Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: May 17, 2017

G. Chen
China Mobile
C. Williams
Consultant
D. Wing
A. Yourtchenko
Cisco Systems, Inc.
November 13, 2016

Happy Eyeballs Extension for Multiple Interfaces draft-ietf-mif-happy-eyeballs-extension-11

Abstract

This memo proposes extensions to the Happy Eyeball's algorithm requirements defined in RFC6555 for use with the multiple provisioning domain architecture. The Happy Eyeballs in MIF would make the selection process smoother by using connectivity tests over pre-filtered interfaces according to defined policy. This would choose the best interface with an automatic fallback mechanism.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on May 17, 2017.

Table of Contents

1. Introduction

The MIF problem statement [RFC6418] describes problems specific for nodes attached to multiple provisioning domains. Specifically, there is a issue description that a node has selected an interface and obtained a valid IP address from the network, but Internet connectivity is not available. This memo intends to address the issue and elaborate more in Section 3.1.

[RFC7556] describes the multiple provisioning domain architecture. It refers to using connectivity tests to validate a Provisioning Domain (PvD). Given a number of implicit/explicit PvDs, plus preferences/policy, what is the process to follow to select the best PvD to use for any given connection. In the event that two or more are deemed to be best, how are the Happy Eyeballs (HE) techniques applied to find the best and deal with resilience. This memo also proposes process requirements using Happy Eyeballs (HE) extensions.

There are a variety of algorithms that can be envisioned. This document describes additional parameters and processes that need to be considered in addition to the HE algorithm requirements defined in [RFC6555] necessary to support multiple interfaces, so that a node with multiple interfaces can select the best path for a particular connection-oriented flow (e.g., TCP, SCTP).

2. Terminology

This document makes use of following terms:

- Happy Eyeballs (HE): specifies requirements for an algorithm that reduces the user-visible connection delay for dual-stack hosts with a single interface per-protocol.

- Happy Eyeballs - Multi-Interface (HE-MIF): Extends the Happy Eyeballs concept to the multiple provisioning domain architecture. It describes additional requirements for algorithms that offer connectivity tests on PVD-aware or non-PVD-aware nodes [RFC7556] to select the best interface for a specific connection request.

3. Use Cases

The section describes scenarios the HE-MIF targeted to use.

3.1. WiFi is broken

Assuming a MIF node has both a 3GPP mobile network interface and a WiFi interface, a common practice would be to always prefer the WiFi connection when the node enters an area with WiFi available. In this

situation, a node might assume that because a valid IP address has been allocated, the WiFi link provides connectivity to destinations through the Internet. However, this might not be the case for several reasons:

- WiFi access-point authentication requirements
- WiFi has no global Internet connectivity
- Instability at layer 2

In order to resolve this problem, the user would need to disable the device's interface preferences, e.g. by disabling the WiFi interface. HE-MIF offers users the possibility of configuring their preferences for the choice of the most suitable network interface to use, such as via setting on their mobile phone.

In this case, users may prefer to wait an appropriate time period for connections to be established over a WiFi path. If no connection can be made it will fall back to attempting the connection over a 3GPP mobile network path.

3.2. Policy Conflict

A node has network access via both WiFi and 3GPP networks. In a mobile network, IPv6-only may be preferable since IPv6 has the potential to be simpler than dual-stack. The WiFi access offers IPv4 only. In this scenario, the combination of source address selection [RFC6724] and preferring the WiFi interface may cause a problem. The transition to IPv6 may mean that IPv6 is the preferred protocol, so the 3GPP interface should be chosen even though it could be considered a suboptimal selection e.g. the WiFi interface likely is less expensive.

4. Happiness Parameters

This section provides input parameter proposal that HE-MIF should catch. Two sets of "Happiness" parameters have been defined. It serves applications and initiates HE-MIF connection tests subsequently. By following the process described below, MIF nodes can select an appropriate interface that best meets the configuration parameters defined by the user. The two sets of "Happiness" parameters are called Hard Set and Soft Set respectively.

4.1. Hard Set

Hard set contains parameters which should be complied with. It helps to select candidate interfaces through which a particular flow should be directed. These should be seen as constraints on the choice, such as provider policies, support for IPv4 or IPv6, and other parameters which would prevent a particular interface and transport from being used by a particular flow. Parameters in the hard set should be easy to use and understand. When several parameters in the hard set are in conflict, the user's preference should be prioritized.

