The Economic Implications of Edge-Directed Routing: A Network Operator's Perspective

Patrick Kwadwo Agyapong*†
pagyapong@cmu.edu
* Instituto Superio Técnico, Lisbon, Portugal

Marvin Sirbu †
sirbu@cmu.edu
† Carnegie Mellon University, Pittsburgh, PA 15213

***Abstract*—Edge-directed routing, a paradigm where sources and sinks of traffic, rather than the network, specify the communication path has recently gained attention as a means to deal effectively with potential conflicts that may arise between various stakeholders in the future Internet. In this paper, we use a simple economic model to show that contrary to current economic thinking, networks can deploy edge-directed routing without raising prices, provided that the service results in a relative increase in external traffic that outweighs the relative increase in costs. However, when edge-directed routing merely results in traffic shift from one path to another, then price increases are required to make it economically viable. Hence, we recommend that edge-directed routing protocols put in place payment mechanisms and support systems in their design to facilitate various pricing strategies.**

***Index Terms*—edge-directed routing; future Internet architecture; economic incentives.**

I. Introduction

Armed with lessons from almost two decades of commercial operation of the Internet, network architects realize the need to build a future Internet that is fundamentally secure and flexible enough to support diverse uses. At the same time, there is increasing awareness that, in addition to providing the right economic incentives to stimulate adoption, any new architecture must possess mechanisms to deal effectively with conflicts that may arise between various stakeholders.

It is not uncommon for an end-host's objectives to conflict with those of its provider. Unfortunately, the original architecture of the Internet lacks effective mechanisms to address such conflicts. In particular, de-facto network control over routing precludes an end-host from utilizing alternate paths to the default network-provider path. Source routing was introduced in IPv4 and IPv6 to allow sources of traffic to override default network paths. However, the majority of routers on the Internet do not support this functionality for a variety of reasons [1], [2]. Consequently, end-hosts resort to ad hoc fixes, such as overlay networks, built on top of the Internet in order to achieve performance and other security goals [3], [4].

One way to deal with conflicts between various stakeholders is to design future Internet architectures to explicitly support choice and/or competition at all levels of the architecture [5], [6]. Designing for choice implies that the architecture allows different stakeholders to express their preferences for different services [5]. On the other hand, designing for competition refers to the ability of stakeholders to "express their preferences for services by different providers" [6]. To illustrate, edge networks can currently express a preference for a first-hop network provider. Nothing prevents an edge from having multiple first-hop providers, who provide the same or different services, if desired. Hence, one can conclude that the Internet is designed for both choice and competition with regards to first-hop providers. However, an edge network has no binding input in the decisions about the intermediate networks used to transport packets, once a first-hop provider has been chosen.

Likewise, the destination or sink has no binding input in how packets are routed to it, even when it has multiple first-hop providers. In order words, the Internet is not designed for choice with regards to the nth-hop provider for both sources and sinks, when $n > 1$. This lack of nth-hop provider choice is the root cause of some of the conflicts between different stakeholders. For instance, the current Internet architecture does not support competition for the provision of intermediary services, such as filtering or virus scanning, by entities that are not first-hop providers because a sink cannot guarantee that all of its packets pass through that intermediary.

The above realization has motivated several architecture proposals that emphasize nth-hop provider choice and competition in routing [7]–[10]. It has also led to the development of a number of standalone routing protocols that provide path choice to end-hosts (e.g., [11]). We refer to these proposals collectively as edge-directed routing protocols.

Previous studies on edge-directed routing primarily focus on two main areas. The first area of work develops protocols to support edge-directed routing capabilities (e.g., [7]–[13]), whereas the second studies the equilibrium properties of networks built around edge-directed routing (e.g., [14]–[17]). Unfortunately, most studies to date have largely ignored incentive compatibility, which is vital to the successful adoption of any architecture. Specifically, previous work fail to provide any useful analysis of the economic incentives of different stakeholders to deploy and use architectures based on nth-hop provider choice.

The few studies in this area focus on understanding how network operators can maintain control over the flow of traffic through their networks when they relinquish control of path choices to the edge. For instance, Masuda and Whang use a linear programming formulation to show that networks can achieve any desirable traffic flow by using route pricing, node pricing or source-sink pricing [18]. They further show that the traffic flow that results from route or node pricing maximizes social welfare for all network participants. Kelly [19] and Laskowski *et al.* [20] obtain similar results.

