

A Survey of Multihoming Technology in Stub Networks: Current Research and Open Issues

Xiaomei Liu and Li Xiao, Michigan State University

Abstract

Multihoming technology can be traced back to the 1980s or even earlier, but it seldom drew much attention from the research community. Recently, with the explosive deployment of multihoming on the Internet, researchers have started to investigate the issues that were raised by multi-homing networks. This article draws an overall picture of multihoming technology, presents current research on multi-homing networks, and discusses the issues that must be studied in the future.

A stub network refers to a network that conveys traffic only to or from its local hosts. Stub networks never carry traffic for which they are neither the source nor the destination. Traditionally, most stub networks have only one connection to the Internet [1]. If this connection fails, the stub network becomes disconnected, and users of the network suffer long delays.

To achieve reliability and redundancy, some companies connect to the Internet with more than one connection. This is called multihoming. Multihoming refers to the technology where a single network has multiple connections to the Internet. When one connection fails, the network still connects to the Internet via other connections. Multihoming can be applied in different layers of the network protocol stack: link layer, network layer, or transport layer. This article investigates only multihoming for a stub network in the network layer; however, studies of multihoming in other areas, such as the transport layer, are also fairly active [2, 3].

For the network layer, multihoming is discussed in Internet Protocol version 4 (IPv4) and Internet Protocol version 6 (IPv6). However, not all the networks in the Internet support IPv6. IPv6 multihoming is still under discussion regarding the formation of a standard. This article focuses on the multihoming in IPv4. Interested readers can check network working group reports [4, 5] or the survey of IPv6 multihoming [6].

Depending on how many upstream Internet service providers (ISP) that the multihoming networks connect (Fig. 1), a network can be classified as one of the two following models [4]:

- A network that connects to a single Internet service provider (ISP) with multiple connections is called a *multi-attached network* (Fig. 1a).
- A network that connects to more than one ISP is called a *multi-homed network* (Fig. 1b). Connecting to more than one ISP, multi-homed networks remove the dependency of stub networks on one ISP: a stub network can rely on another ISP if one ISP fails.

In multi-homed networks, more than one address block can be assigned by different ISPs. This raises address management issues in the stub networks. At the same time, deployment of multihoming technology in a stub network affects the routing

policy of the network. How to handle the connection failure is also an important issue in deploying multihoming technology. The available solutions to deploy multihoming technology are presented later.

Although multihoming was adopted by some companies as far back as the 1980s, this technology was not used widely and had little influence on the Internet [7]. Today, many companies require reliable Internet connections due to increased business requirements and network traffic load. As a result, multi-homed networks have surpassed single-homed networks in number [8, 9]. There is an increasing trend of deploying multihoming, and this trend is likely to continue with the availability of inter-domain traffic engineering tools and techniques [7, 10–12]. With the dramatic increase in the number of multihoming networks, new challenges and opportunities are created.

One problem raised by multihoming is the non-aggregation problem in the inter-AS routing. The Internet is composed of numerous sub-networks or autonomous systems (ASs). AS refers to a network or a group of networks that is controlled by a common network administrator (or group of administrators) on behalf of a single administrative entity. AS is the unit of the Internet routing policy. The routing process between ASs is called inter-AS routing. During the inter-AS routing process, each AS advertises the routes in its routing table to other ASs. An AS route announcement includes an Internet Protocol (IP) prefix followed by a series of AS numbers, for example, “90.16.1.0/20 103:101.” Here 90.16.1.0/20 is an IP prefix of a network and 103:101 are AS numbers that are on the path to arrive at the network of 90.16.1.0/20. Based on the AS number information, the routers can compute the best route to a network. To reduce the advertising overhead and routing table size, an AS may combine different routes with common characteristics into a single route in the route advertisements. Specifically, the router summarizes a set of IP prefixes and uses only one IP prefix to announce the route to the multiple IP prefixes. This process is referred to as route aggregation. For example, a route announcement of “90.16.1.0/16 103” and “90.16.1.0/20 101” may be summarized as “90.16.1.0/16 103:101.”

As a stub network traditionally connects to only one ISP,

the routes advertised by the stub networks with the same upstream AS (these stub networks have IP prefixes inherited from this upstream AS) can be aggregated to one route by this AS. In multi-homed stub networks, a stub network is multihoming to multiple AS and thus may inherit multiple IP prefixes from different upstream ASs, which makes the upstream AS unable to aggregate the IP prefixes advertised by this stub network. This is referred to as a non-aggregation problem. To construct scalable multihoming networks, new protocols were proposed to solve this problem [13].

Another issue investigated by the research community is the performance gain that multihoming technology brings to stub networks. By attaching through multiple connections, multihoming networks not only can improve the reliability of the connection, but also can achieve better performance by carefully routing traffic among their different connections. New route-optimization mechanisms are proposed for stub networks to achieve the largest benefit that multihoming technology can introduce [14–16].

To provide a solid base for further studies on multihoming networks, quantitative measurements of the reliability improvement and performance gain achieved by multihoming technology were investigated [14–16].

This article presents a survey of network models, protocols, and algorithms that have been proposed for multihoming technology thus far. The purpose of the article is to provide a better understanding of multihoming technology and current research to stimulate new research directions in this area. The remainder of the article is organized as follows. The solutions to deploy multihoming technology in stub networks are discussed in the next section, followed by recent research and open research issues. Conclusions are drawn in the last section.

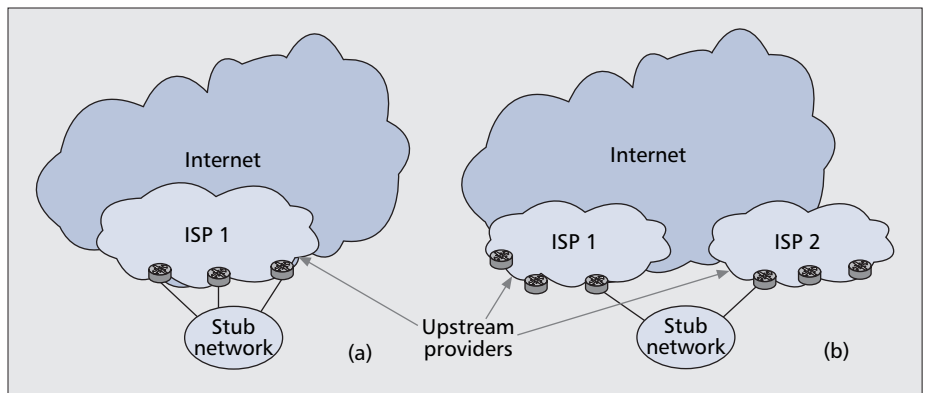
Solutions to Deploy Multihoming Technology

Two main solutions are available to deploy multihoming technology in a stub network: multihoming with Border Gateway Protocol (BGP) only, referred to as BGP multihoming and multihoming with a network address translation (NAT) mechanism, referred to as NAT multihoming. The details of BGP multihoming are presented in the next subsection, followed by the details of NAT multihoming. A comparison of the two solutions is presented at the end of this section.

