Network Working Group
Request for Comments: 4889
Category: Informational

C. Ng
Panasonic Singapore Labs
F. Zhao
UC Davis
M. Watari
KDDI R&D Labs
P. Thubert
Cisco Systems

July 2007

Network Mobility Route Optimization Solution Space Analysis

Status of This Memo

Copyright Notice

Abstract

With current Network Mobility (NEMO) Basic Support, all communications to and from Mobile Network Nodes must go through the Mobile Router and Home Agent (MRHA) tunnel when the mobile network is away. This results in increased length of packet route and increased packet delay in most cases. To overcome these limitations, one might have to turn to Route Optimization (RO) for NEMO. This memo documents various types of Route Optimization in NEMO and explores the benefits and tradeoffs in different aspects of NEMO Route Optimization.

Table of Contents

1. Introduction

Network Mobility Route Optimization Problem Statement [1] describes operational limitations and overheads incurred in a deployment of Network Mobility (NEMO) Basic Support [2], which could be alleviated by a set of NEMO Route Optimization techniques to be defined. The term "Route Optimization" is used in a broader sense than already defined for IPv6 Host Mobility in [3] to loosely refer to any approach that optimizes the transmission of packets between a Mobile Network Node and a Correspondent Node.

Solutions that would fit that general description were continuously proposed since the early days of NEMO, even before the Working Group was formed. Based on that long-standing stream of innovation, this document classifies, at a generic level, the solution space of the possible approaches that could be taken to solve the Route Optimization-related problems for NEMO. The scope of the solutions, the benefits, and the impacts to the existing implementations and deployments are analyzed. This work should serve as a foundation for the NEMO WG to decide where to focus its Route Optimization effort, with a deeper understanding of the relative strengths and weaknesses of each approach.

It should be beneficial for readers to keep in mind the design requirements of NEMO [4]. A point to note is that since this document discusses aspects of Route Optimization, the reader may assume that a mobile network or a mobile host is away when they are mentioned throughout this document, unless it is explicitly specified that they are at home.

1.1. Terminology

It is expected that readers are familiar with terminologies related to mobility in [3] and [5], and NEMO-related terms defined in [6]. In addition, the following Route Optimization-specific terms are used in this document:

Correspondent Router (CR)

This refers to the router that is capable of terminating a Route Optimization session on behalf of a Correspondent Node.

Correspondent Entity (CE)

This refers to the entity that a Mobile Router or Mobile Network Node attempts to establish a Route Optimization session with. Depending on the Route Optimization approach, the Correspondent Entity may be a Correspondent Node or Correspondent Router.

2. Benefits of NEMO Route Optimization

NEMO Route Optimization addresses the problems discussed in [1]. Although a standardized NEMO Route Optimization solution has yet to materialize, one can expect it to show some of the following benefits:

o Shorter Delay

  Route Optimization involves the selection and utilization of a lesser-cost (thus generally shorter and faster) route to be taken for traffic between a Mobile Network Node and its Correspondent Node. Hence, Route Optimization should improve the latency of the data traffic between the two end nodes. This may in turn lead to better overall Quality of Service characteristics, such as reduced jitter and packet loss.

o Reduced Consumption of Overall Network Resources

  Through the selection of a shorter route, the total link utilization for all links used by traffic between the two end nodes should be much lower than that used if Route Optimization is not carried out. This would result in a lighter network load with reduced congestion.

o Reduced Susceptibility to Link Failure

  If a link along the bi-directional tunnel is disrupted, all traffic to and from the mobile network will be affected until IP routing recovers from the failure. An optimized route would conceivably utilize a smaller number of links between the two end nodes. Hence, the probability of a loss of connectivity due to a single point of failure at a link should be lower as compared to the longer non-optimized route.

o Greater Data Efficiency

  Depending on the actual solution for NEMO Route Optimization, the data packets exchanged between two end nodes may not require as many levels of encapsulation as that in NEMO Basic Support. This would mean less packet overheads and higher data efficiency. In particular, avoiding packet fragmentation that may be induced by the multiple levels of tunneling is critical for end-to-end efficiency from the viewpoints of buffering and transport protocols.

- Reduced Processing Delay

  In a nested mobile network, the application of Route Optimization may eliminate the need for multiple encapsulations required by NEMO Basic Support, which may result in less processing delay at the points of encapsulation and decapsulation.

- Avoiding a Bottleneck in the Home Network

  NEMO Route Optimization allows traffic to bypass the Home Agents. Apart from having a more direct route, this also avoids routing traffic via the home network, which may be a potential bottleneck otherwise.

- Avoid the Security Policy Issue

  Security policy may forbid a Mobile Router from tunneling traffic of Visiting Mobile Nodes into the home network of the Mobile Router. Route Optimization can be used to avoid this issue by forwarding traffic from Visiting Mobile Nodes directly to their destinations without going through the home network of the Mobile Router.

  However, it should be taken into consideration that a Route Optimization mechanism may not be an appropriate solution since the Mobile Router may still be held responsible for illegal traffic sent from its Mobile Network Nodes even when Route Optimization is used. In addition, there can be a variety of different policies that might conflict with the deployment of Route Optimization for Visiting Mobile Nodes. Being a policy issue, solving this with a protocol at the policy plane might be more appropriate.

- Avoid the Instability and Stalemate

  [1] described a potential stalemate situation when a Home Agent is nested within a mobile network. Route Optimization may circumvent such stalemate situations by directly forwarding traffic upstream. However, it should be noted that certain Route Optimization schemes may require signaling packets to be first routed via the Home Agent before an optimized route can be established. In such cases, a Route Optimization solution cannot avoid the stalemate.

3. Different Scenarios of NEMO Route Optimization

There are multiple proposals for providing various forms of Route Optimization in the NEMO context. In the following sub-sections, we describe the different scenarios that would require a Route Optimization mechanism and list the potential solutions that have been proposed in that area.

3.1. Non-Nested NEMO Route Optimization

The Non-Nested NEMO Route Optimization involves a Mobile Router sending binding information to a Correspondent Entity. It does not involve nesting of Mobile Routers or Visiting Mobile Nodes. The Correspondent Entity can be a Correspondent Node or a Correspondent Router. The interesting case is when the Correspondent Entity is a Correspondent Router. With the use of Correspondent Router, Route Optimization session is terminated at the Correspondent Router on behalf of the Correspondent Node. As long as the Correspondent Router is located "closer" to the Correspondent Node than the Home Agent of the Mobile Router, the route between Mobile Network Node and the Correspondent Node can be said to be optimized. For this purpose, Correspondent Routers may be deployed to provide an optimal route as illustrated in Figure 1.

```
                  ************************** HAofMR
                *                            #*#
              *                            #*#   +---------------------+
            CN                           #*#     |       LEGEND        |
              o                        #*#       +---------------------+
               o   ###############   #*#         | #: Tunnel           |
                CR ooooooooooooooo MR            | *: NEMO Basic route |
                   ###############  |            | o: Optimized route  |
                                   MNN           +---------------------+
```

Figure 1: MR-CR Optimization

This form of optimization can carry traffic in both directions or independently for the two directions of traffic:

o From MNN to CN

   The Mobile Router locates the Correspondent Router, establishes a tunnel with that Correspondent Router and sets up a route to the Correspondent Node via the Correspondent Router over the tunnel. Traffic to the Correspondent Node would no longer flow through the Home Agent anymore.

o From CN to MNN

The Correspondent Router is on the path of the traffic from the Correspondent Node to the Home Agent. In addition, it has an established tunnel with the current Care-of Address (CoA) of the Mobile Router and is aware of the Mobile Network Prefix(es) managed by the Mobile Router. The Correspondent Router can thus intercept packets going to the mobile network, and forward them to the Mobile Router over the established tunnel.

A straightforward approach to Route Optimization in NEMO is for the Mobile Router to attempt Route Optimization with a Correspondent Entity. This can be viewed as a logical extension to NEMO Basic Support, where the Mobile Router would send Binding Updates containing one or more Mobile Network Prefix options to the Correspondent Entity. The Correspondent Entity, having received the Binding Update, can then set up a bi-directional tunnel with the Mobile Router at the current Care-of Address of the Mobile Router, and inject a route to its routing table so that packets destined for addresses in the Mobile Network Prefix will be routed through the bi-directional tunnel.