However, neither of these studies addresses how such a pricing mechanism is determined, the welfare effects if prices are inaccurately determined, or the information requirements needed to implement the pricing mechanism. It is very likely that the costs due to the information requirements associated with implementing elaborate pricing schemes alone will erode any benefits obtained from edge-directed routing [21]. Additionally, neither study sheds any light on whether ISPs will be better off financially from deploying edge-directed routing, even with perfectly determined route prices. Without any solid understanding of the financial risks, network operators will be reluctant to deploy edge-directed routing protocols.

In this paper, we build a simple economic model to investigate the incentives of a network operator to deploy edge-directed *inter-domain* routing. Specifically, we investigate the conditions under which network operators are better off from deploying edge-directed routing, taking into account a few pricing strategies that are relatively easy to implement. Our model shows that contrary to the arguments usually advocated in literature (e.g., [5], [20]), networks may be able to deploy edge-directed routing without increasing prices, provided that edge-directed routing results in a relative increase in external traffic that is at least equivalent to the relative increase in costs.

On the other hand, we show that if edge-directed routing merely shifts traffic from one path to another, then price increases will be required to make it economically viable. Thus, edge-directed routing protocols must incorporate payment mechanisms to support different pricing strategies. Interestingly, we also show that a flat-rate fee, with very little information requirements, will be enough to make edge-directed routing economically viable under most circumstances.

The rest of the paper is organized as follows. In Section II, we describe our network and cost model. Next, we use the model to derive the conditions that make a network operator better off from deploying edge-directed routing. We follow this with a discussion of real-world deployment issues and the implications for edge-directed routing protocol design in Section III. We conclude and discuss some future research directions in Section IV.

II. Model and Assumptions

Even though edge routing capability can be implemented at either the end-host or autonomous system (AS) level, it may be more practical to implement it at the AS level for scalability reasons. Besides, a large fraction of end-users consider the default network-selected paths sufficient for their needs and will therefore delegate path selection to the AS [11]. Moreover, it is likely that entities who are willing to pay for the capabilities of edge-directed routing consist of enterprises, educational institutions and other large organizations that can be treated as ASes. One could imagine a scenario where such an entity implements edge-directed routing on behalf of end-hosts within its control. Because of these reasons, the discussions which follow assume that edge-directed routing is implemented at the AS level.

We consider a scenario where a transit network provider, Network C, provides connectivity to N customers, who send traffic to various destinations. There exist potentially multiple paths to reach each destination outside Network C. We refer to destinations that are not customers or peers of Network C as external destinations. We consider only a single external destination for simplicity. An extension to multiple destinations is straightforward and leads to the same results. We think of each distinct path to an external destination as a different routing service (RS). We assume that Network C has the potential to offer M different routing services to reach an external destination and represent the RS set by $S = \{S^0, S^1, S^2, \ldots, S^M\}$. We illustrate this in Figure 1. In the figure, Network C has the potential to offer four routing services for its customers to reach Network L, namely $CDFIJL$, $CDFIKL$, $CEHIJL$ and $CEHIKL$.

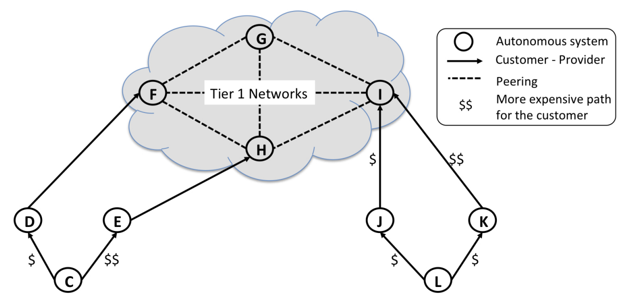

Fig. 1. Network C has the potential to offer several paths to reach a particular destination. Each different path can be considered as a routing service (RS). In the above, Network C has the potential to offer four routing services to reach Network L, namely $CDFIJL$, $CDFIKL$, $CEHIJL$ and $CEHIKL$.

Network C has two options for routing packets. It could advertise a default RS, S^0, and select the actual RS used to route packets on behalf of its customers or allow customers to specify an RS from S. For a rational network provider, S^0 is the path that minimizes its external connectivity costs. In our example in Figure 1, $CDFIKL$ is the default RS to reach Network L from Network C. This follows from the fact that Network I will only propagate path IKL in BGP because it earns more revenues when it routes traffic to L through path IK. In general, Network C can predict its own costs involved in advertising only S^0. However, the costs that Network C incurs when it allows its customers to choose from S are not well understood. In addition, it is not obvious how Network C can recover additional costs that may arise from this routing flexibility. In what follows, we provide some insights into the cost implications of allowing customers to choose from S.