BGP Multihoming

BGP is the default inter-AS routing protocol. It specifies how to exchange the network reachability information (i.e., route advertisements) among different ASs. BGP runs on the edge routers of the AS, called eBGP (external BGP) border routers. By default, BGP chooses the shortest route based on the AS hop counts as well as the preference level specified by each AS. Traditionally, users of stub networks choose to implement multihoming by installing two or more public network connections to the Internet using BGP. Next, we discuss BGP multihoming in more detail.

Address Management — A stub network must have its own AS number and a minimum address space identified by a 24-bit address prefix (noted as a /24 address space) or larger to deploy BGP multihoming. There are two address allocation schemes: provider-independent address (PI address) and



■ Figure 1. Multihoming: a) multi-attached network; b) multihomed network.

provider-assigned address (PA address). Different issues arise when BGP multihoming is deployed in the stub networks adopting different address schemes.

For a multi-homed network with greater than /20 address space, the owner of the stub network can apply for a PI address block directly from an independent authority (e.g., APNIC for North America). With the portable PI address space, the stub network is completely independent from its upstream providers. As a result, routes with PI addresses cannot be aggregated by the upstream ISPs, which eventually increases the BGP routing table overhead further upstream from these ISPs.

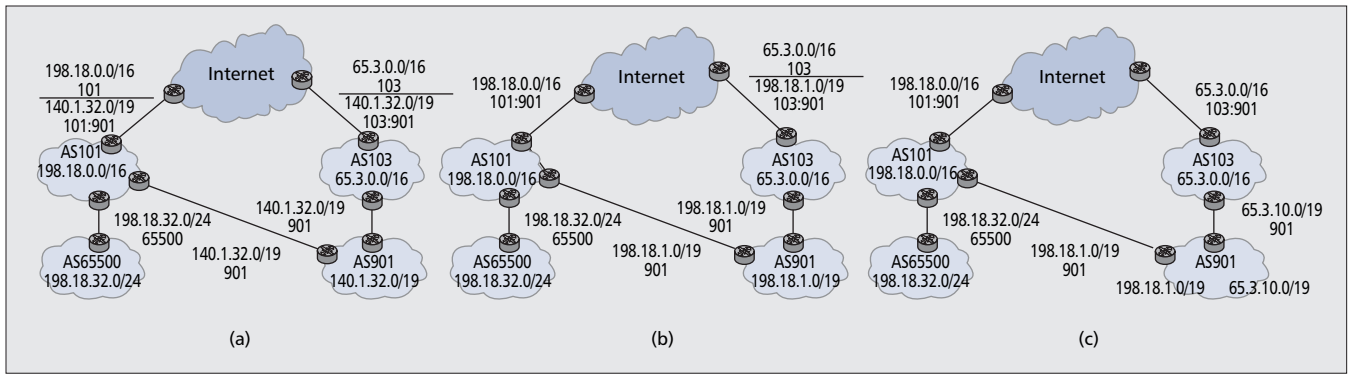
Networks that have less than /20 address space but greater than /24 address space can obtain PA addresses from their upstream ISPs. One address management mechanism based on PA addresses logically separates the whole stub network into several sub-networks according to the connections to the Internet [17]. Each sub-network inherits a separate address prefix from the upstream ISP closest to it. This approach treats sub-networks as separate networks and does not automatically maintain the back-up routes.

In another address management mechanism, the stub network uses only one address block assigned by one of its upstream ISPs, which is the default address block [17]. Other upstream ISPs of the stub network maintain a specific routing table entry for the route associated with the default address block. If the connection between the stub network and the ISP that issues the default address block fails, traffic can be routed to the network along the connections passing by the other upstream ISPs, which have the routing table entry for the route associated with the default address. The overhead here is caused mainly by the update of the routing table entries. In both mechanisms, routes with PA addresses may be aggregated.

Routing Process — Non-aggregated routes are advertised across multihoming networks with PI addresses, and aggregated routes are advertised across multihoming networks with PA addresses. Examples of routing processes of BGP multihoming are shown in Fig. 2. AS65500 is a non-multihoming network whose upstream ISP is AS101. AS901 is multihomed to AS101 and AS103. The route announcement of AS65500 “198.18.32.0/24 65500” is always aggregated by AS101 as 198.18.0.0/16, since 198.18.32.0/24 is a sub-address block of 198.18.0.0/16.

In Fig. 2a, AS901 has a PI address of 140.1.32.0/19. AS901 announces route “140.1.32.0/19 901” to both AS101 and AS103. Since the IP prefix of 140.1.32.0/19 is different from 198.18.0.0/16 (AS101) and 65.3.0.0/16 (AS103), neither AS101 nor AS103 can aggregate this announcement. Accordingly, both AS101 and AS103 must announce a specific route of “140.1.32.0/19 10x:901,” where $x = \{1, 3\}$, for AS901.

In Fig. 2b, AS901 adopts a PA address block 198.18.1.0/19



■ Figure 2. Routing process of BGP multihoming to two ISPs: a) PI address; b) PA address: only adopt one address block; c) PA address: multiple subnetworks.

that is assigned by AS101 and sends out an announcement “198.18.1.0/19 901.” As this address block is assigned by AS101, AS101 can combine the route of “198.18.1.0/19 901” with the one announced by AS65500 and then send out an aggregated route announcement “198.18.0.0/16 101:901.” As 198.18.1.0 is not assigned by AS103, AS103 cannot aggregate this route announcement but must send out a specific route of “198.18.1.0/19 103:901.” According to the BGP protocol, routers forward the packets along the most specific route. The specific route means with respect to the minimum address range. Here the most specific route is “198.18.1.0/19 103:901,” where the address range is /19, vs. /16 in “198.18.0.0/16 101:901.” Accordingly, AS901 will receive all packets via AS103 unless the link between AS901 and AS103 is disconnected. In this case, links between AS901 and AS101 can be used as a back-up link of the link between AS901 and AS103.

