Putting BGP on the Right Path: Better Performance via Next-Hop Routing

Michael Schapira, Yaping Zhu, and Jennifer Rexford
Princeton University
{ms7,yapingz,jrex}@cs.princeton.edu

ABSTRACT

Today's interdomain routing system does not perform well, because the BGP protocol converges slowly, selects paths without regard for performance, does not support multipath routing, and has numerous security vulnerabilities. Rather than adding mechanisms to an already complex protocol, or redesigning interdomain routing from scratch, we propose making BGP *simpler*, and handling issues such as data-plane performance and security where they belong—*outside* the routing protocol. We propose a transition from today's *path-based routing* (where routing decisions depend on the entire AS-PATH) to *next-hop routing*—a solution that selects and exports routes based *only* on the neighboring domain. Based on theoretical and experimental results, we show that next-hop routing leads to significantly better network performance than path-based routing, and is especially effective at preventing the most serious BGP convergence problems, alongside other advantages (including being more amenable to multipath routing and a reduced attack surface).

1. INTRODUCTION

The Border Gateway Protocol (BGP) stitches a single global Internet together out of smaller networks with diverse policy objectives. BGP is plagued by poor network performance, as well as security vulnerabilities, configuration errors, software bugs, and more. BGP does not react to serious performance and reachability problems in the data plane, which are, in fact, often utterly invisible to the routing protocol. BGP route fluctuations can lead to high packet loss, intermittent loss of connectivity, and increased path latency [1]. Remarkably, almost half of the performance problems with VoIP are the result of BGP route fluctuations [2].

The research and standards communities have proposed numerous ways to fix BGP. These proposals generally fall into two categories: (1) enhancements to BGP (*e.g.*, to improve convergence [3, 4] and security [5, 6]); and (2) alternative architectures that take a "clean-slate approach" to interdomain routing (see [7–10]). However, these proposals face serious practical obstacles. On the one hand, enhancements to BGP add mechanisms to an already complex protocol, which can introduce new problems (*e.g.*, configuration errors, protocol convergence issues, and new attack vectors). On the other hand, alternative architectures are difficult, often even impossible, to deploy incrementally.

We present a new way to fix interdomain routing. We introduce a novel routing architecture called "next-hop routing", which is based on two guiding principles: (1) Rather than extending or replacing BGP, *simplifying* the routing protocol by *constraining* how routes are selected and exported; and (2) Handling issues such as data-plane performance and security where they belong—*outside* the routing protocol. We argue that next-hop routing is sufficiently expressive to realize network operators' goals. Our analysis and simulations show that next-hop routing significantly improves performance, without compromising other global objectives.

1.1 Simplicity through next-hop routing

Many of BGP's performance problems relate to how ASes base routing decisions on the sequence of ASes along a path. To avoid the "count-to-infinity" problems that plagued earlier distance-vector protocols, BGP is a *path-vector* protocol that includes the AS-PATH attribute in each route announcement. This allows an AS to quickly detect and avoid paths that contain loops.

However, the AS-PATH is increasingly used for far more than loop detection and, in particular, network operators apply complex regular expressions to the AS-PATH to express local preferences over routes. Unfortunately, routing policies that consider the entire AS-PATH can lead to long convergence time, and can even result in overall protocol divergence. Operational experience and measurement studies indicate that BGP convergence is too slow for interactive applications, *e.g.*, VoIP, multi-player games, and financial transactions. Techniques for reducing convergence delay introduce additional timers and configuration options to the router software, exacerbating an already complex system.

We argue that such reliance on the AS-PATH is more a hindrance than a help in achieving good network performance. Instead, we advocate next-hop routing—that an AS rank and export paths based *solely* on the next-hop AS en route to each destination prefix, and apply

a simple "consistent filtering" rule [11] when exporting routes. That is, we relegate the AS-PATH to its traditional role in loop detection. We show, both analytically and experimentally, that next-hop routing achieves significantly better network performance than traditional path-based routing in two important respects:

Better convergence. We show that next-hop routing converges much more quickly to a "stable" routing configuration than conventional BGP. In addition, the convergence process involves much fewer update messages and forwarding changes. Our simulation results establish that next-hop routing is especially effective at preventing the most serious BGP convergence problems. Our results suggest that next-hop routing could allow ASes to disable mechanisms for reducing convergence delay and rate-limiting the sending of update messages.

More amenable to multipath routing. Today's BGP-speaking routers select a single path for each destination, rather than capitalizing on the path diversity in the Internet. We show that next-hop routing naturally supports multipath routing, leading to many benefits, *e.g.*, improved availability, better recovery from failures, load balancing, customized route selection.

1.2 Path-quality monitoring and security

Today's BGP provides network operators with very unsatisfactory means for handling data-plane performance issues and security threats, as reflected in poor network performance and in serious security breaches. We argue that such important objectives should be handled *outside* the routing protocol [12], and propose specific solutions in two key areas:

Performance-driven routing. Today's BGP-speaking routers prefer shorter AS-PATHs as an indirect way to improve end-to-end performance. However, AS-PATH length only loosely correlates with end-to-end propagation delay and does not reflect other performance metrics of interest, *e.g.*, throughput, latency, and loss [13]. Consequently, BGP is unable to cope with serious performance and reachability problems in the data plane. We argue that, to achieve good network performance, route selection should be based on direct observations of path quality, rather than the length of the AS-PATH.

We propose leveraging two mechanisms: end-to-end monitoring and multipath routing (see above). Neither of these mechanisms is provided by today's BGP. Instead, ASes should adapt next-hop rankings, as well as how traffic is split between multiple "next hops," based on path-quality monitoring in the data plane. We describe techniques for stub ASes, online service providers, and transit ISPs to monitor path performance.

Security. BGP is notoriously vulnerable to attacks [6, 14–17]. This weakness of BGP can be exploited for eavesdropping, tampering, packet dropping, and also for economic reasons (*e.g.*, increasing an AS's revenue). Every year or two, a major incident exposes just how incredibly vulnerable BGP is (*e.g.*, [18]). In addition, today's routers have a bewildering array of configuration options for selecting and exporting routes, making configuration errors [19, 20] and software bugs [21–23] common sources of serious Internet outages.

We believe that relying on BGP for data-plane security is misguided. Instead, these guarantees should be assured in other (end-to-end) ways, such as encryption and authentication. We show that next-hop routing significantly reduces BGP's attack surface. Specifically, next-hop routing makes ASes immune to virtually all AS-PATH-based attacks (*e.g.*, path-shortening attacks [17]). Next-hop routing also removes incentives for rational ASes to manipulate the protocol. We also give theoretical evidence that next-hop routing is more resilient to configuration errors.

1.3 Roadmap

We present next-hop routing in Section 2. Since next-hop routing is a broad routing architecture, our evaluation takes many forms, including (i) theoretical analysis (to study convergence, incentive compatibility, and robustness to misconfiguration), (ii) simulation experiments (to measure convergence time), (iii) protocol design (to illustrate simple ways to support multipath routing and path-quality monitoring), and (iv) qualitative arguments (about the reduced attack surface and techniques for traffic engineering).

