Network Working Group D. Crocker
Internet-Draft Brandenburg InternetWorking
Expires: August 9, 2004 A. Doria
ETRI
February 9, 2004

Framework for Common Endpoint Locator Pools
draft-crocker-celp-00.txt

Status of this Memo

Copyright Notice

Abstract

Classic Internet transport protocols use a single source IP Address and a single destination IP Address, as part of the identification for an individual transport data flow. This is problematic for multihomed hosts and for mobile hosts, collectively needing "multiaddressing" support. The basic goal, then, is to find a method for multiaddressing that makes the smallest possible change to the architecture and to current systems, while maintaining flexibility for different emerging solutions. This draft proposes a framework for methods for creating Common Endpoint Locator Pools that can be used by and among the proposed solutions.

Acknowledgment

Thanks go to those who participated in the original SLAP discussion, many of whose ideas and words on the list have been borrowed for this draft. The contributors include Brian Carpenter, Spenser Dawkins, Christian Huitema, and Pekka Nikander.

Table of Contents

1. Introduction

This draft attempts to capture and expand upon the discussion on Shared Locator Address Pool (SLAP) that was held on the Multi6 mailing list.

Classic Internet transport protocols use a single source IP Address and a single destination IP Address, as part of the identification for an individual transport data flow. For example, TCP includes these in its definition of a connection and its calculation of the header checksum. Hence a classic transport association is tied to a particular IP Address pair. This is problematic for multihomed hosts and for mobile hosts, collectively needing "multiaddressing" support. Both have access to multiple IP Addresses -- a "pool" of topologically-related locators -- but they are prevented from using more than one within an existing transport exchange. For a host to use a different IP Address pair, participants must initiate a new exchange. In the case of TCP, this means a new connection.

In recent years, there have been efforts to overcome many of these limitations, through different approaches at different places in the Internet architecture. Some modify the IP infrastructure, with embedded redirection services. Some define transport enhancements to support a set of locators directly, and some define a layer between classic IP and classic transport. Each of the existing proposals has notable limitations in functionality, implementation, deployment or use. A discussion of the architectural choices and summary of existing multiaddressing projects is in [CHOICE]. The locator schemes offered in these efforts comes in two varieties; transport based and wedge based. In the former, multiaddressing support is made an integral part of a particular transport service. In the wedge based approach, multiaddressing support resides in a functional layer between IP and transport.

Transport-based locator-pool schemes ([SCTP], [DCCP], [TCP-MH]) multiplex their control exchange in with data traffic. Also, the transport layer can naturally obtain information on the quality of different paths. For example, SCTP can perform measurements across several paths simultaneously, and can then map flows on one or another path. TCP-MH can detect that the current path has stopped working well, for example if the frequency of repetition becomes too high, and can decide to try another path. Wedge-layer approaches ([HIP], [LIN6], [MAST], [MIP6]) must conduct their control exchange on a separate logical channel. If a host must do this exchange on a separate channel, with every other host it talks to, the aggregate overhead can be high. Hence transport based schemes have an the potential advantage of saving on number of packets.

On the other hand, wedge based approaches can maintain multiaddressing information across transport associations and can maintain pools with different referential granularity. That is, they can have a pool for all associations between two hosts or they can subdivide into different pools for different activities between the hosts. The logical terminus for these pools with a more narrow scope is called an "endpoint". Hence the next transport activity between two endpoints may well be able to use multiaddressing immediately and with no further administrative overhead. Further, wedge-based locator exchange protocols can be incorporated without necessitating modification to any host's IP or transport modules. Wedge protocols may be invoked at any time, before or during a transport association. A host may initiate and conduct a classic, single IP-pair TCP connection. It may then separately query for remote host support of the wedge protocols and initiate a endpoint locator exchange to be used by that connectivity. Either participant is then free to add or remove endpoint locators. Of course, use of a wedge protocol may instead be performed before a transport context is established, so that future contexts immediately have access to multiple endpoint locators. Because many associations are short, initiation of a wedge protocol can be deferred entirely, choosing to apply it only for persistent associations.

The objective of this framework is to permit benefits for all participants. A wedge layer scheme shares "enhanced" transport service's lower-cost locator pool control exchange largess. A transport scheme shares the control exchange the wedge layer has already done, before the transport association is initiated.

1.1 Terminology

Work on multiaddressing is forcing significant changes and greater precision, in the use of some common networking terminology. In this document, the term "address" is used in the introduction, for its familiarity, and then is restricted to be part of the label "IP Address", to refer to that protocol's locator values. Similarly use of the term "host" is restricted to refer to the assignee of one or more IP Addresses. For discussing multiaddressing pools, the term "endpoint" is used to refer to a pool with potentially smaller scope. In general, this document takes its terminology from [CHOICE].

1.2 Discussion Venue

Discussion and commentary are encouraged about the topics presented in this document. The preferred forum is the <mailto:multi6@ops.ietf.org> mailing list, for which archives and subscription information are available at <http://ietf.org/html.charters/multi6-charter.html>.

