NETEXT Working Group
Internet-Draft
Intended status: Experimental
Expires: September 9, 2010

CJ. Bernardos
UC3M
T. Melia
Alcatel-Lucent Bell Labs
P. Seite
France Telecom
J. Korhonen
Nokia Siemens Networks
March 8, 2010

Multihoming extensions for Proxy Mobile IPv6
draft-bernardos-mif-pmip-02

Abstract

The IETF standardized Proxy Mobile IPv6 (PMIPv6). PMIPv6 enables mobile devices to connect to a PMIPv6 domain and roam across gateways without changing the IP address. PMIPv6 also provides limited multi-homing support to multi-mode mobile devices. The IETF is working on optimizations for PMIPv6. While multi-homing item has been proposed to be part of the approved work, discussions showed there are still many controversial issues to be addressed (i.e. the no-host modification theorem). This document explores solutions for the multi-homing use case aiming at helping PMIPv6 development where possible.

Requirements Language

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 RFC 2119 [RFC2119].

Status of this Memo

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

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/1id-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 September 9, 2010.

Copyright Notice

Table of Contents

1. Introduction

Proxy Mobile IPv6 (PMIPv6), specified in RFC 5213 [RFC5213] and [I-D.ietf-netlmm-pmip6-ipv4-support], provides network based mobility management to hosts connecting to a PMIPv6 domain. PMIPv6 introduces two new functional entities, the Local Mobility Anchor (LMA) and the Mobility Access Gateway (MAG). The MAG is the first layer three hop detecting Mobile Node (MN) attachment and providing IP connectivity. The LMA is the entity assigning one or more Home Network Prefixes (HNPs) and zero or one IPv4 Home Address (IPv4-MN-HoA)to the MN and is the topological anchor for all traffic from/to the MN.

PMIPv6 allows an MN to connect to the same PMIPv6 domain through different interfaces. ID [I-D.devarapalli-netext-multi-interface-support] identifies at least three possible scenarios, namely i) unique prefix per interface, ii) same prefix but different global addresses per interface, iii) shared address across multiple interfaces. The ID further describes issues associated with each scenario. The first two scenarios are similar, and bring similar issues, whereas the third one is more complex to tackle, since it requires to deal with the sharing of the same IP address across different interfaces. This document focuses on the two first scenarios, as depicted in Figure 1. However, if [RFC1918] defined private IPv4 addresses are used as IPv4 Home Addresses, the scenario iii) may happen implicitly. Unless the LMA coordinates private IPv4 Home Addresses across different access technologies and mobility session, then there is a possibility that the same private IPv4 Home Address would be assigned to both if1 and if2 of the MN.

