Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: March 01, 2014

G. Chen
China Mobile
C. Williams
Consultant
D. Wing
A. Yourtchenko
Cisco Systems, Inc.
August 28, 2013

Happy Eyeballs Extension for Multiple Interfaces
draft-ietf-mif-happy-eyeballs-extension-03

Abstract

Currently the interface selection in multi-interface environment is exclusive - only one interface can be used at the time, frequently needing manual intervention. Happy Eyeballs in MIF would make the selection process smoother by using the connectivity checks over a pre-filtered interfaces according to defined policy. This would choose "best" interface with an automatic fallback.

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 March 01, 2014.

Copyright Notice

Table of Contents

1. Introduction

In multiple interface context, the problems raised by hosts with multiple interfaces have been discussed. The MIF problem statement[RFC6418] described the various issues when using a wrong domain selection on a MIF node. Happy Eyeballs (HE) [RFC6555] described how a dual-stack client can determine the functioning path to a dual-stack server. It's using stateful algorithm to help applications quickly determine if IPv6 or IPv4 is the most fast path to connect a server. That is a good method to achieve smart path selection. However, the assumption is a single-homed context. The interaction with multiple interfaces is deferred for further study.

[I-D.anipko-mif-mpvd-arch] has proposed a multiple provisioning domain architecture. This memo has been proposed to extend happy eyeballs algorithm to fit into the multiple interfaces context. Several additional considerations have been elaborated to analyze the user demands and initiate HE-MIF connections. It allows a node with multiple interfaces picking a fast flow path.

2. Problem Statement

The section enumerates several concrete use cases in existing networks.

Case 1: WiFi is broken

- [Scenario] A MIF node has both 3G and WIFI interface. When the node enters a WiFi area, a common practice would always prefer WiFi because it' cheap and fast-speed normally.

- [Problem] User assumes the wifi is working, because the node already got IP address from WiFi. However, he can't run applications due to Internet connectivity being unavailable. This may be an authentication required coming into play, or unstable Layer 2 conditions. In order to figure out the problems, users have to turn off the WiFi manually.

- [Workaround] Users can indicate their desire with some setting on the phone. For instance, they may prefer to wait a little bit of time but not forever. After the timer is expired, users would finally give up the WiFi path and try to establish connection over 3G path. Users may won't want the wait time too short, because the 3G path for most people is more expensive than wifi path.

Case 2: VPN (Virtual Private Network) scenario

- [Scenario] In some cases, a node has multiple interface because of VPN. Users would only have interests to connect a corporate network inside VPN. While, connecting to Internet would work outside the VPN.

- [Problem] That is normally a implementation consideration that unmanaged interface may be considered less trustworthy than managed. It results in trusted interfaces having the highest priority. This setting may steer all traffic to VPN interface. When this is a traffic heading to a corporate site, everything is fine. But sometimes, the connections out to Internet sites may suffer from long-distance path delays.

- [Workaround] It's desirable if routing could be bound to each interface. However, a node following weak host model[RFC1122] takes routing tables as node-scoped. Some sophisticated VPN softwares may configure a specific route setting on each interface to dispatch traffic in a predetermined network environment. As an alternative, It may be useful to perform parallel IP connectivity checks before selecting an interface. Consequently, the fastest interface would be picked up automatically.

Case 3: 3G/LTE tethering scenario

- [Scenario] Many mobile phones are equipped with software to offer tethered Internet access. It shares their Internet connection with another Internet-capable mobile phone or other devices over Wi-Fi.

- [Problem] The WiFi link that tethered phone see is not free WiFi link, i.e. it might be 3G backhaul. The policy of "always WiFi" leads to all traffic being sent over the tethering WiFi. Usually, such tethering WiFi link puts sharing limitation to access nodes. It could cause contention on both that WiFi link and the backhaul 3G link, while it be higher cost than going on the 3G that is built in the handset.

- [Workaround] To solve that, it is necessary for the node to be aware of not only the link layer information, but also services information, like billable or free. That could help to facilitate the execution of the algorithm. Same concern has been documented in Section 4.4 of [RFC6418])

Case 4: Policy Conflict

- [Scenario] A node has WiFi and 3G access simultaneously. In mobile network, IPv6-only may be preferable since IPv6 has the potential to be simpler than dual-stack. WiFi access still remain on IPv4.

- [Problem] The problem is caused by policy confliction. The transition to IPv6 is likely to encourage IPv6 and prefer IPv6[RFC6724]. If the 3G path has IPv6 on it and the WiFi does not, a suboptimal interface might be chosen from the cost saving perspective.

