

# Virtual Multi-Homing: On the Feasibility of Combining Overlay Routing with BGP Routing

Zhi Li<sup>1</sup> and Prasant Mohapatra<sup>1</sup> and Chen-Nee Chuah<sup>2</sup>

<sup>1</sup> Computer Science Department, University of California, Davis  
{lizhi,prasant}@cs.ucdavis.edu

<sup>2</sup> Electrical and Computer Engineering Department, University of California, Davis  
{chuah}@ece.ucdavis.edu

**Abstract.** Although high diversity exists in the Internet topology, the de facto inter-domain routing protocol (BGP-4) cannot fully utilize this to provide satisfactory routing service. On the other hand, the senders have no control over how the packets will be routed. In this paper, we propose a new framework called *Virtual Multi-Homing (VMH)* to achieve source-based path selection and improve inter-domain path diversity. VMH is based on the concept of *Virtual Peering* and *Multi-Homing Overlay*. It differs from previous approaches in the following aspects: (1) It uses overlay to overcome the constraints of BGP routing. A new inter-domain relationship called *Virtual Peering* can be established between two remote ASes over the *Multi-homing Overlay Network (MON)*. This can help ASes efficiently achieve flexible inter-AS relationship and loose source based routing. (2) By interacting with BGP routing and utilizing BGP routing states whenever possible, VMH can use overlay network to achieve scalable inter-domain route discovery and maintenance without introducing duplicate work at overlay layer. (3) VMH is a complementary approach to the existing Internet routing infrastructure and can be incrementally deployed. Through simulation and theoretic analysis, we demonstrate the feasibility of the proposed approach and its effectiveness in exploring path diversity.

## 1 Introduction

Internet is composed of thousands of autonomous systems (ASes), which usually belong to different administrative domains. Each AS can choose its own interior routing protocol, such as OSPF or IS-IS. To send data traffic to hosts located in other ASes, the de facto inter-domain routing protocol, Border Gateway Protocol (BGP-4) [19], is used to exchange connectivity information across ASes.

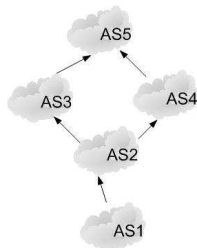
BGP is a policy-based hop-by-hop routing protocol. By manipulating (filtering) the routing information exchange with its neighboring ASes, each AS can apply its own routing policies and the inter-AS relationships (usually provider-customer or peer-to-peer) with neighbors. As a result, even though there are high redundancies [20] in the inter-domain topology, the choice of routes is often limited by the routing policies. In some cases, such as content distribution network (CDN) and service composition overlay network [18], the source AS may want to traverse or avoid specific ASes due to performance issues or potential security risks. Unfortunately, the end users and low-tier ASes currently cannot select their preferred inter-domain paths and control how the packets are routed.

Figure 1 shows an example scenario where AS1 is a customer of its provider AS2, which has two upstream providers, AS3 and AS4. To reach destination AS5, AS2 has two valid inter-domain paths based on the topology: AS2-AS3-AS5 or AS2-AS4-AS5. However, as AS2's customer, AS1 can only use one path

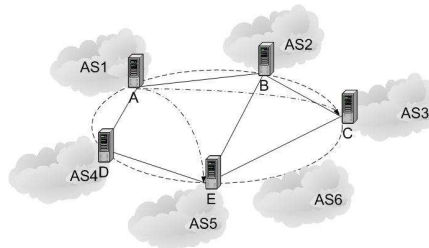
most of the time, which is usually chosen by its provider AS2. For example, the path could be AS1-AS2-AS3-AS5. Even though AS1 may prefer the alternative path, AS1-AS2-AS4-AS5, it cannot enforce its selection under the current routing architecture.

BGP is also known to suffer from slow convergence and instability due to a wide-range of causes. Router mis-configuration or invalid route announcements can destroy network connectivity [14]. To improve path diversity and fault tolerance, some ASes choose to multi-home [6], i.e., connect to more than one upstream service providers. However, multi-homing can only increase the source ASes' flexibility in selecting the immediate next-hop ASes. In addition, the number of providers that an AS can simultaneously subscribe is usually limited due to economic reasons.