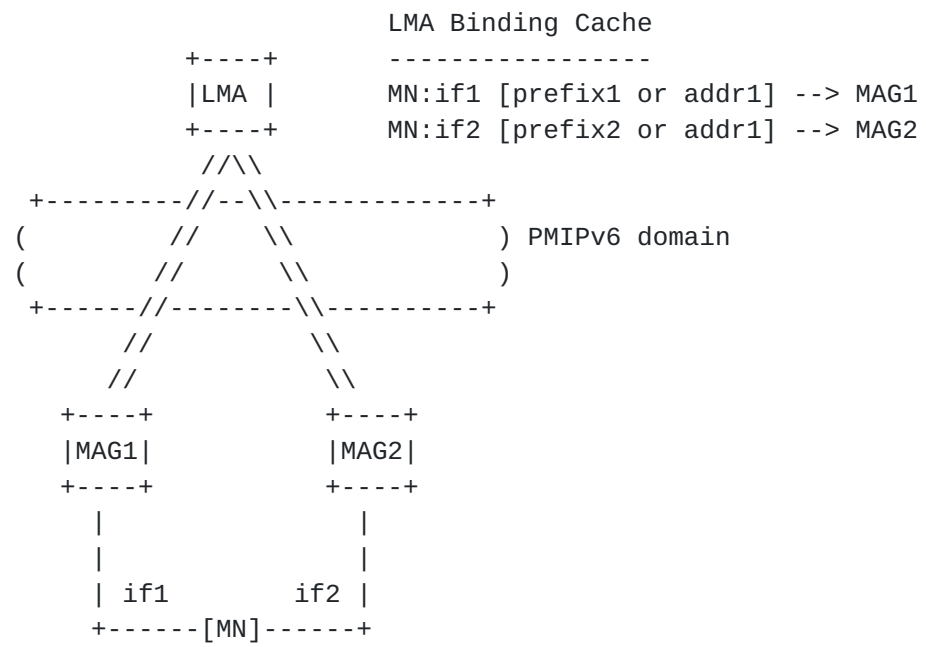

Figure 1: Unique prefix and Unique address per Interface scenarios

The fact is that many (client) hosts currently have the ability to attach to multiple networks simultaneously, and that implies benefits (e.g., enables load balancing, improved connectivity, higher throughput and better reliability, etc.), but also brings some operation issues (e.g., default router selection, address selection, DNS server selection, choice of interface for packet transmission, the treatment of configuration information received from the various networks, etc.). Configuration decisions about how to deal with the different information from each of the interface might have a very strong impact on the connectivity experienced by a node with multiple network interfaces (from now on we refer a node with multiple network interfaces as a MIF node).

In the context of PMIPv6, current specification [RFC5213] does not address the case of a MIF node attaching to a PMIPv6 domain other than stating it is possible. We argue it is important to enable PMIPv6 to bring MIF nodes the advantages related to the simultaneous use of multiple interfaces. Moreover a MIF node could be seen as a not-modified host implementing the right technology for multi-interface handling.

2. MIF scope and PMIPv6

Current scope of MIF nodes as described in [I-D.ietf-mif-problem-statement] only covers the issues of host attaching to multiple networks. The current work is focused on

documenting the system level effects to host IP stacks and identification of gaps between the existing IETF recommendations and existing practice, both for IPv4 and IPv6.

While [I-D.ietf-mif-problem-statement] is not addressing any (neither flow nor host nor network) mobility, a MIF node might find itself connected to a PMIPv6 domain. PMIPv6 should be extended to efficiently support MIF nodes attaching to a PMIPv6 domain, enabling features such as the ones identified in [I-D.jeyatharan-netext-multihoming-ps], e.g., dynamic mobility sessions between different interfaces, allowing traffic to be forwarded to any of the interfaces of a MIF node, not only to the one configured with the destination prefix/address of that traffic).

3. A use case

This section describes a simple use case of a MIF node in a PMIPv6 domain, as an example of a situation where PMIPv6 needs to be extended.

