

December, 2004

Load Sharing in Multihomed Host
[draft-liumin-multi6-loadsharing-00.txt](#)

Status of this memo

This document is an Internet-Draft and is subject to all provisions of [section 3 of RFC 3667](#). By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with [RFC 3668](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on June, 2005.

Copyright Notice

Copyright (C) The Internet Society (2004).

Expires June,2005

[Page 1]

Internet-Draft
Abstract

loadsharing

December 2004

In order to reach the goal of load sharing and traffic engineering in multihoming, there must be some mechanism for the local host to make a selection of the "best" source locator to used. Obviously the selection includes the objective to select a currently viable path. What's more, it also includes the objective to select a path with

larger bandwidth, which is more difficult to judge than reachability. In this memo, we propose a simple mechanism to determine the availability and bandwidth condition between a multihomed host and its ISP. It will help to share the load among multiple paths and provide better performance for the host's traffic. This mechanism could be part of traffic engineering functions, which could be used as the basis of locator selection before initial session establishment and aslo could be used to trigger locator swithes to avoid loss of connectivity, long delay or large loss rate.

Expires June,2005

[Page 2]

Internet-Draft

loadsharing

December 2004

Table of Contents:

1	Introduction.....	4
2	Problem Statement.....	5
3	Scenario Description	6
4	Solution Description	7
4.1	Overview of the Procedure.....	7
4.1.1	Assumption.....	8
4.1.2	List of Available ISP.....	8
4.1.3	Locator Selection.....	9
4.2	Idle Rate Estimation.....	10
5	Security Consideration.....	12
6	IANA Considerations	12
	References.....	13
	Authors' Addresses.....	13
	Intellectual Property and Copyright Statements.....	14

Expires June,2005

[Page 3]

Internet-Draft

loadsharing

December 2004

[1.](#) Introduction

Multiple access links could be used to improve the aggregate bandwidth and the overall availability of Internet connectivity. When the access links are subscribed from different ISPs, this approach is called multihoming. Multihoming is a mechanism by which enterprises can satisfy a number of high-level requirements. As described in [\[6\]](#), with the advent of inexpensive and high-bandwidth broadband

connection technologies for home users, multihoming is poised to emerge from a niche technique for large enterprises and become a dominant technology that underlies the Internet connectivity solutions for small to mid-sized businesses.

[RFC 3582](#) [1] documents some goals that a multihoming approach should attempt to address, among which load sharing and traffic engineering are two important goals. Then the problem is how does the local host make a selection of the "best" source locator to use? Obviously the selection includes the objective to select a currently viable path. What's more, it also includes the objective to select a path with larger bandwidth, short delay or small loss rate, which is more difficult to judge than reachability. At the conceptual level, a multihoming load balancing/sharing system must be able to determine the available bandwidth to a remote subnet through an access link, assign incoming and outgoing Internet traffic to the available access links, and detect failed access links and divert Internet traffic around them. However, in practice, it is very difficult to obtain such a clear picture of the Internet. Therefore, some compromises must be made.

In this memo, we propose a simple mechanism to determine the availability and bandwidth condition between a multihomed host and its ISP. It will help to share the load among multiple paths and provides better performance for the host's traffic in an efficient fashion. This mechanism could be part of traffic engineering functions, which could be used as the basis of locator selection before initial session establishment and also could be used to trigger locator switches to avoid loss of connectivity, long delay or large loss rate.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

Expires June,2005

[Page 4]

Internet-Draft

loadsharing

December 2004

[2.](#) Problem Statement

While multihoming to multiple providers is motivated primarily by a need for link-level and provider-level fault tolerance, there are other goals for subscribers to leverage multihoming. For example, performance to different parts of the network may vary depending on which upstream provider is used. In such situations, careful route selection can significantly improve performance.

