FPKIN: Firewall Public Key Infrastructure for NEMO

Samer Sami Hasan ¹, Zaid Hashim Jaber ²

¹² Department of Computer Science, College of Science, University of Baghdad, 10071, Al-Jadriya, Baghdad, Iraq
Corresponding Author Email: ssami@scbaghdad.com

Abstract

Network mobility (NEMO) is an important requirement for internet networks to reach the goal of ubiquitous connectivity. With NEMO basic support protocols, correspondent entities suffer from a number of limitations and problems that prevent route-optimization procedures to be established between the correspondent nodes and mobile network nodes associated with NEMO. The goal is to alleviate the signaling load and execute the route-optimization steps on behalf of the correspondent entities that are not sophisticated enough to support route optimization. This paper introduces a new architecture that uses firewall as a new entity with new mobility filtering rules and acts as root certificate server supporting PKI infrastructure. The PKI-firewall executes the route-optimization procedure on behalf of these correspondent entities depends on CA distributed to its mobile end nodes. User entities is reachable via optimized path approved by mobile node or user CA As a result of completing the above procedure, performance degradation will be reduced, especially when signaling storm occurs; applying these modifications will increase the security, availability and scalability of NEMO optimization and enable wider NEMO deployment. An analytical model is used to validate the new proposed framework and understand the behavior of this framework under different network scenarios.

Keywords: Network mobility, Route optimization, Public key Infrastructure, firewall, Network performance.

1. Introduction

Our mobile lifestyle is currently reflected in the importance of mobile communications. However, in some situations, devices (or hosts, as we will refer to them) move as a group, for example, when travelers commute in the same train or coach for the same distance. Such cases are not efficiently covered by considering the mobility of individual devices because this involves increased signaling overhead, power consumption and security risks. A more efficient solution is required for the aggregate mobility (or network mobility) of devices using at least one mobile router with secure environments. Nowadays, the NEMO protocol relies on the use of manual symmetric keys for the authentication in its control messages. This procedure is not fitted well to support large number of users. More else, to improve the scalability, the new filtering rules created by firewall with public key infrastructure (FPKIN) is used for authentication among there mobility entities; however, the proposals have a requirement on a mobile entities to perform certificate based (CA) in PKI cryptography operations. Moreover, other registration protocols were proposed, which employ only the minimal use of the CA public keys between a correspondents FPKIN and its entities to avoid the drawback [1]. Besides these the Internet Engineers Task Force (IETF) developed a protocol named Mobile IPv4 (MIP) [2], and for IPv6 communication environments, MIPv6 [3] was developed to support fast and smooth connectivity to the mobile node. Currently, Internet users may own more than one mobile device, and these devices feature multiple interfaces that can be connected to each other as well as to other networks. This includes the set of Internet-connected devices found in vehicles. IETF extends MIPv6 to the design of NEMO BSP [4] to handle node mobility in an aggregate way using a dedicated router, as shown in Fig. 1. In NEMO BSP, there are four main entities, which are defined as follows: Correspondents Node (CN), Mobile Router (MR), Home Agent (HA), and Mobile Network Node (MNN). CN is any IPv6 node that communicates with the MNN. MR is a router that handles all movement transparently for all MNN underneath. HA is a router usually located in the home network of MNN that acts on behalf of the mobile node while away from the home link. The MNN is described as a mobile node that has the ability to move through different networks with seamless connectivity. When MNN leaves its home link and enters a new subnet, it notifies its home agent on its home link. After updating the HA with the new address acquired from the foreign link, which is based on the foreign prefix and called the Care-of Address (CoA), the MNN can then be reached through its HA. In this case, network overheads and handoff latency will be increased due to an insufficient route (i.e., Pinball Routing problem) [5]. The IETF developed an optimization procedure to address this problem. A direct connection is established between the MNN and the CN. To alleviate the performance penalty, Mobile IPv6 includes a mode of operation that allows the mobile node and its peer, a correspondent node (CN), to exchange packets directly, bypassing the home agent completely after the initial setup phase. This mode of operation is called route optimization (RO). When route optimization is used, the mobile node sends its current care-of
address to the correspondent node, using binding update (BU) messages. The correspondent node stores the binding between the home address and care-of address into its Binding Cache [6].