Let us assume that Network C faces a total demand for external connectivity, in packets per second, given by $d = \sum_{i=1}^{N} d_i[S]$, where $d_i[S]$ is the demand for external connectivity from customer i. In the rest of the paper, we use $f[x]$ to indicate that f is a function of x. Unlike most previous work, which assumes that demand is independent of the number of available paths through the network, we explicitly make a provision for this dependency. We believe that the existence of some paths may alter traffic demand for various reasons. For instance, a customer concerned about traffic monitoring will choose to send more traffic when some perceived trustworthy paths become available. We assume that Network C has provisioned its network to support a total demand for external connectivity d.

In Figure 2, we show the costs that Network C incurs to route packets. In the short term, some costs such as, local loop to reach the customer, internal bandwidth within Network C and equipment, do not depend on the volume of external traffic. We denote these costs, in dollars, by K. On the other hand, external connectivity costs depend on both the volume of external traffic, d, and the path, since each path has an associated cost. Thus, we denote the external bandwidth costs, in dollars per byte transferred, by $B\left[d[S]\right]$. Additionally, Network C incurs costs, which increase with M, to distribute and manage S and to account and bill for usage. Finally, we assume that Network C faces a twice differentiable and additive cost function of the form $C\left[B[d[S]], S, K\right] = C\left[B[d[S]]\right] + C\left[S, K\right]$, which satisfies $\frac{\partial C}{\partial d} = C'\left[B[d[S]]\right]|_{d=d_0}$ and $\frac{\partial^2 C}{\partial d^2} = C''\left[B[d[S]]\right]|_{d=d_0} < 0$. The latter captures economies of scale in bandwidth costs.

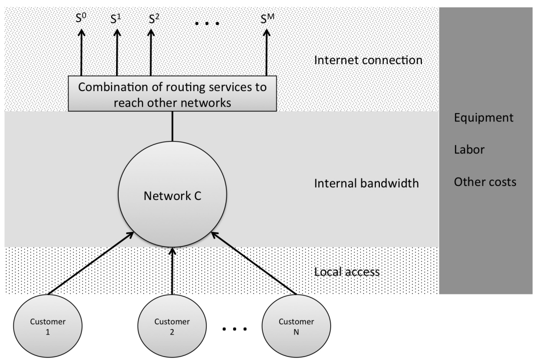

Fig. 2. Network C incurs several costs to provide both external and internal connectivity to its customers. External connectivity costs depend on the path chosen, whereas internal connectivity costs are assumed to be under Network C's control and fixed in the short-term.

A. Network-Controlled Routing

The cost to meet a total demand of d packets per second, in dollars, using S^0 is given by $\text{Cost}^{nc} = C\left[B[d[S^0]], S^0, K\right]$. Due to the presence of both fixed and variable components in the cost structure, we assume that Network C uses a two-part pricing strategy to recover its costs [22]. The choice of a two-part pricing strategy simplifies our analysis, but does not affect our results and is consistent with the interconnection scenario in the Internet [23]. The two-part price consists of a fixed access charge over period T, $E_i[S^0, K]$, in dollars, and a usage-based charge, $W_i\left[d_i[S^0]\right]$, in dollars per byte transferred by customer i. One could think of the fixed component as the flat rate fee paid for internal connectivity. Thus, over a period of T seconds, customer i pays a total of $E_i[S^0, K] + W_i\left[d_i[S^0]\right] d_i[S^0] FT$ dollars for connectivity, where F denotes the average packet size in bytes.