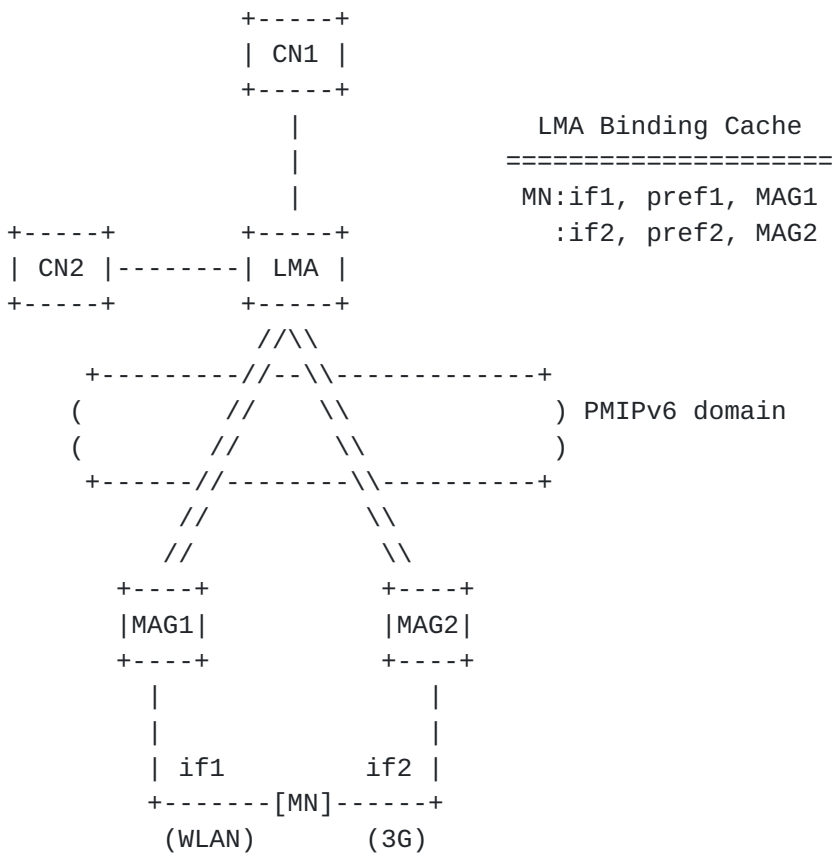

Figure 2: Use case

Figure 2 shows a potential use case of interest involving an MIF mobile node attached to a PMIPv6 domain. The MN is attached to MAG1 through its WLAN interface (if1), and to MAG2 through its 3G interface (if2). Lets consider the case in which each interface has been assigned a different prefix by the LMA (for the sake of simplicity we have left the IPv4 case out of this example). Two different mobility bindings are created in the LMA referring to the MN. In this scenario, if the MN decides to move if1 from MAG1 to a different MAG of the same domain, the PMIPv6 support would take care of ensuring that the same prefix (pref1) is assigned at the new MAG (we assume that there is an L2 identifier for if1 that the new MAG can include in the PBU).

Lets assume for the sake of this example that the MN starts a communication with CN1, using as source IPv6 address (pref1::if1) the one assigned to its WLAN interface (if1), and that it also starts a different communication with CN2, using as source IPv6 address (pref2::if2) the one assigned to its 3G interface (if2). In this scenario, it would be useful to enable the MN be able to receive traffic addressed to pref1::if1 via if2 and vice versa. However, current PMIPv6 specification does not support this. Analogously, it would be also useful to allow the MN send traffic with source address pref1::if1 through if2 and vice versa.

We argue in the next section that PMIPv6 could benefit from MIF outcomes to support the previous scenario while limiting impact on the LMA and MAG operation.

4. Considerations on feasibility and approach overview

We analyse in the next sections the feasibility of the scenario presented in Section 3, by identifying the requirements and changes that would be needed in PMIPv6 to support it. In this version of the document we do not specify with all the required details the solution, but rather concentrate on the concept, with the goal of triggering the discussion within the IETF.

Figure 3 shows in a glimpse the extensions to PMIPv6 required to support the MIF example scenario shown in Section 3.

