INTERNET-DRAFT
Expires March 2003

Marco Liebsch
Yoshihiro Ohba
[Editors]
Tao Zhang
September 2002

Architecture and Protocol framework for Dormant Mode Host Alerting
<draft-liebsch-dmha-framework-00.txt>

Status of this Memo

Abstract

This document discusses and describes an architecture and protocol framework for the design and specification of a generic IP based architecture and protocol for Dormant Mode Host Alerting, aka IP paging. It focuses on the location and operation related interworking of the different functional entities as required to allow for a flexible and scalable deployment of the protocol. Furthermore, protocol transport and security mechanisms are evaluated and proposed to meet respective requirements on a DMHA protocol. This document intends to be a comprehensive guideline for the design of a generic IP paging architecture and protocol.

Table of Contents

1. Introduction 2
1.1 Background information 2
1.2 General Overview 3
2. Terms 4
3. Dormant mode supporting nodes 6
4. Paging area design 6
5. The role of the paging strategy 8
6. Dormant mode location management 10
6.1 General description 10
6.2 Proposal evaluation 10
6.2.1 Mobility proprietary solutions 10
6.2.2 Mobility independent solutions 10
6.2.3 Hybrid solutions 11
6.2.4 Last-hop solutions 12
6.3 3GPP considerations 12
7. DMA function 13
7.1 DMA location and modes 13
7.1.1 Paging trigger packet reception 13
7.1.2 Paging trigger packet capturing 14
7.2 Filtering for DMHA 15
8. PA function 15
8.1 Functional split of the PA function 15
8.2 Addressing of core components 16
8.3 Addressing of last-hop components 17
9. Signaling transport 17
10. Security support 18
10.1 Security threat categories 18
10.2 Consideration for Category 1 attacks 19
10.2.1 Candidate security mechanisms for DMHA 19
10.2.2 Establishing an SA 20
10.3 Consideration for Category 2 attacks 21
11. Availability 22
11.1 Related nodes 22
11.2 Evaluated mechanisms 23
11.3 Proposed solution 23
A References 23
B Author's addresses 25
C Overview on access technologies 26
C.1 802.11 Power Management 26

1. Introduction

1.1 Background information

This document describes an architecture and protocol framework evaluation for the design and specification of a generic IP based

architecture and protocol on Dormant Mode Host Alerting (DMHA), aka IP paging. This framework document is related to various candidate solutions for the issues addressed in [1] and focuses on the location and operation of the different functional entities as identified in [2], required for flexible and scalable deployment of the protocol. Furthermore, protocol transport and security mechanisms are evaluated and proposed to meet the requirements of a generic IP paging protocol, as described in [2].

Various proposals with different characteristics are addressed, discussed and evaluated with respect to meeting the requirements as identified in [2]. Since the generic protocol for IP paging should be independent of the access technology, but deployment of access technology and interface specific support for dormant mode as well as paging should be allowed, the framework has to take appropriate functions into account to allow for mapping the generic IP paging protocol to technology specifics. In case of not getting support on dormant mode and paging from layers below the IP layer on the access link and respective interfaces, the framework should allow integration of appropriate features on IP level, supporting dormant mode and the paging process in an efficient way.

1.2 General Overview

A generic IP paging architecture and protocol targets at an access technology independent solution for finding and re-activating dormant mobile terminals. The generic nature aims at the integration into various IP mobility managed platforms to support heterogeneous access. Since paging on a mobile terminal's access interface is in most cases optimized taking technology specific functions for dormant mode support and the paging procedure into account, the architecture and protocol to be specified for a generic IP paging solution MUST allow for the mapping of the common IP paging protocol to access technology specific functions.

Evaluation of requirements and the attempt to integrate a set of candidate protocols for IP paging into various IP mobility management schemes has shown, that demand on the architecture and the protocol for IP paging varies. Therefore, aiming at a generic solution while meeting requirements on flexibility and scalability demands thorough analysis of different proposals on how to design the architecture for IP paging. In particular the placement and distribution or co-location of individual paging related functional entities, as indicated in [2], is to be evaluated for individual mobility schemes and network characteristics.

The challenge is to find a compromise between complexity and

flexibility, taking the placement of different paging functional entities and interworking between them into consideration. Further issues to be evaluated are demands with respect to availability. In

this regard, the IP paging architecture has to allow approaches for the introduction of redundancy and efficient deployment of fail-over handling mechanisms. A further demand is to allow providers a high degree of flexibility in paging area design, taking co-existence of L2 and L3 paging areas into account.