- [Workaround] Users interests should be well understood and considered before interface selection. The different preconditions may impact subsequent behaviors. Users concern about high-reliability or high-speed or less-cost should make different choice. A flexible mechanism should be provided allow to make smart decision.

3. Happiness Parameters

To solve the problems, this section provides the design proposal for HE-MIF. Two sets of "Happiness" parameters have been defined. It serves upper applications and initiates HE-MIF connections to below level API subsequently. Going through the process, MIF nodes could pick an appropriate interface which would correspond to user demands.

The two sets of "Happiness" parameters are called Hard set and Soft set respectively.

- Hard set: It contains parameters which have mandatory indications that interface behaviour should comply with. This might provide an interface for applications constraints or delivering operator's policies. Basically, parameters in Hard set should be easy-to-use and easy-to-understand. The potential users would directly use those. When several hard parameters were conflicted, user's preference should override.

  - User's preference: users would express preferences which may not have a formally technical language , like "No 3G while roaming", "Only use free WiFi", etc.

  - Operator policies: operators would deliver the customized policies in particular network environments due to geo-location or services regulation considerations. One example in 3GPP network is that operator could deliver policies from access network discovery and selection function (ANDSF).

- Soft set: It's a factor contributing to the best path. The following is considered as for the justification.

  - Next hop: [RFC4191] allows configuration of specific routes to a destination.

  - DNS selection: [RFC6731] could configure nodes with information to indicate DNS server address for a particular namespace.

  - Source address selection: the information provided by [RFC6724] would be considered.

  - Other factors: There is a common practice may impact interface selection, e.g. WiFi is preferable. Such conventional experiences should also be considered.

4. HE Behaviour in MIF

Corresponding to the two sets of parameters, a HE-MIF node may take a two-step approach. One is to do "hard" decision to synthesize policies from different actors (e.g., users and network operator). In a nutshell, that is a filter which will exclude the interfaces from any further consideration. The second is to adjust how we make a connection on multiple interfaces after the filter. It's a sorting behaviour. In the multiple provisioning domain architecture, Provisioning Domain (PVD) selection is performed based on "hard" and "soft" inputs. Connections intend to be initiated on the resultant

PVDs in parallel. Those two steps are described as following sub-sections. It should be noted that HE-MIF doesn't prescribe such two-step model. It will be very specific to particular cases and implementations. For example, if one interface or PVD is left after the first step, the process would be closed.

4.1. First Step, Filter

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

Decision on 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 due to cost saving consideration. Whereas, users may dedicatedly prefer 3G interface to seek high-reliability or security benefits even to actively turn off WiFi interface. The decision on mergence of policies may be made by implementations, by node administrators, even by other standards investigating customer behaviour. However, it's worth to note that a demand from users should be normally considered higher priority than from other actors.

The merged policies would serve as a filter principle doing iterate across the list of all known interfaces. Qualified interface would be selected to sort processing at next step.

4.2. Second Step, Sort

Sort process would guarantee "best" interface selection with fallback capacities. As soon as a node connects to a network at bootstrap or changes to a different network, network connectivity status probes have been performed in some existing implementations, e.g. Windows Vista, Windows 7, Windows Server 2008 and iOS. In the process, a pre-configured URL have been connected to examine a certain answer. 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 HTTP proxy that forces a login or a click-through. The unexamined interfaces should be accounted as "unconnected" . Those interfaces should not join the sort process. For a PVD-aware node, it could instinctively avoid the mismatch of provisionning information. Those status detection behaviors may not be applied to such node.

Two phases normally are involved in a sort process, i.e. name resolving and data session establishing. Parameters in soft set should considered at this stage.

When the node initiates name requests, it should follow the instruction in [RFC6731]if DNS server selection DHCP option is provided. Otherwise, DNS queries would be sent out on multiple interfaces on relevant PVDs in parallel. More discussions of DNS selection in HE-MIF are elaborated at Section 6.3.

Once a peer address was resolved, a connection would be intended to setup. Heading to a destination, a particular interface on relevant PVDs may comply with the configuration of soft parameters , e.g. next hop[RFC4191] , source address selection[RFC6724] or a common practice. A particular interface should be treated with higher priority compared to others. And, it should be choose to initiate the connection in advance. This could avoid thrashing the network, by not (always) making simultaneous connection attempts on multiple interfaces. After making a connection attempt on the preferred interface and failing to establish a connection within a certain time period (see Section 6.2), a HE-MIF implementation will decide to initiate connection attempt using rest of interfaces in parallel. This fallback consideration may make subsequent connection attempts successful on non-preferable interface.