In this memo, we propose a simple mechanism to help the local host make a selection of the "best" source locator to used before initial session establishment and trigger locator swithes if necessary. This mechanism will have the following characteristics:

- o It is simple and easy to deployed.
- o The goal of this mechanism is to help implement load sharing and traffic engineering in multihoming.
- o The mechanism could be implemented as a protocol element within the host or at a site-exit router.
- o The mechanism could determine the availability and bandwidth condition between a multihomed host and its ISP. To decrease the cost, end-to-end bandwidth condition will not be determined.
- o The probing traffic is very little. The probing packets could also be used as keep-alive traffic if there are some tunnles.
- o The idle rate estimation function in the mechanism could reports a test result per second, which could be used to estimate available bandwidth and update the records. This time granularity could be configured as needed.
- o Host could configure the threshold of bandwidth to trigger locator swithes.
- o To provide the user or the application the ability to configure different "Weight" to different transmission technology or access network, for matters of cost, efficiency, politics, etc.
- o After determining the "best" candidate source locator, in order to verify that the locator is working, a Reachability Test exchange is needed. The details of Reachability Test is out of the scope of this draft.

Expires June,2005

[Page 5]

Internet-Draft

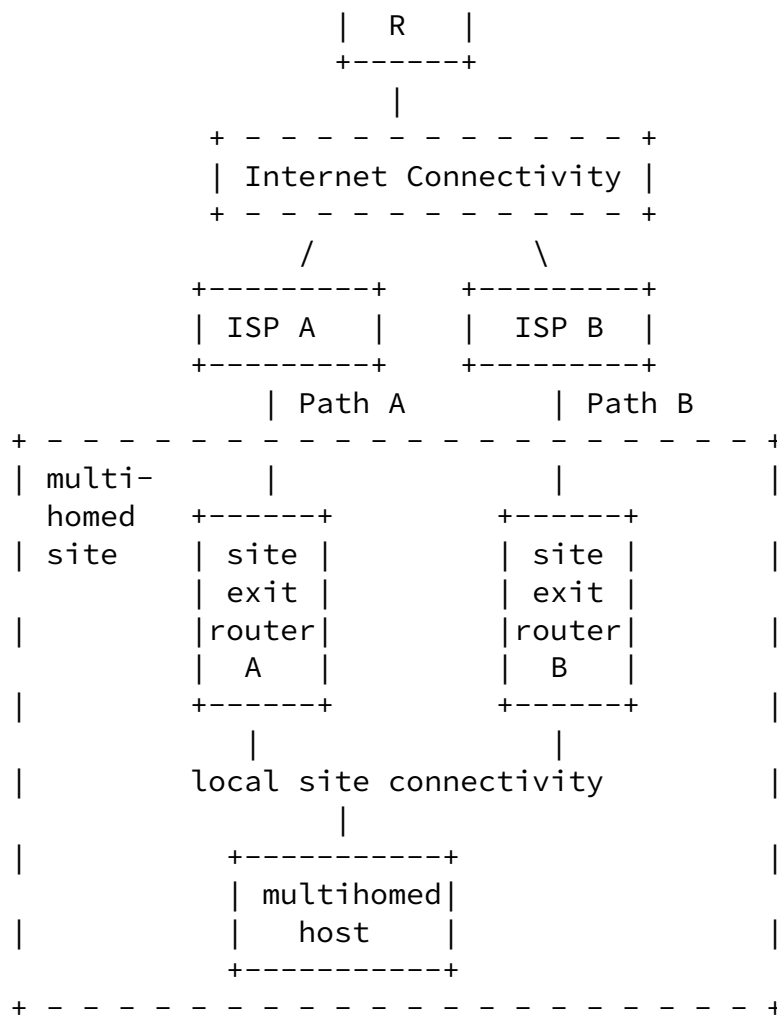
loadsharing

December 2004

[3.](#) Scenario Description

The scenario that we concerned in our solution is the simple multihoming environment showed in Figure 1(Figure 1 in [\[3\]](#)).

```
+-----+
|remote|
| host |
```



The major characteristic of this scenario is that the address space used by, and advertised as reachable by, ISP A is distinct from the address space used by ISP B.

More complex scenarios such as the local multihomed host is communicating with a remote multihomed host, and the local host should have some discretion in the choice of a destination locator is out of the scope of this draft.