In Fig. 2c, AS901 divides itself into two sub-networks and obtains address blocks 198.18.1.0/24 (from AS101) and 65.3.10.0/19 (from AS103) respectively, as mentioned earlier. As these two address blocks are sub-network address blocks of 198.18.0.0/16 (AS101) and 65.3.0.0/16 (AS103) respectively, the routes associated with these two address blocks are aggregated by AS101 and AS103. Note that when an ISP aggregates the route associated with its sub-network, the inbound traffic via this ISP to that sub-network is still sent to the ISP even when the connection between this ISP and the sub-network is broken. In Fig. 2c, for example, the traffic to “65.3.10.0/19 901” is sent to AS103 even when the connection between AS103 and AS901 is disconnected. As we discussed in the last section, the two sub-networks are treated by the upstream ASs as separate networks and do not maintain a back-up route automatically. Due to security reasons, upstream AS may refuse to accept traffic from the stub network with a source address prefix that is not advertised to this AS by the stub network. In this example, AS103 refuses traffic from AS901 with source address prefix 198.18.1.0/19. Therefore, if a connection between a sub-network and the upstream ISP fails, the hosts with addresses based on that ISP become unreachable via inter-domain routing.

Failure Handling Procedure in Multihoming Networks — One approach to handling failure situations is via prepending AS routes advertisement. A stub network injects routes of both primary connections and back-up connections to the BGP routing tables in the Internet. The routes of back-up connections are made longer by the stub network via repeatedly prepending its AS number in the route. The primary route is used in normal situations since it is shorter than back-up routes. When the primary route is down, back-up routes can be used because they are also available in the BGP routing tables.

RFC2260 [18] suggests two failure-handling approaches for BGP multihoming with multiple PA-address prefixes. In the first approach, the eBGP border router of a stub network

advertises to an upstream ISP only the reachability of address prefixes assigned by that ISP in steady state. If the connection to the specific ISP is down, the eBGP border router of the stub network advertises to other upstream ISPs the reachability to the IP address block assigned (to the stub network) by this ISP. This mechanism is shown in Fig. 3. The border router announcement in steady state is shown in Fig. 3a, and the border router announcement in a connection failure situation is shown in Fig. 3b.

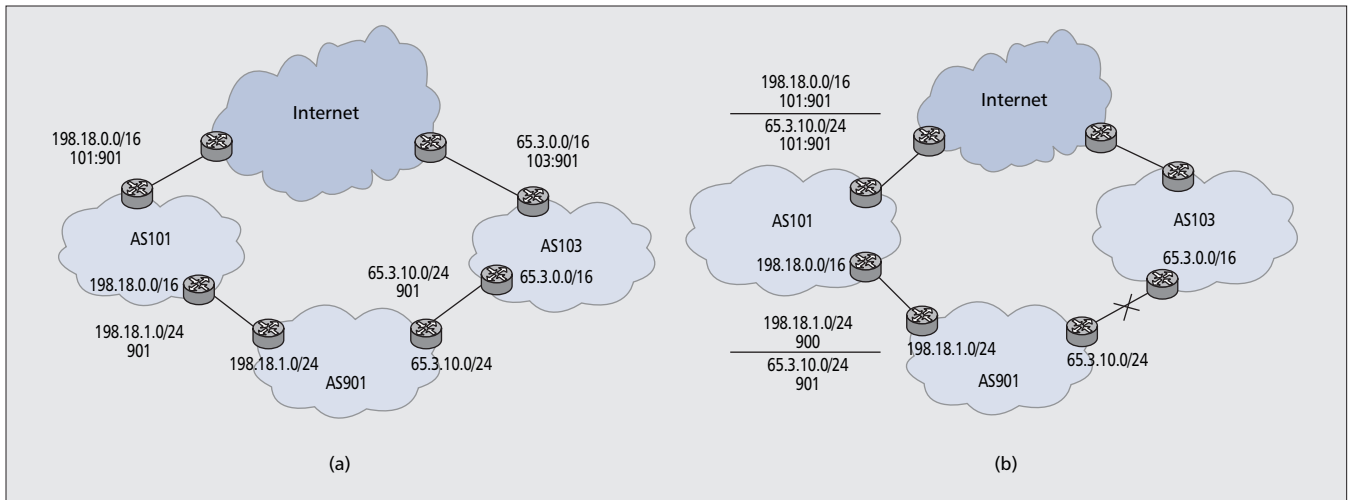
Another way to handle the link failure issue while adopting multiple prefixes from upstream ISPs is via packet encapsulation. In this case, the eBGP router of a stub network not only exchanges information with the provider eBGP router that is directly connected with it, but also with the other provider eBGP routers that connect to the stub network but do not directly connect to this stub-network eBGP router. For example, BR-A in Fig. 4a exchanges the route information with both PR-A, which directly connects with BR-A, and with PR-B, which is not directly connected with BR-A but is connected to AS901 via BR-B. When a link failure occurs between BR-B and PR-B, PR-B encapsulates all the packets that should be sent to BR-B with the IP prefix of BR-A. PR-B then sends these encapsulated packets via other connections of AS 103 to BR-A, which decapsulates the packets and routes them to the hosts inside AS901. This process is shown in Fig. 4b.

Compared to the first mechanism of AS number prepending, the other two methods require the cooperation of upstream ISPs. In other words, stub networks must negotiate with other ISPs to implement these mechanisms. This may increase the operating cost of the stub network. Note that a stub network adopting an AS number-prepend method also may be required to negotiate with its upstream ISP when it advertises to this ISP a “new” IP prefix that is not issued by this ISP. For example, the stub network adopts a PI-address mechanism. However, this is not induced by AS number prepending.