4.1.1. Operator Policy

Operators may deliver the customized policies for a particular network environment because of geo-location or service regulation considerations. One example relevant for 3GPP networks is an operator delivering policies from an Access Network Discovery and Selection function (ANDSF) [TS23.402].

The ANDSF provides a node with policies and network selection information to influence the selection between different access technologies, such as 3GPP mobile networks, WiFi access. The ANDSF can provide the node with three types of information[TS24.302].

- Access network discovery and selection information: it includes a list of access networks available in the vicinity of the node. The information may include the access technology types (e.g. WiFi), network identifiers (e.g. SSID in the case of WiFi) as well as validity conditions (e.g. where and when).

- Inter-System Mobility Policies (ISMPs): they are a set of operator-defined rules and preferences that affect the inter-system mobility decisions, e.g. decisions about whether to use 3GPP mobile network or a WiFi network.

- Inter-System Routing Policies (ISRPs): the node uses ISRPs when it can route IP traffic simultaneously over multiple radio access networks. It could provide routing policies in an IP flow granularity.

4.1.2. User Preference

User's preference: users may express preferences which likely not have a formally technical language , like "No 3/4G while roaming", "Only download applications larger than 20Mb over WiFi", etc. Those information are normally input from User Interface (UI).

4.2. Soft Set

Soft set contains factors which impact the selection of the path across which a particular flow should be transmitted among the available interfaces and transports which meet the hard set requirements described above.

4.2.1. Provisioning Domain Identity

A PVD-aware node uses PvD Identity(PvD-ID) to select a PvD with a matching ID for special-purpose connection requests. The PvD-ID may be generated by the node implicitly or received from the network explicitly. for explicit PvDs, the node could take the parameter from PvD ID Option [I-D.ietf-mif-mpvd-id] via the configuration protocols ([I-D.ietf-mif-mpvd-dhcp-support] or [I-D.ietf-mif-mpvd-ndp-support]). A PVD-aware node may decide to use one preferred PVD or allow the use of multiple PVDs simultaneously for applications. The node behavior should be consistent with MPVD architecture [RFC7556].

4.2.2. DNS Selection

At the name service lookup step, the node has to choose a recursive DNS server to use. A HE-MIF node should take the parameter of RDNSS Selection DHCP Option [RFC6731] to select an interface for a particular namespace.

4.2.3. Next Hop

[RFC4191] allows the configuration of specific routes to a destination. A HE-MIF node should take the parameters of router preference and route information to identify the next hop.

4.2.4. Source Address Selection

For each destination, once the best next hop is found, the node should consider IP prefix and precedence parameter in policy table to select the best source address according to the rule defined in [RFC6724].

4.2.5. Common Practice

There is relevant common practice related to interface selection, e.g. Prefer WiFi over a 3GPP interface, if available. Such conventions should also be considered.

5. HE-MIF Process Requirements

An HE-MIF node may use the two sets of parameters as two steps in the interface selection process. The first step is to use the Hard Set to synthesize policies from different actors (e.g., users or network operators). These hard set parameters will provide a filter which will exclude not qualifying interfaces from any further consideration.

The second step is to influence how a node makes a connection when multiple interfaces still remain in the candidate list after first step. This is essentially sorting behavior. In the multiple provisioning domain architecture, a PVD aware node makes connectivity tests as described in Section 5.3 of [RFC7556]. A PVD agnostic node take other parameters apart from PVD-ID in the Soft Set to proceed the sort process.

The two steps are described in more details in the following sub-sections. It should be noted that HE-MIF does not prescribe such two-step model. It will be very specific to particular cases and implementations. The two step model mainly describes requirements for how to use the hard/soft set.

5.1. First Step, Filter

One goal of the filter is to reconcile multiple selection policies from users or operators. Afterwards, merged demands would be mapped to a set of candidate interfaces, which are judged as qualified.

Decision on the reconciliation of different policies will depend very much on the deployment scenario. An implementation may not be able to determine priority for each policies without explicit configuration provided by users or administrator. For example, an implementation may by default always prefer the WiFi because of cost saving consideration. Whereas, other users may turn off a device's WiFi interface to guarantee use of a 3GPP network interface to assure higher reliability or security.