The definition of Correspondent Router does not limit it to be a fixed router. Here we consider the case where the Correspondent Router is a Mobile Router. Thus, Route Optimization is initiated and performed between a Mobile Router and its peer Mobile Router. Such solutions are often posed with a requirement to leave the Mobile Network Nodes untouched, as with the NEMO Basic Support protocol, and therefore Mobile Routers handle the optimization management on behalf of the Mobile Network Nodes. Thus, providing Route Optimization for a Visiting Mobile Node is often out of scope for such a scenario because such interaction would require extensions to the Mobile IPv6 protocol. This scenario is illustrated in Figure 2.

```
HAofCR ********************************** HAofMR
  #*#                                     #*#
    #*#                                 #*#   +---------------------+
      #*#                             #*#     |       LEGEND        |
        #*#                         #*#       +---------------------+
          #*#   ###############   #*#         | #: Tunnel           |
             CR ooooooooooooooo MR            | *: NEMO Basic route |
             |  ###############  |            | o: Optimized route  |
            MNN2                MNN1          +---------------------+
```

Figure 2: MR-MR Optimization

This form of optimization can carry traffic for both directions identically:

o MNN1 to/from MNN2

The Mobile Router locates the Correspondent Router, establishes a tunnel with that Correspondent Router, and sets up a route to the Mobile Network Node via the Correspondent Router over the tunnel. Traffic to the Mobile Networks Nodes would no longer flow through the Home Agents.

Examples of this approach include Optimized Route Cache (ORC) [7][8] and Path Control Header (PCH) [9].

3.2. Nested Mobility Optimization

Optimization in Nested Mobility targets scenarios where a nesting of mobility management protocols is created (i.e., Mobile IPv6-enabled host inside a mobile network or multiple Mobile Routers that attach behind one another creating a nested mobile network). Note that because Mobile IPv6 defines its own Route Optimization mechanism in its base protocol suite as a standard, collaboration between this and NEMO protocols brings various complexities.

There are two main aspects in providing optimization for Nested Mobility, and they are discussed in the following sub-sections.

3.2.1. Decreasing the Number of Home Agents on the Path

The aim is to remove the sub-optimality of paths caused by multiple tunnels established between multiple Mobile Nodes and their Home Agents. Such a solution will seek to minimize the number of Home Agents along the path, by bypassing some of the Home Agent(s) from the original path. Unlike the scenario where no nesting is formed and only a single Home Agent exists along the path, bypassing one of the many Home Agents can still be effective.

Solutions for Nested Mobility scenarios can usually be divided into two cases based on whether the nesting involves Mobile IPv6 hosts or only involves Mobile Routers. Since Mobile IPv6 defines its own Route Optimization mechanism, providing an optimal path for such hosts will require interaction with the protocol and may require an altering of the messages exchanged during the Return Routability procedure with the Correspondent Node.

An example of this approach include Reverse Routing Header (RRH) [10].

3.2.2. Decreasing the Number of Tunnels

The aim is to reduce the amplification effect of nested tunnels due to the nesting of tunnels between the Visiting Mobile Node and its Home Agent within the tunnel between the parent Mobile Router and the parent Mobile Router's Home Agent. Such a solution will seek to minimize the number of tunnels, possibly by collapsing the amount of tunnels required through some form of signaling between Mobile Nodes, or between Mobile Nodes and their Home Agents, or by using routing headers to route packets through a discovered path. These limit the consequences of the amplification effect of nested tunnels, and at best, the performance of a nested mobile network will be the same as though there were no nesting at all.

Examples of this approach include the Reverse Routing Header (RRH) [10], Access Router Option (ARO) [11], and Nested Path Info (NPI) [12].

3.3. Infrastructure-Based Optimization

An infrastructure-based optimization is an approach where optimization is carried out fully in the infrastructure. One example is to make use of Mobility Anchor Points (MAPs) such as defined in HMIPv6 [13] to optimize routes between themselves. Another example is to make use of proxy Home Agent such as defined in the global Home Agent to Home Agent (HAHA) protocol [14]. A proxy Home Agent acts as a Home Agent for the Mobile Node, and acts as a Mobile Node for the Home Agent, Correspondent Node, Correspondent Router, and other proxies. In particular, the proxy Home Agent terminates the MRHA tunnel and the associated encryption, extracts the packets, and re-encapsulates them to the destination. In this case, proxy Home Agents are distributed in the infrastructure and each Mobile Router binds to the closest proxy. The proxy, in turn, performs a primary binding with a real Home Agent for that Mobile Router. Then, the proxy might establish secondary bindings with other Home Agents or proxies in the infrastructure, in order to improve the end-to-end path. In this case, the proxies discover each other using some form of Next Hop Resolution Protocol, establish a tunnel and exchange the relevant Mobile Network Prefix information in the form of explicit prefix routes.

Alternatively, another approach is to use prefix delegation. Here, each Mobile Router in a nested mobile network is delegated a Mobile Network Prefix from the access router using DHCP Prefix Delegation [15]. Each Mobile Router also autoconfigures its Care-of Address from this delegated prefix. In this way, the Care-of Addresses of each Mobile Router are all formed from an aggregatable address space starting from the access router. This may be used to eliminate the multiple tunnels caused by nesting of Mobile Nodes.

3.4. Intra-NEMO Optimization

A Route Optimization solution may seek to improve the communications between two Mobile Network Nodes within a nested mobile network. This would avoid traffic being injected out of the nested mobile network and route them within the nested mobile network. An example is the optimized route taken between MNN1 and MNN2 in Figure 3 below.

```
                  +--------+  +--------+  +--------+  +--------+
                  | MR2_HA |  | MR3_HA |  | MR4_HA |  | MR5_HA |
                  +------+-+  +---+----+  +---+----+  +-+------+
                          \       |           |        /
           +--------+    +------------------------------+
           | MR1_HA |----|          Internet            |-----CN
           +--------+    +--------------+---------------+
                                        |
                                   +----+----+
                                   |   MR1   |
                                   +----+----+
                                        |
                         ---+-----------+-----------+---
                            |           |           |
                        +---+---+   +---+---+   +---+---+
                        |  MR5  |   |  MR2  |   |  MR4  |
                        +---+---+   +---+---+   +---+---+
                            |           |           |
                         ---+---    +---+---+    ---+---
                           MNN2     |  MR3  |      MNN3
                                    +---+---+
                                        |
                                    ----+----
                                       MNN1
```

Figure 3: An Example of a Nested Mobile Network

One may be able to extend a well-designed NEMO Route Optimization for "Nested Mobility Optimization" (see Section 3.2) to provide for such kind of Intra-NEMO optimization, where, for example in Figure 3, MNN1 is treated as a Correspondent Node by MR5/MNN2, and MNN2 is treated as a Correspondent Node by MR3/MNN1.

Another possibility is for the "Non-Nested NEMO Route Optimization" technique (see Section 3.1) to be applied here. Using the same example of communication between MNN1 and MNN2, both MR3 and MR2 can treat MR5 as Correspondent Routers for MNN2, and MR5 treats MR3 and MR2 as Correspondent Routers for MNN1. An example of this approach is [16], which has the Mobile Routers announce their Mobile Network Prefixes to other Mobile Routers in the same nested Mobile Network.

Yet another approach is to flatten any nested Mobile Network so that all nested Mobile Network Nodes appear to be virtually on the same link. Examples of such approaches include delegating a single prefix to the nested Mobile Network, having Mobile Routers to perform Neighbor Discovery on behalf of their Mobile Network Nodes, and exposing a single prefix over the entire mobile network using a Mobile Ad-Hoc (MANET) protocol. In particular, it might prove useful to develop a new type of MANET, specialized for the NEMO problem, a MANET for NEMO (MANEMO). The MANEMO will optimize the formation of the nested NEMO and maintain inner connectivity, whether or not a connection to the infrastructure can be established.