Alternative routing architectures have been proposed to decouple routing policy from inter-domain routing process, e.g., NIRA [22], Platypus [17], OPCA [4], and RPC [10]. Other solutions are introduced to provide additional flexibility that BGP lacks. For example, RON [7] offers fault-resilient overlay routing and Feedback-based Routing [23] can facilitate source-based path selection. However, these approaches either require the change of the whole Internet routing architecture [22, 17] or are not intended for large-scale deployment [7]. Overlay networks [16] attempt to improve the inter-domain routing performance. However, most existing overlay networks form mesh-like connectivity and use active probing to maintain/update their own routing states, leading to a lot of overhead. As a result, these overlays can only serve destinations restricted within the overlay nodes. To route traffic to a non-overlay node, it would need a separate DNS-like service module to locate the destination overlay node that is in charge of (or close to) the destination IP-address.



**Fig. 1.** Example of Inter-domain Routing Topology.



**Fig. 2.** Example of Virtual Multi-Homing Framework (VMH).

In this paper, we propose a new framework called Virtual Multi-Homing (VMH) to investigate the feasibility of combining the strength of overlay and BGP routing to improve inter-domain routing performances. VMH can facilitate loose source-based path selection and enhance inter-domain routing path diversity without any significant changes in the existing Internet routing architecture. A new inter-AS relationship called *Virtual Peering* is proposed in this paper. A virtual peering relationship can be established between two remote ASes via an overlay path of a multi-homing overlay network (MON). By doing this, an AS can choose to send out its packets via its physical neighboring ASes or its virtual peering providers.

VMH has the following main advantages. First, VMH enables flexible source-based routing by allowing a source AS to send packets through a series of selected overlay nodes within MON. This loose source-based forwarding feature can help to achieve some degree of end-to-end quality-of-service (QoS) assurance that is desired for real-time or mission-critical applications. Second, VMH uses overlay routing technique to enhance inter-domain path diversity that can be used to provide fault tolerance or perform load balancing. It can be used to provide proactive (local) re-routing in the event of transient inter-AS path failures without relying on the global re-routing that can take minutes to converge. Third, unlike most overlay networks that use a totally separate control plane to discover and exchange path reachability information, VMH interacts with BGP to learn about inter-AS connectivity and routing or policies. By associating a subset of the BGP prefixes with a chosen overlay node of MON (e.g., based on proximity to an AS), VMH extends its overlay routing services to a larger range of destination prefixes, even if they are not part of the overlay nodes.

The following part of the paper is organized as follows. We introduce the basic ideas and the framework of VMH in Section 2. Next, we study the stability of multi-hop peering sessions in Section 3 and present some simulation results in Section 4. Section 5 describes the related work and Section 6 concludes the paper.

## 2 The Framework of Virtual Multi-homing (VMH)

VMH is designed to combine overlay and BGP routing to provide efficient inter-domain routing services. Besides improving routing performance, we also investigate the possibility of combining BGP with overlay to provide source based AS-level path selection to a large group of users. Overlay service providers, such as Akamai [5] can utilize this technique to provide wide area inter-domain routing service to customers.

### 2.1 Multi-homing Overlay Network (MON)

In VMH, each AS can have one or more *Multi-Homing Servers (MHS)*. Each MHS can either co-exist with a border router or be located at a separate host. However, a MHS sets up one or more I-BGP sessions with local AS BGP routers so that it can locate its inter-domain BGP paths to other destinations and receives BGP update messages from these BGP routers. A BGP router will consider its attached MHS as one of its I-BGP peers without accepting BGP update messages from the MHS.

The MHSes from different ASes cooperate with each other to form a *Multi-Homing Overlay Network (MON)*. The connections between MHSes are based on the *Overlay Transit* relationships among VMHes. Similar to current inter-domain relationship, *Overlay Transit* is set up between two ASes. *Overlay Transit* determines whether a MHS can forward packets to its neighboring MHSes within the MON. That is, if there is an *Overlay Transit* relationship between two ASes, there is an overlay link connecting the two MHSes in the MON overlay topology.

Similar to RON [7], an overlay routing protocol is running among the MHSes. This provides a resilient overlay routing path connecting each pair of source and destination MHSes. To support resilient overlay path routing, MON provides similar service as RON [7]. In other words, each MHS continuously probes its neighboring MHSes and sends the latest path performance information to every other MHS. If there is an IP-layer path failure or service degradation between any pair of MHSes, MON will detect it and provide an alternative overlay path.

## 2.2 Virtual Peering

Virtual Multi-Homing (VMH) service runs on top of MON. It is based on the concept of *Virtual Peering*, in contrast to the traditional inter-domain relationship that only exists between two physically adjacent ASes.

**Def. 1** *Virtual Peering* is a remote transit service between two remote ASes via MON. There is a virtual BGP peering session (using an overlay path via MON) between these two ASes' MHSes. One end (an MHS) of the session is a Virtual Peering Provider (VPP) while the other end (an MHS) is a Virtual Peering Customer (VPC).