To reflect the situation in the current Internet peering ecosystem, we assume that the market for external connectivity is competitive. We do not make any assumptions about how Network C sets $E_i[S^0, K]$ for each customer, but we assume that $\sum_i^N E_i[S^0, K] \geq C\left[S^0, K\right]$. Given a competitive market scenario, Network C cannot set $W_i\left[d_i[S^0]\right]$ greater than the standalone costs that customer i will incur to provide the same service. Otherwise, customer i will have incentives to contract with another provider or provide the service itself. A rational cost-minimizing customer will choose S^0 when it undertakes to provide its own external connectivity. Therefore, it is reasonable for Network C to set $W_i\left[d_i[S^0]\right]$ as

$$W_i\left[d_i[S^0]\right] = \frac{C\left[B\left[d_i[S^0]\right]\right]}{d_i\left[S^0\right] FT} \ , \qquad (1)$$

where $C\left[B\left[d_i[S^0]\right]\right]$ is the cost to satisfy customer i's external bandwidth demand. In order to capture the fact that Internet transit pricing exhibits economies of scale in volume, we assume that the cost function for external bandwidth is strictly sub-additive, i.e., $C\left[B\left[d[S^0]\right]\right] < \sum_i^N C\left[B\left[d_i[S^0]\right]\right]$. Sub-additivity in external bandwidth costs and $\sum_i^N E_i(S^0, K) \geq C\left[S^0, K\right]$ necessarily implies that Network C obtains a positive profit, that is, $\Pi^{nc} > 0$.

B. Edge-Directed Routing

In edge-directed routing, Network C makes S available to customers and does not impose any restrictions on the path that can be used to reach any destination. We consider a scenario where Network C has provisioned its network to support S based on some *a priori* assumptions about d. We do not address the question of how Network C selects S. Rather, we set out to find the forms that $W_i\left[d_i[S]\right]$ and $E_i[S, K]$ should take, in order to ensure that Network C is not worse off when it offers edge-directed routing based on S.

The cost and revenue for Network C when it offers edge-directed routing are given respectively by $\text{Cost}^{ec} = C\left[B[d[S]], S, K\right]$ and $\text{Revenue}^{ec} = FT\sum_i^N d_i\left[S\right] W_i\left[d_i[S]\right] + \sum_i^N E_i[S, K]$, and the resulting profit is given by Π^{ec}. In order to make edge-directed routing worthwhile, we require that $\Pi^{ec} \geq \Pi^{nc}$. If we consider the simple case where $\sum_i^N E_i[S^0, K] = C\left[S^0, K\right]$, then $\Pi^{ec} \geq \Pi^{nc}$ implies that

$$FT\sum_i^N d_i\left[S\right] W_i\left[d_i[S]\right] + \sum_i^N E_i[S, K] - C\left[B[d[S]], S, K\right] \geq \sum_i^N C\left[B\left[d_i[S^0]\right]\right] - C\left[B\left[d[S^0]\right)\right] . \qquad (2)$$

One could think of potentially infinite ways to set $W_i\left[d_i[S]\right]$ and $E_i[S, K]$ to satisfy (2). In what follows, we consider three pricing strategies and evaluate the conditions under which those strategies satisfy (2). In the first, Network C charges the same prices for both edge-directed and network-controlled routing. In the second, Network C introduces some increment on the fixed price component. In the third, the usage-based component of the price is based on the customer's standalone costs incurred to use the requested RS.

Case a : Keep same pricing as network-controlled routing: In this case, Network C sets $W_i\left[d_i[S]\right] = \frac{C\left[B\left[d_i[S^0]\right]\right]}{d_i[S^0] FT}$ and $E_i[S, K] = E_i[S^0, K]$. The constraint in (2) becomes

$$\sum_i^N C\left[B\left[d_i[S^0]\right]\right]\left(\frac{d_i[S]}{d_i[S^0]} - 1\right) \geq \Delta_B + \Delta_E \ , \qquad (3)$$

where $\Delta_E = C[S,K] - \sum_i^N E_i[S^0,K]$, is the difference between the fixed costs in edge-directed and network-controlled routing and $\Delta_B = C[B[d[S]]] - C\left[B\left[d[S^0]\right]\right]$ is the difference between external bandwidth costs in edge-directed and network-controlled routing.

We see from (3) that the terms on the right side are positive. This follows from the fact that $C[S,K] > C\left[S^0,K\right]$. Also, Δ_B is zero only when all transit links are priced equally and traffic distribution is symmetric on all links with and without edge-directed routing, which is highly unlikely. Given this observation, we immediately recognize one scenario where (3) fails to hold. If the existence of a large set of path choices merely shifts demand from one path to another without increasing the total external demand, then $d_i[S] = d_i[S^0]$, which makes the term on the left side of (3) equal zero. Thus, in such a scenario, Network C is worse off when it deploys edge-directed routing without raising prices.