4. Issues of NEMO Route Optimization

Although Route Optimization can bring benefits as described in Section 2, the scenarios described in Section 3 do so with some tradeoffs. This section explores some general issues that may impact a NEMO Route Optimization mechanism.

4.1. Additional Signaling Overhead

The nodes involved in performing Route Optimization would be expected to exchange additional signaling messages in order to establish Route Optimization. The required amount of signaling depends on the solution, but is likely to exceed the amount required in the home Binding Update procedure defined in NEMO Basic Support. The amount of signaling is likely to increase with the increasing number of Mobile Network Nodes and/or Correspondent Nodes, and may be amplified with nesting of mobile networks. It may scale to unacceptable heights, especially to the resource-scarce mobile node, which typically has limited power, memory, and processing capacity.

This may lead to an issue that impacts NEMO Route Optimization, known as the phenomenon of "Binding Update Storm", or more generally, "Signaling Storm". This occurs when a change in point of attachment of the mobile network is accompanied with a sudden burst of signaling messages, resulting in temporary congestion, packet delays, or even packet loss. This effect will be especially significant for wireless environment where bandwidth is relatively limited.

It is possible to moderate the effect of Signaling Storm by incorporating mechanisms such as spreading the transmissions burst of signaling messages over a longer period of time, or aggregating the signaling messages.

Even so, the amount of signaling required might be overwhelming, since large mobile networks (such as those deployed on a train or plane) may potentially have a large number of flows with a large number of Correspondent Nodes. This might suggest a need to have some adaptive behavior that depends on the amount of signaling required versus the effort needed to tunnel home.

4.2. Increased Protocol Complexity and Processing Load

It is expected that NEMO Route Optimization will be more complicated than NEMO Basic Support. Thus, complexity of nodes that are required to incorporate new functionalities to support NEMO Route Optimization would be higher than those required to provide NEMO Basic Support.

Coupled with the increased complexity, nodes that are involved in the establishment and maintenance of Route Optimization will have to bear the increased processing load. If such nodes are mobile, this may prove to be a significant cost due to the limited power and processing resources such devices usually have.

4.3. Increased Delay during Handoff

Due to the diversity of locations of different nodes that Mobile Network Node may signal with and the complexity of NEMO Route Optimization procedure that may cause several rounds of signaling messages, a NEMO Route Optimization procedure may take a longer time to finish its handoff than that in NEMO Basic Support. This may exacerbate the overall delay during handoffs and further cause performance degradation of the applications running on Mobile Network Nodes.

Another NEMO-specific delay during handoff is that in a nested mobile network, a child Mobile Network Node may need to detect or be notified of the handoff of its parent Mobile Router so that it can begin signaling its own Correspondent Entities. Apart from the compromise of mobility transparency and location privacy (see Section 4.7 and Section 4.8), this mechanism also increases the delay during handoffs.

Some of the solutions for Mobile IPv6, such as Fast Handovers for Mobile IPv6 [17], may be able to alleviate the increase in handoff delay.

4.4. Extending Nodes with New Functionalities

In order to support NEMO Route Optimization, some nodes need to be changed or upgraded. Smaller number of nodes required to be changed will allow for easier adoption of the NEMO Route Optimization solution in the Internet and create less impact on existing Internet infrastructure. The number and the types of nodes involved with new functionalities also affect how much of the route is optimized. In addition, it may also be beneficial to reuse existing protocols (such as Mobile IPv6) as much as possible.

Possible nodes that may be required to change include the following:

- Local Fixed Nodes

  It may prove to be difficult to introduce new functionalities at Local Fixed Nodes, since by definition, any IPv6 node can be a Local Fixed Node. This might mean that only those Local Fixed Nodes that are modified can enjoy the benefits of Route Optimization.

- Visiting Mobile Nodes

  Visiting Mobile Nodes in general should already implement Mobile IPv6 functionalities, and since Mobile IPv6 is a relatively new standard, there is still a considerable window to allow mobile devices to implement new functionalities.

- Mobile Routers

  It is expected that Mobile Routers will implement new functionalities in order to support Route Optimization.

- Access Routers

  Some approaches require access routers, or nodes in the access network, to implement some new functionalities. It may prove to be difficult to do so, since access routers are, in general, standard IPv6 routers.

- Home Agents

  It is relatively easier for new functionalities to be implemented in Home Agents.

- o Correspondent Nodes

  It may prove to be difficult to introduce new functionalities at Correspondent Nodes, since by definition, any IPv6 node can be a Correspondent Node. This might mean that only those Correspondent Nodes that are modified can enjoy the benefits of Route Optimization.

- o Correspondent Routers

  Correspondent Routers are new entities introduced for the purpose of Route Optimization, and therefore new functionalities can be defined as needed.

4.5. Detection of New Functionalities

One issue that is related to the need for new functionalities as described in Section 4.4 is the need to detect the existence of such functionalities. In these cases, a detection mechanism might be helpful to allow the initiator of Route Optimization to detect whether support for the new functionalities is available. Furthermore, it might be advantageous to have a graceful fall back procedure if the required functionalities are unavailable.

4.6. Scalability

Given the same number of nodes, the number of Route Optimization sessions would usually be more than the number of NEMO Basic Support tunnels. If all Route Optimization sessions of a mobile network are maintained by a single node (such as the Mobile Router), this would mean that the single node has to keep track of the states of all Route Optimization sessions. This may lead to scalability issues especially when that single node is a mobile device with limited memory and processing resources.

A similar scalability issue may be faced by a Correspondent Entity as well if it maintains many route-optimized sessions on behalf of a Correspondent Node(s) with a large number of Mobile Routers.

4.7. Mobility Transparency

One advantage of NEMO Basic Support is that the Mobile Network Nodes need not be aware of the actual location and mobility of the mobile network. With some approaches for Route Optimization, it might be necessary to reveal the point of attachment of the Mobile Router to the Mobile Network Nodes. This may mean a tradeoff between mobility transparency and Route Optimization.

4.8. Location Privacy

Without Route Optimization, the Correspondent Nodes are not aware of the actual location and mobility of the mobile network and its Mobile Network Nodes. To achieve Route Optimization, it might be necessary to reveal the point of attachment of the Mobile Router to the Correspondent Nodes. This may mean a tradeoff between location privacy [18] and Route Optimization.

In Mobile IPv6, a mobile node can decide whether or not to perform Route Optimization with a given Correspondent Node. Thus, the mobile node is in control of whether to trade location privacy for an optimized route. In NEMO Route Optimization, if the decision to perform Router Optimization is made by the Mobile Router, it will be difficult for Mobile Network Nodes to control the decision of having this tradeoff.

4.9. Security Consideration

As Mobile Router and Home Agent usually belong to the same administration domain, it is likely that there exists a security association between them, which is leveraged by NEMO Basic Support to conduct the home Binding Update in a secure way. However, NEMO Route Optimization usually involves nodes from different domains (for example, Mobile Router and Correspondent Entity); thus, the existence of such a security association is not a valid assumption in many deployment scenarios. For this reason, the security protection of NEMO Route Optimization signaling message is considered "weaker" than that in NEMO Basic Support. It is expected that some additional security mechanisms are needed to achieve the same or similar level of security as in NEMO Basic Support.

When considering security issues of NEMO Route Optimization, it might be useful to keep in mind some of the security issues considered when Mobile IPv6 Route Optimization was designed as documented in [19].

4.10. Support of Legacy Nodes

NEMO Basic Support is designed so that all legacy Mobile Network Nodes (such as those that are not aware of the mobility of the network they are in, and those that do not understand any mobility protocols) can still reach and be reached from the Internet. Some Route Optimization schemes, however, require that all Mobile Routers implement the same Route Optimization scheme in order for them to operate. Thus, a nested Mobile Router may not be able to achieve Route Optimization if it is attached to a legacy Local Fixed Router.