Different from *Overlay Transit* relationship mentioned above, a *Virtual Peering* provider can send its customer's traffic to any destination in the Internet. A VPC can receive path reachable messages (BGP update messages) from its VPPs in addition to the messages for its local BGP router via I-BGP session. However, a VPC will not send any BGP update messages to its VPPs. This restriction ensures that VMH will not introduce any extra routing messages into the BGP routing system and affect its performance. Based on MON, an AS can subscribe several VPPs in addition to its directly-connected provider(s). It can also subscribe new VPPs on-demand if necessary. A virtual peering BGP session between VPP and VPC usually passes through a resilient overlay path on top of MON. To send a packet to its destination, a source AS can either (a) send a packet via its direct physical providers, or, (b) send the packet using an overlay path to one of its VPPs, which will then forward the packet to its destination. We can see that the traditional multi-homing service can be deemed as a special case of VMH. It can be observed that "virtual peering" can help a source AS to explore inter-domain path diversity, which will potentially improve the inter-domain paths service quality.

For example, in Figure 2, the MON is composed of five MHSes: A, B,...,E, where are located in 5 ASes respectively, AS1, AS2,...,AS5. In this figure, AS1 has two "virtual peering" providers, AS3 and AS5 (denoted by the dashed lines with arrows). AS3 and AS5 also will deem AS1 as one of their customers: setting up BGP sessions between the two corresponding MHSes and sending BGP update messages. In addition, denoted by the solid lines, A also has two *Virtual Transits*, B and D, which can forward A's traffic to other MHSes. For each destination, AS1 not only has the path information it receives from its direct service provider, it also has additional candidate paths via virtual peering provider AS3 and AS5 as well as the MON's topology and overlay links' performance information. If necessary, AS1 can send packets to one its VPPs (such as AS3) using an overlay path via MON. Then, the VPP will forward the packets to destinations. It should be noted that the virtual BGP session and data packets between two MHSes do not necessarily follow the same overlay routing path as discussed later.

## 2.3 Loose Source-based Data Forwarding via VMH

In the VMH framework, a host can send packets via its local MHS to destinations if it detects failures on its default IP-layer paths. For some critical applications, without waiting for the failure notification, a host can send duplicate copies of packets via its local MHS and the default IP-layer paths at the same time.

Once a packet arrives at a MHS, the MHS takes the following steps to locate loose path information (a list of MHSes) to the packet destination. First, it picks a destination MHS from its set of VPPs based on the BGP path information received from VPPs and local BGP router. The selection of a destination MHS is based on the distances between VPPs and the packet destination as well as the

path disjointness between default BGP paths and corresponding inter-domain paths via VPPs.

Second, the source MHS can take one of the following two ways to find the path from itself to the destination MHS: 1) By default, based on the resilient overlay routing protocol, the source MHS can find an overlay path (list of MHSes) with best service quality (least loss rate or shortest delay); Or 2) the source MHS can find an alternate overlay path through MON based on the constraint specified by packet sender or local ISP's routing policy.

After finding the loose forwarding path (LFP) to a destination MHS for a packet based on the above method, the source MHS can encapsulate the original data packets with an additional header which includes the list of MHSes (Loose Forwarding Paths). It can use IP Tunneling [15] or similar methods, such as wide-area relay addressing protocol (WRAP) [8]. When a MHS receives a loose data forwarding packet from one of its neighboring MHSes (its *Overlay Transit* customers), if the current MHS is the destination MHS on the LFP and the packet source is one of its VPCs, it just removes the LFP header part and sends the packet to its destination via the normal IP forwarding path. Otherwise, if the packet comes from one of its *Overlay Transit* customers, it locates the LFP from the packet header part and sends out the packet to the next hop MHS.

We use Figure 2 as an example to explain how VMH can use virtual peering relationship to achieve loose data forwarding. For example, in Figure 2, if A1 want send packets to a destination within AS6 via VMH. Based on the BGP information and the MON topology, it can locates a destination VPP and an LFP to the VPP, such as E and A-D-E. By encapsulating the LFP within packets, the traffic can be transmitted following the LFP via MON. When the data passes through the overlay links(A-D and D-E), they will be transmitted via IP-Tunneling through corresponding IP-layer paths. The destination MHS, E, will then decapsulate the packet headers and send the data to destination via normal IP-layer paths.

As shown in [20], the inter-domain forwarding performance can be greatly improved via one or two intermediate forwarding points, we expect that most of loose forwarding paths will pass through no more than 3 MHSes. As a result, the extra encapsulated header part will not incur too much overhead.