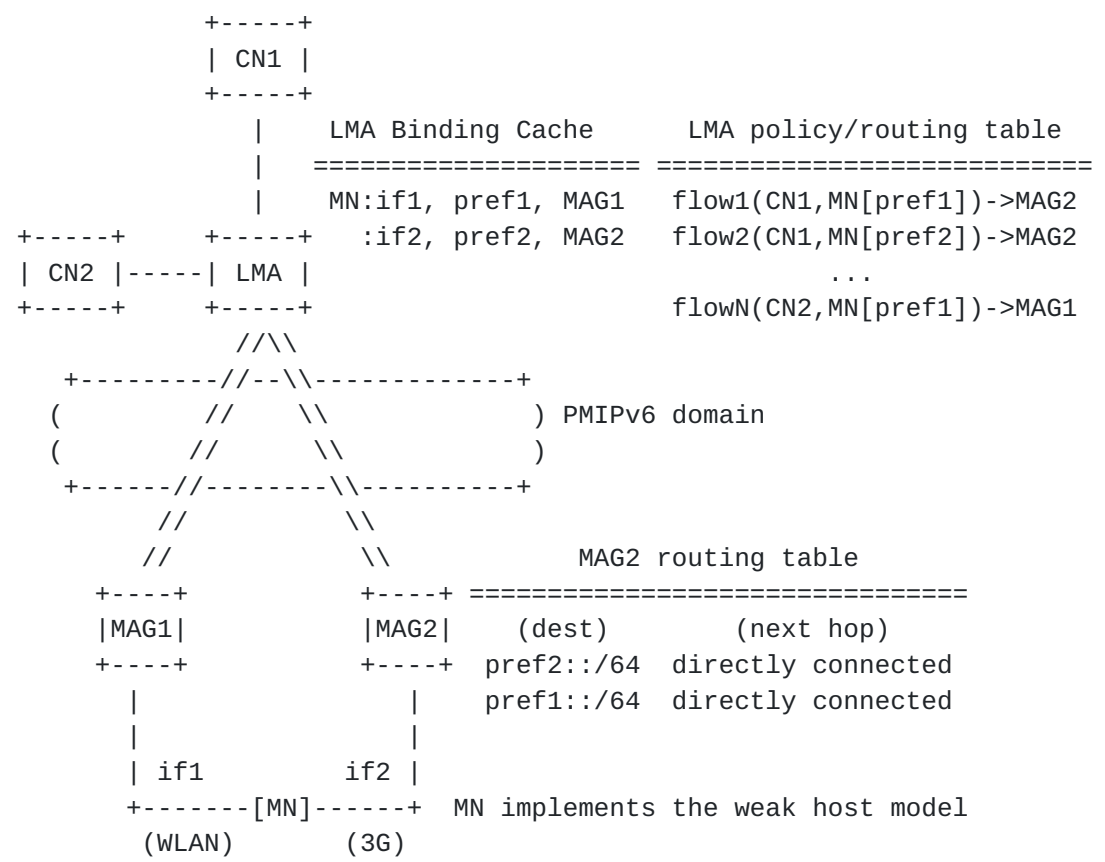

Figure 3: Solution overview

4.1. MN considerations

In order to support the reception of traffic addressed to pref1::if1 at the interface if2, the MN MUST follow the Weak host model [RFC1122], [I-D.thaler-ip-model-evolution]. This model does not limit traffic reception at a host only to IP packets whose destination address matches the IP address assigned to the interface receiving the packets, but allows to receive and process packets whose IP destination address corresponds to that of any of the local interfaces of the host.

By implementing the Weak host model, the MN in Figure 3 would be able to process traffic addressed to any of its IP addresses (i.e., pref1::if1 and pref2::if2), no matter to which interface that traffic arrives to.

We have performed some tests with different operating systems, and the results show that both Linux (tested with Linux-2.6.26) and Mac OS X (tested with Leopard) implements the Weak host model for both IPv4 and IPv6 traffic. We have not performed tests with Windows, but some results have been reported in [I-D.ietf-mif-current-practices].

It should be noted that Windows XP and Windows Server 2003 use the Weak host model for sends and receives for all IPv4 interfaces and the Strong host model for sends and receives for all IPv6 interfaces. This behavior cannot be modified. The Next Generation TCP/IP stack in Windows Vista and Windows Server 2008 supports strong host sends and receives for both IPv4 and IPv6 by default on all interfaces. The stack can be configured to use weak host model.

Generally it should be possible to enable automatic configuration of the weak model during network attachment/entry according to policies configured in the operator's network. Signaling exchanged between the MAG and the LMA (PUB, PBA) needs to be extended to configure the MN (via RS/RA or DHCP) to use the weak host model on a specific interface. As an example according to RFC 5175 [RFC5175] a bit can be assigned in the RA message indicating such option. The access provider could then decide to configure the MAGs to advertise the MN for weak model configuration. Obviously, understanding a new RA/RS bit or a DHCP option would require new functionality in the MN`s IP stack, or at minimum some kind of a networking configuration manager running in a MIF node.

4.2. LMA considerations

The LMA MUST be able to identify all the mobility bindings at its Binding Cache (BC) that refer to the same MN, using the MN-identifier. The LMA SHOULD have an additional policy/routing table. This table is used by the LMA to store and look up information about how to route packets to a certain MN. With current PMIPv6 specification, the LMA decides on the next hop towards a particular MN based only on the destination prefix (that would result on an outgoing tunnel interface to reach the MAG where that prefix is currently reachable). In order to allow the LMA to dynamically decide which is the best path for a certain traffic to reach the MN, a policy/routing table SHOULD be used. By using this table, the LMA would be able to send different flows addressed to the same destination IP address (e.g. pref1::if1) via different MAGs.

4.3. MAG considerations

The MAG MUST support routing packets addressed to MNs locally attached to the MAG, but using a destination prefix or address that is not on-link. In order to do that, the MAG SHOULD be informed by the LMA about the set of IP addresses that the MN has acquired from the PMIPv6 domain, so the MAG can add the required entries on its routing table. The PBA MAY be extended to include such information. The prefixes advertised in the Router Advertisement (RA) sent from the MAG to the MN include only those that would be advertised in case of base RFC 5213 operation without any flow/policy routing

extensions.

4.4. Downlink and Uplink considerations

The extensions outlined in this document would allow an MN to simultaneously receive traffic through all of its interfaces that are attached to the same PMIPv6 domain. Enabling such a feature in the Downlink (DL) makes sense when several access networks are available at the same time, as for example in heterogeneous PMIPv6 domains where several access technologies exhibiting different DL capacities are found (e.g., WLAN and 3G).

Enabling the feature on the Uplink (UL) is also possible. Enabling the network (i.e., the LMA) to have the control on which MN's outgoing interface it used for a certain flow requires changes on the MN side, as well as signaling on the MN-AR interface or configuring explicit routes on the MN using existing host configuration protocols at IP level (e.g. DHCP). Nevertheless, if the decision is on the MN side, this might be easily supported by the solution outlined in this document, by properly configuring the routing and ingress filtering at the MAGs.