One could imagine that a larger set of path choices increases external traffic flowing through Network C, but does not affect total external traffic flowing through all networks. This could happen when customers move traffic away from self-provisioned (or contracted) links to Network C. If the increase in demand is large enough, then it is likely that Network C will be better off when it offers edge-directed routing without raising prices. In the short-term, Network C will obtain a competitive advantage by doing so. However, in the long-term, offering path choices will become a competitive necessity which will drive down industry-wide profits. Still, it is possible that a larger set of path choices increases external traffic flowing through the entire network. Under such a scenario, all networks will eventually find it a competitive necessity to deploy edge-directed routing without increasing prices, provided that the cost increase for doing so are smaller than the value of increased demand resulting from offering edge-directed routing. Unlike the previous case, however, it may be possible for all networks to maintain their profit margins in the long-run if external traffic increases across the entire network.

To obtain some idea about the kind of traffic increase required to make the provision of edge-directed routing justifiable without price hikes, we consider a simple case where all customers of Network C send the same volume of external traffic and pay the same prices for usage and access. Under this special case, constraint (3) can be expressed as

$$\frac{d_i[S]}{d_i[S^0]} \geq 1 + \frac{\phi C\left[B\left[d[S^0]\right], S^0, K\right]}{NC\left[B\left[d_i[S^0]\right]\right]}\ , \tag{4}$$

where $\phi = \frac{\Delta_B + \Delta_E}{C[B[d[S^0]], S^0, K]}$ is the relative change in total costs as a result of edge-directed routing. Furthermore, we can express the total costs for Network C as

$$C\left[B\left[\cdot\right], S^0, K\right] = N\left(C\left[B\left[\cdot\right]\right] + E_i\left[S^0, K\right]\right)(1-\varsigma), \tag{5}$$

where $\varsigma = \frac{\Pi^{nc}}{\text{Revenue}^{nc}}$ is the profit margin that Network C obtains when it undertakes network-controlled routing.

Let us define $\delta = \frac{d_i[S] - d_i[S^0]}{d_i[S^0]}$ as the relative increase in external demand as a result of edge-directed routing. Based on the pricing structure assumed in our model, we can also make the approximation $\frac{C\left[B\left[d_i[S^0]\right]\right] + E_i\left[S^0,K\right]}{C\left[B\left[d_i[S^0]\right]\right]} \approx \frac{C\left[B\left[d[S^0]\right], S^0, K\right]}{C\left[B\left[d[S^0]\right]\right]}$. When changes in total costs are mostly due to changes in external bandwidth costs, which is very reasonable when $d \leq D$, then we can derive δ from (4) and (5) as

$$\delta \geq \frac{\Delta_B}{C\left[B\left[d[S^0]\right]\right]}(1-\varsigma)\ . \tag{6}$$

Equation (6) simply states that when profit margins are close to zero, then the relative increase in external traffic must be at least equal to the relative increase in bandwidth costs in order to justify the deployment of edge-directed routing without a price increase in a well-provisioned network. For instance, when external bandwidth costs increase by 10%, then external traffic must also increase by at least 10% in order to make edge-directed routing worthwhile in the absence of a price increase. It is interesting to note that networks that only possess peering links (Tier 1 networks) do not suffer any adverse economic consequences from edge-directed routing, as long as, it does not reduce traffic flow.

Case b: Increase fixed component of price: In a scenario where the increase in external traffic from S is not large enough to overcome the additional costs associated with edge-directed routing, Network C could increase the fixed access component of the price it charges to customers. This scenario is likely to occur when edge-directed routing leads to significantly higher fixed costs (e.g. due to the need for the network to increase capacity to accommodate traffic) and/or external bandwidth costs. The idea here is that Network C sets $E_i[S,K]$ in such a way that $\sum_i^N E_i[S,K] \geq C[S,K]$, while keeping $W_i\left[d_i[S]\right] = \frac{C\left[B\left[d_i[S^0]\right]\right]}{d_i[S^0]FT}$.

In order to prevent distortion effects in pricing, we consider the case where $\sum_i^N E_i[S,K] = C[S,K]$. In setting $E_i[S,K]$, Network C could distribute the recovery burden equally among all customers or target customers who make use of edge-directed routing. It may be desirable to target the latter, in order to prevent other customers from switching to other network providers. Compared to the earlier discussion, we see that we require a smaller increase in external traffic in order to satisfy (3) when $\Delta_E = 0$.