## 2.4 Discussions

**Deployment Issue:** The setup of *Virtual Peering* relationships and the overlay links (*Overlay Transit*) between MHSes are based on the business relationships. Each MHS will be remitted based on the amount of traffic it sent to neighboring MHSes or destination IPs from its customers. The customer ASes can be remitted by receiving better routing services.

In addition, this approach does not require modification of the whole Internet routing architecture or BGP routing protocol. It can begin its service with a minimum of two virtual peering parties' (one provider and one customer) support. The extra overhead is the packet encapsulation and decapsulation time. Comparing to the existing Internet peering relationship, it need not to set up new physical links between ASes. VMH does not require to changes in the existing BGP routing process.

**Impact on BGP Routing Performance:** One may ask whether VMH will affect existing BGP routing performance. The two aspects we need consider are: BGP routing states and BGP routing performance. *BGP routing states:* from previous sections, we can see that the VMH approach does not introduce

any BGP routing messages to the BGP routing system. It just serves as a client to accept BGP routing messages, similar to most BGP looking class [2]. *Border Gateway Performance*: BGP's routing performance also depends on border gateways' performance (data processing speed). MHSes can either co-exit with border routers or be located at separate hosts and serve as clients to receive BGP update messages. As a result, we believe the default routing performance will not be affected by deploying the VMH framework. However, similar to other overlay networks, VMH will certainly affect the intra-domain and inter-domain traffic distribution. How ISPs can effectively deal with the dynamic traffic shifting caused by overlay networks is an issue need further investigation.

**Security Issue:** Source-based routing is considered as providing malicious users with the capacity to attack network [23]. The attackers can spoof the true sources of packets so that the innocent victim hosts will respond packets to the trusted sources. This approach can result in DDoS attack. In our proposed VMH framework, a *Virtual Peering* session or *Overlay Transit* is set up between two trustworthy MHSes. An MHS will only deliver packets for its *Virtual Peering* or *Overlay Transit* customers. In addition, we can use IPsec to deliver the the data packets between two VMHes. This can facilitate source tracing and prevent DoS attack to some extent without introducing extra security concerns.

In summary, under VMH framework, the end-to-end routing via VMH is a combination of IP and overlay routing. This will result in less redundant work and is inherently "topology-aware" comparing to pure overlay based approaches. For an AS, its *Overlay Transit* ASes' MHSes can forward its packets to other MHSes within MON. However, only part of the MHSes (its Virtual Peering Providers) can forward data packets to packet destinations outside of MON.

### 3 Multi-Hop E-BGP Session Stability Analysis over MON

In VMH, a virtual peering session between two remote virtual peers (MHSes) is an overlay path via MON. The overlay path is an inter-domain path passing across multiple ASes in contrast to regular E-BGP sessions (passing through a physical link) or I-BGP sessions (passing by an intra-domain path).

One may argue that the performance degradation of inter-domain paths may introduce frequent BGP session resets between virtual peers. This will affect the loose forwarding path stability and cause frequent BGP routing message exchanges. As shown in [21], the number of I-BGP peering session resets during time period  $t_c$  is:

$$P_p(t_c) = \begin{cases} 0 & \text{if } t_c \leq T_h - T_k \\ 1 & t_c \geq T_h \\ 1 - (\frac{t_r(i^*) - t_c}{T_k})^2 & \text{Otherwise} \end{cases} \quad (1)$$

where,

$$t_r(i^*) = \begin{cases} T_k + (2^{\lfloor \log_2 \frac{(T_h - T_k + 1)}{R_0} \rfloor} - 1)R_0 & \text{If } T_h \leq (\lfloor \log_2 \frac{R_m}{R_0} \rfloor) * R_0 + T_k \\ T_k + \lfloor \frac{T_h - T_k - (2^{1 + \lfloor \log_2 \frac{R_m}{R_0} \rfloor} - 1) * R_0}{R_m} \rfloor * R_m + (2^{1 + \lfloor \log_2 \frac{R_m}{R_0} \rfloor} - 1) * R_0 & \text{Otherwise} \end{cases} \quad (2)$$

In above expressions,  $T_k$  is BGP Keep Timer value and  $T_h$  is BGP Hold Timer value.  $R_m$  is the maximal TCP timeout value (by default, it is usually 60 seconds.) while  $R_0$  is the average round trip time between the two BGP routers.  $t_c$  is the underlying path failure last time. We can use the same method to evaluate the performance of virtual BGP peering session stability.

