Mutually Controlled Routing with Independent ISPs

Ratul Mahajan
Microsoft Research

David Wetherall
University of Washington
and *Intel Research*

Thomas Anderson
University of Washington

Abstract – We present Wiser, an Internet routing protocol that enables ISPs to jointly control routing in a way that produces efficient end-to-end paths even when they act in their own interests. Wiser is a simple extension of BGP, uses only existing peering contracts for monetary exchange, and can be incrementally deployed. Each ISP selects paths in a way that presents a compromise between its own considerations and those of other ISPs. Done over many routes, this allows each ISP to improve its situation by its own optimization criteria compared to the use of BGP today. We evaluate Wiser using a router-level prototype and simulation on measured ISP topologies. We find that, unlike Internet routing today, Wiser consistently finds routes that are close in efficiency to that of global optimization for metrics such as path length. We further show that the overhead of Wiser is similar to that of BGP in terms of routing messages and computation.

1 Introduction

The Internet is made up of independent ISP networks that cooperate to carry traffic for each other and at the same time compete as business entities. This tension means that no one ISP is able to dictate traffic paths solely according to its best interests, but must instead acquiesce in some manner to the interests of other ISPs. BGP, as commonly used today, codifies one division of control: ISPs usually select the path of outgoing traffic and relinquish control over the path of incoming traffic. This is problematic because it means that ISPs may lack control where needed, e.g., to shift incoming traffic in response to a failure or temporary overload [4, 36]. Studies have also shown that routing paths, while often reasonable, can sometimes be poor from an end-to-end perspective [40, 44, 50].

This state of affairs is not new and has witnessed little real change over the past decade. Lacking effective protocol support, ISPs can mitigate problems through network engineering, e.g., by peering widely to minimize path length inflation in the common case [44], and by manually overriding configurations to handle very poor routes [27]. Newer routing platforms such as RCP [12] can also do a more effective job of optimizing routing within an ISP. However, while valuable, these approaches do not change the nature of the problems that exist.

Our intent is to develop an interdomain routing protocol that addresses these problems at a more basic level. We aim to allow all ISPs to exert control over all routes to as large a degree as possible, while still selecting end-to-end paths that are of high quality. This is a difficult problem and there are very few examples of effective mediation in networks, despite competitive interests having long been identified as an important factor [8].

While it is not a priori obvious that it is possible to succeed at this goal, our earlier work on Nexit [28] suggests that efficient paths are, in fact, a feasible outcome of routing across independent ISPs in realistic network settings. Specifically, Nexit shows it is possible for two ISPs to improve both their individual positions and end-to-end path quality by negotiating and making trades over their set of interconnection choices. So while the "price of anarchy" that measures the inherent inefficiency in multi-party models may be significant in theory [19, 39], Nexit suggests that it may be negligible for real networks.

However, Nexit is far from a complete or practical solution. It focuses on the limited problem of selecting interconnections between two ISPs, with no straightforward extension to routing across more ISPs. And it uses a negotiation mechanism that is much more heavyweight than BGP in terms of message and computational complexity. These factors limit its applicability in practice.

In this paper, we develop an interdomain routing protocol, called Wiser, that is complete and practical in the above senses of running across multiple ISPs and with overheads comparable to BGP. Wiser lets all ISPs exert a degree of control over all paths and produces high-quality paths. We undertake this design in a context that is compatible with independent ISPs: we do not require ISPs to disclose sensitive internal information (such as monetary transit costs) and we do allow ISPs to make decisions according to their best interests and their own optimization criteria (such as a mix of latency and utilization). Wiser paths are also completely policy compliant.

Wiser extends BGP with a simple coordination mechanism that builds on the bilateral ISP contracts that are already in place and is incrementally-deployable across pairs of ISPs. Each downstream ISP tags routing advertisements with costs that are similar to BGP MEDs. Each
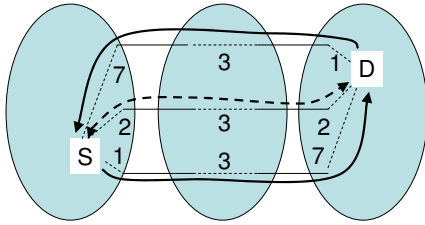
Figure 1: *Unilaterally controlled routing leads to longer paths and higher resource consumption inside ISPs. The solid lines depict early-exit routes. The dashed line depicts a route that is better overall as well as for the left and right ISPs.*

upstream ISP then selects paths with an amended decision process that considers the sum of its own costs and those reported by the downstream ISPs. This lets both upstream and downstream ISPs to exert control over all route choices. Wiser has built-in mechanisms to discourage potential abuse by ISPs. Prototypes in XORP [54] and SSFNet [46] show that its implementation complexity and message overhead are similar to BGP.

For ISPs to adopt this protocol, given no change in monetary compensation, it must be the case that all ISPs improve their position by following this protocol. We find this is so in nearly all cases because ISPs trade small concessions on some routes for larger gains on others.

In our experiments, the end-to-end paths with Wiser are comparable to the best that can be attained with a single entity selecting the entire path using complete information. Over measured ISP topologies, Wiser consistently comes close to the efficiency of globally optimized routing for several measures of interest to users and ISPs. With a latency metric, the most inflated 1% of paths are only 1.5 times longer than ideal, while they are 6 times longer with BGP defaults. With a bandwidth-sensitive metric, we find that Wiser reduces the ISP provisioning level needed to handle interconnection failures by an average of 8% relative to BGP.

These improvements in ISP and overall efficiency seem useful; however, the import of our work is that there exist alternative models for controlling routing among competing ISPs that are practical, policy-compliant, and provide high-quality paths. We consider this to be an issue that has been open since the original Detour study [40]. If we may generalize from our case study of Internet routing, a broader implication is that competing interests in a distributed system can be harnessed with practical protocols and in a way that they do not lead to poor efficiency.

2 Motivation

To see how unilateral control over routing paths can be problematic, consider Figure 1. The middle ISP interconnects to each of the other two ISPs at three locations. The numbers inside the ISPs represent the internal costs

of carrying traffic along paths. We use rough path length as a visual surrogate for cost which will typically include latency and capacity considerations.

With typical business contracts, each ISP can select the route for its outgoing traffic. If each ISP acts individually, its best option is to choose the interconnection that minimizes internal cost. The result is the "early-exit" pattern, shown with solid lines, in which each ISP sends outgoing packets via the nearest interconnection. However, observe that both the left and right ISPs would be better off and end-to-end paths would be better if both ISPs were to use the middle exit. This is shown with the dashed line.

Unfortunately, there is no straightforward way to achieve this routing today. Path cost information can be exchanged between pairs of ISPs with MEDs. But MEDs are not transitive, and even so, their semantics simply change who controls routing; MEDs implement late-exit routing (in which downstream ISPs control the path) and neither provide joint control nor improve routes in the aggregate. Other available mechanisms such as communities have similar shortcomings.

Independent decisions imply that ISPs have poor control over their traffic. This is more pronounced for incoming traffic, for which the available methods are imprecise and coarse-grained at best [4, 36]. It also impacts outgoing traffic, for which ISPs today can only implement early- or late-exit routing but nothing in between. Our conversations with network operators confirm that this can be a major problem in practice.

Several studies have also observed inefficiencies as a result of unilateral control of routing [40, 44, 50]. They find many Internet paths to be slightly inflated and a small fraction to be highly inflated. The latter is most evident when paths traverse a chain of ISPs, since each ISP uses only a local view to select its portion of the route. These studies also find that a primary cause for inflation is how ISPs select their paths today, and other factors such as commercial preferences or inadequate peering are not major contributors [44, 50]. Although manual intervention can fix these routes, it is costly and error-prone. We interpret these studies as motivating a mechanism that consistently (and automatically) produces good paths.

As our example suggests, improving routing requires both sharing and incentives for overall decision making. The left ISP cannot know that the middle interconnection produces a better path without some sharing of information. Given only sharing, the left ISP has little incentive to act for the greater good and use the middle interconnection unless it is compensated, for instance, by the reverse traffic using the middle link as well. But by not reciprocating, the right ISP stands to decrease its costs at the expense of the left ISP. This can easily lead to a standoff. Our work aims to resolve such issues in a mutually-beneficial manner.

3 Problem and Approach

Our goal is to design a practical interdomain routing protocol that enables ISPs to jointly control routing and compute good end-to-end paths while acting in their own interests. In this section, we informally describe the constraints we place on the problem and our approach to it. We provide more detail in the next section.