The mapping of a flow to an interface may be driven by the terminal, the LMA or both:

1. driven by the terminal: the terminal establishes the policy and selects the interface to send packets. The LMA must be aware of the flow/interface mapping policy to keep consistency in routing (the terminal would expect receiving traffic on a specific interface). So the terminal may provide its policy to the LMA.

2. driven by the LMA: the LMA have the control on which MN's outgoing interface is used for a certain flow. In such a case the MN's routing table is updated according to the policy which must be provided to the MN by the LMA.

3. MN driven but assisted by the LMA: the terminal controls the mapping of the flows to the possible interfaces. However the LMA provides some default policies which can be updated by the MN. The policies must be exchanged in both directions (from LMA to MN and vice versa).

4.5. IPv4 considerations

IPv4 Home Addresses work mostly in a similar manner as IPv6 HNPs in the context of PMIPv6 and MIF nodes. Though, a MIF node may by default apply a different host model depending on the IP version.

One problem with IPv4 Home Addresses is the possible use of private IPv4 addresses [RFC1918]. It is possible for a MIF node to configure overlapping public IPv4 Addresses on multiple interfaces. This is not a new issue as it has been possible since the introduction of [RFC1918] and any multi-homed IPv4 node. Still, the host operation is not generally clearly defined in case of multiple overlapping addresses. The only common advice is to avoid overlapping [RFC1918] private IPv4 Home Addresses within PMIPv6 domain, unless the MIF nodes are known to be able to handle such situation gracefully. This situation resembles the scenario iii) of [I-D.devarapalli-netext-multi-interface-support] and therefore is out of scope of this document.

5. Implementation Experience

In this section we report on early implementation experience under Linux OS from a testbed running the solution proposed in this document.

5.1. Test setup

The test-bed is made up of 5 PCs connected according to the scheme of Figure 4 actually the MN's NICs in use are two WLAN cards, and the routes and policies refer to an already established multihoming scenario).