Section 3 shows that next-hop routing significantly reduces convergence time and router overhead. Section 4 shows how next-hop routing greatly simplifies support for multipath routing. Section 5 shows that next-hop routing reduces the attack surface for BGP and provides incentives for rational ASes to participate honestly in the protocol. We present techniques for traffic management (i.e., path-quality monitoring and traffic engineering) in Section 6. We present related work in Section 7 and conclude in Section 8. Proofs are omitted due to space constraints and are available in an extended version of the paper [24].

2. NEXT-HOP ROUTING ARCHITECTURE

We now present our next-hop routing architecture, which consists of two components: (1) three simple rules that constrain how routes are selected and exported in the BGP decision process; and (2) external mechanisms for performance-driven routing based on end-to-end data-plane monitoring, and for security. Our focus in this section is on the first component—the next-hop routing rules. We discuss security and performance-driven routing in Sections 5 and 6, respectively.

In today's Internet, the bilateral business contracts ASes sign with their *immediate* neighbors play a cru-

cial role in determining their routing policies (see [25] and Appendix B). ASes tend to prefer revenue-generating routes through customers over routes (for which they pay) through their providers, and to avoid carrying traffic between providers. Under next-hop routing, ASes rank and export paths based *solely* on the next-hop AS en route to each destination prefix, but an AS has full control over which immediate neighbors carry its traffic and direct traffic through it. Next-hop routing thus allows ASes sufficient expressiveness to realize their business policies. (Indeed, these common routing practices can be realized with next-hop routing.)

2.1 The existing BGP decision process

BGP-speaking routers typically receive multiple routes to the same destination. There are three steps a BGP router uses to process route advertisements:

1. an **import policy** determines which routes should be filtered (and hence eliminated from consideration) and assigns a "local preference";
2. a **decision process** selects the most desirable route;
3. an **export policy** determines which of the neighbors learn the chosen route.

The decision process consists of an ordered list of attributes across which routes are compared. Here, we focus on the following important steps in the BGP decision process (see [26] for the full list of steps):

- **Prefer higher local preference (LocalPref).** The BGP LocalPref attribute enables operators to express rich, engineering- and business-motivated preferences over routes, thus potentially choosing a longer AS-PATH over a shorter route;
- **Prefer shorter routes.** Ties between routes that share the highest LocalPref are broken in favor of routes with the lowest AS-PATH length;
- **Prioritize old routes over new ones.** This decision step is implemented in some routers [27] but not always enabled by default.

2.2 Next-hop routing rules

We now specify our three next-hop routing rules.

Rule I: Use next-hop rankings. Configure rankings of routes based *only* on the immediate next-hop AS en route to each destination (*e.g.*, to prefer customer-learned routes over provider-learned routes). Ties in the rankings of next-hops are permitted.

Rule II: Prioritize current route. To minimize path exploration, when faced with a choice between the "old" (current) route and an equally-good (in terms of next-hop) new one, re-select the old route.

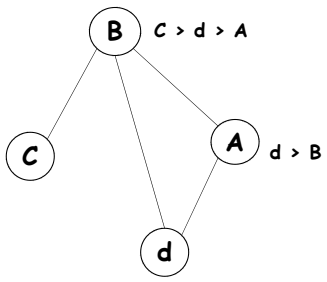


Figure 1: ASes A, B and C wish to send traffic to AS d. All three ASes have next-hop rankings, *e.g.*, A prefers sending traffic directly to d over sending traffic through B. AS B is willing to export route BAd, but not the more preferred route Bd, to C. Hence, if B changes its route from BAd to the Bd, it will stop announcing its route to C and thus disconnect C from the destination d.

Rule III: Consistently export [11]. If a route P is exportable to a neighboring AS i, then so must be all routes that are more highly ranked than P. Intuitively, Consistent Export prevents undesirable phenomena as in Figure 1, where an AS disconnects a neighbor from a destination by selecting a better route for itself.

2.3 Implementing the next-hop routing rules

Network operators can transition to next-hop routing simply by configuring their existing routers in accordance with the next-hop routing rules. This is achieved via the following three actions: (i) applying only next-hop rankings when configuring import policies to assign LocalPref for eBGP-learned routes, (ii) disabling the AS-path length step in the BGP decision process (as supported on most commercial routers), and (iii) selecting eBGP export policies that obey consistent filtering. Thus, an AS can readily deploy next-hop routing without any changes to the underlying equipment and without any support from its neighbors[1].

3. BETTER ROUTING CONVERGENCE

Intuitively, BGP can converge slowly for two main reasons: (1) small and faraway routing changes can lead an AS to select a new next-hop, thus leading to a chain reaction of subsequent routing changes; and (2)

[1]Some ASes cannot freely ignore AS-PATH length in their decision process because of the expectations of their neighbors. In particular, peering contracts often require an AS to export routes of the same AS-PATH length across all peering points with the same neighbor.This "problem" is circular—an AS cannot export paths of inconsistent length because its neighbors are (erroneously, we think) placing too much emphasis on AS-PATH length when selecting routes. If two neighboring ASes agreed to disregard AS-PATH length, these kinds of inconsistencies would no longer matter.

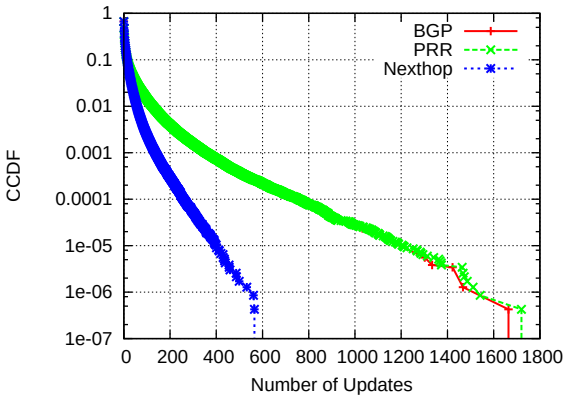


Figure 2: Fraction of non-stub ASes experiencing more than x update messages after a link failure.

inconsistencies between path rankings and route export policies can lead an AS to disconnect other ASes from a destination when selecting a better route for itself, pushing them to seek alternate routes (see Figure 1).

Our next-hop routing rules prevent these scenarios; next-hop rankings guarantee that remote routing changes do not drive an AS to change its next hop; Consistent Export guarantees that when bettering its own route an AS *never* disconnects other ASes from a destination. We now present experimental and theoretical evidence that next-hop routing converges *quickly* to a "stable" routing configuration.

Metrics. We focus on three metrics for measuring performance during the convergence process:

- **# forwarding changes**, that is, the number of times ASes change their choices of next hop during the convergence process (summed across all ASes);
- **# routing changes**, that is, the number of times ASes' routes change during the convergence process (summed across all ASes);
- **# BGP updates** transmitted during the convergence process (summed across all ASes).

3.1 Simulation results

Our experiments show that next-hop routing significantly reduces the number of update messages, routing changes, and forwarding changes, under various network events and vantage points.

