Figure 3, we normalize the number of updates for each type by the total number of updates seen at the same observation point. To avoid biases in the breakdown, we filtered out those observation points which do not advertise the full BGP routing table.

We find that slightly less than 50% of BGP updates occur for *lonely* prefixes, while *delegated* and *deaggregated* get roughly 20% and 27%, respectively. Overall, the fraction of BGP updates for each prefix class is strongly related to the fraction of prefixes in each class (see Figure 3). Thus, no class is responsible for a disproportionate number of updates. Contrary to our expectations, *deaggregated* prefixes do not generate more than their share (see Figure 3) of BGP updates. At a macroscopic level, deaggregated prefixes behave similarly as the other types of prefixes.

However, we note that, if prefix aggregation were enforced, for example by aggregating some of the *delegated* and *deaggregated* prefixes into their least specifics (*top* prefixes), BGP updates could be reduced by up to 50%. Recall from Section III that prefix aggregation would also shrink the routing table by a similar factor. One way to perform aggregation of deaggregated or delegated prefixes is to tag such routes with BGP communities [20], [25] that tell the upstream ASs whether the prefix can or should be aggregated into a less specific entry of the BGP routing table. Given the widespread usage of BGP communities, the impact of such a technique can be large. However, its success depends also very much on the usage of route tagging by edge networks which may be unwilling to do so. And we must take into account that aggregation is basically just not done in the commercial Internet.

C. Prefixes with high update rates

We now turn to a more detailed study of those prefixes for which we can observe a large number of updates within a certain time period. Geoff Huston has repeatedly pointed out that a few prefixes are responsible for a large fraction of BGP updates [2]. To confirm this, we rely on the same BGP update data set as in Section VI-A, and compute the number of updates observed for each prefix. For each prefix, we consider the median number of updates observed across our observation points, in order to avoid biases due to outlying observation points.

Similarly to the results in [2], we find that 98% of prefixes undergo less than one hundred updates during the considered one month period, while a tiny fraction of prefixes are responsible for a huge number of updates.

In the previous section we observed that, when looking at all prefixes, there is no specific class of ASs that announce a disproportionate number of prefixes. An interesting question is whether this property holds also for the subset of prefixes that account for a large number of routing updates. Again, we rely on the median number of observed BGP updates across our observation points for each prefix. We rank the prefixes by decreasing values of the median number of observed updates, and select the subsets corresponding to the top $x\%$ of the ranked prefixes, for $x \in \{0.1, 0.25, 0.5, 1, 2.5, 5, 10, 25, 50\}$. For each subset, we break observed prefix updates down into classes, according to the deaggregation type of the prefix (see Section III-A).

In Figure 10 values at $x = 100$ represent the breakdown of BGP updates into deaggregation classes when considering all prefixes. Values at $x = 1$ represent the breakdown when only considering the 1% of prefixes that generated the highest number of updates.

According to Figure 10, even when restricting the analysis to the prefixes with the highest churn, there are only negligible differences among different deaggregation classes. Moreover, the fraction of BGP updates each class is responsible for is highly correlated with the portion of the routing table that class accounts for (see Figure 3). In particular, we point out that *deaggregated* prefixes do not generate a disproportionate number of BGP updates, even when we consider the small subset of prefixes that are responsible for a huge amount of routing updates.

To check that the way prefixes are classified does not bias the results, we repeated the analysis of this section with the classification introduced in Section IV-A. We do not show these results due to space limitations, but found no statistically significant difference in the behavior of the different prefix classes. Only the “root/MS-complete” class has lower worst-case behavior compared to other classes. Otherwise, the distributions of the number of BGP updates per prefix class are extremely similar.

D. Summary

Our analysis of BGP updates confirms the findings from previous sections: we cannot claim that a specific class of

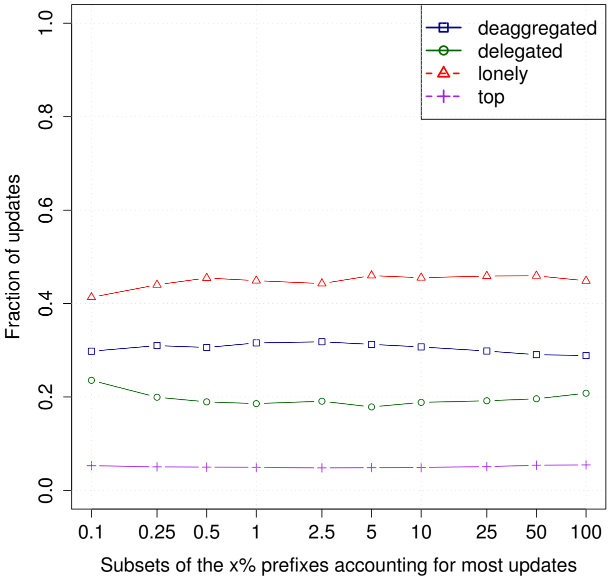


Fig. 10. Prefixes ranked by frequency of updates — Fraction of updates by class. Each point on the *x*-axis corresponds to the subset of the *x*% prefixes for which most updates are observed. The *y*-axis shows the fraction of observed updates by class.

prefixes misbehaves more than others. In particular, deaggregated prefixes cannot be blamed for putting more stress than their fair share on the routing system.