4 Solution Description

Available bandwidth (avail-bw for short) is a crucial metric in capacity provisioning, routing, traffic engineering, QoS management and so on. At the conceptual level, it is also the basis for a multihoming load balancing/sharing system. However, end-to-end avail-bw estimation is a very challenging task, which often needs a long measurement period and a large quantity of probing traffic. As a

result, currently, end-to-end avail-bw estimation still remain at the theoretical research level and impractical for many actual applications. Thus in multihoming environment, it is impossible to determine the available bandwidth from a host to a remote subnet in real time to make a selection of the "best" source locator to used before initial session establishment or to trigger locator swithes when necessary.

In the simple multihoming environment showed in Figure 1, when we only consider the availability and bandwidth condition between a multihomed host and its ISP, we can simplify the scenario and estimate the network condition efficiently with low cost.

[4.1](#) Overview of the Procedure

The procedure can be showed in Figure 2:

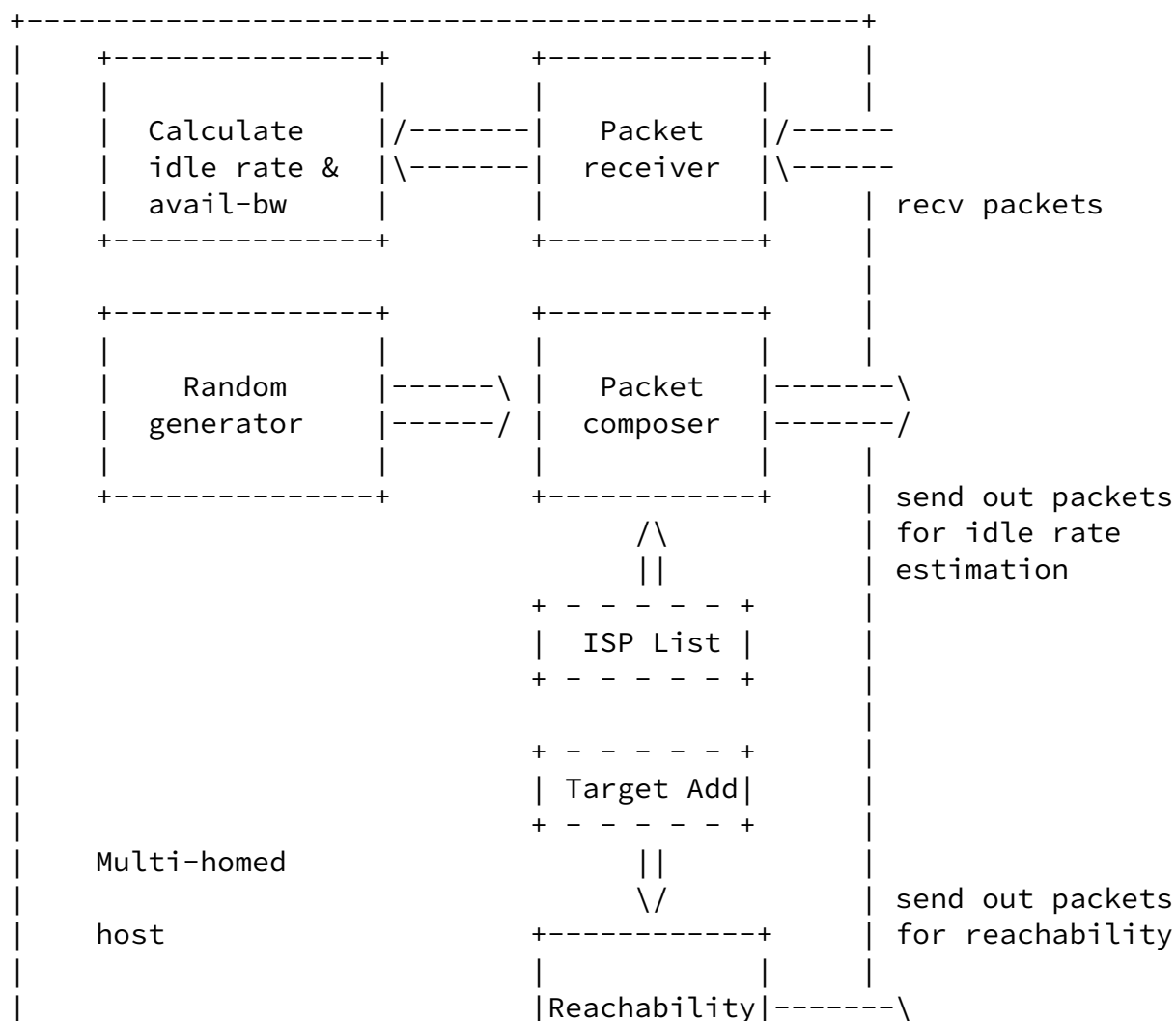
Expires June,2005

[Page 7]

Internet-Draft

loadsharing

December 2004



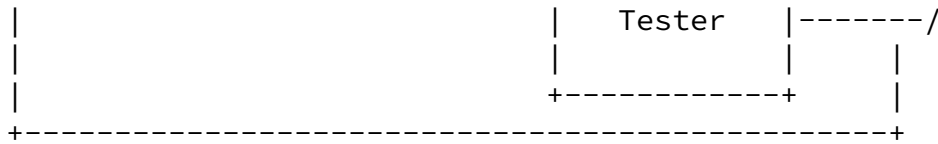


Figure 2. solution procedure

4.1.1 Assumption

We assume that the capacity between a multihomed host and its ISP is known. Generally, capacity is unchanged and could be known from ISP or through some router inquiry software. Thus, if we could estimate the idle rate of the path from the host to its ISP, the avail-bw condition could be indicated to a certain extent.

4.1.2 List of Available ISP

For a multihomed host which has several upstream provider to choose between, we could estimate the idle rate of the path from the host

Expires June,2005

[Page 8]

Internet-Draft

loadsharing

December 2004

to its each upstream provider in real time. For example, in Figure 1, we assume the capacity of Path A is $C(A)$. We could estimate the idle rate of Path A, which could be referred to as $I(A)$. Then the product of $C(A)$ and $I(A)$ could indicate the avail-bw condition of Path A, which could be referred to as $BW(A)$. Similarly, $I(B)$ and $BW(B)$ could also be estimated. Each ISP and its corresponding capacity, idle rate and avail-bw will be stored in the "list of available ISP". Basically, each network path has different cost, performance, access range, and reliability. The user or the application have the ability to configure different "Weight" to different transmission technology or access network. The item "Weight" will also be stored in the "list of available ISP".