For our solution to be practical, it must be incrementally deployable in the existing Internet. Pragmatically, this requires that it be comparable to BGP in terms of traditional overheads such as message complexity, computation, and convergence time. It also leads us to build on the existing business framework in which pairs of ISPs have peering contracts that only coarsely tie traffic levels to monetary compensation.

Handling independent ISPs raises a different set of issues [28]. We consider three constraints to be important:

1. Individual ISPs should improve their position by using our protocol compared to today. This is because we do not assume changes in monetary compensation; without compensation, ISPs will select routes that are to their own benefit. If a new protocol does not improve their position in the aggregate (by finding better paths within their networks), they simply will not run it. We refer to this property as *win-win*. While all more efficient routings will improve the overall situation, not all of them are win-win for each ISP involved, e.g., when one ISP is required to carry traffic further than necessary because it has the better network. Further, we want the win-win property to approximately hold even if one ISP abuses the protocol to try and take advantage of other ISPs.

2. While we need some information sharing, we should not require ISPs to disclose information in forms that they consider sensitive. This includes internal performance metrics and the monetary cost of carrying packets. Such information can be used against ISPs by competitors in the marketplace and is not disclosed today.

3. We should allow ISPs to set their own optimization criteria. In general, ISPs prefer paths that avoid congestion and minimize latency over some combination of their internal network and the performance experienced by their customers. However, different ISPs are bound to optimize paths differently, e.g., depending on their provisioning, monetary costs, the needs of their customers, and other factors of which we cannot be aware. There does not exist a universal metric that works for all ISPs, and past attempts to define such metrics have failed [18].

These constraints rule out known approaches to optimizing interdomain routing (§8). Our approach is to share rough path cost information between ISPs and enable each to improve its routing by its own reckoning. This does not explicitly improve end-to-end paths; rather, better overall paths are a welcome side-effect of better paths for individual ISPs. These paths do depend on the metric each ISP uses and so they cannot, for example, improve end-to-end latency if ISPs optimize strictly for utilization. But in practice, reasonable ISP metrics will factor in both delay and congestion. Our results then show that improving paths for ISPs is sufficient to avoid egregiously bad overall paths, bringing most of the benefit of a more direct but infeasible global optimization [20].

To share information, we extend BGP. ISPs tag advertised routes with costs that are derived from internal path costs. These advertised costs are *agnostic* in that they are simply cardinal numbers whose relative magnitude is significant. They serve to coordinate ISPs without requiring a standard metric or cost derivation methodology. We believe they limit the information that is shared to an acceptable level; they resemble MEDs – ISPs often use (cardinal) IGP costs to set MEDs [29, 30] even though MEDs have ordinal semantics – and not transparent measures that disclose information in defined units, e.g., latency in seconds or monetary cost in cents. It is of course possible that ISPs may attempt to reverse-engineer network properties from agnostic costs, but this is hardly a new capability because even today outsiders can measure ISP networks [45].

To enable ISPs to improve their position, we build on bilateral ISP contracts with a simple mechanism that coordinates the route selections of each ISP with those of its neighbors. An ISP selects paths based on the combination of its own costs and the costs advertised by its neighbor, and in return the neighbor does the same. Both track how costs are used to see that this is so. This mechanism is based on the observation that the interaction between ISPs spans many flows over time. It is not necessary that each ISP come out ahead for each individual flow. Rather, routing can be close to win-win when ISPs take a holistic view of traffic, compromising on some flows in exchange for bigger gains on others.

4 Design of Wiser

We now describe our routing protocol, beginning with the problem formulation.

4.1 ISP Model and Goal

Consider an internetwork of ISPs. Let each ISP associate a cost with each of their internal paths. We assume that each ISP aims to reduce the average cost of carrying traffic by its own measure, i.e., minimize the sum of the cost of its paths weighted by the traffic that they carry. So if $\mathsf{intcost}_I(p)$ is the internal cost of path p in ISP I, and $\mathsf{traffic}_I(p)$ is the rate of traffic carried along it over some period, we have:

$$\mathsf{cost}_I = \min_{paths} \left(\sum_{p \in paths} \mathsf{traffic}_I(p) \times \mathsf{intcost}_I(p) \right)$$
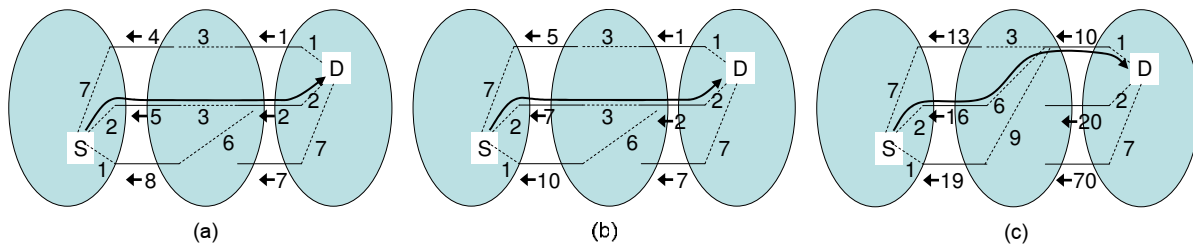
Figure 2: *(a) Traditional shortest path routing with comparable costs. (b)* Wiser *routing with agnostic costs approximates overall shortest path routing. (c) ISPs can unduly bias routing without cost normalization.*

This is close to what IGP protocols such as OSPF and IS-IS achieve today for an intradomain workload with internal costs that are the sum of link weights. The problem we tackle is that BGP does not achieve this goal for an interdomain workload due to the peering point selections that result from the interaction of independent ISPs.

We do not mandate how ISPs set internal costs, since we wish to let them use their preferred method. In our evaluation, we use IGP cost, i.e., the internal cost of a path is the sum of its component link weights. Further, we assume that ISPs act reasonably, for instance, by setting internal costs to favor paths with shorter internal distances and disfavor congestion. To reduce distance, they can simply set the link weights to reflect a measure of delay. To factor in utilization, they can use mappings that assign higher cost to links with a higher utilization [14]. Minimizing the sum of such costs then finds paths that disfavor congestion and are otherwise short.

We also need to measure the overall efficiency across ISPs so that we can assess the benefits of different schemes. If the internetwork were treated as one large ISP, with a single method of assigning costs, we could compute the routing with minimum global cost. However, there is no well-defined optimum if ISPs use different internal cost metrics. In such systems, the best solutions that can be obtained are *Pareto-optimal*. A solution is Pareto-optimal if the cost for any party cannot be reduced without hurting at least one other party. Pareto-optimality rules out solutions with obvious wastage when all parties could be better off. Unilateral routing is Pareto-optimal for individual paths (because each is best from the selector's angle) but not when considering all paths of an ISP in aggregate. We leave the goal of overall efficiency as an informal one. In our evaluation, we look at effects on overall measures of latency and bandwidth to see the impact of independent ISPs, and we compute efficiency only within individual ISPs.

4.2 Lowest-Cost Routing Across ISPs

Wiser adapts lowest-cost routing to the setting of independent ISPs. The above discussion suggests that, if we can put ISP interests aside, we can achieve efficient routing in a simple manner: have ISPs use the same method of assigning costs and run a traditional lowest-cost rout-

ing protocol over them. This is shown in Figure 2(a). Here, a route is computed to send traffic from source S to destination D. Each ISP has internal costs, which for ease of exposition we make static and symmetric in both directions of a link. Each ISP advertises to its neighbor the total cost to reach the destination, which is the sum of internal costs thus far. Each router selects the lowest cost path. The dark line shows the optimal route that is found.

However, a naïve lowest-cost routing protocol that does not handle ISP interests may fail to find such routes for several reasons. Costs may not be comparable and thus cannot simply be added across ISPs. Even so, optimal routes may require one ISP to lose for the greater good; we have argued that this is undesirable without changes in monetary compensation. Moreover, there is nothing to prevent ISPs from biasing their advertised costs to suit their own interests, inflating some and reducing others. Finally, even with reasonable cost advertisements, nothing prevents ISPs from ignoring others' costs while selecting paths.

The following subsections describe how we address these problems to approximate lowest-cost routing with independent ISPs. We use *cost normalization* to handle incomparable metrics, find win-win solutions, and incent ISPs to report costs honestly. We use bounds on the ratio of *usage cost* to incent ISPs to select paths honestly.