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 filmily states. So long as new connections are being attempted by the MIF-node, such an implementation should occasionally make connection attempts using the soft-parameter's preferred interface, as it may have become functional again.

If there are no specific soft-parameters provided, all selected interface on relevant PVDs should be equally treated. The connections would initiate on several interface simultaneously. The goal here is to provide fast connection for users, by quickly attempting to connect using one of interfaces. Afterwards, the node would do the same caching and flushing process as described above.

5. Implementation Framework

The simplest way for the implementation is within the application itself. The mechanism described in the document would not require any specific support from the operating system beyond the commonly available APIs that provide transport service. It could also be

implemented as high-level API approach, linking to MIF-API [I-D.ietf-mif-api-extension]. A number of enhancements could be added, making the use of the high-level APIs much more productive in building applications.

6. Additional Considerations

6.1. Usage Scope

Connection-oriented transports (e.g., TCP, SCTP) could be directly applied as scoped in [RFC6555]. For connectionless transport protocols (e.g., UDP), it was also described "a similar mechanism can be used if the application has request/ response semantics (e.g., as done by Interactive Connectivity Establishment (ICE) to select a working IPv6 or IPv4 media path[RFC6157])."

6.2. Fallback Timeout

When the preferred interface was failed, HE-MIF would trigger fallback process to start connection initiation on several candidate interfaces. It should set a reasonable wait time to comfort user experiences. 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 3G would bring a bill to customers. Considering the reasons, it is recommended to prioritize the input from users(e.g. real customers or applications) through UI(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. 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 due to 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's superior to leave it to 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.

6.3. DNS Selections

In the sort process, HE-MIF prioritizes [RFC6731]inputs to select a proper server. [RFC6731]could help to address following two cases that HE-MIF failed to address.

- A DNS answer may be only valid on a specific provisioning domain, but HE-MIF may not be aware of that mapping because DNS reply may not be 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. HE-MIF treats the DNS answer with fast response as optimal only if the record is valid. That may cause messy for data connections, since NXDOMAIN doesn't provide useful information.

By doing HE-MIF, it 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 4.2). The node will be able to automatically rely upon other interfaces to select right DNS servers by excluding the unexamined interfaces.

It should be noticed that both [RFC6731]and HE-MIF can't fully solve the problems of DNS resolution issues, which was described in Section 2.3 of [RFC6731]. In order to handle the issues, a MIF-node should have PVD-aware capability to explicitly differentiate various provisioning domains.

6.4. Flow Continuity

Interface changing should only happen at the beginning of new session in order to keep flow continuity for ongoing TCP session. Dynamic movement of traffic flows are beyond the scope of this document.

7. IANA Considerations

This memo includes no request to IANA.

8. Security Considerations

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

9. 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 and Ted Lemon for their helpful comments.

10. References

10.1. Normative References

[RFC1122] Braden, R., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, October 1989.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, November 2005.

[RFC6555] Wing, D. and A. Yourtchenko, "Happy Eyeballs: Success with Dual-Stack Hosts", RFC 6555, April 2012.

[RFC6724] Thaler, D., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol Version 6 (IPv6)", RFC 6724, September 2012.

[RFC6731] Savolainen, T., Kato, J., and T. Lemon, "Improved Recursive DNS Server Selection for Multi-Interfaced Nodes", RFC 6731, December 2012.

10.2. Informative References

[I-D.anipko-mif-mpvd-arch] Anipko, D., "Multiple Provisioning Domain Architecture", draft-anipko-mif-mpvd-arch-02 (work in progress), July 2013.

[I-D.ietf-mif-api-extension] Liu, D., Lemon, T., and Z. Cao, "MIF API consideration", draft-ietf-mif-api-extension-03 (work in progress), November 2012.

[RFC5461] Gont, F., "TCP's Reaction to Soft Errors", RFC 5461, February 2009.

[RFC6157] Camarillo, G., El Malki, K., and V. Gurbani, "IPv6 Transition in the Session Initiation Protocol (SIP)", RFC 6157, April 2011.

[RFC6418] Blanchet, M. and P. Seite, "Multiple Interfaces and Provisioning Domains Problem Statement", RFC 6418, November 2011.

Authors' Addresses

Gang Chen
China Mobile
53A,Xibianmennei Ave.,
Xuanwu District,
Beijing 100053
China

Email: phdgang@gmail.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