5. Analysis of Solution Space

As described in Section 3, there are various different approaches to achieve Route Optimization in Network Mobility Support. In this section, we attempt to analyze the vast solution space of NEMO Route Optimization by asking the following questions:

1. Which entities are involved?

2. Who initiates Route Optimization? When?

3. How is Route Optimization capabilities detected?

4. How is the address of the Mobile Network Node represented?

5. How is the Mobile Network Node's address bound to location?

6. How is signaling performed?

7. How is data transmitted?

8. What are the security considerations?

5.1. Which Entities Are Involved?

There are many combinations of entities involved in Route Optimization. When considering the role each entity plays in Route Optimization, one has to bear in mind the considerations described in Section 4.4 and Section 4.5. Below is a list of combinations to be discussed in the following sub-sections:

- Mobile Network Node and Correspondent Node

- Mobile Router and Correspondent Node

- Mobile Router and Correspondent Router

- Entities in the Infrastructure

5.1.1. Mobile Network Node and Correspondent Node

A Mobile Network Node can establish Route Optimization with its Correspondent Node, possibly the same way as a Mobile Node establishes Route Optimization with its Correspondent Node in Mobile IPv6. This would achieve the most optimal route, since the entire end-to-end path is optimized. However, there might be scalability issues since both the Mobile Network Node and the Correspondent Node may need to maintain many Route Optimization sessions. In addition, new functionalities would be required for both the Mobile Network Node and Correspondent Node. For the Mobile Network Node, it needs to be able to manage its mobility, and possibly be aware of the mobility of its upstream Mobile Router(s). For the Correspondent Node, it needs to be able to maintain the bindings sent by the Mobile Network Nodes.

5.1.2. Mobile Router and Correspondent Node

Alternatively, the Mobile Router can establish Route Optimization with a Correspondent Node on behalf of the Mobile Network Node. Since all packets to and from the Mobile Network Node must transit the Mobile Router, this effectively achieves an optimal route for the entire end-to-end path as well. Compared with Section 5.1.1, the scalability issue here may be remedied since it is possible for the Correspondent Node to maintain only one session with the Mobile Router if it communicates with many Mobile Network Nodes associated with the same Mobile Router. Furthermore, with the Mobile Router handling Route Optimization, there is no need for Mobile Network Nodes to implement new functionalities. However, new functionality is likely to be required on the Correspondent Node. An additional point of consideration is the amount of state information the Mobile Router is required to maintain. Traditionally, it has been generally avoided having state information in the routers to increase proportionally with the number of pairs of communicating peers.

5.1.3. Mobile Router and Correspondent Router

Approaches involving Mobile Routers and Correspondent Routers are described in Section 3.1. The advantage of these approaches is that no additional functionality is required for the Correspondent Node and Mobile Network Nodes. In addition, location privacy is relatively preserved, since the current location of the mobile network is only revealed to the Correspondent Router and not to the Correspondent Node (please refer to Section 5.8.3 for more discussions). Furthermore, if the Mobile Router and Correspondent Router exchange prefix information, this approach may scale well since a single Route Optimization session between the Mobile Router and Correspondent Router can achieve Route Optimization between any Mobile Network Node in the mobile network, and any Correspondent Node managed by the Correspondent Router.

The main concern with this approach is the need for a mechanism to allow the Mobile Router to detect the presence of the Correspondent Router (see Section 5.3 for details), and its security impact. Both the Mobile Router and the Correspondent Router need some means to verify the validity of each other. This is discussed in greater detail in Section 5.8.

A deployment consideration with respect to the use of Correspondent Router is the location of the Correspondent Router relative to the Correspondent Node. On one hand, deploying the Correspondent Router nearer to the Correspondent Node would result in a more optimal path. On the other hand, a Correspondent Router that is placed farther away from the Correspondent Node can perform Route Optimization on behalf of more Correspondent Nodes.

5.1.4. Entities in the Infrastructure

Approaches using entities in the infrastructure are described in Section 3.3. The advantages of this approach include, firstly, not requiring new functionalities to be implemented on the Mobile Network Nodes and Correspondent Nodes, and secondly, having most of the complexity shifted to nodes in the infrastructure. However, one main issue with this approach is how the Mobile Router can detect the presence of such entities, and why the Mobile Router should trust these entities. This may be easily addressed if such entity is a Home Agent of the Mobile Router (such as in the global Home Agent to Home Agent protocol [14]). Another concern is that the resulting path may not be a true optimized one, since it depends on the relative positions of the infrastructure entities with respect to the mobile network and the Correspondent Node.

5.2. Who Initiates Route Optimization? When?

Having determined the entities involved in the Route Optimization in the previous sub-section, the next question is which of these entities should initiate the Route Optimization session. Usually, the node that is moving (i.e., Mobile Network Node or Mobile Router) is in the best position to detect its mobility. Thus, in general, it is better for the mobile node to initiate the Route Optimization session in order to handle the topology changes in any kind of mobility pattern and achieve the optimized route promptly. However, when the mobile node is within a nested mobile network, the detection of the mobility of upstream Mobile Routers may need to be conveyed to the nested Mobile Network Node. This might incur longer signaling delay as discussed in Section 4.3.

Some solution may enable the node on the correspondent side to initiate the Route Optimization session in certain situations. For instance, when the Route Optimization state that is already established on the Correspondent Entity is about to expire but the communication is still active, depending on the policy, the Correspondent Entity may initiate a Route Optimization request with the mobile node side.

There is also the question of when Route Optimization should be initiated. Because Route Optimization would certainly incur tradeoffs of various forms, it might not be desirable for Route Optimization to be performed for any kind of traffic. This is, however, implementation specific and policy driven.

A related question is how often signaling messages should be sent to maintain the Route Optimization session. Typically, signaling messages are likely to be sent whenever there are topological changes. The discussion in Section 4.1 should be considered. In addition, a Lifetime value is often used to indicate the period of validity for the Route Optimization session. Signaling messages would have to be sent before the Lifetime value expires in order to maintain the Route Optimization session. The choice of Lifetime value needs to balance between different considerations. On one hand, a short Lifetime value would increase the amount of signaling overhead. On the other hand, a long Lifetime value may expose the Correspondent Entity to the risk of having an obsolete binding cache entry, which creates an opportunity for an attacker to exploit.

5.3. How Is Route Optimization Capability Detected?

The question here is how the initiator of Route Optimization knows whether the Correspondent Entity supports the functionality required to established a Route Optimization session. The usual method is for the initiator to attempt Route Optimization with the Correspondent Entity. Depending on the protocol specifics, the initiator may receive (i) a reply from the Correspondent Entity indicating its capability, (ii) an error message from the Correspondent Entity, or (iii) no response from the Correspondent Entity within a certain time period. This serves as an indication of whether or not the Correspondent Entity supports the required functionality to establish Route Optimization. This form of detection may incur additional delay as a penalty when the Correspondent Entity does not have Route Optimization capability, especially when the Route Optimization mechanism is using in-band signaling.

When the Correspondent Entity is not the Correspondent Node but a Correspondent Router, an immediate question is how its presence can be detected. One approach is for the initiator to send an Internet Control Message Protocol (ICMP) message containing the address of the Correspondent Node to a well-known anycast address reserved for all Correspondent Routers [7][8]. Only the Correspondent Router that is capable of terminating the Route Optimization session on behalf of the Correspondent Node will respond. Another way is to insert a Router Alert Option (RAO) into a packet sent to the Correspondent Node [9]. Any Correspondent Router en route will process the Router Alert Option and send a response to the Mobile Router.

Both approaches need to consider the possibility of multiple Correspondent Routers responding to the initiator, and both approaches will generate additional traffic or processing load to other routers. Furthermore, both approaches have yet to consider how the initiator can verify the authenticity of the Correspondent Routers that responded.

5.4. How is the Address of the Mobile Network Node Represented?

Normally, Route Optimization would mean that a binding between the address of a Mobile Network Node and the location of the mobile network is registered at the Correspondent Entity. Before exploring different ways of binding (see Section 5.5), one must first ask how the address of the Mobile Network Node is represented. Basically, there are two ways to represent the Mobile Network Node's address:

- inferred by the use of the Mobile Network Prefix, or

- explicitly specifying the address of the Mobile Network Node.

Using the Mobile Network Prefix would usually mean that the initiator is the Mobile Router, and has the benefit of binding numerous Mobile Network Nodes with one signaling. However, it also means that if location privacy is compromised, the location privacy of an entire Mobile Network Prefix would be compromised.