Route Optimization typically requires the Mobile Node and Correspondent Node to have certain capabilities, such as the possibility to execute a Return Routability procedure - MN transmitting Home Test Init (HoTI), Care-of Test Init (CoTI) and direct Binding Update messages to CN, with the CN responding with respective Home Test (HoT), Care-of Test (CoT), and Binding Acknowledgement messages to the MN. If the correspondent node is a basic IP node without support for Route Optimization, the MN with support for Route Optimization cannot set up Route Optimization with this CN because RFC 3775 [3] specifies "If a mobile node attempts to set up route optimization with a node with only basic IPv6 support, an ICMP error will signal that the node does not support such optimizations and communications will flow through the home agent".

The nodes involved in performing Route Optimization would be expected to exchange additional signaling messages 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 for signaling is likely to increase with the increasing number of Mobile Network Nodes and/or Correspondent Nodes and may be amplified with the 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 [7]. This may lead to an issue that impacts NEMO Route Optimization known as the phenomenon of "Binding Update Storm", or more generally, "Signaling Storm" due to highly registration signals and with lack of flooding attack. This occurs when a change in the point of attachment of the mobile network is accompanied with a sudden burst in signaling messages, resulting in temporary congestion, packet delays, or even packet loss. This effect will be especially significant for wireless environments where bandwidth is relatively limited. It is possible to moderate the effect of Signaling Storm using the proposed architecture PKI-FPKIN works on behalf of mobile entities to perform authenticated registration as described in next sections.

2. The Proposed FPKIN

In the proposed Firewall Route Optimization NEMO (FPKIN) architecture shown in Fig. 2, the CNs are protected by the modified IPv6 firewall, as shown in one of the firewall scenarios described by Krishnan [8]. The MIPv6 Firewall will accomplish the route optimization on behalf of the corresponding nodes (i.e., standard IPv6 or mobile IPv6 node). If the corresponding entity does not support route optimization, the firewall will start to create optimization functionality on behalf of these nodes. The return routability procedure started [3] between the mobile nodes and the firewall where corresponding entities exist belongs to that firewall. This return routability procedure provides a level of security by reusing the MIPv6 security concepts and uses a cryptographic key to generate a crypto address [9] in line with the firewall and NEMO environments. Currently, there are various types of firewalls [10, 11]. Independent of the adopted methods, firewalls generally look at five parameters of the arriving messages (source IP address, destination IP address, protocol type, source port number and destination port number). Based on these five parameters, packets are dropped or passed through a firewall [12]. The MIPv6 firewall should take into consideration the newly developed stateful filtering rules that allow the packets moving to or from the HA, MR and MNN to pass to or from the CN network without filtering the data messages and the signaling to pass through in accordance with the problems noted in a previous study [13]. This firewall will be supported by the RRP for NEMO to provide NEMO-RO without any modification of the CN’s entities. In addition, route optimization using this scenario can be easily maintained, even though the CN is not sophisticated enough to support the RO without adding new entities to the infrastructure [14] such as correspondent router (CR) in [15-17]. Deploying the correspondent router in NEMO incur additional problems to NEMO BSP optimization [18] such as, discovery of the correspondent routers consumes total handover time, and lack in security considerations between MR-CR and CN-CR. Most internet infrastructures today operate under tiered client/server systems, so most CNs may act as servers under heavy-traffic conditions. This construct will cause major difficulties in updating or modifying various CNs. In a firewall network, where a CN is easy to “plug and play”, applying modifications to a firewall is more reliable and scalable than trying to modify the CN, especially when a firewall protects several correspondent entities.

![Fig.1: NEMO Architecture showing suboptimal path in nested NEMO](image)

![Fig.2: Architecture of FPKIN showing its optimized path](image)
The transmission flow in the proposed FPKIN was started when an incoming packet received from the IP address of the MNN designated to the IP address of the CN attached to the Firewall. The firewall needs to check the incoming IP header to check the packet type as soon as it receives it. The firewall will check the mobility header and the home-of-address option. If the incoming packet does not include both the mobility header and the home-of-address option, then the firewall will forward the incoming packet depending on the normal routing policy without any update. If the incoming packet includes the home-of-address option but does not contain the mobility options, then the firewall will recognize that this message does not belong to the mobility signaling messages; the firewall should then check if its firewall route optimization cache table (FWROCache) depends on the HoA, source and destination address of the received packet. The existence of this entry in the FWROCache indicates that route optimization exists between the two nodes. Then, the firewall will replace the source address (MNN CoA) from the received packet with the MNN HoA, remove the home-of-address option from the packet and send it directly to the correspondent entity behind this firewall as shown in the above figure.