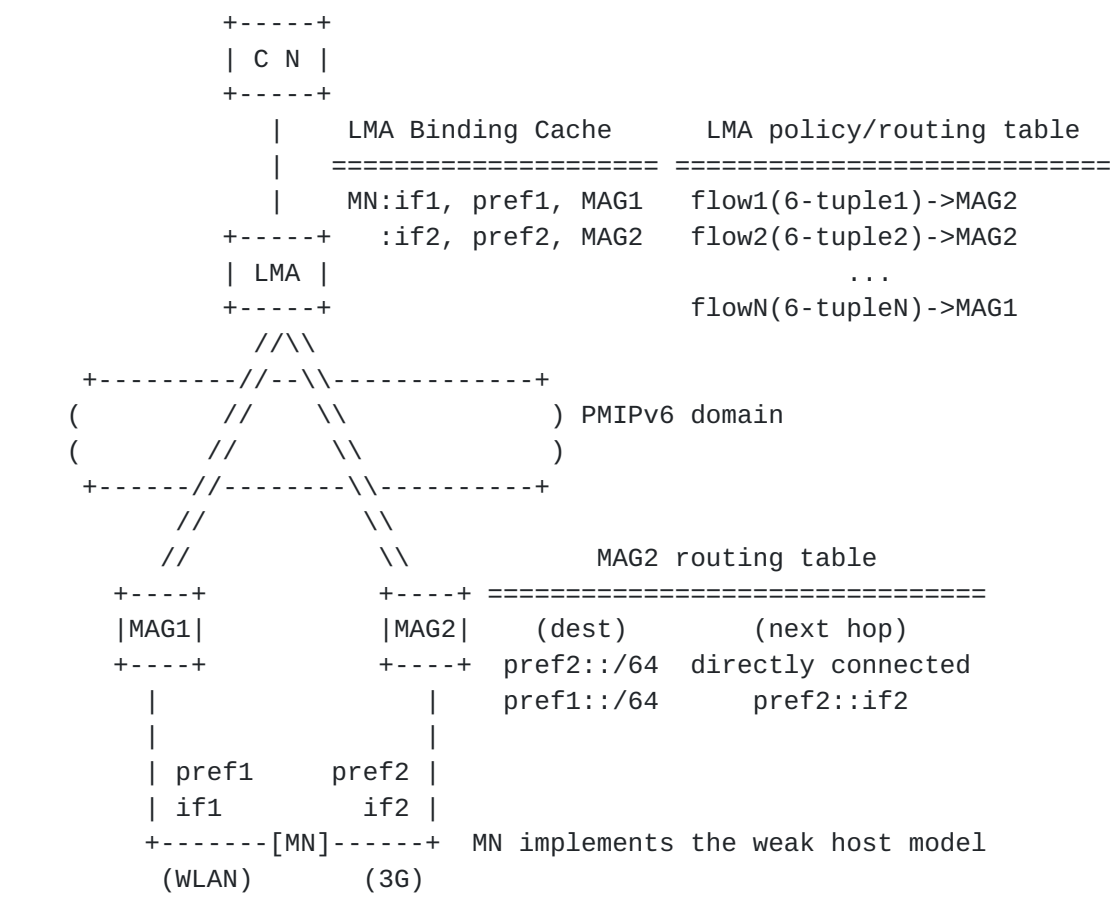

Figure 4: Test setup

5.2. Attachment phase

5.2.1. MAG considerations

Upon receiving an RS from the MN, the MAG checks whether the MN is proxy authorized and consequently runs for authentication. This procedure is replicated by means of a static configuration file that also maps the MN's set of MAC addresses into a unique MN-ID and provides a Multi-homing request indication. As described in Section 4.2, the LMA MUST be able to identify all the mobility bindings at its Binding Cache (BC) that refer to the same MN, using the MN-ID, and this is ensured by filling the PBU's MN-ID and MN-LL-ID options respectively with the MN-ID and MAC address specified in the config file.

One extra option, called MuHo option, is added in the PBU if the config file specifies a multi-homing request. The option format coincides with the Home Network Prefix Option specified in section 8.3 of RFC 5213, but for the option type number, that has to be agreed on (the MuHo option will be fully specified in a subsequent

version of this document). The MAG sends an empty option, indicating that the MN has MuHo capabilities. By means of this option, the MAG is requesting all the prefixes that might have been assigned to other interfaces of the same MN. These prefixes are then obtained by looking for the same option in the PBA message. One option is used for each prefix and multiple options may be present if the MN has already two or more interfaces attached. Once the MAG gains these prefixes it's able to set up downlink and uplink routes for all the MN's interfaces via the the one that's attempting to attach.

5.2.2. LMA considerations

Once the PBU is received by the LMA, if the MuHo option is present, is then processed to look if the registration might be related to other BCEs that belong to the same MN. The LMA stores an extra data structure which entries contain pointers to group together all the BCEs that share the same MN-ID. Every BCE will also have a pointer to its correspondent MuHo entry. In such a way, when retrieving a BCE by looking for a prefix, the LMA is able to find quickly all the prefixes assigned to the interfaces that are already connected to the domain. If the lookup succeeds, the LMA sends a PBA message with one MuHo option for each prefix, otherwise it replies with an empty option.

5.2.3. Miscellaneous considerations

Multiple attachments procedure should work as follows. Every time that the MN attaches a new interface via a new MAG the LMA updates its binding cache accordingly. The LMA should further notify all the previous MAGs about the configured HNPs. To this end the LMA can re-use the binding revocation mechanism to notify the MN that PMIP multi-homing service has been updated. This allows the LMA to propagates all the HNPs across multiple MAGs.

5.3. Flow Management

The test is intended to let the flows be driven only by the LMA, i.e. by the network side. Throughout the following considerations a flow is then intended to come from outside the PMIP domain and addressed to the MN regardless of the nature nor content of the stream itself.

Every packet that passes through the LMA has to be inspected in order to be assigned to a particular flow and then routed according to a flow-policy. Flow management is therefore divided into two steps:

1. Identification of the flows;

2. Routing of the flows;

Netfilter API and ip6tables can be used to accomplish both tasks. Netfilter provides 5 hooks in the routing scheme where a packet can be manipulated or made available for user-space applications (PREROUTING, INPUT, OUTPUT, FORWARDING and POSTROUTING in Figure 5).