Case c: Charge based on actual RS used: When the two pricing schemes described above fail to work, then Network C needs to charge path-based usage prices in order to make edge-directed routing economically viable. This is the conclusion reached in previous studies such as [18]–[20]. Unlike these studies, we go a step further to suggest the forms that the prices could take. For instance, the fixed access charges could be set such that $\sum_i^N E_i[S,K] = C[S,K]$, with the burden preferably shifted to customers who make use of edge-directed routing. If the network can estimate the demand from customers that will result from S, then it can set usage-based prices equal to the standalone costs for the customers to acquire the service. In other words, Network C sets $W_i\left[d[S]\right] = \frac{C[B[d_i[S]]]}{d_i[S]FT}$. Customers who obtain a utility from S will have incentives to pay the increased usage-based price, whereas other customers pay the price for using the default network path.

To summarize, our analysis suggests that edge-directed routing protocols that do not introduce significant fixed costs can be deployed without any increase in prices, provided that they lead to a relative increase in external traffic roughly equivalent to the relative increase in external bandwidth costs. Even when fixed costs increase, networks could recover these costs by raising fixed access charges for customers who utilize edge-directed routing. From the network provider's point of view, such pricing schemes are desirable because they incur minimum overhead to account and bill for usage. Hence, designers must minimize the fixed costs associated with edge-directed routing protocols in order to provide the maximum incentives for networks to deploy them.

Our discussion so far has focused on Network C as the provider of edge-directed routing. In truth, the arguments we have made and the pricing strategies we have discussed apply recursively from Tier 1 Networks ($F, G, H,$ and I in Figure 1) to Network C. These Tier 1 networks offer edge-directed routing to their customers based on an RS set S'. Our results suggest that Tier 1 networks can offer a large RS set, since by definition, they do not pay for transit. The customers may then act as providers to other networks and offer edge-directed routing based on an RS set S, where $S \in S'$. Thus, our model is consistent with future Internet architectures like SCION (see [10]), which relies on a top-down notion of trust (and payment) relationships among networks.

III. Real World Deployment Issues

In this section, we identify key features required to support the deployment of edge-directed routing. After this, we highlight some business opportunities enabled by edge-directed routing and the potential costs to exploit them.

A. Features Required to Support Edge-Directed Routing

We identify five features necessary for commercial deployment of any edge-directed routing protocol namely *knowledge*, *choice*, *enforcement*, *metering* and *verification*.

1) Knowledge: First, edge-directed routing protocols must provide a scalable means for edges to obtain knowledge of a set of secure and policy-complaint paths to a given destination. In particular, the process of path discovery must recognize and address the business needs of ISPs. For example, ISPs may want to control the diversity of alternative paths exposed to the edge. Even though most previous work discuss ways to provide knowledge about paths, more work is needed to design scalable and incentive-compatible knowledge dissemination mechanisms for edge-directed routing.

2) Choice: Secondly, edge-directed routing protocols must equip the source, sink or both with the capability to specify a policy-compliant path. Additionally, there must be mechanisms for intermediate networks on the path to consent to the use of the specified paths. In our model, the latter requirement was achieved by making S a subset of S', but that is not the only way one could meet this requirement.

In order to overcome some of the security issues that plagued source routing in IP networks, we note that the protocol must provide **both** the source and the sink with the joint responsibility to specify a policy compliant end-to-end path. This is the approach taken in proposals such as SCION [10] and ICING [9]. Alternatively, the protocol could insist that sinks find an independent path back to the source. It may be more desirable to give joint responsibility to the source and sink because this allows both parties to effectively deal with scenarios where they have conflicting requirements.

3) Enforcement: Thirdly, the protocol must provide a scalable means for the routers in origin and transit networks to enforce that the ASes to which they belong have approved the use of a specific network resource. This is especially important if networks charge for edge-directed routing, since the network needs to distinguish between paying and non-paying customers. This could be achieved using technical mechanisms such as cryptography embedded within the protocol or with other means outside the protocol such as admission control. This area, which was largely ignored in the past, has attracted some attention recently [9], [10]. For example, SCION uses an *opaque field* in the path construction beacon to account for the use of specific resources [10]. Similarly, ICING uses *proofs of consent* (PoCs), issued by networks on the path, as a means to enforce the use of network resources [9].