Furthermore, if the incoming packet includes the home-of-address option and does not contain the mobility header, a null value is returned while matching the firewall cache entries. The received firewall should ignore this packet because it may be an error in the routing packet.

Moreover, a new proposed architecture for setting up the secure communication dynamically is needed in the NEMO Procedure. In order to achieve scalable authentication a FPKIN will be realized in the Firewall filtering rules and activates the root certificate service in the correspondent's gateway. IPsec technology is effective for the mutual authentication method. However, the configuration of IPsec security and manual associations (SAs) is not scalable to the all Correspondent node because a not all mobile entities provide MIP in advance. As the method for establishing key authentications for CNs the digital signature method based on the CA model (PKI) is used [19].

In this proposed architecture, the mutual authentication between a mobile network node and a FPKIN is realized by verifying a communication digital signature and its key certificate between partners. The key certificate and signature are exchanged in the key exchange procedure. Also, a FPKIN and a MNN both holds a key pair (Private Key and public key). The signature, which is hashed and data coded by the private key, can be verified by only the sender's public key in the key certificate. The key certificate is issued by a certificate authority (CA) and identifies the owner of the public key. The CA's digital signature in the key certificate can be decoded only by the CA's public key. If this decoded data is equal to the hashed data of the contents of the key certificate, the key certificate can be verified and the public key's owner can be identified.

3. Performance Evaluation

In this section, the advantage of FPKIN is shown by analyzing and implementing its efficiency in solving NEMO problems with a large number of prefixes and CNs.

3.1. Total Handoff Delay

An analytical model for the NEMO BSP based on route optimization has been developed. The network topology considered for analysis is illustrated in Fig 5. For simplicity, we consider the same number of hops (d x-y) between connected entities. Moreover, all costs are deemed symmetric, i.e. $T_{MR-HA} = T_{HA-MR}$.

For NEMO BSP, the handoff latency is derived by the following formula:

$$h_{NEMO} = t_{L2} + t_{RD} + t_{DAD} + t_r + t_{RR}$$

where

$t_{L2}$ is the Link layer switching delay
$t_{RD}$ is the round trip delay for router discovery
$t_{DAD}$ is the delay for the duplicate address detection procedure
$t_r$ is the MR CoA registration time with its HA. $t_r$ calculated using the following formula:

$$t_r = \left( T_{BA}^{BU} + T_{BA}^{MR-HAMR} + T_{BA}^{BA} \right)$$

$T_{x-y}$ is the one-way transmission delay between two nodes, X and Y, over wired and wireless link that depends on packet size (p), link delay ($\lambda_{\text{wireless}}$, $\lambda_{\text{wired}}$), bandwidths ($R_{\text{wireless}}$, $R_{\text{wired}}$), the probability of wireless link failure ($P_{\text{fail}}$), the number of hops between X and Y ($d_{x-y}$), and queuing delay at each router hop ($D_{\text{queue}}$)[20].

$$t_{RR}$$ is the delay for the NEMO return routability procedure as written in [21, 22], calculated as follows:
the total signaling cost for NEMO BSP \( C_{NEMO} \) could be written as a sum of BU cost \( C_{BU} \), BR cost \( C_{BR} \), and packet delivery cost \( C_{PD} \).

\[
C_{NEMO}^{BU} = C_{BU} + C_{BR} + C_{PD}
\]  

(9)

Where the transmission cost of control packet between nodes X and Y is written as

\[
C_{X,Y} = \lambda d_{X,Y} s_{c}
\]  

(10)

The binding update signaling cost for NEMO BSP is given by:

\[
C_{BU} = \gamma (4C_{MR,AR} + 2A_{proc} + C_{HACN})
\]  

(11)

Where \( C_{HACN} \) is the cost of binding both HA and all CNs.

\[
C_{HACN} = 2(C_{MR,HA} + N_{c}C_{MR,CN})
\]  

(12)

\[
+ HA_{proc} + N_{c}C_{proc} + C_{BRP}
\]

\[
C_{BRP}, \text{ is the cost consumed by return routability procedure and calculated as given by:}
\]

\[
C_{BRP} = 2(C_{MR,HA} + N_{c}C_{HA,CN} + N_{c}C_{MR,CN})
\]  

(13)

\[
+ HA_{proc} + N_{c}C_{proc}
\]

The binding refresh message is typically used when binding lifetime near to be expire. Thus BR signaling cost is given by:

\[
C_{BR} = 2\left(\frac{1}{\sigma} C_{MR,HA}\right) + 2\left(\frac{1}{\nu} N_{c}C_{MR,CN}\right)
\]  

(14)

Where \( \sigma \) and \( \nu \) represent HA binding lifetime and CN binding lifetime, respectively.