3.1.1 Simulation setup

Topology. We use the Cyclops [28] AS-level Internet topology on Jan 1, 2010. Each AS is represented by one router, and the links between ASes are annotated with business relationship. The topology contains a total of 33,976 ASes, of which 4,670 are non-stubs, that is, ASes that have no customers of their own, and the rest are stubs. The topology contains 54,786 customer-provider links and 43,888 peer-peer links.

Protocols. We evaluate BGP (standard decision process), PRR (Prefer Recent Route) [27] where the final tie-breaking step prefers the current best route over new routes, and next-hop routing. For all these protocols, we follow the Gao-Rexford import and export conditions [25] (see Appendix B), where ASes prioritize customer-learned routes over peer-learned routes over provider-learned routes. BGP and PRR prefer shorter routes to longer ones within each category (customer-/peer-/provider-learned). All three protocols obey the Consistent Export requirement. Comparing the three protocols allows us to quantify the importance of next-hop ranking (only in next-hop routing) and prioritizing the current route (in both next-hop ranking and PRR).

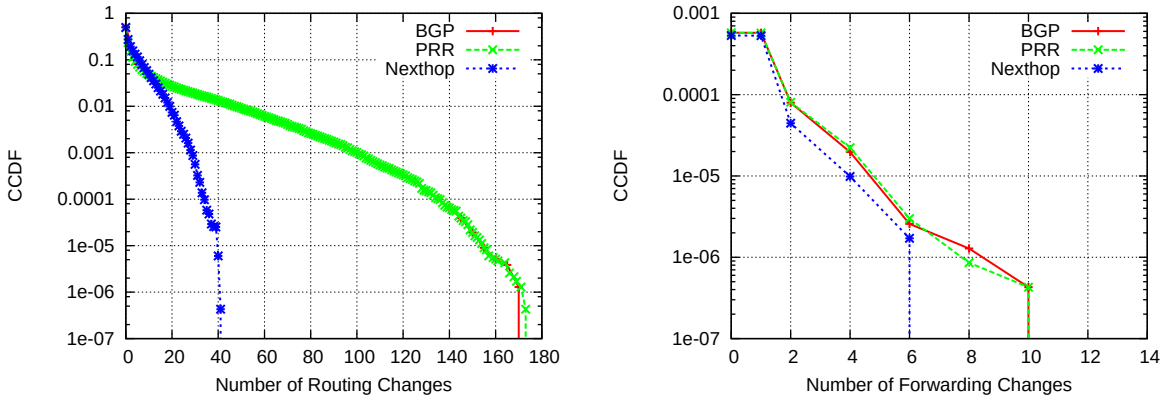
Simulator. We use the C-BGP [29] simulator. Note that C-BGP does not support the MRAI (Minimum Route Advertisement Interval) timer, which rate-limits the updates sent on each session. Although the BGP RFC recommends the default MRAI of 30 seconds [30], router vendors [31] and the IETF [32] advocate using much smaller MRAIs (e.g., a few seconds) or removing the timer entirely, due to the performance requirements of interactive applications. Thus, our simulation results without MRAI settings should be a reasonable estimation of forwarding changes, routing changes, and BGP update messages as the Internet moves away from large MRAI timers. In fact, our experimental and theoretical results below suggest that next-hop routing could allow ASes to remove the MRAI timer entirely. Since C-BGP does not produce accurate estimates of convergence *time*, we do not present results for this metric.

Events. We consider three events: prefix announcement, link failure, and link recovery. We first inject a prefix from a randomly selected multi-homed stub AS, next randomly fail a link between the stub AS and one of its providers, and then recover the failed link. We wait for the routing protocol to converge after each step, before moving on to the next step. For each experiment, we compute all the metrics at selected vantage points (see below). We repeat this experiment for 500 randomly-chosen multi-homed stub ASes.

Vantage points. We observe the number of update messages, routing changes, and forwarding changes from all 4,670 non-stub ASes and from randomly chosen 5,000 stub ASes as vantage points.

3.1.2 Convergence results

We now present our results for the failure of a link connecting an AS to the rest of the Internet. Our results



(a) # routing changes (b) # forwarding changes

Figure 3: Fraction of non-stub ASes experiencing more than x routing changes/forwarding changes after a link failure

for the other events we evaluate are similar.

Comparison: BGP updates at non-stubs. Figure 2 plots the distribution of the number of update messages seen at non-stub ASes. Around one-third of the ASes on the Internet see no routing changes after a link failure. For the ASes which experience a routing changes, the average number of updates received at each AS is 11.7 for BGP, 11.8 for PRR, and 8.0 for next-hop BGP. Since many ASes see little or no effects after any event, we plot the complementary cumulative distribution function (CCDF) to focus on ASes that experience many update messages.

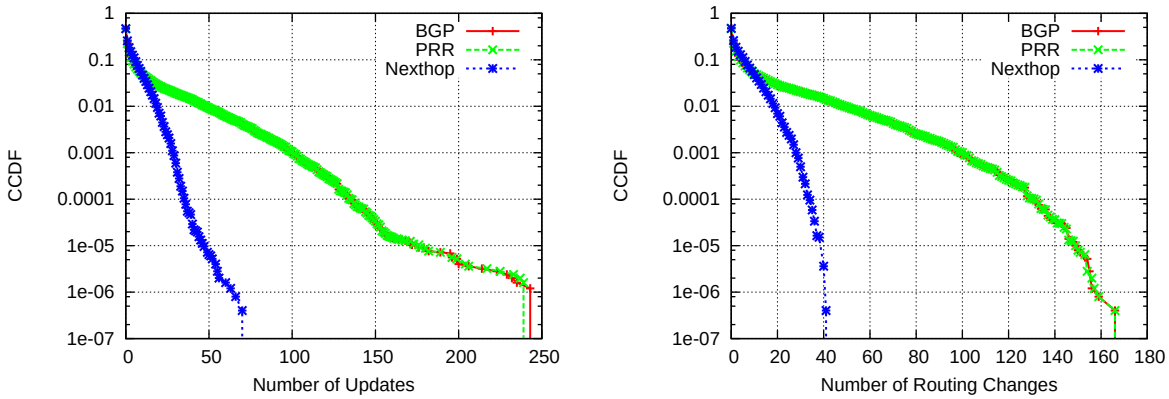
Under BGP (upper curve in Figure 2), some non-stub ASes receive thousands of update messages. Interestingly, the larger, better-connected ASes tend to receive more update messages, presumably because they have many more neighbors exporting routes to them. The most-affected non-stub AS receives 1698 update messages, and the most-affected stub receives 244. PRR slightly reduces this number (middle curve in Figure 2, which mostly overlaps with the BGP curve, except at the tail), by avoiding unnecessary path exploration. Next-hop routing leads to significant improvement (bottom curve). The much smaller number of update messages suggests that next-hop routing would allow ASes to remove the MRAI timer (as is the case in our experiments) without introducing a large number of messages (see also Theorem 3.1 below).

Recall that next-hop routing protocol is composed of three parts: next-hop ranking, prefer old route, and consistent export. While the step of consistent export is applied in all protocols of BGP, PRR, and next-hop, the gap between the curves of BGP and PRR in Figure 2 shows the minor benefits of the PRR step in next-hop routing. By contrast, the gap between the curves of PRR and next-hop in the figure clearly demonstrates the improvement by using next-hop ranking.