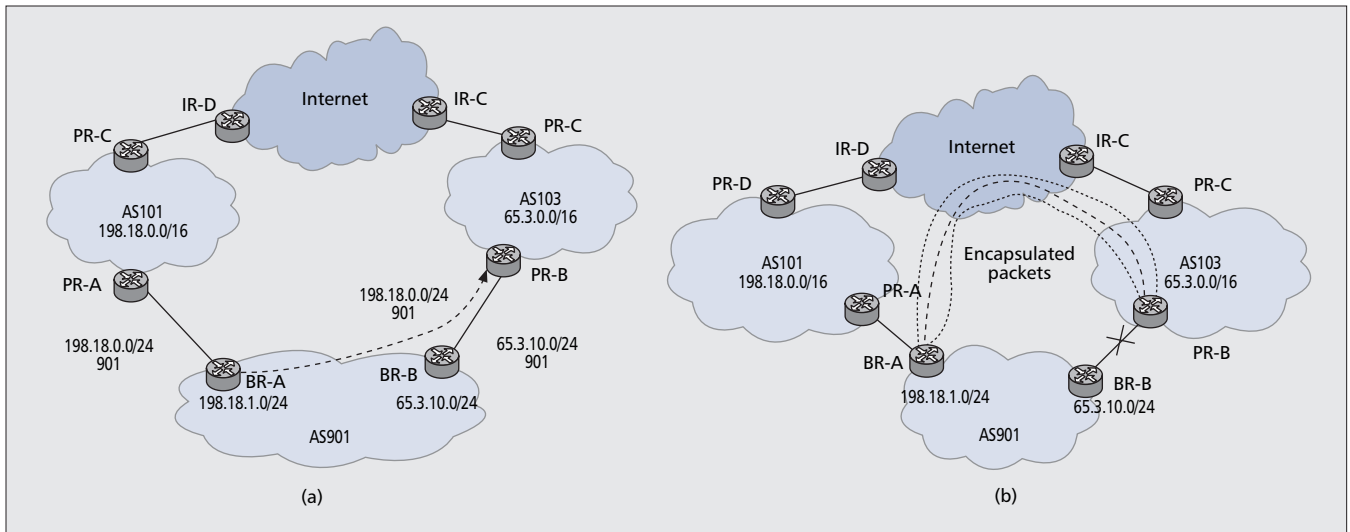
NAT Multihoming

The networking group of the Internet Engineering Task Force (IETF) suggested that local networks adopt NAT to reuse the IP addresses inside their networks [19]. NAT functions to translate between the public Internet address and the internal local network address. With this address-mapping feature, NAT is extended to help implement multihoming. NAT multihoming has no requirement regarding the size of the network. Nor does NAT multihoming have a requirement for the AS number. NAT also can be used in the stub networks with BGP functionality. Here we focus on NAT multihoming networks without BGP functionality.

Address Management — In most networks adopting NAT multihoming, the address blocks of the stub networks are assigned by upstream ISPs. Stub networks that adopt NAT



■ Figure 3. a) Border router advertises only to ASs the reachability of the address prefix to which it is assigned; b) border router advertises address prefix reachability when the AS that assigns this address prefix is unavailable.



■ Figure 4. a) BR-A exchanges information with both PR-A and PR-B (not connected with BR-A) and so BR-B; b) PR-B encapsulates the packets and sends to BR-A when failure occurs.

multihoming can use the inherent functionality of NAT to manage the assigned address blocks. The hosts in a NAT multihoming stub network share the network addresses among themselves. NAT maps address blocks assigned by different upstream ISPs to the internal address space of the stub network behind it. An address map is stored in the NAT router in advance. During the data transfer process, the NAT router translates IP addresses in the packet into internal or external addresses according to the map. NAT separates the address space of the stub network from the public Internet address space. The host addresses inside the stub network do not need to be rearranged if the upstream ISPs change.

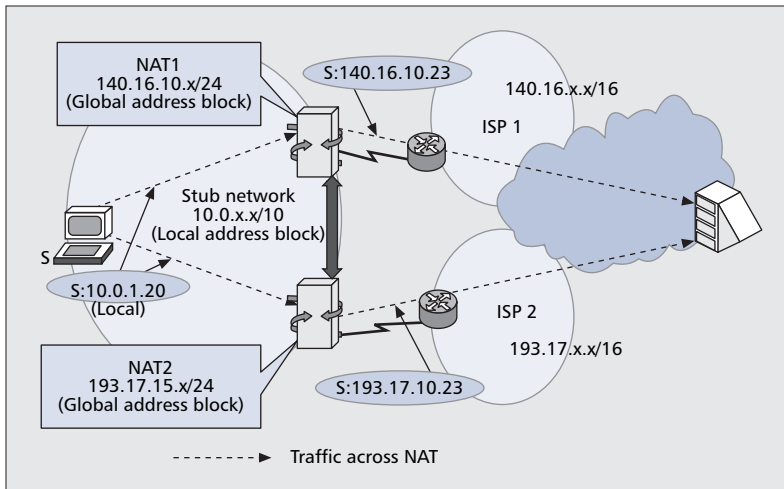
Routing Process — If the stub network that adopts NAT multihoming does not have an AS number, this network cannot control the BGP route advertisement process outside the local network. At the same time, because the address blocks of a NAT multihoming stub network are assigned by its upstream ISP, NAT multihoming does not create the non-aggregation problem. Mainly, a NAT router controls the routing of the traffic across the border of stub networks.

Figure 5 shows an example of NAT multihoming. Host S is assigned a local address of 10.0.1.20. A local address is used

to locate hosts within the stub network behind a NAT and cannot be seen by the hosts outside the stub network. If S sends a packet via the NAT1 router, its address in that packet is replaced by a global address 140.16.10.23. A global address belongs to an address block that is assigned by an upstream ISP to the stub network. Therefore, the global address can be seen by the hosts outside the stub network. Here, the global address 140.16.10.23 belongs to the sub-network address block of ISP1. If S sends a packet via the NAT2 router, its address in that packet is replaced by a global address 193.17.10.23, which belongs to the sub-network address block of ISP2.

In this example, we can observe that a host in a NAT multihoming network can send traffic by either of the two NAT routers of the network. A packet sent from the same host in the NAT multihoming network could have two different source addresses, which are aggregated addresses obtained from upstream ISP.

Failure Handling Process — For the NAT multihoming networks that do not adopt BGP and are not involved in the inter-AS routing process, the failure handling process can be accomplished by setting the traffic mapping rules in the NAT router. However, after a NAT mapping rule is set for the host



■ Figure 5. NAT multihoming.

IP address that associated with a traffic flow, this mapping rule cannot be automatically updated after a failure. This causes traffic loss during the failure-handling process. Another method is to use NAT together with the DNS server to reduce the effect of connection failure. In this situation, a host in the NAT network is bound with more than one IP address. If one ISP is not available, the IP address from another ISP is returned. However, traffic loss can still occur with this method.

Comparison of BGP Multihoming and NAT Multihoming

BGP multihoming and NAT multihoming are different in address management, routing process control, and the failure-handling process. In addition, BGP is the standard Internet inter-AS protocol; whereas NAT is introduced as a functionality to map local IP addresses inside a network with the global IP addresses outside the network. Therefore, BGP and NAT themselves also create differences between BGP multihoming and NAT multihoming.