As a tool for autonomous parties, how Wiser will be used is likely to differ across ISPs. Our intent is to design it such that its works reasonably well by default for common situations and anticipated uses. Similarly, our mechanisms are not intended to completely prevent unwanted behavior. In fact, there is a fundamental conflict between efficiency and being provably cheat-proof [7, 33]. We favor efficiency because we expect honesty to be the common case. Competitors tend to act honestly when they seek mutual gains over a default contract [38], and even today ISPs cooperate using mechanism that are not cheatproof. Cheating tends to be a poor strategy in long-term relationships such as those between ISPs because it risks long-term harm [1, 31].

4.3 Cost Normalization

In Wiser, ISPs normalize the costs received from neighbors so that they are comparable to internal costs. Each

ISP scales the costs it receives from a neighbor such that the sum of the costs received from the neighbor equals the sum of the costs advertised to the neighbor. This requires border routers to share information in order to determine the totals; we discuss how it may be implemented in §5. Once a normalization factor has been computed, ISPs can add their internal costs to normalized external costs to propagate routes. Routers then select the path with lowest total cost and advertise it upstream, as before.

If $\mathsf{advcost}_{I \to N}(P, d)$ is the cost advertised for destination d by ISP I to N over peering link P then we have:

$$\mathsf{totaladv}_{I \to N} = \sum_{P,d} \mathsf{advcost}_{I \to N}(P, d)$$

$$\mathsf{normal}_{N \to I} = \frac{\mathsf{totaladv}_{I \to N}}{\mathsf{totaladv}_{N \to I}}$$

$$\mathsf{advcost}_{I \to N'}(P', d) = \min_{P,N} \Big(\mathsf{intcost}_I \left(P'P \right) +$$

$$\mathsf{normal}_{N \to I} \times \mathsf{advcost}_{N \to I}(P, d) \Big)$$

Above, $\mathsf{normal}_{N \to I}$ is the normalization factor that ISP I applies to the costs it receives from neighboring ISP N. The final equation is the propagation rule: each ISP advertises to its neighbors its best cost route: the minimum sum of its internal cost and the normalized cost it receives from its neighbors (that are most preferred).

Figure 2(b) shows how normalization approximates efficient overall routes. The total advertised cost from the right ISP to the middle ISP is 10. For simplicity, we have not shown the routing advertisements for S which travel from left to right. Assuming symmetry, the costs from the middle to the right ISP will be the same as those advertised from the middle to the left ISP, which have a total of 22. The normalization factor for costs coming in from the right ISP to the middle one is thus $\frac{22}{10}$. The top interconnection between these two ISPs, for example, is normalized to roughly 2 ($1 \times \frac{22}{10}$), and is propagated as 5 after adding 3 for the internal path costs. While the advertised costs are somewhat different than lowest-cost routing, the same globally shortest path is selected.

Normalization brings two key benefits as well as allowing costs to be compared across different ISPs. (In contrast, MEDs received from different ISPs are semantically incomparable, which can lead to instabilities [29] and other practical problems [47].) First, it limits biases that a dishonest ISP can cause by manipulating costs. Without normalization an ISP could trivially control routing by scaling its costs. For instance Figure 2(c) shows what happens if the right ISP inflates its costs by a factor of ten and the other ISPs continue to minimize the sum of costs. The new path is unfavorable to the middle ISP and has worse overall properties. Normalization nullifies the impact of such inflation.

The bias of relative changes in advertised costs is also limited because increasing the relative values of some costs implies decreasing the relative values of others. We

find that even with complete knowledge, relative changes gain little for the downstream and inflict little damage on the upstream compared to not running Wiser (§6.3). And in the more realistic case of partial information, the outcome of manipulation can be hard to predict and possibly inferior for the downstream. For instance, consider Figure 2(b) and let the right ISP increase the relative cost of the middle link with the goal of causing the traffic to use the top link. It must decrease the relative cost of the other two links. By doing so, the right ISP may inadvertently cause traffic to use the more expensive bottom link (as the internal costs of the middle ISP in that direction are not directly advertised to the right ISP). The combination of limited gain and uncertainty discourages manipulations.

A second benefit of normalization is that it approximates win-win routing by making path selection sensitive to the concerns of both ISPs. For instance, in a scenario where one ISP has costs in the range [0-10] and the other ISP has costs in the range [0-1000], shortest path routing without normalization will strongly favor the ISP whose costs dominate. Normalization gives the ISPs an equal-footing on which to compare their costs and benefits.

Treating a neighboring ISP as an equal is an approximation to an optimal solution that we have found to work well. This is most likely because of the structure of the ISP marketplace in which peer relationships, at least where there is routing choice, often occur between rough equals. As with peering contracts, however, it would be easy for ISPs to use different weightings if they wanted to account for significant asymmetries, e.g., a higher weight may be assigned to (paying) customers or ISPs that send more traffic. Similarly, instead of computing it based on current advertisements, some ISPs may choose to have a normalization factor based on longer-term averages. Our experiments use an equally weighted normalization based on current advertisements.

4.4 Bounds on the Ratio of Usage Cost

Given that ISPs are encouraged to honestly advertise costs, we would like to also incent them to respect those costs in making their routing choices. In the absence of such an incentive, an upstream ISP might undermine the protocol by selecting locally optimal paths. For instance, the left ISP in Figure 2 might select the bottom interconnection regardless of the advertised costs. It is not straightforward for the downstream ISPs to catch this behavior since their expensive links may be an appropriate choice for some upstream sources.

We use a cost usage ratio to encourage ISPs to honestly select paths. It is inspired by current contractual practices, in which peer ISPs set a traffic exchange ratio that involves no money transfer in the common case [34]. Both upstream and downstream ISPs independently track how the upstream is sending traffic to the downstream.

They keep a running total of usage costs, which we define to be the sum of the advertised cost of a route multiplied by the rate of traffic that is sent along it.

The average usage cost, which is the total usage cost divided by the total rate of traffic, will vary with ISP path selections. With honest path selection, the average usage cost will be low because the upstream ISP tends to avoid paths that are costly for the downstream ISP. In contrast, if the upstream ISP is dishonest, the average usage cost will be higher. Wiser leverages this expected behavior by adding a clause to the contract between ISPs that stipulates a bound on the ratio of the average usage cost to the average advertised cost. Using the same notation as before, we have:

$$\text{avgusage}_{I \to N} = \frac{\sum_{P,d} \text{advcost}_{N \to I}(P, d) \times \text{traffic}(P, d)}{\sum_{P,d} \text{traffic}(P, d)}$$

$$\text{avgadv}_{N \to I} = \frac{\sum_{P,d} \text{advcost}_{N \to I}}{\sum_{P,d} 1}$$

$$\text{ratio}_{I \to N} = \frac{\text{avgusage}_{I \to N}}{\text{avgadv}_{N \to I}}$$

Above, $\text{avgusage}_{I \to N}$ is the average usage cost of ISP I sending traffic to ISP N and $\text{avgadv}_{N \to I}$ is the usage cost that will result if path selections are randomized over the advertised costs of ISP N. The final equation gives the quantity that is checked against the contractual bound which is determined by ISPs based on their situation.

ISPs have the flexibility to make individual decisions within the bound, but are incented to use advertised costs in their overall route selection to stay within the contract. Usage costs also incent an ISP to honestly propagate the received costs in its own advertised costs; if it fails to do so, its upstreams may select paths that are more costly and increase its usage costs. Finally, other contractual clauses are also possible, e.g., a provider might charge a customer based on the measured ratio. We leave their exploration for future work.

5 Implementation

We have implemented Wiser as an extension to BGP on two independent platforms, XORP [54] and SSFNet [46]. We made the following changes to BGP:

- As well as BGP attributes, routing messages carry agnostic costs using a new optional, non-transitive route attribute. (A new community attribute could also be used.) To compute these costs, we use the IGP metric as the internal component. This is similar to the way that ISPs use IGP costs as the basis for MEDs [29, 30]. It would be straightforward to accept costs via a different channel to accommodate ISPs that do not have IGP costs available, e.g., because they use MPLS, or prefer other costs.

- Border routers keep track of the sums of the advertised and received agnostic costs for each neighbor and periodically share them with the other border routers of

| 1. Highest local preference |
| **2. Lowest Wiser cost** |
| 3. Shortest AS-path length |
| 4. Lowest origin type |
| 5. eBGP-learned routes over iBGP-learned |
| 6. Lowest IGP cost to egress router |
| 7. Lowest router ID of the BGP speaker |

Table 1: *The routing decision process with* Wiser *is similar to that with BGP except for an additional filter (Step 2) that is based on* Wiser *cost.*

the same ISP. Information from all the routers is aggregated to compute the normalization factor, which is the ratio of the incoming to outgoing costs summed across all border routers. We leverage the existing iBGP mesh and route reflection mechanisms but new platforms such as RCP [12] can also be easily adapted for this purpose. Intra-ISP partitions do not pose a major problem; routers can either continue using the pre-partition factor or compute it based on the subset of reachable routers.