VII. RELATED WORK

A large number of research projects [26]–[29] which are concerned with routing architectures for the future Internet, motivate their work with alarming numbers of the growth of routing table size or rates of routing updates. In general such numbers can be obtained from various sources, including [2], [4], [8], [9], [30]. Probably, the most prominent source of information is the Potaroo web page [2], [4]. Based on results from [2], Huston concluded that the Internet grows “at a rate which has, at a minimum, doubled in size each year” [31].

In the past, many publications studied the evolution of the Internet topology, mainly with a focus on the AS-level structure, e.g., [16], [22], [32]. One general insight, relevant for this work, is that the Internet is predominantly growing at the “edge”, i.e. the number of enterprise, access, or content networks is increasing faster than the core.

Moreover, efforts have been made to quantify the growth of routing tables, to understand the underlying causes, and to make predictions for the future. Bu et al. [3] explore the degree to which factors such as multi-homing or address fragmentation contribute to routing table size. While their study does not include data on either allocations or routing dynamics, others [33] explore how IPv4 address allocation practices affect the BGP table growth. Meng et al. [1] take the step of correlating information about allocated blocks with what is actually announced in BGP routing tables. Compared to our classification of Section III-A and Section IV-A, their analysis is more coarse-grained, since [1] solely distinguishes between *covered* and *covering* prefixes. Finally, work has been done to explore BGP routing dynamics, e.g., [6], and to identify prefixes that significantly contribute to routing churn [23]. Despite the importance of interdomain traffic engineering, to the best of our knowledge we are the first to quantify the popularity of AS-path prepending and scoped advertisements based on observed routing data.

According to recent findings [34] certain proportions between different measures are invariant in spite of the ongoing growth of the Internet. For example, an estimation of the lower bound on the number of active public IP addresses appears to be consistent with a square law relationship with both the number routing table entries and the number of ASs. The findings of this paper are similar in that certain patterns in the Internet, such as usage of allocated blocks, have not and do not significantly change in the long term.

VIII. CONCLUSION

There is widespread belief in a high and fast growing number of ASs that deaggregate prefixes, e.g., for the purpose of traffic engineering [1]. Moreover, researchers often blame specific classes of ASs for generating a disproportionate amount of BGP updates.

Our primary objective is to challenge such widespread assumptions (“myths”) and *not* just to confirm previous findings [1]–[3]. Surprisingly, we find severe discrepancies between existing myths and reality. According to our results, there is no trend towards more aggressive prefix deaggregation or traffic engineering over time. With respect to update dynamics, deaggregated prefixes generally do not generate a disproportionate number of BGP updates with respect to their share of the BGP routing table.

We stress that our work should not be used as an argument to deny that there exists a problem with address space deaggregation. Instead, what our results reveal is that the global impact of “bad guys” on Internet address space deaggregation and BGP churn has not changed for the worse in recent years. These problems have been present for a long time and have not become more prevalent recently. This is a far more positive picture of the behavior of IPv4 address space and its dynamics compared to the alarming discourses typically heard [1], [2], [4].

Our findings indicate that the growth of routing table and BGP dynamics is due to the growing edge of the Internet. In this light, core and edge separation, as advocated by clean-slate approaches [7], could be an approach worth exploring to improve routing dynamics.

ACKNOWLEDGMENT

We are particularly grateful to Randy Bush, Pierre Francois, and Olaf Maennel for insightful comments and feedback about earlier versions of this work.

The research results presented herein are partly funded by a grant from Deutsche Telekom Laboratories, from Trilogy (`http://www.trilogy-project.eu`), a research project (ICT-216372) partially funded by the European Community under its Seventh Framework Programme, as well as from G-lab (`http://www.german-lab.de`), a research project (support code 01 BK 0805) funded by the Federal Ministry of Education and Research of the Federal Republic of Germany. The views expressed here are those of the author(s) only. The European Commission is not liable for any use that may be made of the information in this document.

REFERENCES

[1] X. Meng, Z. Xu, B. Zhang, G. Huston, S. Lu, and L. Zhang, "IPv4 Address Allocation and the BGP Routing Table Evolution," *ACM CCR*, vol. 35, no. 1, pp. 71–80, 2005.

[2] G. Huston, "BGP Routing Table Analysis Reports," http://bgp.potaroo.net/.

[3] T. Bu, L. Gao, and D. Towsley, "On Characterizing BGP Routing Table Growth," *Computer Networks*, vol. 45, no. 1, pp. 45–54, 2004.