Comparison: routing/forwarding changes at non-stubs. Next-hop routing also greatly reduces the number of routing and forwarding changes for non-stub ASes (see Figure 3). For the ASes which experience changes, the average number of routing changes is 4.99 for BGP, 5.03 for PRR, and only 3.95 for next-hop BGP. We get similar results for the average number of forwarding changes: 2.36 for BGP, 2.37 for PRR, and 2.20 for next-hop routing. In fact, none of the non-stub ASes experience more than *six* forwarding changes under next-hop routing for these experiments. In contrast, ASes experience as many as *ten* forwarding changes when preferring a shorter AS-PATH over staying with the same next-hop AS, as in PRR.

Comparison: stubs. Figure 4 plots the distribution of the number of update messages and routing changes as seen at stub ASes. Observe that next-hop routing once again leads to significant improvement over both BGP and PRR.

The 0.1% position. The CCDF plots, while useful for illustrating the diverse experiences of ASes during convergence, are unwieldy for head-to-head comparisons of the protocols across different events and convergence metrics. Yet, with such a skewed distribution, the mean and the median are not especially meaningful. So, we focus on the experience of the AS at the 0.1% position in the CCDF plots, allowing us to focus on the (non-stub) ASes that are affected the most link failures without allowing one outlier AS to bias the results. We summarize the data in bar charts in Figure 5. The bars for "Link Failure" correspond to the $y = 0.001$ position in Figures 3(a) and 4(a), respectively. We do not plot the number of forwarding changes since, for most experiments, the 0.1th-percentile AS experienced at most one forwarding change. We get similar results by comparing the number of update messages and routing changes



(a) # updates (b) # routing changes

Figure 4: Fraction of stub ASes experiencing more than x update messages/routing changes after a link failure

during various events for the stub ASes.

Conclusions. Our results show that next-hop routing offers reductions in the number of update messages and routing changes across a range of network events. These results also illustrate how next-hop routing is especially effective at preventing the most serious convergence problems—where an AS experiences hundreds of routing changes and tens of forwarding changes. This not only reduces the performance disruptions experienced by the data traffic, but also significantly reduces the overhead for disseminating BGP update messages.

3.2 Theoretical results

Our next-hop routing rules imply the existence of a stable state in the network [11], but this does not guarantee BGP convergence to a stable state or that such convergence be fast. Indeed, even in the natural "Gao-Rexford framework" (see Appendix B), that captures common ASes' routing practices and where BGP convergence to a stable state is guaranteed, convergence might involve *exponentially many* forwarding changes, routing changes and BGP updates [33].

We complement our simulation results with theoretical evidence that next-hop routing significantly reduces the number of forwarding changes, routing changes, and update messages. We use the standard model for analyzing BGP dynamics put forth in [34]. See Appendix A for an explanation of the model. We show that in the Gao-Rexford framework, if ASes obey the next-hop routing rules[2], fast ASes, convergence to a stable state is guaranteed (that is, convergence involves a *polynomial*, not exponential, number of forwarding and routing changes and of BGP updates) even in the worst-case. We point out that, to the best of our knowledge, this is the first result to establish fast BGP convergence in the standard model of BGP dynamics.

THEOREM 3.1. *In a network of $|N|$ ASes and $|L|$ communication links, if the Gao-Rexford conditions hold, then convergence of next-hop routing to a stable state requires at most*

- $O(|L|^2)$ *forwarding changes (in total, across all routers);*
- $O(|N||L|^2)$ *routing changes;*
- $O(|L|^3)$ *BGP update messages.*

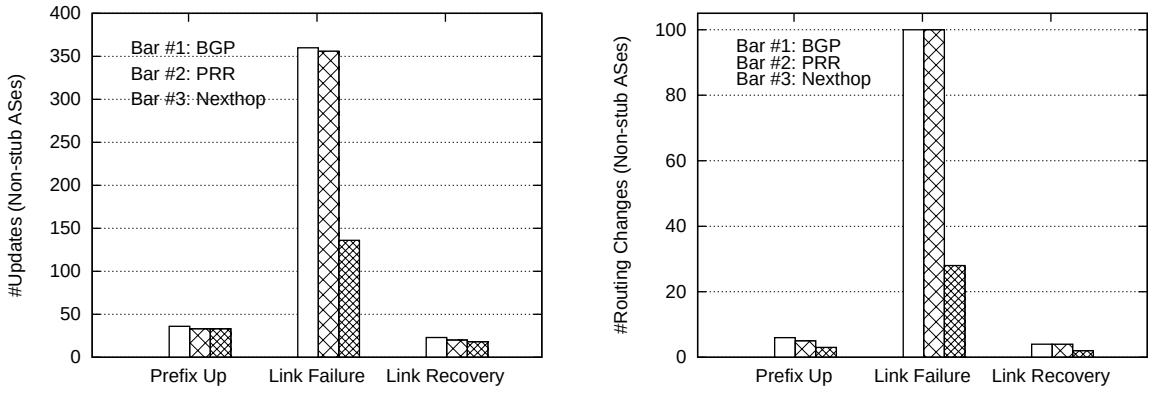
This holds for all timings of router activations/update message arrivals.

The proof appears in the extended version of the paper [24]. We stress that, from a practical perspective, the fact that the above bounds hold for every timing of router activations/update-message arrivals implies that fast convergence with next-hop routing is achieved regardless of the MRAI timer settings and, in particular, even if MRAI timers are removed entirely.

4. MULTIPATH INTERDOMAIN ROUTING

Recently, there has been a surge of interest in multipath interdomain routing (see, *e.g.*, [3, 9, 35–37]). We now discuss how next-hop routing can make multipath routing more scalable and incrementally deployable. We then extend the theoretical results for convergence and incentive compatibility of next-hop routing to a specific multipath routing setting.

[2]The Gao-Rexford conditions do not imply next-hop routing (and vice versa); even if the Preference Condition holds, there are no restrictions on the rankings of routes within each business category (customer-learned, peer-/provider-learned); even if the Export Condition holds, an AS can export a route to a neighboring AS, but not export a "better route" to the same AS. However, by focusing on relationships with immediate neighbors, the Gao-Rexford conditions go naturally together with our next-hop routing rules.



(a) # update messages (b) # routing changes

Figure 5: Number of update messages and routing changes for non-stub ASes

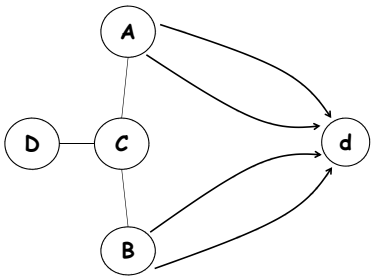


Figure 6: ASes A and B both announce two routes to AS C, and C then announces four routes to AS D, and so on. Clearly, this can result in state explosion.

4.1 Making multipath routing scalable

Unfortunately, naive BGP-based multipath routing schemes can be unscalable, due to the need to disseminate and store multiple routes. For the topology in Figure 6, consider the naive multipath protocol where nodes announce all available routes to neighboring nodes. Observe that ASes A and B, that each have two routes to the destination d, will both announce two routes to AS C. AS C will then announce four routes to AS D, and so on. This can easily result in state explosion and in excessive transmission of update messages.