The decision on mergence of policies may be made by implementations, or by node administrators. However, it's worth to note that a demand from users should be normally considered higher priority than from other actors.

The merged policies serve as a filter which is iterated across the list of available interfaces. Qualified interfaces are selected and the proceed to the second step.

5.2. Second Step, Sort

5.2.1. Interface Validation

The Sort process aims to select the best interface and provide fallback capacities. As stated in [RFC7556], a PVD-aware node shall perform connectivity tests and, only after validation of the PVD, consider using it to serve application connections requests. In current implementations, some nodes already implement this, e.g., by trying to reach a dedicated web server (see Section 3.1.2 [RFC6419]). If anything is abnormal, it assumes there is a proxy on the path. This status detection is recommended to be used in HE-MIF to detect DNS interception or an HTTP proxy that forces a login or a click-through. Unexamined PVDs or interfaces should be accounted as "unconnected". It should not join the sort process.

5.2.2. Name Resolution

Name resolution is executed on the validated interfaces. Before the requests are initiated, it should check if there is a matching PVD ID for the destination name. A PVD agnostic node may request DNS server selection DHCP option [RFC6731] for interface selection guidance. Those information may weight a particular interface to be preferred to others sending resolving requests. If the node can't find useful information in the Soft Set, DNS queries would be sent out on multiple interfaces in parallel to maximize chances for connectivity. Some additional discussions of DNS selection consideration of HE-MIF are described in Section 7.3.

5.2.3. Connection Establishment

Once a destination address was resolved, a connection is to be setup. For the given destination address, a PVD-aware node selects a next-hop and source address associated with that PVD in the name resolution process. A PVD agnostic node may receive certain next hop in a RA message [RFC4191], the node selects best source address according to the rules [RFC6724].

The interface identified by the source address should be treated to initiate the connection prior to others. This could avoid thrashing the network, by not making simultaneous connection attempts on multiple interfaces. After making a connection attempt on the preferred pairs and failing to establish a connection within a certain time period (see Section 7.2), a HE-MIF implementation will decide to initiate connection attempt using rest of interfaces in parallel. This fallback consideration will make subsequent connection attempts successful on non-preferable interfaces.

The node would cache information regarding the outcome of each connection attempt. Cache entries would be flushed periodically. A system-defined timeout may take place to age the state. Maximum on the order of 10 minutes defined in [RFC6555] is recommended to keep the interface state changes synchronizing with IP family states.

If there is no specific Soft Set provided, all selected interfaces should be treated equally. for a node implementing multipath transports (for example, Multipath TCP (MPTCP) [RFC6182]), the interfaces could be treated as valid to perform subsequent multipath process, such as starting subflow. A node only supporting single physical transport would initiate on several interface simultaneously. The goal here is to provide the most fast connection for users, by quickly attempting to connect using each candidate interface. Afterwards, the node would do the same caching and flushing process as described above.

6. Implementation Framework

The simplest way to implement the processes described in this document is within the application itself. This would not require any specific support from the operating system beyond the commonly available APIs that provide transport service. It could also be implemented using a high-level API approach, linking to the MIF-API [I-D.ietf-mif-api-extension].

7. Additional Considerations

7.1. Usage Scope

Connection-oriented transports (e.g., TCP, SCTP) are directly applied as scoped in [RFC6555]. For connectionless transport protocols (e.g., UDP), a similar mechanism can be used if the application has request/response semantics. Further investigations are out of the document scope.

7.2. Fallback Timeout

When the preferred interface was failed, HE-MIF would trigger a fallback process to start connection initiation on several candidate interfaces. A period of time should be set to invalidate the interface and fallback to others. Aggressive timeouts may achieve quick interface handover, but at the cost of traffic that may be chargeable on certain networks, e.g. the handover from WiFi to 3GPP networks brings a charge to customers. Considering the reasons, it is recommended to prioritize the input from users (e.g., real customers or applications) through user interface. For default-setting on a system, a hard error [RFC1122] in replied ICMP could

serve as a trigger for the fallback process. When the ICMP soft error is present or non-response was received, it's recommended that the timeout should be large enough to allow connection retransmission. [RFC1122] states that such timer must be at least 3 minutes to provide TCP retransmission. However, several minutes delay may not inappropriate for user experiences. A widespread practice [RFC5461] sets 75 seconds to optimize connection process.