On the other hand, using the Mobile Network Node's address would mean that either the initiator is the Mobile Network Node itself or the Mobile Router is initiating Route Optimization on behalf of the Mobile Network Node. Initiation by the Mobile Network Node itself means that the Mobile Network Node must have new functionalities implemented, while initiation by the Mobile Router means that the Mobile Router must maintain some Route Optimization states for each Mobile Network Node.

5.5. How Is the Mobile Network Node's Address Bound to Location?

In order for route to be optimized, it is generally necessary for the Correspondent Entity to create a binding between the address and the location of the Mobile Network Node. This can be done in the following ways:

- binding the address to the location of the parent Mobile Router,

- binding the address to a sequence of upstream Mobile Routers, and

- binding the address to the location of the root Mobile Router.

These are described in the following sub-sections.

5.5.1. Binding to the Location of Parent Mobile Router

By binding the address of Mobile Network Node to the location of its parent Mobile Router, the Correspondent Entity would know how to reach the Mobile Network Node via the current location of the parent Mobile Router. This can be done by:

- Binding Update with Mobile Network Prefix

  This can be viewed as a logical extension to NEMO Basic Support, where the Mobile Router would send binding updates containing one or more Mobile Network Prefix options to the Correspondent Entity. The Correspondent Entity having received the Binding Update, can then set up a bi-directional tunnel with the Mobile Router at the current Care-of Address of the Mobile Router, and inject a route to its routing table so that packets destined for addresses in the Mobile Network Prefix would be routed through the bi-directional tunnel.

  Note that in this case, the address of the Mobile Network Node is implied by the Mobile Network Prefix (see Section 5.4).

- Sending Information of Parent Mobile Router

  This involves the Mobile Network Node sending the information of its Mobile Router to the Correspondent Entity, thus allowing the Correspondent Entity to establish a binding between the address of the Mobile Network Node to the location of the parent Mobile Router. An example of such an approach would be [11].

- Mobile Router as a Proxy

  Another approach is for the parent Mobile Router to act as a "proxy" for its Mobile Network Nodes. In this case, the Mobile Router uses the standard Mobile IPv6 Route Optimization procedure to bind the address of a Mobile Network Node to the Mobile Router's Care-of Address. For instance, when the Mobile Network Node is a Local Fixed Node without Mobile IPv6 Route Optimization functionality, the Mobile Router may initiate the Return Routability procedure with a Correspondent Node on behalf of the Local Fixed Node. An example of such an approach would be [20][21][22].

  On the other hand, if the Mobile Network Node is a Visiting Mobile Node, it might be necessary for the Visiting Mobile Node to delegate the rights of Route Optimization signaling to the Mobile

Router (see [23] for an example of such delegation). With this delegation, either the Visiting Mobile Network Node or the Mobile Router can initiate the Return Routability procedure with the Correspondent Node. For the case where the Return Routability procedure is initiated by the Visiting Mobile Node, the Mobile Router will have to transparently alter the content of the Return Routability signaling messages so that packets sent from the Correspondent Node to the Visiting Node will be routed to the Care-of Address of the Mobile Router once Route Optimization is established. The case where the Return Routability procedure is initiated by the Mobile Router is similar to the case where the Mobile Network Node is a Local Fixed Node.

For all of the approaches listed above, when the Mobile Network Node is deeply nested within a Mobile Network, the Correspondent Entity would need to gather Binding Updates from all the upstream Mobile Routers in order to build the complete route to reach the Mobile Network Node. This increases the complexity of the Correspondent Entity, as the Correspondent Entity may need to perform multiple binding cache look-ups before it can construct the complete route.

Other than increasing the complexity of the Correspondent Entity, these approaches may incur extra signaling overhead and delay for a nested Mobile Network Node. For instance, every Mobile Router on the upstream of the Mobile Network Node needs to send Binding Updates to the Correspondent Entity. If this is done by the upstream Mobile Routers independently, it may incur additional signaling overhead. Also, since each Binding Update takes a finite amount of time to reach and be processed by the Correspondent Entity, the delay from the time an optimized route is changed till the time the change is registered on the Correspondent Entity will increase proportionally with the number of Mobile Routers on the upstream of the Mobile Network Node (i.e., the level of nesting of the Mobile Network Node).

For "Binding Update with Mobile Network Prefix" and "Sending Information of Parent Mobile Router", new functionality is required at the Correspondent Entity, whereas "Mobile Router as a Proxy" keeps the functionality of the Correspondent Entity the same as a Mobile IPv6 Correspondent Node. However, this is done at an expense of the Mobile Routers, since in "Mobile Router as a Proxy", the Mobile Router must maintain state information for every Route Optimization session its Mobile Network Nodes have. Furthermore, in some cases, the Mobile Router needs to look beyond the standard IPv6 headers for ingress and egress packets, and alter the packet contents appropriately (this may impact end-to-end integrity, see 5.8.2).

One advantage shared by all the approaches listed here is that only mobility protocol is affected. In other words, no modification is required on other existing protocols (such as Neighbor Discovery). There is also no additional requirement on existing infrastructure (such as the access network).

In addition, having upstream Mobile Routers send Binding Updates independently means that the Correspondent Entity can use the same binding cache entries of upstream Mobile Routers to construct the complete route to two Mobile Network Nodes that have common upstream Mobile Routers. This may translate to lower memory consumption since the Correspondent Entity need not store one complete route per Mobile Network Node when it is having Route Optimization sessions with multiple Mobile Network Nodes from the same mobile network.

5.5.2. Binding to a Sequence of Upstream Mobile Routers

For a nested Mobile Network Node, it might be more worthwhile to bind its address to the sequence of points of attachment of upstream Mobile Routers. In this way, the Correspondent Entity can build a complete sequence of points of attachment from a single transmission of the binding information. Examples using this approach are [10] and [12].

Different from Section 5.5.1, this approach constructs the complete route to a specific Mobile Network Node at the mobile network side, thus offering the opportunity to reduce the signaling overhead. Since the complete route is conveyed to the Correspondent Entity in a single transmission, it is possible to reduce the delay from the time an optimized route is changed till the time the change is registered on the Correspondent Entity to its minimum.

One question that immediately comes to mind is how the Mobile Network Node gets hold of the sequence of locations of its upstream Mobile Routers. This is usually achieved by having such information inserted as special options in the Router Advertisement messages advertised by upstream Mobile Routers. To do so, not only must a Mobile Router advertise its current location to its Mobile Network Nodes, it must also relay information embedded in Router Advertisement messages it has received from its upstream Mobile Routers. This might imply a compromise of the mobility transparency of a mobile network (see Section 4.7). In addition, it also means that whenever an upstream Mobile Router changes its point of attachment, all downstream Mobile Network Nodes must perform Route Optimization signaling again, possibly leading to a "Signaling Storm" (see Section 4.1).

A different method of conveying locations of upstream Mobile Routers is (such as used in [10]) where upstream Mobile Routers insert their current point of attachment into a Reverse Routing Header embedded within a packet sent by the Mobile Network Node. This may raise security concerns that will be discussed later in Section 5.8.2.

In order for a Correspondent Entity to bind the address of a Mobile Network Node to a sequence of locations of upstream Mobile Routers, new functionalities need to be implemented on the Correspondent Entity. The Correspondent Entity also needs to store the complete sequence of locations of upstream Mobile Routers for every Mobile Network Node. This may demand more memory compared to Section 5.5.1 if the same Correspondent Entity has a lot of Route Optimization sessions with Mobile Network Nodes from the same nested Mobile Network. In addition, some amount of modifications or extension to existing protocols is also required, such as a new type of IPv6 routing header or a new option in the Router Advertisement message.

5.5.3. Binding to the Location of Root Mobile Router

A third approach is to bind the address of the Mobile Network Node to the location of the root Mobile Router, regardless of how deeply nested the Mobile Network Node is within a nested Mobile Network. Whenever the Correspondent Entity needs to forward a packet to the Mobile Network Node, it only needs to forward the packet to this point of attachment. The mobile network will figure out how to forward the packet to the Mobile Network Node by itself. This kind of approach can be viewed as flattening the structure of a nested Mobile Network, so that it seems to the Correspondent Entity that every node in the Mobile Network is attached to the Internet at the same network segment.