4.1.3 Locator Selection

Local attempts to initiate a communication with remote hosts should take into account the current connectivity state in undertaking locator selection and setting up initial locator sets. Before initiate a new session, locator selection processor should search the "list of available ISP" to do some load sharing. When the session does not have specific requirement of bandwidth, the locator selection processor could implement flexible load sharing policy based on the items in "list of available ISP". However, if the session has specific requirement of bandwidth, only ISPs with sufficient avail-bw will be considered for load sharing. If no ISP has sufficient avail-bw, different policies could be performed. For example, one feedback could be sent to corresponding applications if

needed; the access control element could refuse the new session; or the locator selection processor could choose the ISP with the largest avail-bw as the candidate source locator.

After determining the "best" candidate source locator, in order to verify that the locator is working, a Reachability Test exchange is needed. This can be used to check if the locator that is chosen is working properly. The design and implementation of Reachability Test is out of the scope of this draft.

If the Reachability Test succeeds, the candidate source locator will become the selected initial locator for this session. Because the idle rate and avail-bw items in the "list of available ISP" are updated in real time, the change of the network conditions could be monitored in an assured time scope. Host could configure the threshold for these metrics to trigger locator swithes, thus avoid any loss of end-to-end connectivity or QoS on the selected initial locator.

Expires June,2005

[Page 9]

Internet-Draft

loadsharing

December 2004

[4.2](#) Idle Rate Estimation

In the procedure of locator selection described in last section, the kernel problem is how to estimate the idle rate of the path from the host to its ISP in real time. The fundamental idea in our proposal is to send very small packets at random moment and calculate the proportion of minimum delay in total test samples.