4) Metering: In addition to enforcement, the network needs to measure and track the amount of traffic that customers send in order to implement usage-based pricing. The information requirements associated with different pricing strategies will determine the viability of offering services based on that pricing model. In general, path-based pricing will require the most information to implement. Thus, ISPs will be reluctant to offer services based on this pricing model unless the information costs are negligible. However, designing metering schemes with negligible information costs poses a huge challenge. To date, most of the literature fail to address this important issue.

5) Verification: Whereas enforcement allows the network provider to account for resource use, verification enables the customer to attest that a requested edge-directed routing service was delivered. This requirement is particularly important when the customer pays for the requested service. It has been shown by Laskowski and Chuang that, at a minimum, *contractible monitors*, defined as “a distributed algorithmic mechanism that runs on the network graph, and outputs, to specific nodes, proofs about current or past network behavior that can serve as input to a contract”, must be present in order to induce networks to deploy innovative technologies [24]. Most edge-directed routing protocols lack any features for verification, thereby dampening any incentives for networks to deploy and use them.

Based on the above discussion, we reckon that edge-directed routing protocols that support all the above features could lead to significant packet and communication overhead. In ICING for instance, a random packet is expected to have about 45% packet overhead [9]. Further, the designers estimate that ICING will add about 23% overhead to total traffic, which is quite significant [9]. Indeed, the extra usage cost associated with path choice may well be due to the increase in overhead, which means that the network provider has to acquire and pay more for external bandwidth, even when all path choices cost the same per byte. Hence, an analysis of typical communication overhead incurred to provide the features above for proposed protocols will prove very useful to understand the viability of some business opportunities.

B. Business Opportunities

A secure and policy-compliant edge-directed routing protocol could enable providers to deploy and monetize new and enhanced routing capabilities. Business models based on charging a premium for edge-directed routing require mechanisms to distinguish between paying and non-paying customers. They also depend on the existence of verification mechanisms or other notions of trust between the network provider and the customer [24].

Unlike network-controlled routing, edge-directed routing guarantees nth-hop provider choice for $n > 1$. This capability opens up the possibility to purchase network services such as filtering, DDoS protection and virus scanning from entities other than the first-hop network provider [7], [9]. Without edge-directed routing, edges either have to purchase such services from the first-hop provider or through other providers that use BGP tricks to appear as a first-hop provider. The main limitation with purchasing middlebox services in this way is that edges are consigned to only a single middlebox service provider and cannot selectively choose which packets avoid the middlebox.

On the contrary, edge-directed routing allows a sink to simply specify a path that contains the middlebox provider for a subset of packets that requires the use of the service. The sink could even contract with multiple middlebox providers for different classes of services and dynamically decide which packets go through which middlebox service. This approach provides greater flexibility and choice to edge networks. Nonetheless, this business opportunity imposes new costs that could instigate new conflicts among Internet stakeholders and needs to be investigated further.

IV. Conclusion and Future Work

In this paper, we have shown that it may be possible for network providers to deploy edge-directed routing without raising prices, provided that it results in a relative increase in external traffic that is at least equivalent to the relative increase in costs. However, when edge-directed routing merely shifts the same external traffic from one path to another, then a price increase is required to justify its deployment. For this, we have shown that a flat-rate increase, which requires relatively little information to implement, will suffice in most cases. Nevertheless, it is important that edge-directed routing protocols provide payment mechanisms to support various pricing strategies. In addition, we have identified some essential features needed to support real-world deployment of edge-directed routing. These features include mechanisms to provide *knowledge*, *choice*, *enforcement*, *metering* and *verification*.

It will be interesting to investigate the relative magnitude of the welfare loss that results from using our simple pricing strategy, as opposed to route or node pricing. We plan to explore this area in our future work. We also hope to further explore the potential business opportunities that edge-directed routing could open up and identify the challenges that must be overcome in order to exploit these opportunities.

Acknowledgment

Support for this work was provided by the Fundação para a Ciência e a Tecnologia (FCT) through the CMU-Portugal Program under award SFRH/BD/33507/2008 and by NSF under award number CNS-1040801. The authors would like to thank the XIA Project Group at Carnegie Mellon University for providing useful feedback during the early stages of this work and the anonymous reviewers for their insightful comments, which helped to improve the clarity of the paper.

References

[1] S. M. Bellovin. Security Problems in the TCP/IP Protocol Suite. *SIGCOMM Comp. Comm. Rev.*, 19:32–48, Apr. 1989.