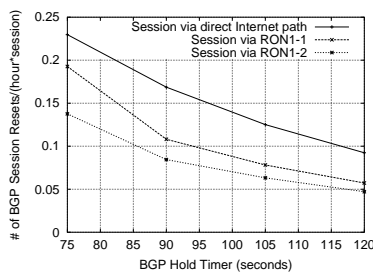
As shown in [7], overlay networks can detect and recover inter-domain routing failures and improve end-to-end routing performance. As a result, it can also improve the stability of virtual peering BGP sessions. To verify this, we have run simulation by analyzing the data set collected by RON [7] test-bed, which includes 12 nodes in different Internet locations. RON1 includes the data collected from 3:41am March 21 to 19:55 pm March 23, 2001. From the data, we can retrieve three different paths' performance information connecting each pair of overlay nodes: 1) default IP-layer path performance; 2) overlay path performance via latency-optimized overlay (data set RON1-1); 3) path performance via loss-optimized overlay (data set RON1-2). RON uses the following method to detect path failures: each overlay node sends probing packets to other nodes every 12 seconds. The timeout value for probing packets is 3 seconds. If a node notices a probing packet is timed out, it will consecutively send out 3 more probing packets. If all the four probing packets are timed out, it deems the path as failed. We follow this method to analyze the data failure events. As it is a conservative method, the results should be the upper bound number of real path failures.

We use the following equation to find the average number of resets for each virtual peering BGP session:

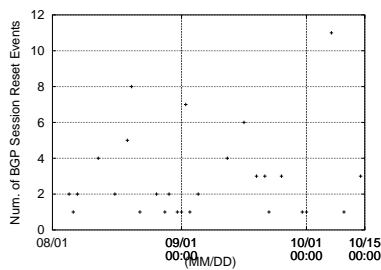
$$\frac{\int P_p(t_c) * N(t_c) dt_c}{\text{time(hours)} * \text{overlay link num}(= 12 * 11)} \quad (3)$$

where  $N(T_c)$  is the number of failures which last  $T_c$  seconds.

During the simulation, we set the BGP Keep Timer value as 30 seconds. By analyzing the data, we can get the average RTT values for these three scenarios: 0.107 (regular IP-layer paths), 0.1015 (delay optimized overlay paths) and 0.1217 second (loss-rate optimized overlay paths). We vary the value of BGP Hold Timer value and obtain the simulation results as shown in Figure 3. From the results, we can observe that the overlays can greatly increase the virtual peering session stability by reducing the number of reset events to half of the value as via normal IP-layer paths. Comparing between RON1-1 and RON1-2, we can see that the loss-rate optimized overlay has better performance than latency-optimized overlay in improving the BGP session stability. Another trend observed from the figure is that the increase in the value of  $T_h$  can greatly improve the virtual peering session stability.



**Fig. 3.** BGP Holding Time vs. Num. of Resets.



**Fig. 4.** Num. of BGP Session Resets Between Aug.1 and Oct. 15 2001.

From Figure 3, under normal BGP setup value of  $T_h$  90 seconds, the average per-hour peering session resets is around 0.3. This is much higher than expected. This may be because of our failure estimation method, which is a conservative approach as mentioned earlier. We believe that in reality, the value should be much less. To verify this, we analyze the real Internet BGP routing data from RIPE [2]. RIPE uses multi-hop (via inter-domain paths) BGP peering sessions to collect BGP routing data from various locations of the Internet. It can be deemed as a kind of virtual peering. RCC01 is peering 12 remote providers' BGP routers. Among the 12 BGP sessions, two of them pass one AS-level hop inter-domain path, three of them pass through two AS-level hops, five of them pass through three AS-level hops, and three of them pass through four or more AS-level hops inter-domain paths. We analyze the RIPE RCC01 data from August 1st 2002 to Oct.15 2002. There are two reasons for which we choose this data: 1) During this time, we can obtain partial BGP routing log information; 2)  $T_h$  (90 seconds) and  $T_k$  (30 seconds) setups are same as normal BGP operation. By analyzing this data, we can get the real upper bound of virtual peering session stability via normal inter-domain paths, which is also the upper bound of virtual peering session stability via MON. The number of BGP session resets for each day are shown in Figure 4. Based on the numbers, the average per-day resets for a node with 12 virtual peers via normal BGP paths is around one. The maximal reset events per day is around 11. This result is the number of multi-hop BGP session reset events via normal inter-domain path. As shown in Figure 3, overlay networks can greatly improve the multi-hop BGP session stability, we believe that the virtual peering BGP session is feasible via MON. In addition, we also can utilize the approaches proposed in [21] to improve the virtual peering session stability, such as shorten  $T_k$ , use larger value of  $T_h$ , or modify TCP retransmission behaviors.