- When a border router receives routes from a neighbor, it normalizes their advertised costs by multiplying them by the normalization factor for that neighbor. Only normalized costs are propagated within the ISP.

- Each router uses the modified BGP decision process shown in Table 1 to select routes. Compared to BGP, it includes an additional step that selects routes based on the Wiser cost, which is the sum of the normalized received cost and the internal cost. This step comes after considering local preferences that implement commercial policies (e.g., prefer customers over providers). It comes before AS-path, which it effectively replaces, and before internal cost, so that decisions factor in non-local costs.

- When the normalization factor for a neighbor changes significantly (in response to major routing changes), the border routers re-evaluate routes received from that ISP. This is similar to what happens today when IGP costs change. However, unlike IGP cost changes, this re-evaluation can be done in the background, while other tasks are processed with a higher priority. This is because exactly one router, which received the route directly from the neighboring ISP, is responsible for normalizing the costs of any given route. Until that router applies the change, the other routers do not realize that the cost of the route has changed and have a consistent view of it.

- Finally, to verify a neighbor's path selection behavior, border routers log information required to compute the incoming usage costs. This includes the amount of incoming traffic and the announced cost for each destination prefix. A sampling mechanism similar to NetFlow is used for this purpose. Periodically, the estimates are logged to disk and reset. ISPs collect this information from all of their border routers and check if the usage ratio is below the contractual threshold. Similar logging

is implemented to compute outgoing usage costs, which helps to cross-check if a neighbor claims that this ISP has exceeded the threshold. We have not yet implemented usage cost logging in XORP.

The modifications to BGP above can be deployed incrementally by pairs of ISPs to improve routing between them. Moreover, all routers within an ISP need not be upgraded simultaneously. In a partially upgraded state, the normalization factor can be statically programmed at the upgraded routers, before transitioning to a dynamic computation. (Done this way, some care needs to be taken regarding the consistency of forwarding paths.)

6 Evaluation

Our evaluation of Wiser considers the following questions:

1. How efficient is Wiser and is it win-win? For the topologies and cost models that we study, we show that Wiser is win-win and its efficiency comes close to ideal routing that globally optimizes the internetwork based on complete information.

2. What is the overhead of running Wiser? We show that the overhead of Wiser is similar to BGP in terms of implementation complexity, routing message and computational requirements. Its convergence time is acceptable even in response to major failures.

3. How robust is Wiser to cheating? For the strategies that we study, we show that Wiser is robust in that it limits the gain for dishonest ISPs and the loss for honest ISPs.

4. What factors facilitate efficient routing with Wiser? While previous evaluations use inputs based on the current Internet, this part uses synthetic cost and topology models to understand situations under which Wiser will be effective. We show that it produces efficient end-to-end paths as long as ISP objectives include relevant factors and that its efficiency is higher when the costs of ISPs that interconnect in multiple places are similar.

The answer to the questions above depend on many aspects of ISP networks, some of which are hard to model. To focus on realistic rather than theoretical best- or worst-case bounds, we combine measured data with models based on known properties of the Internet. As measured input, we use an internetwork topology of 65 ISPs and their interconnections [44]. These ISPs have diverse sizes and geographical presence. A node in an ISP topology corresponds to a city where the ISP has a point of presence (PoP). We compute paths over these topologies (rather than use separately measured paths) so that our results are less influenced by measurement errors. The models that we use depend on the experiment but a common one is approximating propagation delay of a link by the geographic distance between the two end-point cities [35]. Our results are of course limited to the topologies, models, and ISPs behaviors that we study.

Efficiency (§6.1)	Similar ISP objectives	Path length
		Bandwidth provisioning
	Diverse ISP Objectives	Path length and bandwidth
		Inferred weights
Overhead (§6.2)	Implementation complexity	
	Routing msgs. and convergence	Load independent costs
		Load sensitive costs
	Computational requirements	Normal workloads
		Highly dynamic workloads
Robustness to cheating (§6.3)	Dishonest cost disclosure	
	Dishonest path selection	
	Dishonest cost propagation	
Understanding efficiency (§6.4)	Impact of topology	
	Impact of ISPs' objectives	
	Comparison with Nexit	
	ISP costs for individual flows	

Table 2: *Experiments with* Wiser. *The shaded cells correspond to experiments not presented in this paper.*

To compute the routing produced by Wiser, we use our XORP and SSFNet prototypes as well as a custom, high-level simulator that does not model message passing. These three engines have different scaling properties because they model different levels of detail. We use the custom simulator to study efficiency and robustness to cheating, and the prototypes to study overhead.

Table 2 provides an overview of the experiments we have conducted. We present only a subset of results in the following sections due to space limitations.

6.1 Efficiency

The efficiency of global routing can be measured in several ways. We study scenarios with both similar and diverse ISP objectives. For the former, we first consider end-to-end path length, since long paths degrade application performance. This metric assumes that the network capacity is well-matched to traffic and ISPs are primarily interested in minimizing the distance a packet travels inside their networks. We then consider bandwidth provisioning required by ISPs to avoid overload when the traffic and capacity are no longer well-matched, e.g., due to a failure. We use inferred link weights for scenarios with diverse ISP objectives, since they capture the choices of different ISPs.

6.1.1 Path length

We compare the path lengths produced by Wiser to two other routing methods: *global* and *unilateral*. The former minimizes path length using global information on the lengths of network links. It is not feasible in practice due to ISP autonomy issues, and we study it to understand the cost of this autonomy. *Unilateral* mimics the
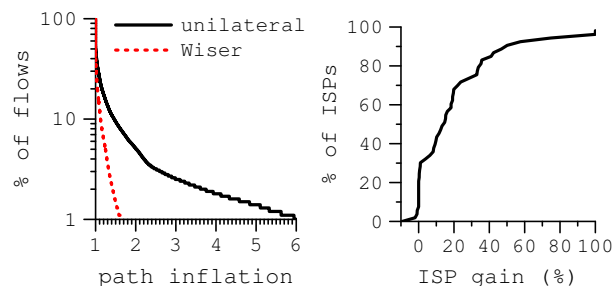
Figure 3: Wiser *produces efficient routing paths and is win-win.* Left: *The CCDF (in log scale) of path inflation.* Right: *The CDF of gain for individual ISPs with* Wiser.

common BGP policies of shortest AS-path and early-exit. With Wiser, ISPs use internal distance as the basis for assigning costs to internal paths. This is a rough measure of the resources consumed inside the network; minimizing it is the motivation for early-exit routing. All three routing methods follow common commercial policies [34, 48], e.g., ISPs do not provide transit to peers and providers. Traffic in this experiment consists of a flow between each pair of PoPs.

We find that, unlike *unilateral*, end-to-end path lengths with Wiser come close to that of *global*. The average path length with Wiser is only 4% higher than *global*, while it is 13% higher with *unilateral*. This improvement is useful, though it suggests that the common case path length in the Internet today is acceptable.

We find the more significant difference between *unilateral* and Wiser to be the distribution of path lengths of individual flows. The left graph in Figure 3 shows the path length inflation with Wiser and *unilateral* compared to *global* as a complimentary cumulative distribution function (CCDF): the y-value is the percentage of flows inflated by at least the corresponding x-value. An inflation of two implies that the path length was doubled.

We see that, while half of the paths are not inflated at all, some are highly inflated with *unilateral*: 5% are inflated by more than a factor of 2 and 1% by more than a factor of 6. In terms of absolute inflation, 5% of the paths are inflated by over 40 ms and 1% by over 70 ms. High inflation is not limited to short paths or intercontinental paths: even when we consider only paths that are longer than 20 ms with unilateral routing or only paths within the USA, the worst 1% are still inflated by a similar factor. Applications using overly long paths will experience high latencies unless the paths are (manually) fixed.

To better understand how overly long paths arise, consider two examples. The first example involves two large ISPs in the USA that have national presence but interconnect largely along the two coasts. For traffic going from the middle of the country in one ISP to the east coast inside another, with *unilateral*, the source ISP picks an interconnection that is closest to the source. It happens

to be an interconnection on the west coast. Hence, the source ISP takes the traffic to the west coast and the destination ISP brings it to the east coast. With *global*, an east coast interconnection is employed, leading to a path that is roughly three times shorter.