Being the standard Internet inter-AS protocol, BGP guarantees the uniqueness of the host IP address and provides the largest support for the upper level applications. On the other hand, NAT removes the uniqueness of the IP address (since multiple hosts in the network are mapped into one address by NAT) and does not support all of the upper level applications. For example, NAT does not fully support non-client-server applications such as peer-to-peer file sharing systems and multi-party conference systems. Neither does NAT support the transmission of packets incurred by the applications based on the IP address, such as the IP address-based endpoint authentication in Voice over IP (VoIP) applications. For such applications, if only one of the two parties is behind the NAT, the common way to solve the problem is to add a middle server between two parties that makes the party who is not behind the NAT send a request to the middle server first [20, 21]. For the case of a requestor behind the NAT of one AS and the service provider behind the NAT of another AS, no effective solution as yet exists. To illustrate these issues, we summarize the differences together with the advantages and disadvantages of BGP multihoming and NAT multihoming in Table 1. We can observe in Table 1 that BGP can provide the largest support to the upper level applications and can control the routing process outside local networks; whereas NAT has no requirement in network size and does not induce the non-aggregation problem. Therefore, BGP is recommended for large organizations, which may be required to provide support

to all upper level applications. NAT is recommended for small to mid-size organizations that are not required to be involved in the routing process outside the local network.

Research on Multihoming Networks

The increasing popularity of adopting the multihoming technology leads to challenges in existing networks. This section surveys the current research progress on multihoming networks. A summary of the on-going research of multihoming networks is presented in Table 2.

Measurement Studies of Multihoming Networks

A series of measurement studies were deployed recently to investigate the benefits that multihoming technology can bring to stub networks.

These measurements require a comparison of the performance of the multihoming networks to that of non-multihoming networks. As we mentioned before, a large portion of existing networks are multihomed, and it is difficult to change the network router configurations to fit the requirement of these measurement studies. Therefore, most of the research studies in this area adopt the method of a combination of real world measurement and emulation instead of the pure on-site measurement of multihoming and non-multihoming networks.

The studies thus far can be divided into the following two categories: studies to investigate the reliability and the availability of the network and studies to investigate the performance of the available network services.

Reliability Improvement — Achieving additional reliability and availability is the primary motivation for adopting a multihoming technology. Recent studies show that multihoming greatly improves availability. In the emulation based on 68 servers allocated in 17 cities across the United States, it was observed that in the ideal case, where providers can be picked up by users with no limitation, the two-multihoming network can improve availability of the non-multihoming network by about 9 percent (from 91 percent to 99.85 percent) [14]. The three-multihoming network can further improve performance slightly more than that of the two-multihoming network. An n -multihoming network here refers to a network multihomed to n upstream ISPs.

Reliability is achieved in multihoming networks by the redundancy of the traffic path: if one connection of a multihoming network is broken, traffic can be rerouted to other connections. This redundancy is impacted by the path diversity of the Internet. Researchers believe that path diversity gives an upper bound of the gains that can be achieved by optimizing the existing multihoming network. Measuring the path diversity of the Internet helps to calculate how much reliability a multihoming network can achieve.

Han et al. [16] observed that more than 80 percent of paths in multihoming networks have at least one overlap. This is an upper bound in terms of increasing the availability gain of multihoming networks. They also found that increasing the amount of upstream providers contributes to reliability improvement, but marginal improvement is achieved in the case where the number of upstream providers is greater than three.

Studies show that with the same number of upstream providers, choosing a different set of providers can achieve a different availability gain in multihoming networks [15, 16]. Akella et al. [15] check the redundancy by two metrics: one metric indicates the number of the paths from the stub network to a series of dedicated remote web nodes; another one

		BGP multihoming	NAT multihoming
Address management	Advantages	Guarantees the uniqueness of the host IP address Provides largest support to upper level applications	No requirement on network size and AS number Separate stub network address space from public address space
	Disadvantages	Requires AS number and a minimum /24 address block Stub network address needs to be re-assigned if ISP changes	Removes uniqueness of IP address Difficult to support non-client server applications
Routing policy	Advantages	Able to control routing process outside local networks	Avoids non-aggregation problem
	Disadvantages	May cause non-aggregation problem	Not able to control routing process outside local networks
Failure handling process	Implementation mechanism	Implement failure handling via BGP advertisement	Implement failure handling via NAT mapping
	Disadvantages	Needs cooperation of upstream ISPs	Established connections using the predefined mapping rules that cannot be updated after a failure
Recommended usage		Large organizations; organizations that want to control the flow of the traffic outside the local networks Networks provide support to all upper level applications	Small to mid-size organizations that do not want to be involved in route control Networks mainly provide support for client-server applications

■ Table 1. BGP multihoming and NAT multihoming.

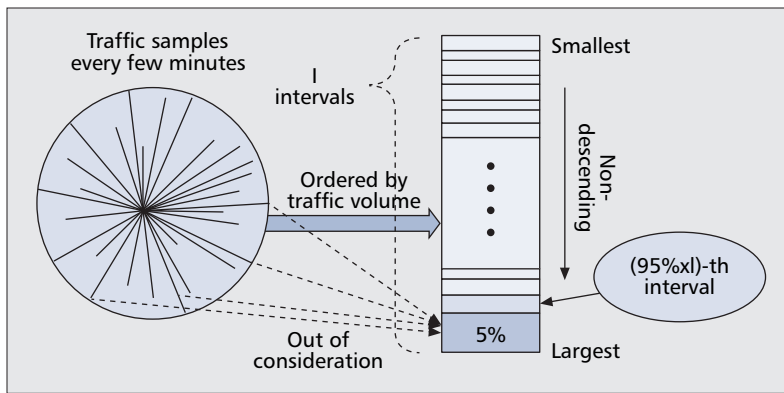
Measurement	Reliability	9–40% improvement (in terms of path availability);
	Performance	25–40% improvement (in terms of link delay or response time reduction);
Route optimization	Link performance monitoring	Active mode: send out probes to destination; may cause outbound traffic
		Passive mode: samples monitored links periodically; can not distinguish no traffic situation from “link down” event
	Traffic assignment	Classified by assignment granularity: packet by packet link assignment, connection by connection link assignment, and session by session link assignment
		Classified by traffic type: inbound traffic, outbound traffic
Economic cost	Traffic submission: optimizes financial cost of traffic submission	
	ISP subscription: chooses financial cost optimized ISPs	
Network scalability	SIMPLER	AS only reports non-aggregatable address to upstream ISPs on demand; use NAT to handle sub-network address management

■ Table 2. Summary of multihoming research.

indicates the overlap of the end-to-end paths from the stub network to a series of dedicated remote web nodes. They show that careful selection of upstream providers can achieve at least 25 percent more reliability than the random selection. Han [16] points out that whether a set of upstream providers can achieve higher reliability than other sets of upstream providers depends largely on the path destination.