## 4 Performance Evaluation of VMH (Virtual Multi-Homing)

In this section, we study VMH's performance in terms of exploring end-to-end routing path diversity through simulations.

### 4.1 Simulation Setup and Performance Evaluation Metrics

Our simulation is based on the real Internet inter-domain topology provided by [13]. As shown in [13], the Internet is composed of different tiers of ASes: dense cores (Tier 1), transit cores (Tier 2), outer cores (Tier 3), small regional ISPs (Tier 4), and customers (Tier 5). The inter-AS relationships can be categorized as: Peer-to-Peer and Customer-Provider relationships. To facilitate our simulation, we prune out Tier 5 ASes, which left a topology with 2473 ASes. For each simulation, we randomly construct a MON whose size varies from 25 to 200. For each simulation, a virtual peering customer is randomly chosen from the Tier 4 ASes with only one physical provider. We also vary the number of the customer's Virtual Peering Providers (VPPs) from one to eight.

We evaluate the following two performance metrics of VMH:

- End-to-end path availability (*EEPA*): It is defined as the possibility of finding a path between a source and destination pair for a given inter-domain link failure ratio. If AS S has  $n$  loose data forwarding paths via *VPPs* and one default BGP path to destination D, denoted as  $P_1, P_2, \dots, P_n, P_{n+1}$ . Suppose  $LINK_i$  is the distinct number of inter-domain links for path  $P_i$ . If  $p$  is the

link failure ratio, the EEPA between S and D is defined as:

$$EEPA(S, D) = 1 - \prod_{i=1}^{n+1} (1 - (1 - p)^{LINK_i}) \quad (4)$$

- AS-level Path Penalty: This metric is based on a default BGP path’s corresponding most disjoint loose forwarding path via VMH. It is defined as the number of AS-level hops of the most disjoint loose forwarding path via VMH divided by the number of hops of the corresponding BGP path.

#### 4.2 Selection of Virtual Peering Providers (VPPs)

The VMH’s performance for a customer depends on how the customer chooses its VPPs. Since Tier 1 ASes may provide extraordinary benefits, during simulation, the VPPs are selected from Tier 2 or lower-tier ASes.

A direct method to select VPPs is to consider the potential VPPs’ EEPA performance for each destination. Then, we can select a set of VPPs that can provide the best overall EEPA to all the destinations. If the network size is  $n$  and the MON size is  $m$ ,  $W$  is the total number of VPPs a customer wants to choose, the computation complexity is at least  $O(n*(m*log^m)^{W+1})$ . Although the virtual peers selection could be an off-line activity, this method may not always meet our requirement. Since our simulation is not to compare the performance of different VPP selection methods, we use a heuristic-based VPP selection method. The basic theory under this heuristic selection algorithm is: the tier 1 providers can directly reach most of the destinations. If a candidate VPP has totally different list of tier 1 providers as that of the source AS, it has a better chance to provide most disjoint paths from the customer to any destinations, which can enhance the end-to-end path availability. This method is formalized in Algorithm 1.

---

#### Algorithm 1 VPP-Select

---

```

Retrieve the list of Tier 1 providers for each candidate VPP
Rank the candidate VPPs based on the number of distinct tier 1 providers comparing
to the customer AS
If there is a tie between two candidate VPPs, break the tie based on the following
method
  For each candidate VPP
  {
    Set the destination set ( $D_{set}$ ) as the tier 1 providers of the candidate VPPs
    For each destination in  $D_{set}$ 
      Find the BGP path from the customer AS to the destination
      Find the other path from the source AS to the destination via the candidate
VPP
      Get EEPA for these pair of paths
      Average the EEPA for this candidate VPP
    }
  break the tie based the average EEPA value
Return the first  $W$  ASes in the list

```

---

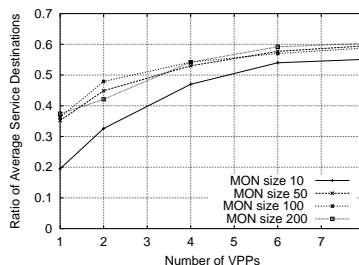
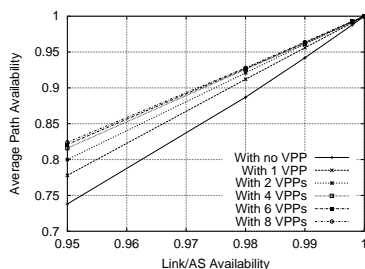
As there are around 20 tier 1 providers in the current Internet topology, this heuristic algorithm can greatly reduce the computation complexity.