The second example involves traffic going from the southeast to the east coast of the USA between two ISPs that do not directly interconnect. With *unilateral*, the source ISP transfers the packets to an intermediate ISP that does not interconnect with the destination ISP on the east coast, which makes the traffic traverse an interconnection on the west coast before returning to the east coast. With *global*, the chosen intermediate ISP interconnects with the destination ISP at the east coast itself. The resulting path is roughly five times shorter.

The graph shows that Wiser can automatically fix such overly long paths: 5% of the paths are inflated by a factor of 1.2 and the worst 1% by only a factor of 1.5. This gain in efficiency stems directly from joint control over routing. Unlike *unilateral*, by combining inputs from all ISPs, Wiser can avoid long paths that, while slightly favorable for the source ISP, are very long end-to-end.

The right graph in Figure 3 shows that improvement in end-to-end paths with Wiser does not require individual ISPs to suffer for the global good, i.e., it is close to win-win as we desire. It plots the CDF of gain for individual ISPs with Wiser, measured as the average reduction in distance, relative to *unilateral*, that a packet travels inside the ISP's network. In terms of adoption incentives, this measure ignores the improvement in performance for customers and the reduction in operational cost from not having to manually fix poor routes. Almost no ISP loses by running Wiser and many ISPs gain, and thus ISPs have an adoption incentive. The graph shows that a handful of ISPs do lose a little. These are small, edge ISPs for whom the changed routing pattern represents a minor loss according to our measure. We find that the overall quality of routing is not impacted even if such ISPs chose to not run Wiser. Alternatively, they can also negotiate a different normalization factor with their neighbors.

6.1.2 Bandwidth provisioning

To be free of congestion under dynamic conditions, ISPs can either highly provision their networks or dynamically alleviate overload. The former approach is common today because ISPs do not have proper control over routing. But Wiser can help ISPs with the latter, and thus reduce provisioning, by signaling to their neighbors to send less traffic along certain paths.

To assess this benefit, we must use models based on known properties of Internet routing. Provisioning is hard to evaluate because it is affected by factors such as link capacities and workloads for which there is almost no public information. We use models that are similar to previous work [21, 28]. We use a gravity approach to model
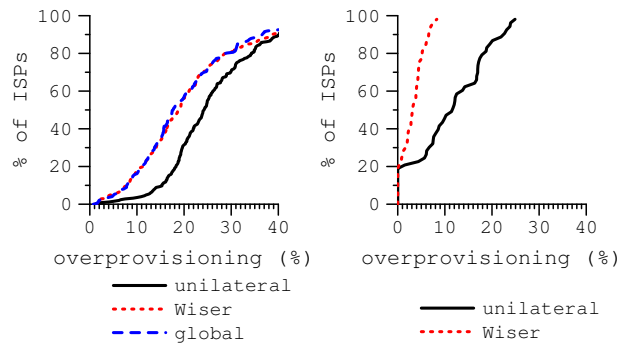
Figure 4: Wiser *reduces overprovisioning level.* Left: *Pairs of ISPs.* Right: *The entire internetwork topology.*

Figure 5: Wiser *produces efficient routing paths with inferred link weights.* Left: *The CCDF of path inflation.* Right: *The CDF of gain for individual ISPs with* Wiser.

traffic between two PoPs as proportional to the product of the population of their host cities [49, 55]. We model link capacities as proportional to the stable load on the link, since in steady-state a well-designed network tends to be roughly matched to its traffic [55]. We simulate dynamic conditions by failing interconnections between ISPs, as these failures can cause congestion today [23, 24]; we leave the study of other perturbations such as internal failures for future work.

We measure efficiency using the *overprovisioning level*. For a link, this is the maximum additional load, compared to its stable load, that it carries across all simulated failures. The overprovisioning level for an ISP is the weighted average of the overprovisioning levels of its links. The weight of a link is its stable load, to reflect that doubling the capacity is costlier for higher-capacity links.

We compare the same three routing methods. *Global* minimizes the overprovisioning level across the entire internetwork and is computed by solving a linear programming problem. For computational tractability, we allow fractional routing which provides a lower-bound on any protocol with non-fractional routing. To illustrate the benefit of Wiser, we set costs to be the product of static and dynamic factors. The static factor is its length. The dynamic factor varies linearly with load but is updated only when the load changes by more than 10%. *Unilateral* is computed as before and is load-insensitive; it is uncommon for ISPs to automatically respond to overload because that may hurt neighbors. Techniques that let ISPs respond without hurting others have limited efficacy [11].

Due to computational limits, we could not compute *global* for the entire topology. We divide our results into two parts. First, we compare all three routing methods over subsets of the overall topology. This illustrates how close to *global* Wiser can get. Then, we compare *unilateral* and Wiser over the entire topology.

The left graph in Figure 4 shows the results for subset topologies which are pairs of adjacent ISPs with traffic flowing between all pairs of PoPs in the two ISPs. There
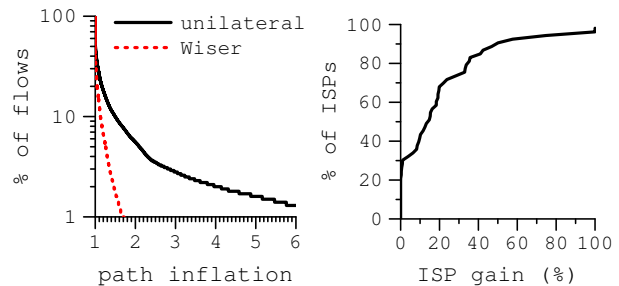
are two points for each ISP pair. We find that, unlike *unilateral*, Wiser closely approximates *global* to the extent that their lines are visually indistinguishable. Relative to the *global*, the average overprovisioning level is 0% with Wiser and 7% with *unilateral*. The right graph shows the results for the entire topology. Traffic consists of flows between a randomly selected 10% of all possible PoP pairs. We simulate the failure of each interconnection between tier-1 ISPs. There are over 400 such interconnections in the dataset. As is the case for the smaller topologies, the overprovisioning with Wiser is much less than that with *unilateral*. The average difference is 8%. While ISPs usually upgrade capacity by bigger factors, this difference may translate into significant monetary savings if they need to upgrade less often.

Though not shown in the graphs, we find that Wiser is win-win for this measure as well, i.e., the overprovisioning level of individual ISPs does not increase compared to *unilateral*.

6.1.3 Inferred link weights

Finally, we study the efficiency of Wiser when ISPs have diverse objectives. We model this using link weights inferred for each ISP [44]. The shortest path routing produced by these weights is consistent with the routing observed inside ISPs, though these weights are not necessarily identical to what the ISP uses [44]. We assume that these weights capture the existing internal objectives of ISPs at the time the topologies and weights were inferred.

The left graph in Figure 5 shows the efficiency of Wiser and *unilateral* with inferred link weights. It plots the multiplicative inflation in path length relative to the length of *global* (as computed in §6.1.1). Unlike *unilateral*, Wiser comes close to *global*. The worst 1% of the paths are inflated by a factor of 7 with *unilateral* but by only 1.7 with Wiser. The right graph shows that Wiser is win-win for this cost model as well. It plots the gain for individual ISPs, measured as the average reduction in path *weight* relative to *unilateral*.
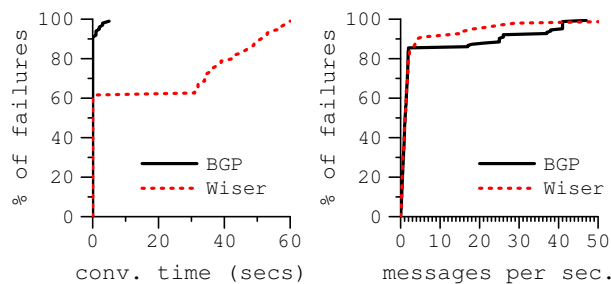
Figure 6: *The convergence time and routing message overhead of* Wiser. Left: *The CDF of convergence time.* Right: *The CDF of maximum rate of routing messages.*

6.2 Overhead

We now study the overhead of Wiser relative to BGP using our XORP and SSFNet prototypes.

Implementation complexity   Using lines of code as a rough measure of implementation complexity, we find that our XORP and SSFNet prototypes add only 3% and 6% lines to their respective BGP implementations.