We show that next-hop routing is more amenable to multipath than path-based routing. The key observation is that under next-hop routing, a node need not learn a neighboring node's multiple paths, but merely learn enough to avoid loops. If AS C in Figure 6 has a next-hop ranking of routes then, to enable C to detect loops, AS A (or B) can merely send C an (unordered) *list* of all the ASes its (multiple) routes traverse. BGP allows the aggregation of routes into one such AS-SET [38], that summarizes the AS-PATH attributes of all the individual routes. Thus, BGP route aggregation, used to keep BGP routing tables manageable in other contexts, can also be used to greatly mitigate the cost of multipath next-hop routing.

Hence, next-hop routing lowers the barrier for making multipath routing a reality. Capitalizing on multipath routing can yield the following benefits:

Availability. Multipath routing increases the likelihood that an AS has at least one working path.

Failure recovery. An AS with multiple next-hops can react *immediately* to a failure in one outgoing link by sending traffic along another (and not wait for the protocol to re-converge).

Load balancing. An AS could have multiple next-hops (with equal local preference) and decide whether and how much traffic to direct through each next-hop (*e.g.*, based on data-plane monitoring). Conventional techniques, such as hashing on fields in the IP header, can ensure that packets of the same flow traverse the same path, to prevent out-of-order packet delivery.

4.2 Neighbor-specific next-hop routing

While under BGP a router is restricted to selecting a single route (for each destination prefix), an AS may want to select different routes for different neighboring ASes for economic reasons (*e.g.*, an ISP could offer different services to customers [37]) or operational reasons (*e.g.*, to increase resiliency to failures [3]). Neighbor-Specific BGP [37] is a recent proposed extension to BGP that allows a router to select a different (single) route for each neighbor, that is, customize route selection on a *per-neighbor* basis. Surprisingly, the additional flexibility in route selection under NS-BGP also improves global network stability [37].

We now show that our theoretical result for conver-

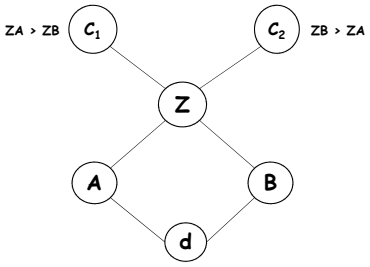


Figure 7: NS-BGP with next-hop routing: C_1 prefers ZA over ZB as its first two hops en route to d. C_1 prefers ZB over ZA as its first two hops en route to d. Z is an ISP running NS-BGP and can thus use the next-hop ranking $A > B$ for choosing routes for C_1, and the next-hop ranking $B > A$ for choosing routes for C_2.

gence of next-hop routing extends to NS-BGP. Specifically, we consider next-hop routing with NS-BGP, that is, routing with NS-BGP such that (1) an AS is restricted to only using next-hop rankings to select routes for neighboring ASes (possibly using different rankings for different neighbors); and (2) an AS consistently exports with respect to each neighboring AS. See Figure 7 for an illustration of next-hop routing with NS-BGP. We prove the analogues of Theorem 3.1 (see Section 3) in this multipath setting. See Appendix B for an explanation of the Gao-Rexford conditions.

THEOREM 4.1. *In a network of $|N|$ ASes and $|L|$ links, if the Gao-Rexford Topology Condition and Export Condition hold, then convergence of next-hop routing with NS-BGP to a stable state involves a total number of at most*

- $O(|L|^2)$ *forwarding changes;*
- $O(|N||L|^2)$ *routing changes;*
- $O(|N||L|^3)$ *BGP update messages.*

This holds for all initial states of the system and for all timings of router activations/update message arrivals.

5. SECURITY

We now present three advantages of next-hop routing over path-based routing with BGP, which render some dangerous attacks against BGP ineffective or less harmful: (1) significantly reduced attack surface; (2) improved resilience to configuration errors; and (3) incentive compatibility. We then explain how additional security mechanisms can be used to secure the routing system from the remaining forms of attacks.

5.1 Reduced attack surface

Some of the most dangerous attacks on BGP are based on "lying" about the length of an AS's path. Indeed, an AS can attract much traffic by announcing a shorter route than that it really has [17], thus launching so called blackhole or interception attacks for tampering, dropping packets, eavesdropping, *etc.*

Under next-hop routing, ASes do not consider the AS-PATH when making routing decisions—beyond the first hop, which cannot be forged!—and so an AS no longer benefits from such attacks. Thus, next-hop routing significantly reduces BGP's attack surface. In fact, the only effective AS-PATH-based attacks on next-hop routing are attacks that trigger BGP's loop detection mechanism (that are clearly also effective against path-based routing with BGP).

5.2 Less vulnerable to configuration errors

We consider a threat model that captures some common configuration errors. Today's AS networks typically comprise commercial routers, and so it is reasonable to consider configuration errors that lead ASes (routers) to rank and export routes in unintended and undesirable ways (as opposed to running new router software). In our threat model, the "adversary" can choose a single AS A and then change that AS's routing policy, that is, A's ranking of routes and route export rules. We impose no restrictions on how the adversary changes A's routing policy.

We restrict our attention to the following simple question: "When can the adversary eliminate all stable states from the network?". (See Appendix A for a model of BGP dynamics and a definition of stable states.) That is, we wish to understand when the adversary can make the network inherently unstable by changing the routing policy of a single AS. While the existence of a stable state does not guarantee protocol convergence to a stable state, understanding when a stable state even exists in the network is an important first step on the path to understanding protocols dynamics in the presence of unwanted behavior. We present the following theorem.

THEOREM 5.1. *If all ASes in a network obey the next-hop routing rules then a stable state is guaranteed to exist in every routing system that can be obtained from changing the routing policy of a single AS.*

5.3 Incentive compatible

BGP is not incentive compatible, in the sense that ASes might have incentive not to participate honestly in the routing protocol [15, 16, 39]. In contrast, [15] shows that next-hop routing is incentive-compatible. Specifically, in the Gao-Rexford framework (see Appendix B), an AS cannot get a "better" next-hop en route a destination by "deviating" from BGP (*e.g.*, announcing bogus

routes, reporting inconsistent information to neighboring ASes, *etc.*).

THEOREM 5.2. *[11, 15] If the Gao-Rexford conditions hold, then next-hop routing is incentive compatible.*

To illustrate Theorem 5.2 we revisit the example in Figure 1. Observe that in the unique stable routing state ASes A, B, and C send traffic along the routes Ad, Bd, and CBd, respectively. Hence, BGP is guaranteed to converge to a routing state where each AS directs traffic via its most preferred *feasible* next-hop AS. (Observe that, while B would prefer using C as a next-hop, no route from C to d that does not traverse B exists, and hence B cannot hope to send traffic through C.) [15] shows that this is true for general network topologies (see also [11]).