#### 4.3 Simulation Results

Figure 5 shows the average end-to-end path availability for a customer AS passing through a MON of 100 nodes. For a customer AS, without any VPP, the

customer only has one BGP path provided by its direct service provider. From the simulation figure, we can observe that the average EEPA can be greatly improved by VMH. VMH can provide the customer more flexibility to select loose forwarding path via MON. The EEPA can be improved with the increasing number of VPPs. However, when the number of VPPs is more than 4, the improvement diminishes. This is because the VMH’s performance is also restricted by the underlying physical topology and the composition of MON, which together determine the maximum benefit VMH can provide to the customer ASes.

Similarly, VMH also cannot provide a customer with resilient source-based routing paths (alternate AS-level paths) to all the destinations. Figure 6 shows the average resilient serviced destination ratios for a VMH customer. From the figure, we can observe that both MON size and the number of VPPs determine the number of serviced destination paths. Restricted by the underlying physical topology, only a selected set of destination paths can be benefited. This means that careful selection of VPPs is important for a VMH customer.



**Fig. 5.** Path Availability (MON size 100). **Fig. 6.** Average Serviced Destination Paths.

The average inter-domain AS-level path penalty via VMH for various sizes of MON is presented in Figure 7. As defined above, the inter-domain path penalty is defined as the number of ASes in the most-disjoint AS-level path via VMH divided by the number of the corresponding BGP path ASes. From the results, we can see that the value of the average path penalty for different MONs is almost the same, around 2.0. However, as we can see from Figure 8, a larger size MON can provide service to more destination paths with less penalty. The penalty value varies from 1.5 to 2.7. Most of the path penalty values are around 2.0. In our simulation, we assume that BGP always takes shortest inter-domain paths constrained by the inter-AS relationships: peer-to-peer or customer-and-provider. In reality, each AS can add different preferences in choosing a route, such as optimizing intra-domain traffic, inter-domain traffic engineering, which usually results in non-shortest paths. Based on this fact, we believe that if VMH is deployed in Internet, the average penalty will be decreased.

In summary, from Figure 5 to Figure 8, we can see that VMH can effectively improve the customer ASes’ routing path diversity through VMH. The metric performance of EEPA shows that VMH can provide more chances for the source ASes to overcome the BGP routing anomalies and optimize the inter-domain routing performance. However, the detailed performance will depend on the selection of VPPs, their locations, and the size of MON.

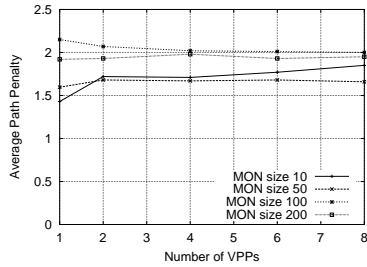


Fig. 7. Average Path Penalty vs. Num. of VPPs.

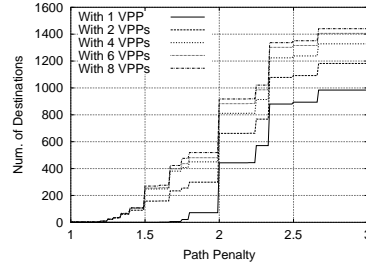


Fig. 8. CDF for Average Path Penalty.

## 5 Related Work

In Inter-Domain Routing Architecture [9], Estrin *et al* proposed that inter-domain routing architecture should contain a hop-by-hop node routing (NR) component and a source demand routing (SDR) component. Similar to SDR, VMH is also designed to facilitate source based route selection. Using a different approach, VMH is based on MON to overcome the routing constraints from BGP and to set up flexible inter-domain peering relationships. In addition, it is more scalable and easier to for deployment.

Similarly, Platypus [17] also intends to make use of Internet diversities and achieve source-based path selection. It is an authenticated source based routing system, built on the concept of "network capabilities". Network capacities allow for accountable, fine-grained path selection by using cryptographical method to enforce policy compliance at each hop along a source path. Feedback-based Routing [23] is another source based inter-domain routing protocol. However, it assumes that we need to completely change the current inter-domain routing infrastructure and ASes would like to expose all its local routing policy information. These two assumptions generally do not hold, which limit the widespread deployment of this approach.

Resilient Overlay Network[7] is an overlay based approach to achieve source based path selection. The authors have shown that it can achieve good performance in terms of source-based fine path selection and overcome BGP routing anomalies. However, it does not scale to large users and the routing is usually restricted to the overlay nodes.

In paper [11], the authors performed extensive experiments to monitor the Internet path availability and locate the path failure points. The results show that 56% of path failures can be recovered via one-hop source routing. The results prove the potential performance benefits of our VMH framework from another perspective.