Convergence time and routing messages   We evaluate the convergence time and routing message overhead of Wiser in response to routing perturbations. We use SSFNet for these experiments. Due to memory limitations, we consider only the "core" of the internetwork topology, which includes roughly 300 nodes that belong to tier-1 ISPs and have multiple neighbors. We believe that our results are reflective of the overall topology because our measures depend heavily on the core [25].

We perturb routing with failures of the interconnections between tier-1 ISPs as in §6.1.2. These are significant events, and most changes will have a smaller impact on routing. One interconnection fails per trial. All costs are static, which corresponds to overprovisioned networks.

The left graph in Figure 6 plots the CDF of the time it takes for the routing to converge after the link fails. There is one point for each simulated failure. Connectivity is restored in both Wiser and BGP as soon as the failure is discovered but the convergence time can differ. For 60% of the failures, the convergence time of Wiser is similar to that of BGP. It is higher for 40% of them.

BGP routers advertise only reachability, and so the routing converges soon after the routers attached to the failed link withdraw routes that use that link and announce new routes. With Wiser, if the normalization factor changes significantly, as it may for a major failure, more routing messages can follow the initial announcements. The delay in this case is dominated by the MRAI (minimum route advertisement interval with a default of 30 seconds) timer of BGP which determines the minimum gap between routing messages sent to neighbor.

These convergence times seem acceptable to us for major changes, especially since connectivity is restored

quickly. Even faster convergence could be obtained with lower values of MRAI, as advocated by some [15], or by ignoring MRAI for normalization factor changes.

The right graph in Figure 6 plots the CDF of the maximum message rate experienced by any router in the topology. We count messages from when the link fails to when the routing converges. We see that the routing message overhead of Wiser is comparable to that of BGP. It is slightly lower probably due to its longer convergence time for a subset of the cases.

Computational requirements   We use XORP to study the computational requirements of Wiser for typical workloads. Routers that run Wiser need to track normalization factors and usage costs, as well as BGP-related responsibilities. To measure this added burden, we feed in a log of routing messages collected by a RouteViews server from forty-one diverse routers in the Internet to a machine (2.2 GHz, 3.8GB RAM) that runs Wiser. Because the logs do not contain Wiser costs, we attach a randomly generated cost in the integral range [1..10] to each message. So that the routing tables fit in memory, we randomly select ten out of the forty-one message sources in a trial. We conduct five trials each for logs from two different days and find that the computational load of Wiser is only 15-25% higher than BGP.

6.3 Robustness to Cheating

To study potential for abuse, we consider a form of cheating that we consider plausible: a dishonest ISP manipulates its advertised costs and favors its own costs in path selection to try to reduce the average internal cost of the traffic it carries. There are bound to be other motivations to cheat and forms of misbehavior of which we are not aware, but we must leave these for future study.

In our experiment, we consider pairs of ISPs that interconnect in multiple places. This allows us to study the average cost of carrying traffic in isolation because the overall traffic in each ISP network stays constant. Traffic is composed of a unit flow between each pair of PoPs. ISPs aim to minimize internal distance. One ISP in the pair is honest and the other is dishonest. We assume that the dishonest ISP has complete information about traffic and the other ISP's internal costs of sending traffic (which are never directly disclosed). This overestimates the abilities of the cheater because there will be uncertainty in this information in practice.

The dishonest ISP changes the relative values of advertised costs and uses a reduced normalization factor to favor its internal costs when it selects paths. Computing advertised costs that give the dishonest ISP the most gain within the normalization constraint is NP-hard (because it similar to bin packing). We use hill climbing to approximate these costs. Similarly, we use binary search to find the lowest normalization factor that still satisfies the
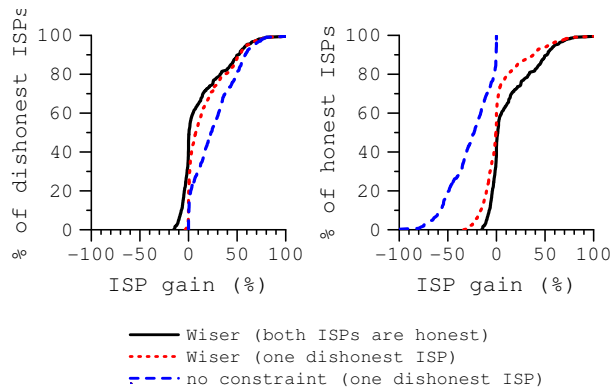
Figure 7: Wiser *limits the gain for dishonest ISPs and the loss for honest ISPs.* Left: *The CDF of gain for dishonest ISPs.* Right: *The CDF of gain for honest ISPs.*

Figure 8: *Efficiency of* Wiser *and gain for individual ISPs with two synthetic cost models.* Left: *The CCDF of path inflation.* Right: *The CDF of gain for individual ISPs.*

bound on the usage cost ratio. We choose a value of 0.8 for this bound for illustrative purposes.

Figure 7 shows the impact of cheating with this strategy. The graphs plot the CDF of gain for the dishonest and honest ISPs, where the gain is measured as the reduction in average distance relative to *unilateral*. For comparison, we also plot the gains for a scenario where both ISPs honestly implement Wiser and for a scenario where the dishonest ISP can cheat arbitrarily because the normalization and usage cost ratio constraints do not exist. We see that with Wiser the curves for one dishonest ISP are close to the case where both ISPs are honest. This implies that Wiser limits the gain for the dishonest ISP and the loss for the honest ISP. We also see higher gain and loss when the constraints imposed by Wiser are removed, which further indicates their effectiveness.

6.4 Understanding the Efficiency of Wiser

We now switch from evaluating Wiser over realistic inputs to explore in more general terms the ISP cost models and topologies for which it can provide efficient routing.

6.4.1 Impact of ISP cost models

We experiment with synthetic cost models to understand under what circumstances efficient end-to-end paths are produced. We know that Wiser produces efficient routes with inferred link weights that model ISPs' costs today. But it will not necessarily do so for arbitrary models of internal ISP costs. To probe this issue, we first consider a scenario where we assume that each ISP has an unknown (to us) objective, and randomly assign costs to each link from a finite range. The solid curves in Figure 8 show the results. The left graph plots the CCDF of path inflation relative to *global*, as computed in § 6.1.1. The right graph plots the individual gain for ISPs, measured as the average reduction in cost of carrying traffic. The graphs show that the quality of end-to-end paths with random costs assignment is poor even though Wiser individually benefits all ISPs.
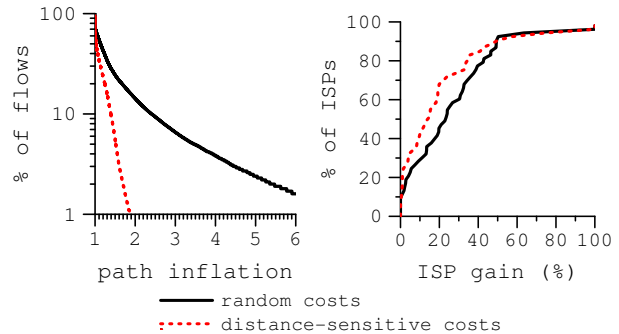
However, ISPs objectives are not arbitrary but influenced by measures of interest to them and their users, e.g., all reasonable objectives are likely to reflect path length to some extent. To evaluate Wiser in this more realistic scenario, we assume that the cost of a link inside an ISP is $\frac{1}{2}L + rL$, where L is the length of the link and r, which is a random number in the range $[0..1]$, captures the unknown components of the ISP objective, scaled to match the length component. Figure 8 shows that the end-to-end paths with these "distance-sensitive costs" are efficient and ISPs individually benefit as well.

6.4.2 Impact of topology

We now use an analytic model to understand the topological characteristics that make Wiser effective. To make this tractable, we restrict ourselves to the two-ISP base case. Generalizing the arguments presented below lead to similar inferences for the multi-ISP case [26].

Consider two ISPs, ISP-1 and ISP-W, that interconnect in N places. We model the internal ISP topology as a fully-connected mesh in which the cost of transporting a packet between two nodes is drawn from a uniform random distribution in the range $[0..1]$ for ISP-1 and $[0..W]$ for ISP-W ($W \geq 1$). W captures ISP heterogeneity: a higher W stems from a higher average cost of carrying packets inside ISP-W. The cost of transporting a packet across the two ISPs is the sum of the costs incurred inside each. This assumes that the ISPs' costs have been mapped to comparable units. The expected costs in this model with different routing methods is shown in Table 3. Their derivation is outlined in [26]. Our model is simplistic, e.g., paths between pairs of nodes are not truly independent of other paths, but we find its results to be consistent with our other experiments. It is also arguably more realistic than the only other analytic model of two-ISP routing of which we are aware [19] because it captures factors such as N and W that we show to be important.