5.4 End-to-end security mechanisms

Today, an AS can use BGP policies to avoid paths that travel through undesirable ASes. For example, the U.S. government may want to avoid directing their traffic through other countries. Similarly, an AS may wish to avoid ASes known to perform censorship, conduct wiretapping, or offer poor performance. This is achieved by applying regular expressions to the AS-PATH to assign lower preference to routes that contain the undesirable ASes, or filtering these routes entirely. Under next-hop routing, an AS can no longer specify BGP routing policies that avoid remote undesirable ASes or countries rendering such "AS-avoiding policies" impossible.

While next-hop routing renders "AS-avoiding policies" impossible, we point out that these kinds of policies come with no guarantees; the AS-PATH lists the sequence of ASes that propagated the BGP announcement, not the path the *data packets* necessarily traverse (and these can differ even for benign reasons). We believe that relying on BGP for data-plane security is *misguided*. It is precisely when issues like confidentiality and integrity are involved that the matter should not be left to chance, or misplaced trust. Instead, we argue that these guarantees should be assured in other (end-to-end) ways, such as encryption and authentication, *e.g.*, using the mechanisms in [12].

5.5 In the horizon

We have shown that next-hop routing renders some dangerous attacks against BGP (almost all AS-PATH-based attacks, *e.g.*, path-shortening attacks) ineffective, while making other attacks less harmful, or non-beneficial for the attacker. We have also shown how end-to-end mechanisms, which can be incrementally deployed, can be used to improve data-plane security. Combining these measures will create a routing system that is significantly more secure than today's BGP. However, next-hop routing alone does not provide full protection against all attacks. Specifically, next-hop routing is still vulnerable to "prefix hijacking" attacks, where the attacker announces an IP prefix which it does not own, as well as to attacks that trigger BGP's loop-detection mechanism.

Over the past decade the standards and research communities have devoted much effort to securing BGP against prefix hijacking and more sophisticated attacks. We are finally witnessing the initial deployment of the Resource Public Key Infrastructure (RPKI)—a cryptographic root-of-trust for Internet routing that authoritatively maps ASes to their IP prefixes and public keys. RPKI will enable an AS to authenticate that the origin of a route announcement indeed owns the announced prefix—a property called "origin authentication". In addition, there is much debate about the deployment of mechanisms for AS-level path validation (*e.g.*, S-BGP, soBGP, BGPsec), which will enable an AS to verify that an announced route actually exists (and was announced to the announcer), thus also preventing attacks against BGP's loop-detection mechanism.

6. TRAFFIC MANAGEMENT

Today's routers prefer shorter AS-PATHs as an indirect way to improve end-to-end performance and avoid selecting backup routes. In this section, we discuss how network operators can more naturally achieve their traffic-management goals *without* relying on the AS-PATH. First, we discuss how to rank next-hop ASes based on measurements of end-to-end path performance. Then, we discuss how operators can balance load by ranking next-hop ASes on a per-prefix basis and tagging backup paths with a BGP community attribute.

6.1 Performance-driven routing

Next-hop routing can lead to longer paths than conventional BGP. To understand the impact on AS-PATH length, we analyze the differences in path lengths for different ASes across a range of "T-up," "T-long," and "T-short" events. Our experiments show that most ASes (68.7%–89.9%, depending on the event) have the same AS-PATH length under BGP and next-hop routing, and most other ASes experience just one extra hop. Still, a non-trivial fraction of ASes see even longer paths. While these paths may perform reasonably well, some ASes may indeed experience worse performance. As a result, we believe that a static next-hop ranking should *not* be the only factor in routing decisions.

Given AS-PATH length only loosely correlates with path performance, we argue that routers should make decisions based on measurements of path quality. Routers could adjust (i) the ranking of the next-hop ASes and (ii) the splitting of traffic over multiple next-hop ASes with the same rank, based on the performance metrics (e.g., throughput, latency, or loss) of interest. Different kinds of ASes may select different techniques for moni-

toring path performance, such as:

Multihomed stub ASes: The many multi-homed stub ASes can apply existing measurement techniques for intelligent route control supported in commercial routers (e.g., [40]). Round-trip performance measurements are relatively easy to collect, since the stub AS sees traffic in both the forward and reverse directions.

Online service provider: Online service providers, such as Google and Microsoft, can easily monitor end-to-end performance for their clients (e.g., by logging TCP-level RTT statistics at their servers [41]).

Transit providers: Performance monitoring is more difficult for transit providers, since most traffic does not start or end in their networks, and asymmetric routing may cause them to see one direction of traffic but not the other. Still, passive flow measurements can be used to infer performance [42]. Transit providers can also measure performance directly by using hash-based sampling [43] to sample both directions of the traffic—for the subset of flows that traverse the AS in both directions. This technique also ensures that AS measures just the downstream portion of the path to the destination, rather the part between the source and the ISP.

In all three cases, ASes can measure the performance of multiple paths by using "route injection" [44] to direct a small portion of traffic on each alternate path to continuously track performance. In addition, ASes can use active probing to monitor alternate paths, rather than directing "real" customer traffic over these paths.

Based on the performance measurements, an AS can adapt how it directs traffic over multiple next-hops. Designing and analyzing effective adaptive load-balancing techniques is a rich research topic in it own right. Recent work suggests that it is indeed possible to design stable and efficient controllers for adapting the flow of traffic over multiple paths [45–48]. We plan to study this issue in greater depth as part of future work.

6.2 Interdomain traffic engineering

In addition to selecting paths with good performance, operators rely on interdomain path selection to balance load on their network links. Operators perform traffic engineering by measuring traffic volumes and selecting paths that optimize the flow of traffic over different links. Today, some operators use the AS-PATH to aid in traffic engineering. We believe other approaches (that do not rely on the AS-PATH) are more appropriate:

Outbound traffic engineering: Operators balance load on edge links to other ASes by adjusting the policies that assign "local preference" to BGP routes. This is still possible under next-hop routing, with the restriction that the preferences depend only the immediate neighbor rather than subsequent hops in the AS-PATH. As such, network operators cannot use "regular expressions" on the AS-PATH to (say) direct some traffic through an alternate egress point based on whether the *second* AS in the path is even or odd [49]. Policies based on regular expressions are arguably quite clumsy, and may not be widely used in practice. Instead, network operators could select different next-hop rankings for different (groups of) destination prefixes as a way to direct some traffic over other paths.

Inbound traffic engineering: Controlling the flow of *incoming* traffic is notoriously difficult, since Internet routing is destination-based. Today, some ASes use "AS prepending"—artificially adding extra hops to make the AS-PATH look longer—to make routes through them look less desirable to others. This is often used as a (somewhat clumsy) way to signify a "backup" route. Observe that next-hop routing, by removing AS-PATH length from the decision process, makes AS prepending ineffective. Instead, network operators could use BGP communities to signal backup routes to neighboring ASes [50]; if these ASes export the "signal" to their neighbors, other ASes can also give lower ranking to backup paths. Perhaps more importantly, path-quality monitoring allows ASes to make decisions based on performance, leading naturally to routing decisions that avoid placing excessive load on lower-bandwidth backup links. Still, just as with today's BGP, inbound traffic engineering remains somewhat of a "black art."