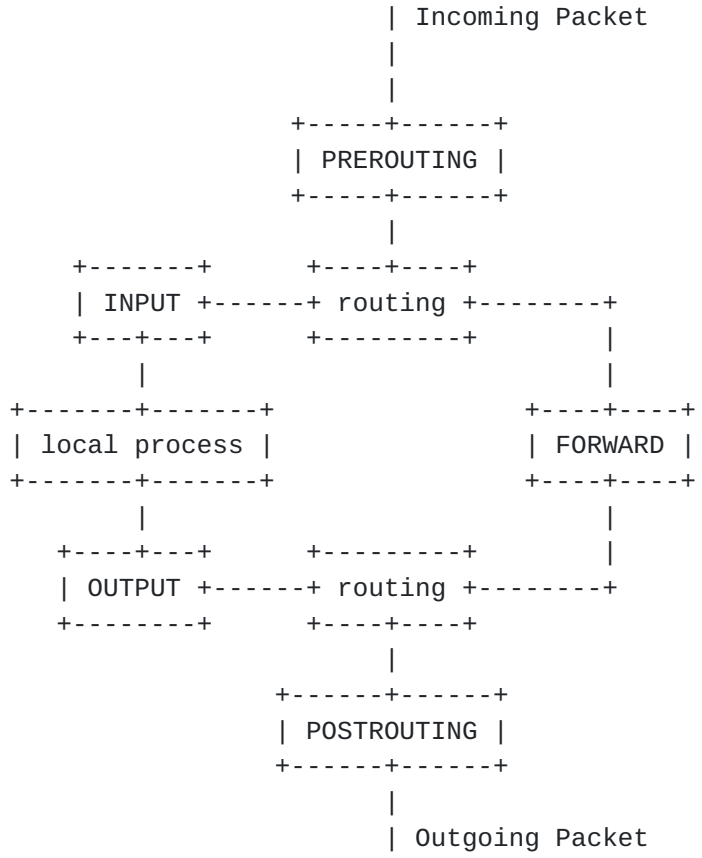

Figure 5: Flow management

5.3.1. Flow identification

Using PREROUTING hook and NFQUEUE policy, ip6tables passes packets to a user-space application that performs both tasks mentioned before and is detached from the genuine LMA implementation. Once the packet is made available to user-space, the first operation consists in extracting the following parameters from it:

- source address;
- destination address;

- source port;
- destination port;
- IPv6 flow label;
- L4 protocol type.

Each flow is singled out by this tuple of parameters and the tuple is mapped into a flow identifier that univocally identifies the stream. A data structure stores the active flows and the associated identifiers, so, as second operation, a lookup over the active flows is performed. Then any packet can alternatively be assigned into an already existing flow, or trigger a new flow generation.

Before the packets leave user-space, the last process' operation is appending a mark containing the associated flow-ID. The mark does not modify the packet's content and it is automatically removed by netfilter when the packet leaves the routing scheme in figure. Even if multiple connections are set up between the same end-points (i.e. the same couple of destination/source addresses), different flows still remain distinguishable (and therefore managed) as far as one of the above parameters changes.

5.3.2. Flow routing

Linux kernel is able to manage up to 256 different routing tables, that may contain contrasting routes. When a packet has to be routed, usually table 254 (or MAIN) is inspected, but routing rules can be added to decide which is the most suitable table for a given packet.

As an example, rules are used to perform source-based routing, since a rule can specify a certain routing table for all the packets that match a given source address. This is the "from" rule-type. Source routing is applied in the MAGs, because every packet coming from the MN must be forwarded through the tunnel.

On the LMA we use the "fwmark" rule-type, instead of the "from" rule-type in the manner explained below. This rule-type forces to inspect a given table if the packet matches the mark that is appended by the user-space netfilter-based process described before. A table is then created for each existing tunnel with just one route that forces to use that tunnel for all destinations, and this table is pointed to by a set of rules looking for different marks. Since every packet is inspected and marked, at this point it is possible to route them according to a given routing table, and therefore forwarded through a desired tunnel, by switching the table the rule points to.

5.4. Extensions on the MN

It was implicitly assumed that a CN outside (or not) the PMIP domain only knows the MN's address that was first acquired by the MN itself. This assumption, in addition to all considerations made in this document, gives consistency to the test as far as we consider the MN to have one public address/interface and a bunch of private addresses/interfaces. In such way, a flow (whether the connection was started by the MN or not) will always be first transmitted over the public interface and then eventually moved upon a LMA decision (this also provides a weak tolerance towards the full multi-homing issue mentioned in Section 5.2.3).

Anyway, it's always possible to use in downlink a desired MN's interface since the MN behaves as weak host. For the uplink, whether the connection is started by a local or a remote process, the MN will transmit through the interface that guarantees to reach the destination by means of a default or specific route. If the connection is started by the CN then the answers will carry as source the address specified in the incoming packet's destination, otherwise the packets will have as source the address assigned to the transmitting interface. In no case the MN can start a connection through an interface carrying the address of another interface.