This document intends to be a guideline for the design of a generic IP paging architecture and protocol. The various issues addressed in individual sections should be evaluated more and more, aiming at a significant framework document for design recommendations. The implementation and comparison of various protocol proposals as well as an evaluation by means of simulations and theoretical approaches is intended to find the best way to go for the architecture's and protocol's design.

2. Terms

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

In terms of IP based paging, this document tries to be as close as possible to the paging problem statement [1]. To resolve ambiguities, the following terms are defined for the context of this document. In addition, please be referred to [4].

Access Router
: a router that provides access to a (potential) L2 connection to the mobile node

Active State
: mobile node state, mainly characterized by the following characteristics:

1. an established L2 connection
2. the communication peer of mobile node knows the currently configured IP address of the MN
3. the L2 connection is used to carry L3 data traffic

Black Hole
: packet loss without any control of the system

Connection
: established L2 connectivity, providing transport service for L3 traffic

L2

layer 2

L3
layer 3

Idle State
mobile node state, mainly characterized by:

1. less frequent location updates

2. network-maintained location information is less accurate than in Active State

3. the MN is not currently involved in an ongoing L3 communication

IPsec
Security Architecture for the Internet Protocol (RFC 2401)

Location
smallest indivisible unit at which a mobile node can be located (e.g. a single base station)

MN
mobile node

Paging Area
in this context used as general term to describe a set of 2 or more locations likely to be polled in the paging process; see also Predefined Paging Area

Paging Strategy
the order and mode of how a set of Locations is polled to locate an idle mobile node

PAI
Paging Area Identifier, used to identify a Static Paging Area

Polling Cycle
single step of a paging search phase

Predefined Paging Area
statically grouped set of locations, manually configured by an operator for paging purposes

Static Paging Area
see Predefined Paging Area

Unreachable Mobile Node
MN, whose location can not be resolved

for the purpose of routing L3 data traffic to it

3. Dormant mode supporting nodes

There are three kinds of IP nodes that may support dormant mode.

Dormant mobile hosts support both dormant mode and IP mobility (i.e., Mobile IP, Mobile IPv6 and/or micro-mobility). A dormant mobile host may move from one IP subnet to another without updating its mobility bindings needed for routing IP packets from mobility management agents to the host. In other words, L3 routability is not guaranteed for dormant mobile hosts. So additional signaling mechanism (i.e., DMHA) is necessary for realizing locatability in order to search the dormant mobile hosts and notify them of the arrival of data packets so that they can update their mobility bindings and perform IP mobility signaling to recover L3 routability.

Dormant stationary hosts support dormant mode but do not support IP mobility. A dormant stationary host is not supposed to change the IP subnet or even location in some L2 as long as it is in dormant mode. However, this does not mean that L3 routability is guaranteed for dormant hosts. For example, if a traffic channel and a signaling channel are provided separately with different frequencies, time-slots or codes, and a dormant stationary host does not listen to the traffic channel, locatability needs to be provided over the signaling channel in order for the dormant host to "switch-on" the traffic channel when it is alerted. Or when L2 dormant mode support is available, some locatability needs to be provided at L3 by maintaining information that is necessary to utilize the L2 dormant mode service. For example, although an 802.11 access point maintains a list of the MAC addresses of active and dormant stations associated with it, there must be some mechanism to maintain, as such information described above, the mapping between the MAC address and IP address of a dormant host, since 802.11 itself does not care about IP address.

Dormant routers support both dormant mode and IP packet forwarding. A dormant router is different from a dormant host in that a previous hop router of the dormant router triggers alerting based on the next hop IP address of an incoming packet not on the destination IP address. When a dormant router that wakes up to receive an incoming packet then may trigger another alerting to forward the packet, where the node to be alerted may either a dormant host or a dormant router. A dormant router may be stationary or mobile.

4. Paging area design

Paging areas can be static or dynamic. A static paging area does not

change unless re-configured by the network operator manually or via network management system. A dynamic paging area may change dynamically in response to changing network dynamics, such as the changing geographical distribution of the mobile user population and the changing mobility pattern of the mobile users.