Performance Improvement — Multihoming technology is also used to improve network performance. A measurement study to investigate performance improvement for multihoming networks is presented in [15]. In this article, the authors divide

the stub networks into two major types according to the direction of the mainstream traffic: data providers, which mainly send data out of the stub networks; and enterprises, which mainly receive data from outside the stub network. The performance of a data provider is measured by the delay estimates over each link of the provider when the remote clients fetch a designated embedded object from a Web site hosted by this data provider. Specifically, the delay estimates are based on how long a connection between the remote host and the local host lasts. The performance of the enterprises is measured using the response time, which is defined as the time it takes for the hosts in the enterprise networks to



■ Figure 6. 95th percentile billing.

retrieve a complete Web page from the remote server.

In both cases, significant performance gain can be achieved by multihoming the stub network to two upstream providers; 25 percent for a multihomed enterprise and more than 40 percent for a multihomed data provider. Increasing the number of upstream ISPs also improves performance gain. However, as is also true with the improvement study of reliability, no additional performance gain is achieved if the number of the upstream ISPs is greater than four. Proper selection of the upstream ISP, which is explored in the research of route optimization and is discussed next, also contributes to the improvement of network performance. Choosing an unqualified upstream ISP can downgrade multihoming network performance by as much as 40 percent.

Route Optimization in Multihoming Networks

Route optimization refers to the operation of a stub network to distribute the traffic among its multiple connections to the Internet. Route optimization makes better use of these connections and lowers the cost. Research studies in [22–26] explore the route optimization in multihoming networks and propose several route-optimization mechanisms. Akella [22] evaluates and discusses the effect of different link performance-monitoring strategies such as the paths to choose for monitoring and how frequently monitor programs should send probes to the links. Their work is discussed in detail in the section “Reliability Improvement.” Guo *et al.* [24] discuss the potential design issues in route-optimization mechanisms, such as how to decide whether a link is available and how to assign traffic to links. We further discuss their work in the “Traffic Assignment” section. Lee [25] proposes a method to reduce the probes in the link-performance monitoring process of the route optimization. Their work is discussed next. Goldenberg [23] and Wang [26] propose scheduling methodologies in the link selection and the upstream provider selection respectively. Their works will be discussed further in a later section.

The issues that must be addressed in route optimization include how to leverage the traffic load among multiple connections, also known as the load balance problem in multihoming, and how to choose the most qualified upstream providers. Common performance metrics to evaluate the performance of route-optimization mechanisms include end-to-end path delay, link throughput, and cost/price per connection/link. To construct an effective route-optimization system, the following issues must be studied:

- How to effectively monitor the link performance so that correct information can be provided for the system to fairly distribute the network load among links. An effective monitoring system must enable correct and prompt feedback.
- Based on the obtained data from monitoring, how to distribute traffic among multiple connections to optimize network performance.

- Cost/price limitation is a common constraint in multihoming networks. What does this constraint bring to the existing route-optimization system of multihoming networks?

Performance Monitoring — Monitoring link performance is a very important component of route optimization. Link performance data provided by monitoring is used as a reference both in the load balancing process to decide traffic distribution and in the process of choosing the upstream providers. The monitor programs can be classified as *active monitor programs* and *passive monitor programs*. In the active mode, the monitor program sends probes to the destinations

via the monitored links. In the passive mode, the monitor program samples the monitored links periodically. Active monitor mode programs incur out-of-bound traffic and may be considered illegal in some applications; whereas passive monitor mode programs may not be able to check the fluctuation of link performance in all situations. For instance, it is hard for passive monitor programs alone to distinguish between the *link down* event and the situation where there is no traffic on the link. Therefore, active monitor and passive monitoring approaches should be used in tandem to obtain accurate performance data for the provider links.

Both active and passive mode programs use traffic throughput and round trip delay as the main metrics to evaluate the performance of the links. Two issues must be covered in the monitor programs.

One issue is how to choose the monitored paths. Link performance monitoring is accomplished by monitoring each path passing through the link. It is possible that too many paths require monitoring if the paths for all the destinations of traffic across the network are selected. Most monitoring programs pick up only the important paths. Research shows that AS paths for many destinations from a stub network follow the same path to the Internet core and then diverge to their own destination [25]. Since path segments from the core to the destinations can be determined only by the remote routers where traffic merges, active probes can be sent only to the merged routers for these destinations instead of all the destinations. By sending active probes to the merged routers, the cost of distributing probes can be reduced to 4 percent of the original cost [25].

Another issue for the design of a monitor program is the sampling/probe frequency. Sampling/probing too frequently incurs heavy overhead, but the sample information may not be fresh enough if the time interval between two samples/probes is too large. Studies show that, contrary to intuition, high sampling frequency downgrades the route-optimization module performance; whereas low sampling frequency cannot capture link failure efficiently [22]. Sampling frequencies of about one or two times per minute are suggested in [22].

Traffic Assignment — Based on the link performance data collected in the link monitoring process, traffic can be assigned among the multiple connections of the network. Traffic assignment can be divided into the following three categories according to the assignment granularity level: packet-by-packet link assignment, connection-by-connection link assignment, and session-by-session link assignment. Packet-by-packet link assignment is only available for BGP multihoming networks whose internal hosts have public IP addresses that are covered by the BGP routers of the ISP [24]. Connection-by-connection link assignment is available for the NAT based multihoming network. Session-by-session link assignment can be used for the applications that require more than one connection. For

the last two categories, traffic assignment can be either stateless or stateful. For the stateless assignment, NAT can decide the link ID of the traffic based on the connection/session ID of the packets. For the stateful assignment, NAT must maintain a table to record the link ID of the traffic flow.

Two types of traffic — inbound traffic and outbound traffic — must be assigned among the connections of the upstream providers. It is easy to control flows of outbound traffic but hard to control the flows of inbound traffic. For BGP multihoming, the only mechanism currently available to balance the inbound traffic is by modifying the route advertisement. For example, announce a longer route by prepending or split the stub network into several sub-networks and announce the reachability of each sub-network to the corresponding upstream ISP. However, these methods only statically split traffic among different connections, and the failover time of BGP routing is generally tens of minutes. They do not realize *dynamic* traffic distribution/load balancing.