In addition, traffic engineering greatly benefits from multipath routing—something next-hop routing supports much more naturally than conventional BGP, as discussed earlier in Section 4. An AS can easily split traffic over *multiple* next-hops leading to the same destination; adjusting the fraction of traffic assigned to each next-hop is a much finer-grain approach to traffic engineering than selecting a single path for each destination prefix.

7. RELATED WORK

Interdomain routing has been an active research area, ever since early work identified thorny problems with BGP [34, 51, 52]. Since then, many papers have characterized BGP's behavior (both theoretically [15, 25, 39], and via measurement [1, 19]), and designed techniques for managing BGP (to detect configuration mistakes [20] and automate traffic engineering [49]). Other research has proposed extensions to BGP (particularly to improve security [5,6] and convergence [3,4], or support multipath routing [3,9,12,35–37]), or designed new interdomain routing architectures [7–10]. We do not attempt to provide a comprehensive overview.

Our work is inspired by recent theoretical work [11, 15,16] on BGP policy restrictions that lead to desirable global properties. We explore whether we can make *reality* look more like those models.

This work contains contributions that did not appear in a preliminary version of this paper [53]. Specifically, the current paper makes the following new contribu-

tions: extensive experimental analysis of next-hop routing, the theoretical model and result for configuration errors, the theoretical analysis of neighbor-specific next-hop routing, and discussion of several important topics (*e.g.*, data-plane monitoring and traffic engineering).

8. CONCLUSION

BGP suffers from many serious problems. We argued that *simplifying* BGP is an attractive alternative to extending or replacing BGP. We proposed next-hop routing—three simple constraints on how routes are selected and exported, combined with external mechanisms for data-plane monitoring and security. Next-hop routing allows ASes sufficient expressiveness to realize their business and engineering objectives, while sidestepping some of BGP's major problems (*e.g.*, slow convergence, large attack surface, incentives to "lie", difficulty in supporting multipath routing, and more).

Our work leaves several interesting directions for future research. We plan to explore how to remove the AS-PATH attribute entirely, that is, how not to rely on the AS-PATH even for loop detection purposes (*e.g.*, by detecting forwarding loops through data-plane monitoring). We also plan to further investigate the stability and efficiency of adaptive interdomain routing based on measurements of data-plane performance.

9. REFERENCES

[1] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian, "Delayed Internet routing convergence," *IEEE/ACM Trans. Networking*, vol. 9, no. 3, pp. 293–306, 2001.

[2] N. Kushman, S. Kandula, and D. Katabi, "Can you hear me now?! it must be BGP," *ACM SIGCOMM Computer Communications Review*, Apr. 2007.

[3] N. Kushman, S. Kandula, D. Katabi, and B. Maggs, "R-BGP: Staying connected in a connected world," in *Proc. USENIX NSDI*, 2007.

[4] P. B. Godfrey, M. Caesar, I. Haken, Y. Singer, S. Shenker, and I. Stoica, "Stabilizing route selection in BGP," tech. rep., University of Illinois at Urbana-Champaign, Jan. 2010. `http://hdl.handle.net/2142/14841`.

[5] S. Kent, C. Lynn, and K. Seo, "Secure border gateway protocol (S-BGP)," *Journal on Selected Areas in Communications*, no. 18(4), pp. 582–592, 2000.

[6] K. Butler, T. Farley, P. McDaniel, and J. Rexford, "A survey of BGP security issues and solutions," *Proceedings of the IEEE*, January 2010.

[7] L. Subramanian, M. Caesar, C. T. Ee, M. Handley, M. Mao, S. Shenker, and I. Stoica, "HLP: A next generation inter-domain routing protocol," in *Proc. ACM SIGCOMM*, pp. 13–24, ACM, 2005.

[8] X. Yang, D. Clark, and A. W. Berger, "NIRA: A New Inter-domain Routing Architecture," *IEEE/ACM Trans. Networking*, vol. 15, no. 4, pp. 775–788, 2007.

[9] P. B. Godfrey, I. Ganichev, S. Shenker, and I. Stoica, "Pathlet routing," in *Proc. ACM SIGCOMM*, 2009.

[10] J. P. John, E. Katz-Bassett, A. Krishnamurthy, T. Anderson, and A. Venkatarmani, "Consensus routing: The Internet as a distributed system," in *Proc. Networked Systems Design and Implementation*, Apr. 2008.

[11] J. Feigenbaum, V. Ramachandran, and M. Schapira, "Incentive-compatible interdomain routing," in *Proc. ACM Electronic Commerce*, pp. 130–139, 2006.

[12] D. Wendlandt, I. Avramopoulos, D. Andersen, and J. Rexford, "Don't secure routing protocols, secure data delivery," in *Proc. Hotnets V*, 2006.

[13] P. R. McManus, "A passive system for server selection within mirrored resource environments using AS path length heuristics," April 1999. `http://www.ducksong.com/patrick/proximate.pdf`.

[14] O. Nordstrom and C. Dovrolis, "Beware of BGP attacks," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 2, pp. 1–8, 2004.

[15] J. Feigenbaum, M. Schapira, and S. Shenker, "Distributed algorithmic mechanism design," in *Algorithmic Game Theory* (N. Nisan, T. Roughgarden, E. Tardos, and V. Vazirani, eds.), ch. 14, Cambridge University Press, 2007.

[16] S. Goldberg, S. Halevi, A. D. Jaggard, V. Ramachandran, and R. N. Wright, "Rationality and traffic attraction: Incentives for honest path announcements in BGP," in *Proc. ACM SIGCOMM*, pp. 267–278, 2008.

[17] S. Goldberg, M. Schapira, P. Hummon, and J. Rexford, "How secure are secure interdomain routing protocols?," in *Proc. ACM SIGCOMM*, 2010.

[18] Rensys Blog, "Pakistan hijacks YouTube." `http://www.renesys.com/blog/2008/02/pakistan_hijacks_youtube_1.shtml`.

[19] R. Mahajan, D. Wetherall, and T. Anderson, "Understanding BGP misconfiguration," in *Proc. ACM SIGCOMM*, 2002.

[20] N. Feamster and H. Balakrishnan, "Detecting BGP configuration faults with static analysis," in *Proc. Networked Systems Design and Implementation*, 2005.

[21] Z. Yin, M. Caesar, and Y. Zhou, "Towards understanding bugs in open source router software," *ACM SIGCOMM CCR*, 2010.

[22] Renesys, "Longer is not always better." `http://www.renesys.com/blog/2009/02/longer-is-not-better.shtml`.

[23] Renesys, "AfNOG takes byte out of Internet." `http://www.renesys.com/blog/2009/05/byte-me.shtml`.

[24] M. Schapira, Y. Zhu, and J. Rexford, "Next-hop routing," June 2011. `http://www.cs.princeton.edu/~jrex/nexthop.pdf`.

[25] L. Gao and J. Rexford, "Stable Internet routing without global coordination," *IEEE/ACM Trans. on Networking*, vol. 9, no. 6, pp. 681–692, 2001.

[26] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)." RFC 4271, Jan. 2006.

[27] "BGP best path selection algorithm." `http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a0080094431.shtml`.