There are various approaches to achieve this:

o Prefix Delegation

  Here, each Mobile Router in a nested mobile network is delegated a Mobile Network Prefix from the access router (such as using Dynamic Host Configuration Protocol (DHCP) Prefix Delegation [15]). Each Mobile Router also autoconfigures its Care-of Address from this delegated prefix. In this way, the Care-of Addresses of Mobile Routers are all from an aggregatable address space starting from the access router. A Mobile Network Node with Mobile IPv6 functionality may also autoconfigure its Care-of Address from this delegated prefix, and use standard Mobile IPv6 mechanism's to bind its Home Address to this Care-of Address.

  Examples of this approach include [24], [25], and [26].

  This approach has the advantage of keeping the implementations of Correspondent Nodes unchanged. However, it requires the access router (or some other entity within the access network) and Mobile Router to possess prefix delegation functionality, and also maintain information on what prefix is delegated to which node. How to efficiently assign a subset of Mobile Network Prefix to child Mobile Routers could be an issue because Mobile Network Nodes may dynamically join and leave with an unpredictable pattern. In addition, a change in the point of attachment of the root Mobile Router will also require every nested Mobile Router (and possibly Visiting Mobile Nodes) to change their Care-of Addresses and delegated prefixes. These will cause a burst of Binding Updates and prefix delegation activities where every Mobile Router and every Visiting Mobile Node start sending Binding Updates to their Correspondent Entities.

o Neighbor Discovery Proxy

This approach (such as [27] and [28]) achieves Route Optimization by having the Mobile Router act as a Neighbor Discovery [29] proxy for its Mobile Network Nodes. The Mobile Router will configure a Care-of Address from the network prefix advertised by its access router, and also relay this prefix to its subnets. When a Mobile Network Node configures an address from this prefix, the Mobile Router will act as a Neighbor Discovery proxy on its behalf. In this way, the entire mobile network and its access network form a logical multilink subnet, thus eliminating any nesting.

This approach has the advantage of keeping the implementations of Correspondent Nodes unchanged. However, it requires the root Mobile Router to act as a Neighbor Discovery proxy for all the Mobile Network Nodes that are directly or indirectly attached to it. This increases the processing load of the root Mobile Router. In addition, a change in the point of attachment of the root Mobile Router will require every nested Mobile Router (and possibly Visiting Mobile Nodes) to change their Care-of Addresses. Not only will this cause a burst of Binding Updates where every Mobile Router and every Visiting Mobile Node start sending Binding Updates to their Correspondent Entities, it will also cause a burst of Duplicate Address Discovery messages to be exchanged between the mobile network and the access network. Furthermore, Route Optimization for Local Fixed Nodes is not possible without new functionalities implemented on the Local Fixed Nodes.

o Hierarchical Registrations

Hierarchical Registration involves Mobile Network Nodes (including nested Mobile Routers) registering themselves with either their parent Mobile Routers or the root Mobile Router itself. After registrations, Mobile Network Nodes would tunnel packets directly to the upstream Mobile Router they register with. At the root Mobile Router, packets tunneled from sub-Mobile Routers or Mobile Network Nodes are tunneled directly to the Correspondent Entities, thus avoiding nested tunneling.

One form of such an approach uses the principle of Hierarchical Mobile IPv6 [13], where the root Mobile Router acts as a Mobility Anchor Point. It is also possible for each parent Mobile Router to act as Mobility Anchor Points for its child Mobile Routers, thus forming a hierarchy of Mobility Anchor Points. One can also view these Mobility Anchor Points as local Home Agents, thus forming a cascade of mobile Home Agents. In this way, each Mobile Router terminates its tunnel at its parent Mobile Router. Hence, although there are equal numbers of tunnels as the level of nestings, there is no tunnel encapsulated within another.

Examples of this approach include [30], [31], [32], and [33].

An advantage of this approach is that the functionalities of the Correspondent Nodes are unchanged.

- Mobile Ad-Hoc Routing

  It is possible for nodes within a mobile network to use Mobile Ad-hoc routing for packet-forwarding between nodes in the same mobile network. An approach of doing so might involve a router acting as a gateway for connecting nodes in the mobile network to the global Internet. All nodes in the mobile network would configure their Care-of Addresses from one or more prefixes advertised by that gateway, while their parent Mobile Routers use Mobile Ad-hoc routing to forward packets to that gateway or other destinations inside the mobile network.

One advantage that is common to all the approaches listed above is that local mobility of a Mobile Network Node within a nested mobile network is hidden from the Correspondent Entity.

5.6. How Is Signaling Performed?

In general, Route Optimization signaling can be done either in-plane, off-plane, or both. In-plane signaling involves embedding signaling information into headers of data packets. A good example of in-plane signaling would be Reverse Routing Header [10]. Off-plane signaling uses dedicated signaling packets rather than embedding signaling information into headers of data packets. Proposals involving the sending of Binding Updates fall into this category.

The advantage of in-plane signaling is that any change in the mobile network topology can be rapidly propagated to the Correspondent Entity as long as there is a continuous stream of data to be transmitted. However, this might incur a substantial overhead on the data packets. Off-plane signaling, on the other hand, sends signaling messages independently from the data packet. This has the advantage of reducing the signaling overhead in situations where there are relatively fewer topological changes to the mobile network. However, data packet transmission may be disrupted while off-plane signaling takes place.

An entirely different method of signaling makes use of upper-layer protocols to establish the bindings between the address of a Mobile Network Node and the location of the mobile network. Such binding information can then be passed down to the IP layer to insert the appropriate entry in the Binding Cache or routing table. An example of such a mechanism is [34], which uses the Session Initiation Protocol (SIP) to relay binding information.

5.7. How Is Data Transmitted?

With Route Optimization established, one remaining question to be answered is how data packets can be routed to follow the optimized route. There are the following possible approaches:

- Encapsulations

  One way to route packets through the optimized path is to use IP-in-IP encapsulations [35]. In this way, the original packet can be tunneled to the location bound to the address of the Mobile Network Node using the normal routing infrastructure. Depending on how the location is bound to the address of the Mobile Network Node, the number of encapsulations required might vary.

  For instance, if the Correspondent Entity knows the full sequence of points of attachment, it might be necessary for there to be multiple encapsulations in order to forward the data packet through each point of attachment. This may lead to the need for multiple tunnels and extra packet header overhead. It is possible to alleviate this by using Robust Header Compression techniques [36][37][38] to compress the multiple tunnel packet headers.

- Routing Headers

  A second way to route packets through the optimized path is to use routing headers. This is useful especially for the case where the Correspondent Entity knows the sequence of locations of upstream Mobile Routers (see Section 5.5.2), since a routing header can contain multiple intermediate destinations. Each intermediate destination corresponds to a point of attachment bound to the address of the Mobile Network Node.

This requires the use of a new Routing Header type, or possibly an extension of the Type 2 Routing Header as defined by Mobile IPv6 to contain multiple addresses instead of only one.

o Routing Entries in Parent Mobile Routers

Yet another way is for parent Mobile Routers to install routing entries in their routing table that will route Route Optimized packets differently, most likely based on source address routing. This usually applies to approaches described in Section 5.5.3. For instance, the Prefix Delegation approach [24][25][26] would require parent Mobile Routers to route packets differently if the source address belongs to the prefix delegated from the access network.

5.8. What Are the Security Considerations?

5.8.1. Security Considerations of Address Binding

The most important security consideration in Route Optimization is certainly the security risks a Correspondent Entity is exposed to by creating a binding between the address of a Mobile Network Node and the specified location(s) of the mobile network. Generally, it is assumed that the Correspondent Entity and Mobile Network Node do not share any pre-existing security association. However, the Correspondent Entity must have some ways of verifying the authenticity of the binding specified, else it will be susceptible to various attacks described in [19], such as snooping (sending packets meant for a Mobile Network Node to an attacker) or denial-of-service (DoS) (flooding a victim with packets meant for a Mobile Network Node) attacks.