More optimal timer may be expected. The particular setting will be very specific to implementations and cases. The memo didn't try to provide a concrete value because of following concerns.

- RTT (Round-Trip Time) on different interfaces may vary quite a lot. A particular value of timeout may not accurately help to make a decision that this interface doesn't work at all. On the contrary, it may cause a misjudgment on a interface, which is not very fast. In order to compensate the issues, the timeout setting based on past experiences of a particular interface may help to make a fair decision. Whereas, it's going beyond the capability of Happy Eyeballs [RFC6555]. Therefore, it leaves a particular implementation.

- In some cases, fast interface may not be treated as "best". For example, a interface could be evaluated in the principle of bandwidth-delay, termed "Bandwidth-Delay-Product ". Happy Eyeballs measures only connection speed. That is, how quickly a TCP connection is established . It does not measure bandwidth. If the fallback has to take various factors into account and make balanced decision, it's better to resort to a specific context and implementation.

7.3. DNS Selections

During the Sort process, HE-MIF prioritizes PVD-ID match or [RFC6731] inputs to select a proper server. It could help to address following two cases.

- A DNS answer may be only valid for a specific provisioning domain, but the DNS resolver may not be aware of that because the DNS reply is not kept with the provisioning from which the answer comes. The situation may become worse if asking internal name with public address response or asking public name with private address answers.

- Some FQDNs can be resolvable only by sending queries to the right server (e.g., intranet services). Otherwise, a response with NXDOMAIN is replied. Fast response is treated as optimal only if

the record is valid. That may cause messy for data connections, since NXDOMAIN doesn't provide useful information.

HE-MIF can help to solve the issues of DNS interception with captive portal. The DNS server modified and replied the answer with the IP address of captive portal rather than the intended destination address. In those cases, TCP connection may succeed, but Internet connectivity is not available. It results in lack of service unless user has authenticated. HE-MIF recommended using network connectivity status probes to examine a pre-configured URL for detecting DNS interception on the path (see more in Section 5.2). The node will be able to automatically rely upon other interfaces to select right DNS servers by excluding the unexamined interfaces.

7.4. Flow Continuity

[I-D.deng-mif-api-session-continuity-guide] describes session continuity guidance for application developers. The flow continuity topic is beyond this document scope.

7.5. Interworking with Happy Eyeball

HE-MIF process could cooperate with HE [RFC6555]. HE is executed on an interface which is selected to make connection establishment (see Section 5.2.3). for example, a node following PvD policy to pick a interface and make both IPv4/IPv6 connection attempts in consistent with HE requirements. The interface state management in HE-MIF is designed to synchronize with IP family states. It could facilitate the HE executions.

7.6. Multipath Applicability

Some nodes may support transports that provide an abstraction of a single connection, aggregating multiple underlying connections. Multipath TCP (MPTCP) [RFC6182] is an example of such a transport protocol. For connections provided by such transports, a node may leverage the "happiness" parameters and process on the underlying connections. Following the HE-MIF requirements, each connection could be performed consistently with user/operator's preference and corresponding provisioning domain information.

8. IANA Considerations

This memo does not include any IANA requests.

9. Security Considerations

The security consideration is following the statement in [RFC6555] and [RFC6418].

10. Acknowledgements

The authors would like to thank Margaret Wasserman, Hui Deng, Erik Kline, Stuart Cheshire, Teemu Savolainen, Jonne Soininen, Simon Perreault, Zhen Cao, Dmitry Anipko, Ted Lemon, Daniel Migault, Russ White and Bing Liu for their helpful comments.

Many thanks to Ralph Droms, Ian Farrer, Jouni Korhonen, Mirja Khlewind and Suresh Krishnan for their detailed reviews.

11. References

11.1. Normative References

[RFC1122] Braden, R., Ed., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, DOI 10.17487/RFC1122, October 1989, <http://www.rfc-editor.org/info/rfc1122>.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, DOI 10.17487/RFC4191, November 2005, <http://www.rfc-editor.org/info/rfc4191>.