[4] T. Bates, P. Smith, and G. Huston, "CIDR Report," `http://www.cidr-report.org/as2.0/`.

[5] C. Labovitz, G. Malan, and F. Jahanian, "Origins of Internet Routing Instability," in *Proc. IEEE INFOCOM*, 1999.

[6] J. Li, M. Guidero, Z. Wu, E. Purpus, and T. Ehrenkranz, "BGP Routing Dynamics Revisited," *ACM CCR*, vol. 37, no. 2, pp. 5–16, 2007.

[7] A. Feldmann, "Internet Clean-Slate Design: What and Why?," *ACM CCR*, vol. 37, no. 3, pp. 59–64, 2007.

[8] "RIPE Routing Information Service," `http://www.ripe.net/ris/`.

[9] "University of Oregon Route Views Project," `http://www.routeviews.org/`.

[10] "Regional Internet Registries allocations statistics files," `ftp://ftp.ripe.net/pub/stats`.

[11] K. Hubbard, M. Kosters, D. Conrad, D. Karrenberg, and J. Postel, "Internet Registry IP Allocation Guidelines," IETF RFC2050, 2000.

[12] T. Hain, "A Pragmatic Report on IPv4 Address Space Consumption," *The Internet Protocol Journal*, vol. 8, no. 3, 2005.

[13] Y. Rekhter and T. Li, "An Architecture for IP Address Allocation with CIDR," IETF RFC1518, September 1993.

[14] V. Fuller, T. Li, J. Yu, and K. Varadhan, "Classless Inter-Domain Routing (CIDR): an Address Assignment and Aggregation Strategy," IETF RFC1519, September 1993.

[15] P. Smith, J. Martin, and R. Bush, "Use of Final /8," `http://www.apnic.net/policy/proposals/prop-062-v002.html`.

[16] A. Dhamdhere and C. Dovrolis, "Ten Years in the Evolution of the Internet Ecosystem," in *Proc. ACM IMC*, 2008, pp. 183–196.

[17] L. Subramanian, S. Agarwal, J. Rexford, and R. Katz, "Characterizing the Internet Hierarchy from Multiple Vantage Points," in *Proc. IEEE INFOCOM*, 2002.

[18] F. Wang and L. Gao, "Inferring and Characterizing Internet Routing Policies," in *Proc. ACM IMC*, 2003.

[19] G. Di Battista, M. Patrignani, and M. Pizzonia, "Computing the Types of the Relationships Between Autonomous Systems," in *Proc. IEEE INFOCOM*, 2003.

[20] B. Donnet and O. Bonaventure, "On BGP Communities," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, 2008.

[21] B. Quoitin, S. Uhlig, C. Pelsser, L. Swinnen, and O.Bonaventure, "Interdomain Traffic Engineering with BGP," *IEEE Communication Magazine*, 2003.

[22] R. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang, "In Search of the Elusive Ground Truth: The Internet's AS-level Connectivity Structure," in *Proc. ACM SIGMETRICS*, 2008.

[23] B. Zhang R. Oliveira, R. Izhak-Ratzin and L. Zhang, "Measurement of Highly Active Prefixes in BGP," in *Proc. IEEE GLOBECOM*, 2005.

[24] D. Jen, M. Meisel, H. Yan, D. Massey, L. Wang, B. Zhang, and L. Zhang, "Towards A New Internet Routing Architecture: Arguments for Separating Edges from Transit Core," in *Proc. ACM Workshop on Hot Topics in Networks*, 2008.

[25] R. Chandra, P. Traina, and T. Li, "BGP Communities Attribute," IETF RFC1997, August 1996.

[26] EU Framework Programme 7 Integrated Project, "Trilogy," `http://www.nets-find.net`.

[27] NSF, "NeTS Initiative, Future Internet Design," `http://www.nets-find.net`.

[28] "Global Environment for Network Innovations," `http://www.geni.net/`.

[29] "Internet Research Task Force – Routing Research Group," `http://tools.ietf.org/group/irtf/trac/wiki/RoutingResearchGroup`.

[30] "Cooperation for Internet Data Analysis (CAIDA)," `http://caida.org/research/topology`.

[31] G. Huston, "Analyzing the Internet BGP Routing Table," *The Internet Protocol Journal*, vol. 4, no. 1, 2002.

[32] R. Oliviera, B. Zhang, and L. Zhang, "Observing the Evolution of Internet AS Topology," in *Proc. ACM SIGCOMM*, 2007.

[33] Z. Xu, X. Meng, L. Zhang, S. Lu, and C. Wittbordt, "Impact of IPv4 Address Space Allocation Practice on BGP Routing Table Growth," in *Proc. of IEEE 18th Annual Workshop on Computer Communications (CCW)*, 2003.

[34] B. Carpenter, "Observed Relationships between Size Measures of the Internet," *ACM CCR*, vol. 39, no. 2, pp. 5–12, 2009.