When the binding is performed between the address of the Mobile Network Node and one Care-of Address (possibly of the Mobile Router; see Section 5.5.1 and Section 5.5.3), the standard Return Routability procedure specified in Mobile IPv6 might be sufficient to provide a reasonable degree of assurance to the Correspondent Entity. This also allows the Correspondent Entity to re-use existing implementations. But in other situations, an extension to the Return Routability procedure might be necessary.

For instance, consider the case where the Mobile Router sends a Binding Update containing Mobile Network Prefix information to the Correspondent Entity (see Section 5.5.1). Although the Return

Routability procedure allows the Correspondent Entity to verify that the Care-of and Home Addresses of the Mobile Router are indeed collocated, it does not allow the Correspondent Entity to verify the validity of the Mobile Network Prefix. If the Correspondent Entity accepts the binding without verification, it will be exposed to attacks where the attacker tricks the Correspondent Entity into forwarding packets destined for a mobile network to the attacker (snooping) or victim (DoS); [39] discusses this security threat further.

The need to verify the validity of network prefixes is not constrained to Correspondent Entities. In approaches that involve the Correspondent Routers (see Section 5.1.3), there have been suggestions for the Correspondent Router to advertise the network prefix(es) of Correspondent Nodes that the Correspondent Router is capable of terminating Route Optimization on behalf of to Mobile Network Nodes. In such cases, the Mobile Network Nodes also need a mechanism to check the authenticity of such claims. Even if the Correspondent Routers do not advertise the network prefix, the Mobile Network Nodes also have the need to verify that the Correspondent Router is indeed a valid Correspondent Router for a given Correspondent Node.

In Section 5.5.2, the registration signaling involves sending the information about one or more upstream Mobile Routers. The Correspondent Entity (or Home Agent) must also have the means to verify such information. Again, the standard Return Routability procedure as defined in [3] is inadequate here, as it is not designed to verify the reachability of an address over a series of upstream routers. An extension such as attaching a routing header to the Care-of Test (CoT) message to verify the authenticity of the locations of upstream Mobile Routers is likely to be needed. The risk, however, is not confined to Correspondent Entities. The Mobile Network Nodes are also under the threat of receiving false information from their upstream Mobile Routers, which they might pass to Correspondent Entities (this also implies that Correspondent Entities cannot rely on any security associations they have with the Mobile Network Nodes to establish the validity of address bindings). There are some considerations that this kind of on-path threat exists in the current Internet anyway especially when no (or weak) end-to-end protection is used.

All these concerns over the authenticity of addresses might suggest that perhaps a more radical and robust approach is necessary. This is currently under extensive study in various Working Groups of the IETF, and many related documents might be of interest here. For instance, in Secure Neighbor Discovery (SEND) [40], Cryptographically Generated Addresses (CGAs) [41] could be used to establish the ownership of Care-of Addresses. [42] employs the Home Agent to check the signaling messages sent by Mobile Routers to provide a way for Correspondent Entities to verify the authenticity of Mobile Network Prefixes specified. [18] documents various proposed enhancements to the Mobile IPv6 Route Optimization mechanism that might be applied to NEMO Route Optimization as well, such as [43], which allows the Correspondent Entity to authenticate a certain operator's Home Agent by verifying the associated certificate. The Host Identity Protocol (HIP) [44] with end-host mobility considerations [45] may be extended for NEMO Route Optimization as well.

In addition, interested readers might want to refer to [46], which discussed the general problem of making Route Optimization in NEMO secure and explored some possible solution schemes. There is also a proposed mechanism in [23] for Mobile Network Node to delegate some rights to their Mobile Routers, which may be used to allow the Mobile Routers to prove their authenticities to Correspondent Entities when establishing Route Optimization sessions on behalf of the Mobile Network Nodes.

5.8.2. End-to-End Integrity

In some of the approaches, such as "Mobile Router as a Proxy" in Section 5.5.1, the Mobile Router sends messages using the Mobile Network Node's address as the source address. This is done mainly to achieve zero new functionalities required at the Correspondent Entities and the Mobile Network Nodes. However, adopting such a strategy may interfere with existing or future protocols, most particularly security-related protocols. This is especially true when the Mobile Router needs to make changes to packets sent by Mobile Network Nodes. In a sense, these approaches break the end-to-end integrity of packets. A related concern is that this kind of approach may also require the Mobile Router to inspect the packet contents sent to/by Mobile Network Nodes. This may prove to be difficult or impossible if such contents are encrypted.

The concern over end-to-end integrity arises for the use of a Reverse Routing Header (see Section 5.5.2) too, since Mobile Routers would insert new contents to the header of packets sent by downstream Mobile Network Nodes. This makes it difficult for Mobile Network Nodes to protect the end-to-end integrity of such information with security associations.

5.8.3. Location Privacy

Another security-related concern is the issue of location privacy. This document currently does not consider the location privacy threats caused by an on-path eavesdropper. For more information on that aspect, please refer to [18]. Instead, we consider the following three aspects to location privacy:

- Revelation of Location to Correspondent Entity

  Route optimization is achieved by creating a binding between the address of the Mobile Network Node and the current location of the Mobile Network. It is thus inevitable that the location of the Mobile Network Node be revealed to the Correspondent Entity. The concern may be alleviated if the Correspondent Entity is not the Correspondent Node, since this implies that the actual traffic end point (i.e., the Correspondent Node) would remain ignorant of the current location of the Mobile Network Node.

- Degree of Revelation

  With network mobility, the degree of location exposure varies, especially when one considers nested mobile networks. For instance, for approaches that bind the address of the Mobile Network Node to the location of the root Mobile Router (see Section 5.5.3), only the topmost point of attachment of the mobile network is revealed to the Correspondent Entity. For approaches such as those described in Section 5.5.1 and Section 5.5.2, more information (such as Mobile Network Prefixes and current locations of upstream Mobile Routers) is revealed. Techniques such as exposing only locally-scoped addresses of intermediate upstream mobile routers to Correspondent Entities may be used to reduce the degree of revelation.

- Control of the Revelation

  When Route Optimization is initiated by the Mobile Network Node itself, it is in control of whether or not to sacrifice location privacy for an optimized route. However, if it is the Mobile Router that initiates Route Optimization (e.g., "Binding Update with Mobile Network Prefix" and "Mobile Router as a Proxy" in Section 5.5.1), then control is taken away from the Mobile Network Node. An additional signaling mechanism between the Mobile Network Node and its Mobile Router can be used in this case to prevent the Mobile Router from attempting Route Optimization for a given traffic stream.

6. Conclusion

The problem space of Route Optimization in the NEMO context is multifold and can be split into several work areas. It will be critical, though, that the solution to a given piece of the puzzle be compatible and integrated smoothly with others. With this in mind, this document attempts to present a detailed and in-depth analysis of the NEMO Route Optimization solution space by first describing the benefits a Route Optimization solution is expected to bring, then illustrating the different scenarios in which a Route Optimization solution applies, and next presenting some issues a Route Optimization solution might face. We have also asked ourselves some of the basic questions about a Route Optimization solution. By investigating different possible answers to these questions, we have explored different aspects to a Route Optimization solution. The intent of this work is to enhance our common understanding of the Route Optimization problem and solution space.

7. Security Considerations

This is an informational document that analyzes the solution space of NEMO Route Optimization. Security considerations of different approaches are described in the relevant sections throughout this document. Particularly, please refer to Section 4.9 for a brief discussion of the security concern with respect to Route Optimization in general, and Section 5.8 for a more detailed analysis of the various Route Optimization approaches.

8. Acknowledgments

The authors wish to thank the co-authors of previous versions from which this document is derived: Marco Molteni, Paik Eun-Kyoung, Hiroyuki Ohnishi, Felix Wu, and Souhwan Jung. In addition, sincere appreciation is also extended to Jari Arkko, Carlos Jesus Bernardos, Greg Daley, Thierry Ernst, T.J. Kniveton, Erik Nordmark, Alexandru Petrescu, Hesham Soliman, Ryuji Wakikawa, and Patrick Wetterwald for their various contributions.

9. References

9.1. Normative References

[1] Ng, C., Thubert, P., Watari, M., and F. Zhao, "Network Mobility Route Optimization Problem Statement", RFC 4888, July 2007.