[28] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the internet AS-level topology," *ACM SIGCOMM CCR, special issue on Internet Vital Statistics*, 2005.

[29] B. Quoitin and S. Uhlig, "Modeling the routing of an autonomous system with C-BGP," *IEEE Network*, vol. 19, no. 6, 2005.

[30] Y. Rekhter, T. Li, and S. Hares, "A Border Gateway Protocol 4 (BGP-4)." RFC 4271, January 2006.

[31] Juniper, "Out-delay." `https://www.juniper.net/techpubs/software/junos/junos57/swconfig57-routing/html/bgp-summary32.html`.

[32] P. Jakma, "Revisions to the BGP 'Minimum Route Advertisement Interval'." Internet Draft draft-ietf-idr-mrai-dep-03, October 2010.

[33] A. Fabrikant, U. Syed, and J. Rexford, "There's something about MRAI: Timing diversity exponentially worsens BGP convergence," in *IEEE INFOCOM*, 2011.

[34] T. G. Griffin, F. B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Trans. Netw.*, vol. 10, no. 2, pp. 232–243, 2002.

[35] W. Xu and J. Rexford, "MIRO: Multi-path interdomain routing," in *Proc. ACM SIGCOMM*, pp. 171–182, 2006.

[36] M. Motiwala, M. Elmore, N. Feamster, and S. Vempala,

"Path Splicing," in *Proc. ACM SIGCOMM*, Aug. 2008.
[37] Y. Wang, M. Schapira, and J. Rexford, "Neighbor-Specific BGP: More flexible routing policies while improving global stability," in *Proc. ACM SIGMETRICS*, 2009.
[38] "Understanding route aggregation in BGP." `http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a0080094826.shtml`.
[39] H. Levin, M. Schapira, and A. Zohar, "Interdomain routing and games," in *Proc. ACM STOC*, pp. 57–66, 2008.
[40] Cisco, "Cisco IOS Optimized Edge Routing Overview." `http://www.cisco.com/en/US/docs/ios/oer/configuration/guide/oer-overview.html`, 2007.
[41] M. Szymaniak, D. Presotto, G. Pierre, and M. V. Steen, "Practical large-scale latency estimation," *Computer Networks*, 2008.
[42] A. Gerber, J. Pang, O. Spatscheck, and S. Venkataraman, "Representative speed tests for free: Estimating achievable download speed from passive measurements," in *Proc. Internet Measurement Conference*, Nov. 2010.
[43] N. G. Duffield and M. Grossglauser, "Trajectory sampling for direct traffic observation," in *Proc. ACM SIGCOMM*, 2000.
[44] Z. Zhang, M. Zhang, A. Greenberg, Y. C. Hu, R. Mahajan, and B. Christian, "Optimizing cost and performance in online service provider networks," in *Proc. NSDI*, 2010.
[45] S. Kandula, D. Katabi, B. Davie, and A. Charny, "Walking the tightrope: Responsive yet stable traffic engineering," in *Proc. ACM SIGCOMM*, pp. 253–264, 2005.
[46] A. Kvalbein, C. Dovrolis, and C. Muthu, "Multipath load-adaptive routing: Putting the emphasis on robustness and simplicity," in *IEEE ICNP*, 2009.
[47] D. Wischik, C. Raiciu, A. Greenhalgh, and M. Handley, "Design, implementation and evaluation of congestion control for multipath TCP," in *Proc. Networked Systems Design and Implementation*, March/April 2011.
[48] J. He, M. Suchara, M. Bresler, J. Rexford, and M. Chiang, "Rethinking internet traffic management: From multiple decompositions to a practical protocol," in *Proc. CoNext*, December 2007.
[49] N. Feamster, J. Borkenhagen, and J. Rexford, "Guidelines for interdomain traffic engineering," *ACM SIGCOMM Computer Communication Review*, vol. 33, no. 5, pp. 19–30, 2003.
[50] E. Chen and T. Bates, "RFC 1998: An application of the BGP community attribute in multi-home routing," August 1996.
[51] K. Varadhan, R. Govindan, and D. Estrin, "Persistent route oscillations in inter-domain routing," *Computer Networks*, vol. 32, pp. 1–16, March 2000.
[52] C. Labovitz, R. Malan, and F. Jahanian, "Internet Routing Instability," *IEEE/ACM Trans. Networking*, vol. 6, no. 5, pp. 515–526, 1998.
[53] M. Schapira, Y. Zhu, and J. Rexford, "Putting BGP on the right path: A case for next-hop routing," in *Proc. HotNets*, Oct. 2010.

APPENDIX

A. MODELING BGP DYNAMICS

We use the standard model for analyzing BGP dynamics put forth in [34]. The reader is referred to [34] for further details.

Network and policies. The network is defined by an *AS graph* $G = (N, L)$, where N represents the set of ASes, and L represents physical communication links between ASes. N consists of n *source-nodes* $\{1, \ldots, n\}$ and a unique *destination node* d. P^i denotes the set of "permitted" simple (noncyclic) routes from i to d in G. Each source-node i has a *ranking function* $\leq_i$, that defines an order over P^i. We allow ties between two routes in P^i only if they share the same next hop. The *routing policy* of each node i consists of $\leq_i$ and of i's *import* and *export policies*.[3]

Protocol dynamics. Under BGP at routing tree to the destination d is built, hop-by-hop, as knowledge about how to reach d propagates through the network. The process is initialized when d announces itself to its neighbors by sending update messages. From this moment forth, every active node establishes a route to d by repeatedly choosing the best route that is announced to it by its neighbors (according to $<_i$) and announcing this route to its neighbors (according to its export policy). The network is assumed to be *asynchronous*; ASes can act at different times and BGP updates can be arbitrarily delayed.

Stable states. Informally, a stable state is a global configuration that once reached remains unchanged. Formally, a **stable state** is an n-tuple of routes in G, $R_1, ..., R_n$, such that: (1) If for two nodes $i \neq j$ it holds that j is on R_i, then it must hold that $R_j \subset R_i$ (that is, R_j is a suffix of R_i). (2) If there is a link $(i, j) \in L$, and $R_i \neq (i, j)R_j$, then $(i, j)R_j <_i R_i$. It is easy to show [34] that a stable state is always in the form of a tree rooted in d.

B. GAO-REXFORD FRAMEWORK

In the Gao-Rexford framework, neighboring ASes have one of two business relationships: *customer-provider* and *peering*. [25] presents three conditions that are naturally induced by the business relationships between ASes and proves that these conditions imply guaranteed BGP convergence to a stable state. The three Gao-Rexford conditions are the following:

- **Topology Condition:** there should be no customer-provider cycles in the AS hierarchy digraph.
- **Preference Condition:** an AS should prioritize customer-learned routes over peer- and provider-learned routes.
- **Export Condition:** an AS should not export peer-/provider-learned routes to other peers and providers.

THEOREM B.1. *[25] If the Gao-Rexford conditions hold for a network then BGP convergence to a stable state is guaranteed.*

[3]The import and export policies can be folded into the routing policies, by modifying the preferences so that paths that are filtered out have the lowest possible value. Thus, we do not explicitly model these.