[RFC6555] Wing, D. and A. Yourtchenko, "Happy Eyeballs: Success with Dual-Stack Hosts", RFC 6555, DOI 10.17487/RFC6555, April 2012, <http://www.rfc-editor.org/info/rfc6555>.

[RFC6724] Thaler, D., Ed., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol Version 6 (IPv6)", RFC 6724, DOI 10.17487/RFC6724, September 2012, <http://www.rfc-editor.org/info/rfc6724>.

[RFC6731] Savolainen, T., Kato, J., and T. Lemon, "Improved Recursive DNS Server Selection for Multi-Interfaced Nodes", RFC 6731, DOI 10.17487/RFC6731, December 2012, <http://www.rfc-editor.org/info/rfc6731>.

[TS23.402] 3rd Generation Partnership Project, 3GPP., "Architecture enhancements for non-3GPP accesses v8.8.0", December 2009.

[TS24.302]
3rd Generation Partnership Project, 3GPP., "Access to the 3GPP Evolved Packet Core (EPC) via non-3GPP access networks v14.0.0", June 2016.

11.2. Informative References

[I-D.deng-mif-api-session-continuity-guide]
Deng, H., Krishnan, S., Lemon, T., and M. Wasserman, "Guide for application developers on session continuity by using MIF API", draft-deng-mif-api-session-continuity-guide-04 (work in progress), July 2014.

[I-D.ietf-mif-api-extension]
Liu, D., Lemon, T., Ismailov, Y., and Z. Cao, "MIF API consideration", draft-ietf-mif-api-extension-05 (work in progress), February 2014.

[I-D.ietf-mif-mpvd-dhcp-support]
Krishnan, S., Korhonen, J., and S. Bhandari, "Support for multiple provisioning domains in DHCPv6", draft-ietf-mif-mpvd-dhcp-support-02 (work in progress), October 2015.

[I-D.ietf-mif-mpvd-id]
Krishnan, S., Korhonen, J., Bhandari, S., and S. Gundavelli, "Identification of provisioning domains", draft-ietf-mif-mpvd-id-02 (work in progress), October 2015.

[I-D.ietf-mif-mpvd-ndp-support]
Korhonen, J., Krishnan, S., and S. Gundavelli, "Support for multiple provisioning domains in IPv6 Neighbor Discovery Protocol", draft-ietf-mif-mpvd-ndp-support-03 (work in progress), February 2016.

[RFC5461] Gont, F., "TCP's Reaction to Soft Errors", RFC 5461, DOI 10.17487/RFC5461, February 2009, <http://www.rfc-editor.org/info/rfc5461>.

[RFC6182] Ford, A., Raiciu, C., Handley, M., Barre, S., and J. Iyengar, "Architectural Guidelines for Multipath TCP Development", RFC 6182, DOI 10.17487/RFC6182, March 2011, <http://www.rfc-editor.org/info/rfc6182>.

[RFC6418] Blanchet, M. and P. Seite, "Multiple Interfaces and Provisioning Domains Problem Statement", RFC 6418, DOI 10.17487/RFC6418, November 2011, <http://www.rfc-editor.org/info/rfc6418>.

[RFC6419] Wasserman, M. and P. Seite, "Current Practices for Multiple-Interface Hosts", RFC 6419, DOI 10.17487/RFC6419, November 2011, <http://www.rfc-editor.org/info/rfc6419>.

[RFC7556] Anipko, D., Ed., "Multiple Provisioning Domain Architecture", RFC 7556, DOI 10.17487/RFC7556, June 2015, <http://www.rfc-editor.org/info/rfc7556>.

Authors' Addresses

Gang Chen
China Mobile
29, Jinrong Avenue
Xicheng District,
Beijing 100033
China

Email: phdgang@gmail.com, chengang@chinamobile.com

Carl Williams
Consultant
El Camino Real
Palo Alto, CA 94306
USA

Email: carlw@mcsr-labs.org

Dan Wing
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
USA

Email: dwing@cisco.com

Andrew Yourtchenko
Cisco Systems, Inc.
De Kleetlaan, 7
Diegem B-1831
Belgium

Email: ayourtch@cisco.com