[2] Jake Edge. IPv6 Source Routing: History Repeats Itself. http://lwn.net/Articles/232781/, May 2007.

[3] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris. Resilient Overlay Networks. *SIGOPS Oper. Syst. Rev.*, 35:131–145, Oct. 2001.

[4] RFC 2547. BGP/MPLS VPNs. Technical report, Mar. 1999.

[5] D. Clark, J. Wroclawski, K. Sollins, and R. Braden. Tussle in Cyberspace: Defining Tomorrows Internet. In *Proc. of ACM SIGCOMM*, pages 347–356, Aug. 2002.

[6] J. Chuang. Loci of Competition for Future Internet Architectures. *IEEE Comm. Mag.*, 49(7):38 –43, Jul. 2011.

[7] M. Walfish, J. Stribling, M. Krohn, H. Balakrishnan, R. Morris, and S. Shenker. Middleboxes no Longer Considered Harmful. In *Proc. of the 6th Conf. on Oper. Sys. Des. & Impl. - Volume 6*, 2004.

[8] X. Yang, D. Clark, and A. Berger. NIRA: A New Inter-Domain Routing Architecture. *IEEE Trans. Net.*, 15(4):775 –788, Aug. 2007.

[9] J. Naous, A. Seehra, M. Walfish, D. Mazieres, A. Nicolosi, and S. Shenker. The Design and Implementation of a Policy Framework for the Future Internet. Technical Report TR-09-28, The University of Texas at Austin, Sep. 2009.

[10] X. Zhang, H. Hsiao, G. Hasker, H. Chan, A. Perrig, and D. Andersen. SCION: Scalability, Control, and Isolation On Next-Generation Networks. Technical Report CMU-CyLab-10-020, Mar. 2011.

[11] W. Xu and J. Rexford. MIRO: Multi-path Interdomain Routing. *SIGCOMM Comp. Comm. Rev.*, 36:171–182, Aug. 2006.

[12] B. Raghavan and A. Snoeren. A System for Authenticated Policy-Compliant Routing. *SIGCOMM Comp. Comm. Rev.*, 34:167–178, Aug. 2004.

[13] P. Godfrey, I. Ganichev, S. Shenker, and I. Stoica. Pathlet Routing. *SIGCOMM Comp. Comm. Rev.*, 39:111–122, Aug. 2009.

[14] A. Orda, R. Rom, and N. Shimkin. Competitive Routing in Multiuser Communication Networks. *IEEE Trans. Net.*, 1:510–521, Oct. 1993.

[15] K. Park, M. Sitharam, and S. Chen. Quality of Service Provision in Non-cooperative Networks with Diverse Service Requirements. *Decision Support Systems*, 28(1 -2):101 – 122, 2000.

[16] T. Roughgarden and É. Tardos. How Bad is Selfish Routing? *Journal of the ACM*, 49:236–259, Mar. 2002.

[17] L. Qiu, Y. Yang, Y. Zhang, and S. Shenker. On selfish routing in internet-like environments. *IEEE Trans. Net.*, 14(4):725 –738, Aug. 2006.

[18] Y. Masuda and S. Whang. Capacity Management in Decentralized Networks. *Manage. Sci.*, 48:1628–1634, Dec. 2002.

[19] F. P. Kelly. Charging and Rate control for Elastic Traffic. *Eur. Trans. Tel.*, 8(1):33–37, 1997.

[20] P. Laskowski, B. Johnson, and J. Chuang. User-Directed Routing: From Theory, Towards Practice. NetEcon '08, pages 1–6, 2008.

[21] D. Levinson and A. Odlyzko. Too Expensive to Meter: the Influence of Transaction Costs in Transportation and Communication. *Phil. Trans. R Soc. A*, 366(1872):2033–2046, 2008.

[22] B. M. Mitchell and I. Vogelsang. *Telecommunications Pricing: Theory and Practice*. Cambridge University Press, 1991.

[23] B. Briscoe and S. Rudkin. Commercial Models for IP Quality of Service Interconnect. *BT Technology Journal*, 23(2):171–195, Apr. 2005.

[24] P. Laskowski and J. Chuang. Network Monitors and Contracting Systems: Competition and Innovation. *SIGCOMM Comp. Comm. Rev.*, 36:183–194, Aug. 2006.