We study the efficiency of the different routing methods as a function of W. We use cost inflation relative

Uni-lateral	$C_{unilateral}(N,W) = \frac{(W+1)(3+N)}{4(N+1)}$ $C^1_{unilateral}(N,W) = \frac{N+3}{4(N+1)}$
Global	$C_{global}(N,W) = \frac{N}{3W} \, _2F1\left(\frac{3}{2}, 1-N, \frac{5}{2}, \frac{1}{2W}\right)$ $+ \frac{N}{2(2W)^N}\left(\frac{(2W+1)((2W-1)^N-1)}{N} - \right.$ $\left.\frac{(2W-1)^{N+1}-1}{N+1}\right) + \frac{2WN+W+1}{(2N+1)(2W)^N}$ $C^1_{global}(N,W) = \frac{N}{6W} \, _2F1\left(\frac{3}{2}, 1-N, \frac{5}{2}, \frac{1}{2W}\right)$ $+ \frac{(2W-1)^N-1}{2(2W)^N} + \frac{N+1}{(2N+1)(2W)^N}$ where $_2F1(a,b,c,z) = \sum_{k=0}^{\infty} \frac{\binom{a}{k}\binom{b}{k}z^k}{\binom{c}{k}k!}$.
Wiser	$C_{Wiser}(N,W) = \frac{(W+1)C_{global}(1)}{2}$ $C^1_{Wiser}(N,W) = C_{global}(N,1)/2$

Table 3: *Expected routing costs with various routing methods. $C_m(N,W)$ is the total cost of routing using method m, and $C^1_m(N,W)$ is the cost incurred by ISP-1. The cost incurred by ISP-W is the difference of the two.*

Figure 9: *Cost inflation (left) and gain for individual ISPs (right) as a function of W with $N=6$.*

to *global* as the measure of efficiency. This captures the average inflation, not the worst-case inflation. As such, it underestimates the benefit of Wiser by discounting the impact of egregiously bad cases.

The left graph in Figure 9 plots cost inflation as a function of W, where we have selected $N=6$ to provide an example. Wiser is always more efficient than *unilateral*. It comes close to *global* for low values of W but is less efficient as W increases. To investigate this effect, the right graph plots the gain of individual ISPs with Wiser and *global*. Gain is computed as the reduction in cost relative to *unilateral*. Both ISPs gain equally with Wiser, but ISP-W gains at the expense of ISP-1 with *global*. Because ISP-W's costs are higher, globally optimal will sacrifice ISP-1's interests to the greater good. Thus, Wiser enables ISPs to cooperate without losing but the overall efficiency is less when ISPs' costs are very diverse. That the efficiency of Wiser comes close to that of *global* for realistic topologies (§6.1) suggests that the costs of ISPs that interconnect in multiple places are roughly similar for the metrics that we study.

6.5 Summary of Results

We end this section with a summary of results. For the topologies and metrics we studied, we showed that joint control of routing in Wiser comes close to ideal routing that globally optimizes end-to-end paths based on complete information. For the path length metric, while the worst 1% of that paths are inflated by a factor of 6 with today's routing practices, they are inflated only by a factor of 1.5 with Wiser. Wiser also reduces the bandwidth provisioning required by ISPs to handle the dynamic conditions that we simulated by 8% on average. We explored other ISP objectives and found that Wiser continues to
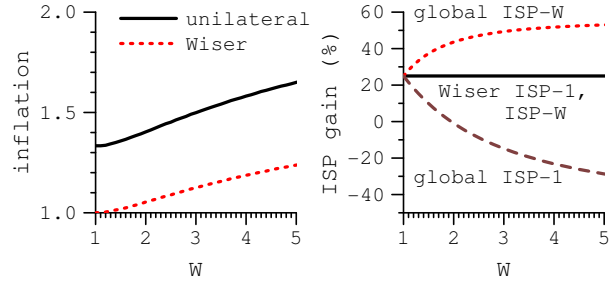
produce efficient paths as long as ISPs' objectives include factors that influence end-to-end paths, as is typically the case today. We found the overhead of Wiser to be similar to BGP in terms of implementation complexity, routing message and computational requirements. For the strategies that we studied, cost normalization and usage cost ratio constraints limit the gain for dishonest ISPs and the loss for honest ISPs. Finally, our analysis showed that the efficiency of Wiser depends on the similarity in ISPs' costs, but it is always higher than unilateral routing and, unlike optimal routing, stays win-win.

7 Discussion

The most surprising aspect of our work is perhaps that the simple mechanisms of Wiser are so effective in our experiments. Wiser obtains high levels of efficiency by combining costs over neighboring pairs of ISPs. This suggests that it is neither necessary nor particularly advantageous to construct more complex costs that are meaningful across larger groups of ISPs, e.g., global currency. This is somewhat surprising because currencies with a larger scope could allow better multi-way trades, in the same way that bigger markets tend to be more efficient. But it confers a significant practical advantage: It is much simpler to implement costs between pairs of ISPs because this mirrors the contractual structure of the Internet.

To see whether even simpler approaches would be equally efficient, we studied two alternate approaches. First, we ran Wiser with ordinal (rather than cardinal) costs to disclose less information. This is an interesting point in the design space because MEDs have ordinal semantics. We found that efficiency with Wiser using ordinal costs was little better than unilateral routing. Cardinal costs can lead to greater efficiency by using relative magnitudes to identify path changes that lead to a small loss for one ISP but a bigger gain for another. Second, we ran a variant of Wiser for pairs of flows. Wiser takes a holistic view of the traffic exchanged between two ISPs and is efficient; at the other extreme, unilateral routing considers each flow in isolation and can be inefficient. Pairs

of flows that go in opposite directions are a natural intermediate point. However, we found the efficiency of this approach to be little better than unilateral routing.

Taken together, the observations above suggest that Wiser occupies an attractive point in the design space: more complex approaches gain little efficiency; and simpler approaches lose efficiency.

Finally, one aspect of our design that we have mostly deferred to future work is stability with different cost functions. Provably stable dynamic routing in large networks is an open research question even under non-strategic behavior [6, 22, 41, 52]. Existing research provides guidelines for assigning costs in a way that enhances routing stability [3, 6, 41, 53] and which apply to our setting. We also note that information sharing and non-greedy decision making may enhance stability because it discourages ISPs from making changes that adversely affect each other.

8 Related Work

Much recent work has highlighted that BGP provides poor control over routing and computes inefficient paths [4, 11, 19, 36, 40, 44, 50]. Our work shows how these problems can be addressed while being consistent with ISP interests. Stability of BGP under different commercial policies has also been scrutinized [16]. Wiser finds paths within commercial preferences and thus neither helps nor hurts on that account except for removing MED-induced oscillations [29].

Wiser builds on our earlier work on Nexit, a framework by which two ISPs can negotiate routes [28]. Unlike Nexit, Wiser handles the general case of multiple ISPs and has a much lower overhead. It accomplishes this while preserving ISP interests in the same manner and disclosing similar amount of information.

Other work has explored optimizing interdomain routing using existing BGP knobs [11, 17, 37, 51]. While this helps, it has limited effectiveness because ISPs lack visibility outside their own network and have little incentive to suffer for the greater good. Wiser tackles these root problems directly. At the other extreme, work on interdomain quality-of-service (QoS) requires full cooperation between ISPs, including the disclosure of sensitive information and agreement on the optimization metric [10]. Wiser eschews this high degree of cooperation to preserve ISP interests and facilitate deployment.

Researchers have also explored monetary payments for carrying traffic along specified routes [2, 32]. This requires ISPs to disclose monetary costs at a fine granularity; approaches based on mechanism design [13] can encourage the disclosure of true costs via strategy-proof (but not budget-balanced) mechanisms. Regardless, monetary costs are difficult to compute [42], can leak sensitive information that can be used to undercut the market outside of the mechanism, and are not compatible with the current "customer pays" charging model that is independent of the direction of traffic. Wiser retains the current charging model to be practical, and our results also suggest that monetary costs are not necessary for better efficiency.

Finally, Wiser is similar in spirit to other work on competitive interests. Like BitTorrent [9], Wiser uses bilateral coordination and favors practicality over the prevention of cheating [43]. Like work on load management in federated systems [5], Wiser leverages offline and bilateral contracts to simplify online operation.