2. Basic Proposal

The basic proposal can be described in the following three propositions:

1. An endpoint runs endpoint Locator Pools (LP) as a resource shared among different consumer services at the endpoint -- for example, a wedge layer service and an enhanced transport service -- potentially with multiple maintainers. LPs might vary in their granularity. Bigger grains makes it more likely that the pool will be shared. A pool that is the finest grain (Connection) can't be shared, of course.
2. A Common Endpoint Locator Address Pool (CELP) capability is used by the different maintenance services, over different communication channels (for example, multiplexed on the transport channel, versus over an independent control channel.) The maintenance services also might use different "configurations", such as peer-to-peer exchange, versus third-party agent mediation.
3. Enhanced transport services access the pool directly. Legacy transport services access the pool via the IP wedge layer service. If an enhanced transport is one of the participants, then there really is no need for a wedge-layer service to conduct an exchange. This saves packets. Hence the wedge layer serves to use the information provided by the transport scheme and apply it to legacy transport and application services.

3. Issues

3.1 Variable Granularity

One transport activity may wish (or need) to share its multiaddressing fate with another. Or it may wish to avoid the problems with that sharing, and tolerate the limitations that ensue. These choices can be achieved by having LPs with different scope. At the widest scope, all multiaddressing between a host-pair is shared by all transport associations between them. Hence, the common pool is characterized by:

    {local, remote}

For finer-grained sharing, the pool can be characterized with a variety of additional attributes. For example, and IPv6 flow might be used:

    {local, remote, flow label}

or all activity for a specific application service might share the

pool. This could be characterized by:

{local, remote, IP protocol, Well-known port}

or a single transport association might wish to have its own pool, as characterized by the classic connection identifier:

{local, remote, IP protocol, local port, remote port}

3.2 Security Threats

As described in [THREATS] there are redirection threats that may occur when multihoming is used. A CELP solution will need to respond to those threats as well to any exacerbation of DOS and other flooding attacks.

3.3 Changes elsewhere in the architecture

In order to support a CELP solution, will other entities in the network need to modified? For example will changes be required in DNS, network management, or trace mechanisms? Additionally, there is the need to determine whether third-party mediation services are required or even warranted.

3.4 Pool synchronization

If several protocols are 'maintaining' the pool there arises a concern about synchronization of the state information in the pool. One consideration is the creation of a protocol to maintain CELP state. It is unclear whether this will be necessary.

3.5 Endpoint Congestion Management

With a CELP mechanism, the transport may not see the locator information, instead seeing only an identifier. However, differential congestion control is needed at the locator level. This indicates that bringing congestion control into the core of the pool mechanism as part of creating pools may be necessary.

3.6 Coordination with other efforts

Given the number of efforts in the IETF at the moment that are suggesting modification to the transport layer and below, it is necessary to coordinate with these efforts.

References

[CHOICE] Crocker, D., "Choices for Support of Multiaddressing", draft-crocker-mast-analysis-00.txt (work in progress), Oct 2003.

[DCCP] Handley, M., Floyd, S. and J. Padhye, "Datagram Congestion Control Protocol (DCCP)", draft-ietf-dccp-spec-05.txt (work in progress), Oct 2003.

[HIP] Moskowitz, R., Nikander, P. and Henderson, "Host Identity Protocol", draft-moskowitz-hip-08.txt (work in progress), Oct 2003.

[LIN6] Teraoka, F., Ishiyama, M. and M. Kunishi, "LIN6: A Solution to Mobility and Multi-Homing in IPv6", draft-teraoka-ipng-lin6-02.txt (work in progress), June 2003.

[MAST] Crocker, D., "MULTIPLE ADDRESS SERVICE FOR TRANSPORT (MAST): AN EXTENDED PROPOSAL", draft-crocker-mast-proposal-01.txt (work in progress), Sept 2003.

[MIP6] Johnson, D., Perkins, C. and J. Arkko, "Mobility Support in IPv6", draft-ietf-mobileip-ipv6-24.txt (work in progress), June 2003.

[SCTP] Stewart, R., Xie, Q., Schwarzbauer, K., Morneault, K., Sharp, C., Taylor, T., Rytina, I., Kalla, M., Zhang, L. and V. Paxson, "Stream Control Transmission Protocol", RFC 2960 RFC 2960, Oct 2000.

[TCP-MH] Matsumoto, A., Kozuka, M., Fujikawa, K. and Y. Okabe, "TCP Multi-Home Options", draft-arifumi-tcp-mh-00.txt (work in progress), Oct 2003.

[THREATS] Nordmark, E. and T. Li, "", draft-nordmark-multi6-threats-00.txt (work in progress), Oct 2003.

Authors' Addresses

D. Crocker
Brandenburg InternetWorking
675 Spruce Drive
Sunnyvale 94086
USA

Phone: +1.408.246.8253
EMail: dcrocker@brandenburg.com
URI: http://brandenburg.com/

Avri Doria
ETRI

EMail: avri@acm.org
URI: psg.com/~avri

Intellectual Property Statement

Full Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.