Inbound-traffic assignment is easier for NAT multihoming. With the help of NAT, inbound traffic that is incurred by the request that originated from inside the stub network is routed to the same connection that issued the request. For the inbound traffic that is incurred by the request that originated from outside the stub network, many existing multi-link load-balance systems can be used to solve the problem because NAT implements address sharing among the hosts of the stub network. One popular method is to use DNS together with NAT, where DNS returns the IP address associated with the link selected, based on its performance data and predefined assignment rules.

Economic Cost Consideration — In addition to network performance, another important issue in route optimization is the economic cost of the different links from different ISPs. Optimizing only technique metrics may trigger a high economic cost. To avoid high economic charges, a general rule is to designate the connection that is inexpensive but less reliable as the primary connection to carry the most traffic and to designate the connection that is expensive but more reliable as its back-up connection.

In the real world, several price structures, such as total volume-based charging and percentile-based charging are available for ISPs to charge traffic volumes in multihoming networks. Among them, 95th percentile billing (shown in Fig. 6) is the most common price structure for the burst traffic [27]. In 95th percentile billing, ISPs sample the traffic every few minutes. At the end of the billing cycle, the ISPs sort the collected sampling data in a non-descending order and charge the customer's traffic based on the $(95 \text{ percent} \times I)$ -th sorted interval, where I is the total number of time intervals of the charging period.

Goldenberg et al. [23] proposed a series of route-optimization mechanisms under the 95th percentile billing price structure. With the constraints of economic cost, the route-optimization problem can be classified further into the following two sub-problems: how to minimize the price of the traffic sent to and received from the upstream providers and with the given price, how to distribute the traffic to obtain the best performance.

For the first sub-problem, the lower bound of the price can be achieved if the traffic intervals that must be assigned to each ISP do not exceed the highest 5 percent of peak-traffic intervals of the charging period, that is, $((1-95 \text{ percent}) \times I)$ -th sorted intervals. If the traffic intervals that must be submitted to ISPs exceed the highest 5 percent of peak-time intervals, traffic assignment can be constructed with a greedy algorithm. This algorithm assigns traffic to such providers in each time interval that a minimal total cost from the beginning of the charging period to this time interval can be achieved. In prac-

tice, more factors must be considered, such as when the capacity of an ISP link is limited, and the customer traffic is not given beforehand. A series of mechanisms based on the aforementioned greedy algorithm also are proposed to meet the extra requirements.

Given the solution of the first sub-problem, the second problem of traffic balancing can be solved by the mechanisms discussed earlier. Specifically, if the traffic in a time interval can be assigned in multiple links with the same cost, then the link is selected based on its performance.

Besides minimizing price with the price structure of given ISPs, stub networks can also choose ISPs that they prefer based on the price given by each ISP. How to choose the proper ISPs is called the *ISP-subscription problem*, and a solution for it, based on the greedy algorithm, is proposed in article [23].

The relationship of customer demands for network and ISP profits is investigated in paper [26]. If customers chose ISPs based only on price, they would always choose the ISP with the lowest price, and eventually ISPs would be unable to earn any profit. However, in reality, customers choose ISP based on both price and performance, which eventually will improve the overall performance of the Internet.

Construct Scalable Multihoming Networks

Most AS only export the routing advertisements of their customers to their providers, peers, and other customers [28]. Thus, the routing table entries in a non-multihoming scenario are aggregated based on a hierarchy of customers and providers. However, this is not true in multihoming networks due to the popularity of non-aggregative address prefixes, and the situation is even worse with the evolution of the Internet toward a densely interconnected mesh network [29]. This not only increases the routing table size but also leads to heavier traffic overhead and longer failover time.

SIMPLER [13] tries to force address prefix aggregation over the entire network under the scenario of multihoming. In SIMPLER, each AS should inherit address sub-blocks from its upstream provider. A NAT mechanism is adopted to map multiple address prefixes to one address space. AS only report a non-aggregative address prefix to their upstream providers on demand, such as when the link associated with that prefix fails. The non-aggregative address-prefix routing information is withdrawn as long as the link is recovered. In this way, the hierarchy structure of a customer and its providers is maintained to the largest extent. This approach constrains the usage of non-aggregative prefixes and thus the routing table size. However, it turns to NAT to solve the problem of assigning multiple addresses.

Open Research Issues in Multihoming Networks

As we discuss in previous sections and show in Table 2, current research in multihoming networks focuses on the following three areas: measurements deployed in multihoming networks to quantitatively investigate network performance, route optimization of multihoming networks to achieve the greatest gain, and new mechanisms to construct scalable multihoming networks. There are many open research issues in these areas, such as the following:

- It is observed that a multihoming network is more reliable and achieves better performance than a non-multihoming network. However, it is still unknown whether the performance of the multihoming network is good enough that one can replace a single high-speed (fat pipe) connection in the stub network with several lower speed connections, for example, using multiple DSL/cable modem lines instead of

a T1 line. Ultimately, how long can existing local-access-network systems satisfy continuously increasing user requirements without major changes? More measurement studies must be undertaken on different types of access connections to answer this question.

- Studies show that optimizing route selection is critical for improving the performance of the multihoming network. Poor route selections downgrade the performance of the multihoming network to 40 percent. Selecting routes for outbound traffic is relatively easier. While mechanisms to implement load balance for outbound traffic with and without NAT are available, no mechanism is available to implement load balance for inbound traffic without NAT. As NAT was originally proposed in the context of client-server applications, it does not support non-client-server applications. Even with enhancements, non-client-server applications still can fail if both parties are behind the NAT.
- Widespread use of multihoming in the Internet would increase the size of the BGP routing table and further slow the failover time of the data links. There are no effective solutions for this problem yet. The mechanism proposed by [13] still relies on NAT, which as in the case of route optimization, can hinder the non-client-server applications.
- Providing quality of service (QoS) is an important task for routing to avoid wasting network resources and fulfill user requirements. Most previous research studies on QoS routing are based on non-multihoming networks. The popularity of multihoming technology gives rise to new issues in QoS routing. New mechanisms must be proposed to build QoS routing on the multihoming network.

Multihoming also brings challenges for network management and security. For example, IP addresses of hosts in a multihoming network may be reassigned frequently or inherit from more than one upstream ISP. This increases the complexity of tracing the traffic related to these hosts. Attacks targeted at destroying address-mapping mechanisms also can be introduced.

Conclusion

This article reviewed the existing solutions for deploying multihoming technology in stub networks, reported the recent research progress, and discussed open areas in multihoming networks.