It would thus be possible to provide a method to add, delete or change a per-host route whenever we would like to switch interface for a given connection.

The lacks of this solution resides in:

- the impossibility to manage multiple connections over different interfaces between the same end-points (i.e. the same couple of source/destination addresses);

- starting a connection with an undesired address. This problem could be overcame by using the netfilter-based application in the MN too, in such a way to either "reflect" packets through the same interface that received the flow those packets belong to (if the connection was started remotely), or force the application-layer processes to choose the source address according to table MAIN but then actually route packets inspecting another routing table (if the connection is started by the MN).

A different approach is adopted when using a virtual interface. Linux kernel provides a module called "bonding" to group together several interfaces (or "slaves", in the module's terminology) that will have the same L2 and L3 address and figure out to be as just one interface. The module offers the possibility to select the

transmitting slave according to pre-configured policies. Unfortunately these policies do not cover the scope of flow management so the module has to be extended to allow an external input to select the transmitting slave.

As second configuration we suppose the MN to use a virtual interface, so that each interface will have the same MAC address and the same prefix will be assigned. In this case the LMA will store the same BCE for all the interfaces, and conflicts may arise. Indeed, when the LMA receives a PBU from a MAG for a Proxy Registration, it may find that the BCE already exists with a different CoA, as that CoA is the address of the MAG to which another interface of the same MN is attached. This may be interpreted as an unexpected handover if the handoff indicator field is not properly set. The Access Technology Type in conjunction with a new value of the HI field in the PBU might be used to avoid conflicts in the registration. The correct behavior for the LMA would lead to a new tunnel creation in order to allow the MN to be reached via all the MAGs to which the MN's interfaces are attached. That's why the BCE format must be extended too to contain multiple CoAs and tunnel identifiers.

6. IANA Considerations

MuHo option, TBD.

7. Security Considerations

None.

8. Acknowledgements

The authors would like to thank Fabio Giust for his work on the implementation of the mechanism described on this document.

The authors would like to thank Paulo Ferrer and Marco Liebsch for their comments and discussion on this document.

The research of Carlos J. Bernardos leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n. 214994 (CARMEN project) and also from the Ministry of Science and Innovation of Spain, under the QUARTET project (TIN2009-13992-C02-01).

9. References

9.1. Normative References

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

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5175] Haberman, B. and R. Hinden, "IPv6 Router Advertisement Flags Option", RFC 5175, March 2008.

[RFC5213] Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, August 2008.

9.2. Informative References

[I-D.devarapalli-netext-multi-interface-support] Devarapalli, V., Kant, N., Lim, H., and C. Vogt, "Multiple Interface Support with Proxy Mobile IPv6", draft-devarapalli-netext-multi-interface-support-00 (work in progress), March 2009.

[I-D.ietf-mif-current-practices] Wasserman, M., "Current Practices for Multiple Interface Hosts", draft-ietf-mif-current-practices-00 (work in progress), October 2009.

[I-D.ietf-mif-problem-statement] Blanchet, M. and P. Seite, "Multiple Interfaces Problem Statement", draft-ietf-mif-problem-statement-01 (work in progress), October 2009.

[I-D.ietf-netlmm-pmip6-ipv4-support] Wakikawa, R. and S. Gundavelli, "IPv4 Support for Proxy Mobile IPv6", draft-ietf-netlmm-pmip6-ipv4-support-18 (work in progress), February 2010.

[I-D.jeyatharan-netext-multihoming-ps] Jeyatharan, M. and C. Ng, "Multihoming Problem Statement in NetLMM", draft-jeyatharan-netext-multihoming-ps-01 (work in progress), March 2009.

[I-D.thaler-ip-model-evolution] Thaler, D., "Evolution of the IP Model",

draft-thaler-ip-model-evolution-01 (work in progress), July 2008.

Authors' Addresses

Carlos J. Bernardos
Universidad Carlos III de Madrid
Av. Universidad, 30
Leganes, Madrid 28911
Spain

Phone: +34 91624 6236
Email: cjbc@it.uc3m.es
URI: http://www.it.uc3m.es/cjbc/

Telemaco Melia
Alcatel-Lucent Bell Labs

Email: Telemaco.Melia@alcatel-lucent.com

Pierrick Seite
France Telecom

Email: pierrick.seite@orange-ftgroup.com

Jouni Korhonen
Nokia Siemens Networks

Email: jouni.korhonen@nsn.com