The packet delivery cost incurs while session continuity of MR. Also it can be defined as a liner combination of packet loss cost \( C_{loss} \) and the cost of packet tunneling \( C_{tun} \). Let \( \kappa \) and \( \mu \) be weighting factor, where \( \kappa + \mu = 1 \), which underline dropping effect and tunneling effect.

\[
C_{PD} = \kappa C_{loss} + \mu C_{tun}
\]  

(15)

Where \( C_{tun} = 0 \), due to no forwarding during MR handoff.

Let \( s_{c} \) and \( s_{d} \) are size of control packet and data packet, respectively and \( \eta = \frac{s_{d}}{s_{c}} \), then the cost of packet loss is given by:

3.2. Total Signaling Cost

Total signaling cost arising from mobility management schemes should be considered while analysis the performance of wireless networks. In next generation of wireless networks, there are two types of location update signaling. The first one resulted from MR’s crossing and the other happens when the MR’s binding life time is expiring. To distinguish between them, the first one refers to binding update (BU) message and the second one refers to binding refresh (BR) message. In addition, the packet delivery of data consumes the network resources then additional signaling cost is produced. Thus
\[ C_{\text{loss}} = \gamma \eta (C_{\text{PAR,MR}} + C_{\text{proc, MR}}) \\
( t_L + t_{RD} + t_{DAD} + t_r + N_c C_{\text{proc}} + t_{RBP} ) \quad (16) \]

The total signaling cost for FPKIN and NEMO-CR can be calculated as follows:

\[ C_{\text{FRON}}^{\text{NEMO}} = C_{\text{MR}}^{\text{NEMO}} + N_c F_{\text{proc}} \quad (17) \]

Where the MR assumed that number of CN=1 according to firewall functionality. Moreover, \( F_{\text{proc}} \) is divided into the mapping table lookup cost and the routing cost inside the firewall.

For NEMO-CR:

\[ C_{\text{MR}}^{\text{NEMO-CR}} = C_{\text{MR}}^{\text{NEMO}} + C_{\text{CRD}} \quad (18) \]

Where the cost for correspondent router discovery ( \( C_{\text{CRD}} \) ) is divided into the cost of the MR consumed to discover a correspondent router and the cost of the correspondent node to discover the correspondent router. In addition, additional cost incurred by the MR and CNs to verify and validate the CR.

The total handoff delay depicted in Fig. 6, as a function of number of CNs. We observe that total handoff delay increased in NEMO BSP while increasing the number of CNs. However, the NEMO-CR consumes higher handoff time than other schemes when one or two CN peer MNN due to CR discovery costs. But when the number of CNs increased, NEMO-CR is significantly less affected by increasing of CNs. FPKIN is quite bit affected by increasing the CNs.

Fig. 7 show the relationship between the packet arrival rate and the total packet loss. When large number of CNs used (Nc=200) with average number of MNNs inside the MR. The result shows that the packet loss is decreased when the packet arrival rate is also decreased for all schemes. However, the FPKIN is significantly less affected by increasing the arrival rate.

In order to evaluate the signaling cost when the MR change its point of attachment, we assume that up to 10 MNNs are presented and that each MNN is connected to up to 10 CNs. Fig. 8. shows the total signaling cost as a function of number of CNs. In FPKIN, lower signaling cost is observed compared with the other schemes. And the advantage of FPKIN is more obvious. That is because the functionality of new architecture which saves more signaling.

4. Conclusions

The NEMO basic support protocol supports the movement and changes in the point of attachment of entire networks. This solution suffers from a number of limitations and problems that affect network performance, such as signaling overhead, memory overhead, header overhead and network delay. Due to these limitations, correspondent nodes may not be supported in optimized NEMO architectures. To overcome and alleviate the performance penalty, we have designed and evaluated the new network architecture with...
new light weight firewall filtering rules supported with PKI structure and mechanisms. This new architecture FPKIN supports the route-optimization procedure to enhance reachability, manageability, conservation of bandwidth and network performance in an aggregated way. In a firewalled network, CNs are easy to “plug and play”. Applying modifications to the firewall is more reliable and secure way than trying to modify the CN, especially when a firewall protects several correspondent entities using its PKI domain. As future work, we envision implementing the FPKIN architecture with a network simulator (NS-2) and comparing its performance with that of the RO-NEMO architecture.

References