Huston [12] first suggested to use alternate routing through overlay networks or replace BGP with a new inter-domain routing protocol to address the inter-domain routing problems. Using similar idea, OPCA [4] uses an overlay-based policy control architecture on top of BGP to achieve fast route failure and traffic engineering. RCP [10] proposes to completely separate inter-domain routing functionalities from routers.

## 6 Conclusion & Future Work

In this paper, we propose a new approach called Virtual Multi-Homing (VMH) to achieve source-based inter-domain routing and explore inter-domain path di-

versity by combining BGP routing with overlay routing. VMH is based on a new inter-AS relationship (*Virtual Peering*) and *Multi-Homing Overlay Network (MON)*. In contrast to previous overlay approaches, the proposed method utilizes BGP routing states whenever possible, hence reducing the possibility of introducing duplicate work at overlay layer. This approach can be easily deployed without impacting the existing BGP routing performance. We performed some preliminary simulations to investigate VMH's performance and VMH virtual peering BGP session stability. The results have shown that VMH is feasible provide desirable performance to end users.

In our future work, we plan to deploy test-bed on top of Planet-Lab [1] to investigate VMH's performance. In addition, we want to explore how different virtual peering provider selection methods impact VMH service performance.

## References

1. Planetary network testbed, <http://www.planet-lab.org>.
2. RIPE. <http://www.ripe.net/>.
3. University of Oregon RouteViews project. <http://www.routeviews.org/>.
4. S. Agarwal, C. Chuah, and R. H. Katz. Opca: Robust interdomain policy routing and traffic control. *IEEE Openarch*, 2003.
5. Akamai Corporation. <http://www.akamai.com>.
6. A. Akella, B. Maggs, S. Seshan, A. Shaikh, and R. Sitaraman. A measurement-based analysis of multihoming. In *ACM SIGCOMM' 03*, Aug. 2003.
7. D. G. Andersen, H. Balakrishnan, M. F. Kaashoek, and R. Morris. Resilient overlay network. In *18th ACM SOSP*, Oct. 2001.
8. D. Cheriton and M. Gritter. Triad: A scalable deployable nat-based internet architecture, 2000. <http://www.wds.stanford.edu/papers/triad/triad.html>.
9. D. Estrin, Y. Rekhter, and S. Hotz. Scalable inter-domain routing architecture. In *Proc. ACM SIGCOMM*, 1992.
10. N. Feamster, H. Balakrishnan, J. Rexford, A. Shaikh, and J. Merwe. The case for separating routing from routers. In *ACM SIGCOMM FDNA workshop'04*, Sep. 2004.
11. K. P. Gummadi, H. V. Madhyastha, S. D. Gribble, H. M. Levy, and David Wetherall. Improving the reliability of internet paths with one-hop source routing. In *USENIX: OSDI '04*, December 2004.
12. Geoff Huston. Architecture Requirements for Inter-domain Routing in the Internet. Internet draft 01, internet architecture board, May 2001.
13. L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz. Characterizing the Internet hierarchy from multiple vantage points. *Proc. IEEE INFOCOM*, 2002.
14. R. Mahajan, D. Wetherall, and T. Anderson. Understanding bgp misconfiguration. In *ACM SIGCOMM 2002.*, 2002.
15. C. Perkins. IP Encapsulation within IP, ietf rfc 2993, 1996.
16. L. Qiu, Y. R. Yang, Y. Zhang, and S. Shenker. On selfish routing in internet-like environments. In *ACM SIGCOMM*, 2003.
17. B. Raghavan and A. C. Snoeren. A system for authenticated policy-compliant routing. In *ACM SIGCOMM'04*, Sep. 2004.
18. Bhaskaran Raman and Randy H. Katz. Load balancing and stability issues in algorithms for service composition. In *IEEE INFOCOM 2003*.
19. Y. Rekhter and T. Li. A Border Gateway Protocol 4 (BGP-4), ietf rfc 1771, 1995.
20. S. Savage, A. Collins, E. Hoffman, J. Snell, and T.E.Anderson. The end-to-end effects of internet path selection. In *SIGCOMM*, 1999.
21. L. Xiao and K. Nahrstedt. Reliability models and evaluation of internal bgp networks. In *IEEE INFOCOM*, March 2004.
22. X. Yang. NIRA: A New Internet Routing Architecture. In *ACM Sigcomm FDNA workshop*, 2003.
23. D. Zhu, M. Gritter, and D. Cheriton. Feedback based routing. In *ACM HotNets-I*, 2002.