9 Conclusions

Our work shows that, at least for Internet routing, competing interests can be harnessed using practical protocols and without significant loss in efficiency. Wiser enables ISPs to jointly control routing and find good end-to-end, policy-compliant paths while allowing them to improve their own routing by their own reckoning. It builds on the existing contractual framework, does not require new monetary exchange, and is incrementally deployable. We evaluated Wiser via simulation over measured topologies and with XORP and SSFNet prototypes. In our experiments, Wiser was win-win and its efficiency came close to that of routes that were globally optimized with complete information. It was especially useful in automatically improving the tail of the paths which can be overly long with current routing methods. The overhead of Wiser was similar to BGP, and the built-in checks and balances encouraged ISPs to use it honestly.

Our evaluation suggests that Wiser is a promising way to provide more control to ISPs while increasing the efficiency of routing at the same time. To better understand its benefit in practice, in the future, we will conduct a more detailed investigation into models of ISPs' behaviors and the extent and frequency of routing inefficiencies that can be fixed with Wiser.

We hope that lessons from our work will prove useful in other contexts. Normalization may be broadly useful because it enables trading where there is no common basis for assigning values to commodities. Similarly, mechanisms such as usage cost ratios that reduce the degrees of freedom of individual parties may help to address other conflicts between efficiency and incentive compatibility. And the use of offline contractual clauses, which has received little attention in research, appears to be a powerful method to simplify online operation.

References

[1] M. Afergan. Using repeated games to design incentive-based routing systems. In *INFOCOM*, 2006.

[2] M. Afergan and J. Wroclawski. On the benefits and feasibility of incentive based routing infrastructure. In *PINS Workshop*, 2004.

[3] G. Apostolopoulos, *et al.* Quality of service based routing: A performance perspective. In *SIGCOMM*, 1998.

[4] D. O. Awduche, *et al.* Overview and principles of Internet traffic engineering. RFC-3272, 2002.

[5] M. Balazinska, H. Balakrishnan, and M. Stonebraker. Contract-based load management in federated distributed systems. In *NSDI*, 2004.

[6] A. Basu, A. Liu, and S. Ramanathan. Routing using potentials: A dynamic traffic aware routing algorithm. In *SIGCOMM*, 2003.

[7] S. J. Brams. *Negotiation Games: Applying game theory to bargaining and arbitration*. Routeledge, 1990.

[8] D. Clark. The design philosophy of the DARPA Internet protocols. In *SIGCOMM*, 1988.

[9] B. Cohen. Incentives build robustness in BitTorrent. In *1st Workshop on Economics of Peer-to-Peer Systems*, 2003.

[10] E. S. Crawley, *et al.* A framework for QoS-based routing in the Internet. RFC-2386, 1998.

[11] N. Feamster, J. Borkenhagen, and J. Rexford. Guidelines for interdomain traffic engineering. *CCR*, 33(5), 2003.

[12] N. Feamster, *et al.* The case for separating routing from routers. In *FDNA Workshop*, 2004.

[13] J. Feigenbaum, *et al.* A BGP-based mechanism for lowest-cost routing. In *PODC*, 2002.

[14] B. Fortz and M. Thorup. Internet traffic engineering by optimizing OSPF weights. In *INFOCOM*, 2000.

[15] T. Griffin and B. Premore. An experimental analysis of BGP convergence time. In *ICNP*, 2001.

[16] T. Griffin and G. T. Wilfong. An analysis of BGP convergence properties. In *SIGCOMM*, 1999.

[17] Internap Flow Control Xcelerator. http://www.internap.com/product/technology/fcx/.

[18] P. Jacob and B. Davie. Technical challenges in the delivery of interprovider QoS. *IEEE Communications Magazine*, 43(6), 2005.

[19] R. Johari and J. N. Tsitsiklis. Routing and peering in a competitive Internet. T.R. P-2570, MIT LIDS, 2003.

[20] K. Johnson, *et al.* The measured performance of content distribution networks. In *Int'l Web Caching and Content Delivery Workshop*, 2000.

[21] S. Kandula, *et al.* Walking the tightrope: Responsive yet stable traffic engineering. In *SIGCOMM*, 2005.

[22] A. Khanna and J. Zinky. The revised ARPANET routing metric. In *SIGCOMM*, 1989.

[23] B. Kruglov. Re: Cogent and level3 peering issues. NANOG mailing list archives: http://www.merit.edu/mail.archives/nanog/2002-12/msg00379.html, 2002.

[24] S. M. Kusiak. Re: Congestion peering C&W <-> @home. NANOG mailing list archives: http://www.merit.edu/mail.archives/nanog/2001-11/msg00282.html, 2001.

[25] C. Labovitz, *et al.* An experimental study of delayed Internet routing convergence. In *SIGCOMM*, 2000.

[26] R. Mahajan. *Practical and Efficient Internet Routing with Competing Interests*. Ph.D. thesis, University of Washington, 2005.

[27] R. Mahajan, D. Wetherall, and T. Anderson. Understanding BGP misconfiguration. In *SIGCOMM*, 2002.

[28] R. Mahajan, D. Wetherall, and T. Anderson. Negotiation-based routing between neighboring domains. In *NSDI*, 2005.

[29] D. McPherson and V. Gill. BGP MED considerations. Internet draft, 2005.

[30] CA*net routing policy. http://www.canarie.ca/canet4/services/c4_routing_policy.pdf, 2003.

[31] R. Miller. Legal battle ended for AT&T, MCI. InternetNews.com, 2004. http://www.internetnews.com/xSP/article.php/3316751.

[32] R. Mortier and I. Pratt. Incentive based inter-domain routeing. In *ICQT Workshop*, 2003.

[33] R. B. Myerson and M. A. Satterthwaite. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29(2), 1983. Cited in Brams [7].

[34] W. B. Norton. Internet service providers and peering. Equinix whitepaper, version 2.5, 2001. http://www.equinix.com/pdf/whitepapers/PeeringWP.2.pdf.

[35] V. N. Padmanabhan and L. Subramanian. An investigation of geographic mapping techniques for Internet hosts. In *SIGCOMM*, 2001.

[36] B. Quoitin, *et al.* Interdomain traffic engineering with bgp. *IEEE Communications Magazine*, 41(5), 2003.

[37] B. Quoitin, *et al.* Interdomain traffic engineering with redistribution communities. *Computer Communications Journal*, 27(4), 2004.

[38] H. Raiffa. *The art and science of negotiation*. Harvard University Press, 1982.

[39] T. Roughgarden and E. Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2), 2002.

[40] S. Savage, *et al.* The end-to-end effects of Internet path selection. In *SIGCOMM*, 1999.

[41] A. Shaikh, J. Rexford, and K. G. Shin. Load-sensitive routing of long-lived IP flows. In *SIGCOMM*, 1999.

[42] S. Shenker, *et al.* Pricing in computer networks: Reshaping the research agenda. *CCR*, 26(2), 1996.

[43] J. Shneidman, D. C. Parkes, and L. Massoulie. Faithfulness in Internet algorithms. In *PINS Workshop*, 2004.

[44] N. Spring, R. Mahajan, and T. Anderson. Quantifying the causes of path inflation. In *SIGCOMM*, 2003.

[45] N. Spring, *et al.* Measuring ISP topologies with Rocketfuel. *IEEE/ACM ToN*, 12(1), 2004.

[46] Scalable simulation framework. http://www.ssfnet.org/.

[47] S. Stickland. Utilising upstream MED values. NANOG mailing list archives: http://www.merit.edu/mail.archives/nanog/2005-03/msg00400.html, 2005.

[48] L. Subramanian, *et al.* Characterizing the Internet hierarchy from multiple vantage points. In *INFOCOM*, 2002.

[49] N. Taft, *et al.* Understanding traffic dynamics at a backbone PoP. In *SPIE ITCOM*, 2001.

[50] H. Tangmunarunkit, *et al.* The impact of routing policy on Internet paths. In *INFOCOM*, 2001.

[51] S. Uhlig and O. Bonaventure. Designing BGP-based outbound traffic engineering techniques for stub ASes. *CCR*, 34(5), 2004.

[52] Z. Wang and J. Crowcroft. Analysis of shortest-path routing algorithms in a dynamic routing network environment. *CCR*, 22(2), 1992.

[53] Z. Wang and J. Crowcroft. Quality-of-service routing for supporting multimedia applications. *IEEE JSAC*, 14(7), 1996.

[54] XORP: Open source IP router. http://www.xorp.org/.

[55] Y. Zhang, *et al.* Fast accurate computation of large-scale IP traffic matrices from link loads. In *SIGMETRICS*, 2003.