For one probing packet, if its queuing delay from the host to the ISP is zero, we could consider the path is in idle state at the moment. After sending a number of random probing packets, we can estimate the idle rate in this period.

Probing packets could be ICMP echo request packets or UDP packets. The measurement method is similar. The only difference is that the measurement must adopt the C/S architecture when using UDP probing packets, while ICMP measurement could be implemented only in the multihomed host, thus need no modification in the ISP. However, this method relies on the routers in ISP handling ICMP packets and timely delivery of acknowledgement.

In order to decrease the overhead of probing traffic, the probing packet should be as small as possible. In addition, considering that if the interval between probing packets is too short, the front packet will have great influence on the latter one, the interval

should be large enough.

In fact, what we can measure is the transmission delay, not queuing delay. The transmission delay is the sum of two components: a fixed propagation delay and a variable queuing delay. We assume the observed minimum delay in all probing packets as the propagation delay. In other words, we assume the minimum delay indicate the path is in idle state at the sending time. Certainly, such an assumption may result in error in case of a very congestion network, where none probing packet has been transmitted without queuing delay because the avail-bw is zero. But we can take some action to eliminate such error. For example, we can record the number of probing packets received by the idle rate estimation function. When the avail-bw is extremely small, there will be packet loss. So if the number of accepted packets is significantly different from the prospective value, we can draw the conclusion that the idle rate/avail-bw approximates zero, thereby avoid the error in the assumption.

For each measurement, the transmission delay of each probing packet will be saved in an array. When this measurement is over, the number of accepted packets will be compared with the sending number. If the number of accepted packets is significantly different from the

prospective value, the idle rate/avail-bw approximates zero. Otherwise, search the array and find the minimum value, and calculate the arisen times of this value. We can calculate the proportion of the minimum result in total test samples and get the idle rate in the measurement period.

The idle rate estimation function could reports a test result per second, which could be used to estimate available bandwidth and update the records in "list of available ISP". This time granularity could be configured as needed. At the host startup, the function calculates the avail-bw from all probe traffic. When the probe traffic becomes more than 50 packets, the function calculates avail-bw from the last 50 packets. The number of probing packets in calculation could also be configured and modified.

The probing packets could be used as keep-alive traffic if there are some tunnels between the host and its ISP. For this purpose, it needs an estimate of the "lifetime" parameter used in the tunneling technique.

Expires June,2005

[Page 11]

[5](#) Security Consideration

Because the idle rate estimation is implemented between host and its ISP, it will not open up a host on either end of a communication to a time-based attack. The authentication for host will follow the specific ISP's security consideration and existing mechanisms. Thus the security considerations for this proposal should be the same as described in [\[4\]](#).

[6](#) IANA Considerations

This document has no associated IANA actions.

Expires June,2005

[Page 12]

Normative References

- [1] Abley, J., Black, B. and V. Gill, "Goals for IPv6

Site-Multihoming Architectures", [RFC 3582](#), August 2003.

Informative References

- [2] Lear, E., "Things MULTI6 Developers should think about", [draft-ietf-multi6-things-to-think-about-00.txt](#)(Work in progress), June 2004.
- [3] Huston, G., "Architectural Approaches to Multi-Homing for IPv6", [draft-ietf-multi6-architecture-02.txt](#)(Work in progress), October 2004.
- [4] Nordmark, E. and T. Li, "Threats relating to IPv6 multi-homing solutions", [draft-ietf-multi6-multihoming-threats-01.txt](#)(Work in progress), July 2004.
- [5] Hinden, R., "IPv6 Host to Router Load Sharing", [draft-ietf-ipv6-host-load-sharing-02](#) (work in progress), May 2004.
- [6] Guo, F. and Chen, J., et al. "Experiences in Building A Multihoming Load Balancing System", INFOCOM 2004.

Authors' Addresses

Min Liu
Institute of Computing Technology
Chinese Academy of Sciences
Box 2704, Beijing, 100080 PRC
Email: liumin@ict.ac.cn

Expires June,2005

[Page 13]

Internet-Draft

loadsharing

December 2004

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of

such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

Expires June,2005

[Page 14]