[2] Devarapalli, V., Wakikawa, R., Petrescu, A., and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol", RFC 3963, January 2005.

[3] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

[4] Ernst, T., "Network Mobility Support Goals and Requirements", RFC 4886, July 2007.

[5] Manner, J. and M. Kojo, "Mobility Related Terminology", RFC 3753, June 2004.

[6] Ernst, T. and H-Y. Lach, "Network Mobility Support Terminology", RFC 4885, July 2007.

9.2. Informative References

[7] Wakikawa, R., Koshiba, S., Uehara, K., and J. Murai, "ORC: Optimized Route Cache Management Protocol for Network Mobility", 10th International Conference on Telecommunications, vol 2, pp 1194-1200, February 2003.

[8] Wakikawa, R. and M. Watari, "Optimized Route Cache Protocol (ORC)", Work in Progress, November 2004.

[9] Na, J., Cho, S., Kim, C., Lee, S., Kang, H., and C. Koo, "Route Optimization Scheme based on Path Control Header", Work in Progress, April 2004.

[10] Thubert, P. and M. Molteni, "IPv6 Reverse Routing Header and its application to Mobile Networks", Work in Progress, February 2007.

[11] Ng, C. and T. Tanaka, "Securing Nested Tunnels Optimization with Access Router Option", Work in Progress, July 2004.

[12] Na, J., Cho, S., Kim, C., Lee, S., Kang, H., and C. Koo, "Secure Nested Tunnels Optimization using Nested Path Information", Work in Progress, September 2003.

[13] Soliman, H., Castelluccia, C., El Malki, K., and L. Bellier, "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)", RFC 4140, August 2005.

[14] Thubert, P., Wakikawa, R., and V. Devarapalli, "Global HA to HA protocol", Work in Progress, September 2006.

[15] Troan, O. and R. Droms, "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC 3633, December 2003.

[16] Baek, S., Yoo, J., Kwon, T., Paik, E., and M. Nam, "Routing Optimization in the same nested mobile network", Work in Progress, October 2005.

[17] Koodli, R., "Fast Handovers for Mobile IPv6", RFC 4068, July 2005.

[18] Vogt, C. and J. Arkko, "A Taxonomy and Analysis of Enhancements to Mobile IPv6 Route Optimization", RFC 4651, February 2007.

[19] Nikander, P., Arkko, J., Aura, T., Montenegro, G., and E. Nordmark, "Mobile IP Version 6 Route Optimization Security Design Background", RFC 4225, December 2005.

[20] Bernardos, C., Bagnulo, M., and M. Calderon, "MIRON: MIPv6 Route Optimization for NEMO", 4th Workshop on Applications and Services in Wireless Network, Online: http://www.it.uc3m.es/cjbc/papers/miron_aswn2004.pdf, August 2004.

[21] Calderon, M., Bernardos, C., Bagnulo, M., Soto, I., and A. Oliva, "Design and Experimental Evaluation of a Route Optimisation Solution for NEMO", IEEE Journal on Selected Areas in Communications (J-SAC), vol 24, no 9, September 2006.

[22] Bernardos, C., Bagnulo, M., Calderon, M., and I. Soto, "Mobile IPv6 Route Optimisation for Network Mobility (MIRON)", Work in Progress, July 2005.

[23] Ylitalo, J., "Securing Route Optimization in NEMO", Workshop of 12th Network and Distributed System Security Syposuim, NDSS Workshop 2005, online: http://www.isoc.org/isoc/conferences/ndss/05/workshop/ylitalo.pdf, February 2005.

[24] Perera, E., Lee, K., Kim, H., and J. Park, "Extended Network Mobility Support", Work in Progress, July 2003.

[25] Lee, K., Park, J., and H. Kim, "Route Optimization for Mobile Nodes in Mobile Network based on Prefix Delegation", 58th IEEE Vehicular Technology Conference, vol 3, pp 2035-2038, October 2003.

[26] Lee, K., Jeong, J., Park, J., and H. Kim, "Route Optimization for Mobile Nodes in Mobile Network based on Prefix Delegation", Work in Progress, February 2004.

[27] Jeong, J., Lee, K., Park, J., and H. Kim, "Route Optimization based on ND-Proxy for Mobile Nodes in IPv6 Mobile Network", 59th IEEE Vehicular Technology Conference, vol 5, pp 2461-2465, May 2004.

[28] Jeong, J., Lee, K., Kim, H., and J. Park, "ND-Proxy based Route Optimization for Mobile Nodes in Mobile Network", Work in Progress, February 2004.

[29] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[30] Kang, H., Kim, K., Han, S., Lee, K., and J. Park, "Route Optimization for Mobile Network by Using Bi-directional Between Home Agent and Top Level Mobile Router", Work in Progress, June 2003.

[31] Lee, D., Lim, K., and M. Kim, "Hierarchical FRoute Optimization for Nested Mobile Network", 18th Int'l Conf on Advance Information Networking and Applications, vol 1, pp 225-229, 2004.

[32] Takagi, Y., Ohnishi, H., Sakitani, K., Baba, K., and S. Shimojo, "Route Optimization Methods for Network Mobility with Mobile IPv6", IEICE Trans. on Comms, vol E87-B, no 3, pp 480-489, March 2004.

[33] Ohnishi, H., Sakitani, K., and Y. Takagi, "HMIP based Route optimization method in a mobile network", Work in Progress, October 2003.

[34] Lee, C., Zheng, J., and C. HUang, "SIP-based Network Mobility (SIP-NEMO) Route Optimization (RO)", Work in Progress, October 2006.

[35] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[36] Bormann, C., Burmeister, C., Degermark, M., Fukushima, H., Hannu, H., Jonsson, L-E., Hakenberg, R., Koren, T., Le, K., Liu, Z., Martensson, A., Miyazaki, A., Svanbro, K., Wiebke, T., Yoshimura, T., and H. Zheng, "RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed", RFC 3095, July 2001.

[37] Jonsson, L-E., "RObust Header Compression (ROHC): Terminology and Channel Mapping Examples", RFC 3759, April 2004.

[38] Minaburo, A., Paik, E., Toutain, L., and J. Bonnin, "ROHC (Robust Header Compression) in NEMO network", Work in Progress, July 2005.

[39] Ng, C. and J. Hirano, "Extending Return Routability Procedure for Network Prefix (RRNP)", Work in Progress, October 2004.

[40] Arkko, J., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.

[41] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, March 2005.

[42] Zhao, F., Wu, F., and S. Jung, "Extensions to Return Routability Test in MIP6", Work in Progress, February 2005.

[43] Bao, F., Deng, R., Qiu, Y., and J. Zhou, "Certificate-based Binding Update Protocol (CBU)", Work in Progress, March 2005.

[44] Moskowitz, R., Nikander, P., Jokela, P., and T. Henderson, "Host Identity Protocol", Work in Progress, April 2007.

[45] Henderson, T., "End-Host Mobility and Multihoming with the Host Identity Protocol", Work in Progress, March 2007.

[46] Calderon, M., Bernardos, C., Bagnulo, M., and I. Soto, "Securing Route Optimization in NEMO", Third International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, WIOPT 2005, pages 248-254, April 2005.

Authors' Addresses

Chan-Wah Ng
Panasonic Singapore Laboratories Pte Ltd
Blk 1022 Tai Seng Ave #06-3530
Tai Seng Industrial Estate, Singapore 534415
SG

Phone: +65 65505420
EMail: chanwah.ng@sg.panasonic.com

Fan Zhao
University of California Davis
One Shields Avenue
Davis, CA 95616
US

Phone: +1 530 752 3128
EMail: fanzhao@ucdavis.edu

Masafumi Watari
KDDI R&D Laboratories Inc.
2-1-15 Ohara
Fujimino, Saitama 356-8502
JAPAN

EMail: watari@kddilabs.jp

Pascal Thubert
Cisco Systems
Village d'Entreprises Green Side
400, Avenue de Roumanille
Batiment T3, Biot - Sophia Antipolis 06410
FRANCE

EMail: pthubert@cisco.com

Full Copyright Statement

Intellectual Property

Acknowledgement

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