The existing research focuses on three areas: quantitatively measuring the performance gain of multihoming networks, constructing new multihoming route-optimization algorithms to improve the performance of current multihoming technology, and preventing multihoming from producing heavy overhead and constructing scalable multihoming networks.

Progress has been made in these areas, but there are still many open research issues. More measurement studies must be performed to investigate the performance gain of multihoming technologies over other stub network technologies. Effective route-optimization systems should be proposed to implement real load balancing in BGP multihoming. The non-aggregation problem in BGP multihoming still must be solved. Multihoming also raises new issues in other network research areas such as QoS routing and security.

Acknowledgments

The authors would like to thank the anonymous referees for their critical and constructive comments on this article. They would also like to thank Kim Thompson for reading the article and her suggestions. This work is supported in part by the U.S. National Science Foundation CCF-0514078, CNS-0549006, and CNS-0551464.

References

- [1] J. Stewart, *BGP4: Inter-Domain Routing in the Internet*, Addison Wesley, 1999.
- [2] J. R. Iyengar, P. D. Amer, and R. Stewart, "Receive Buffer Blocking in Concurrent Multipath Transfer," presented at *GLOBECOM*, St. Louis, Mo., 2005.
- [3] A. Caro, P. D. Amer, and R. Stewart, "Retransmission Policies with Transport Layer Multihoming," presented at *ICON*, Sydney, Australia, 2003.
- [4] J. Abley, B. Black, and V. Gill, "RFC 3582: Goals for IPv6 Site-Multihoming Architectures," 2003.
- [5] E. Nordmark and T. Li, "RFC4218: Threats Relating to IPv6 Multihoming Solutions," 2005.
- [6] C. Launois and M. Bagnulo, "The Paths Towards IPv6 Multihoming," *IEEE Commun. Surveys and Tutorials*, vol. 8, 2006.
- [7] D. Allen, "NPN: Multihoming and Route Optimization: Finding the Best Way Home," *IEEE Network Mag.*, 2002.
- [8] J. Kim, S. Bahk, and H. Lee, "A Connection Management Protocol for Stateful Inspection Firewalls in Multi-Homed Networks," presented at *IEEE Int'l. Conf. Commun.*, 2004.
- [9] A. Dhamdhere and C. Dovrolis, "ISP and Egress Path Selection for Multihomed Networks," presented at *INFOCOM*, 2006.
- [10] J. Bartlett, "Optimizing Multi-homed Connections," *Business Commun. Review*, 2002.
- [11] N. Feamster, J. Borkenhagen, and J. Rexford, "Guidelines for the Interdomain Traffic Engineering," *SIGCOMM Comp. Commun. Review*, vol. 33, 2003, pp. 19-30.
- [12] B. Quoitin, S. Uhlig, and O. Bonaventure, "Using Redistribution Communities for Interdomain Traffic Engineering," *Lecture Notes in Computer Science*, vol. 2511, 2002, pp. 125-34.
- [13] R. Gummadi and R. Govindan, "Practical Routing-Layer Support for Scalable Multihoming," presented at *INFOCOM*, Miami, FL, 2005.
- [14] A. Akella *et al.*, "A Comparison of Overlay Routing and Multihoming Route Control," presented at *SIGCOMM*, Portland, Ore., 2004.
- [15] A. Akella and A. S. R. Sitaraman, "A Measurement-Based Analysis of Multihoming," presented at *SIGCOMM*, Karlsruhe, Germany, 2003.
- [16] J. Han and F. Jahanian, "Impact of Path Diversity on Multi-homed and Overlay Networks," presented at *DSN*, 2004.
- [17] Y. Rekhter and T. Li, "RFC1518: An Architecture for IP Address Allocation with CIDR," 1993.
- [18] T. Bates and Y. Rekhter, "RFC2260: Scalable Support for Multi-homed Multi-provider Connectivity," 1998.
- [19] K. Egevang and P. Francis, "RFC1631: The IP Network Address Translator (NAT)," 1994.
- [20] B. Ford, P. Srisuresh, and D. Kegel, "Peer-to-Peer Communication Across Network Address Translators," presented at *USENIX Annual Technical Conf.*, Anaheim, CA, 2003.
- [21] M. Holdrege and P. Srisuresh, "RFC3027: Protocol Complications with the IP Network Address Translator," 2001.
- [22] A. Akella, S. Seshan, and A. Shaikh, "Multihoming Performance Benefits: An Experimental Evaluation of Practical Enterprise Strategies," presented at *USENIX Annual Technical Conf.*, Boston, MA, 2004.
- [23] D. K. Goldenberg *et al.*, "Optimizing Cost and Performance for Multihoming," presented at *SIGCOMM*, Portland, Ore., 2004.
- [24] F. Guo *et al.*, "Experiences in Building a Multihoming Load Balancing System," presented at *IEEE INFOCOM*, 2004.
- [25] S. Lee, Z. Zhang, and S. Nelakuditi, "Exploiting AS Hierarchy for Scalable Route Selection in Multi-homed Stub Networks," presented at *ACM SIGCOMM Internet Measurement Conf.*, Taormina, Italy, 2004.
- [26] H. Wang *et al.*, "Optimal ISP Subscription for Internet Multihoming: Algorithm Design and Implication Analysis," presented at *INFOCOM*, Miami, FL, 2005.
- [27] M. Ripeanu, "Peer-to-Peer Architecture Case Study: Gnutella Network," presented at *Int'l. Conf. Peer-to-Peer Computing*, 2001.
- [28] L. Subramanian *et al.*, "Characterizing the Internet Hierarchy from Multiple Vantage Points," presented at *INFOCOM*, New York, NY, 2002.
- [29] G. Huston, "The Unreliable Internet," *Broadband Satellite Column*, 2001.

Biographies

XIAOMEI LIU (liuxiao@cse.msu.edu) received a B.S. degree in electronics and information system technology from East China Normal University in 1996 and an M.S. degree in computer engineering from the University of Toledo in 2000. She is currently a Ph.D. student at Michigan State University. Her research interests include distributed operating systems and computer networks. She is a student member of the IEEE and the IEEE Computer Society.

LI XIAO (lixiao@cse.msu.edu) received her B.S. and M.S. degrees in computer science from Northwestern Polytechnic University, China and her Ph.D. degree in computer science from the College of William and Mary in 2002. She is an assistant professor of computer science and engineering at Michigan State University. Her research interests are in the areas of distributed and Internet systems, overlay systems and applications, and sensor networks. She is a member of the ACM, the IEEE, the IEEE Computer Society, and IEEE Women in Engineering.