Paging area design is to determine how paging areas should be constructed and modified. Paging area design could have significant impact on the signaling overheads incurred for location update and paging. For example, small paging areas could increase the frequency at which mobile hosts have to update their locations, but may reduce the paging signaling overhead incurred by the network to locate the mobile host. Dynamic paging areas, if designed properly, could reduce the overall paging-related signaling overhead (including the signaling overheads for location update and paging). However, supporting dynamic paging areas may need complex algorithms, software and interactions among network nodes for maintaining and modifying the paging areas dynamically. Therefore, a key issue in the design of paging protocol is to allow a network operator to achieve a proper balance between paging-related signaling overheads and system complexity by supporting a proper degree of flexibility in paging area design.

An IP layer paging area is a set of IP layer network attachment points. When designing IP layer paging areas, the following issues that are specific to IP networks also need to be addressed:

1. How should an IP layer paging area be mapped into IP subnets?
2. How should an IP layer paging area be mapped into layer 2 paging areas?
3. How to identify an IP paging area?
4. How to route packets to an IP paging area?
5. How to disburse packets within an IP paging area to the dormant mobiles?

A straightforward IP layer paging area design is to map IP paging areas one-to-one to the underlying IP subnets. That is, every subnet makes a separate IP paging area. This however, could force blanket paging to the entire network in many cases. For example, in today's enterprise wireless networks, one single IP subnet is often used to support all the mobile users in one location. In many such existing networks, hundreds of mobile users are supported on the same subnet that has high capacity enabled by, e.g., switched LAN technologies. In such networks, if the IP subnet is a single IP paging area, all paging messages could be broadcast to all active and dormant mobile users on the subnet. This could lead to a significant waste of the scarce power resources on both the active and dormant mobiles. Therefore, an IP paging protocol should allow flexible design of

paging areas. For example, it should allow an IP paging area to span only part of one IP subnet or across multiple IP subnets. Such a flexible IP paging protocol will allow a network operator to

customize the design of the paging areas to fit its own business requirements.

While it is important to support flexible paging area designs, it is equally important to reduce the potentially high protocol and network complexity associated with supporting such flexible paging area designs. Take supporting dynamic paging area for example. When paging areas are changed dynamically, the number and the shapes of the paging areas may change accordingly. These changing paging areas need to be identified at the IP layer. When the paging areas covering a specific location changes, the changes need to be communicated to the network entities that are responsible for advertising the paging area identities to the mobiles. The mobiles in these affected locations also need to be informed of the changes in a timely manner. These requirements could significantly increase the complexity of a paging protocol if the protocol is not designed properly. Support for flexible paging area design may also impact the architecture of a paging protocol because it impacts the locations and operations of paging agents. For example, if a dedicated paging agent is used for each paging area and N paging agents are configured in a network, then no more than N paging areas can be supported simultaneously. Therefore, intelligent methodologies need to be developed to make the IP paging protocol simple and yet effective in supporting the flexibility needed in paging area design.

When a paging area contains multiple IP access routers, a paging message can be sent to these access routers either via unicast or multicast. Disbursing of paging messages over the wireless media may be carried out in any way available in each particular wireless medium.

5. The role of the paging strategy

Many paging strategies exist today. They can be classified into the following categories:

o Blanket paging:
A paging message is broadcast simultaneously to all dormant mobiles inside a paging area. Blanket paging is used in most of today's wireless networks. Its main advantage is that it generates the least paging latency. The potential drawback, however, is that broadcasting paging messages to all mobiles in a large paging area could consume a significant amount of scarce resource, including power on all the mobiles in the paging area and the radio bandwidth in the radio network.

o Sequential paging:

With this strategy, a large paging area is divided into small paging sub-areas. Paging messages are first sent to a subset of the

paging sub-areas where the network believes the mobile is most likely to be. If the target mobile is not found in one sub-area, subsequent paging messages will be sent to another sub-area. This process continues until the entire paging area is searched or the target mobile is located. Different techniques may be used to determine how to divide a large paging area into smaller paging sub-areas and which sub-areas should be searched first.

o Individual paging:
   With this strategy, an individualized paging area is maintained dynamically for each mobile host. To page a mobile host, paging messages will be sent to the mobile host's individualized paging area.

o Other:
   Other paging strategies also exist that cannot be clearly classified into any of the categories mentioned above. For example, Geographic position-based paging uses the geographical position of a mobile to determine where a paging message should be send. Group paging sends paging messages to a group of dormant mobiles at a time instead of one individual mobile each time.

The choice of paging strategy has a significant impact on the paging performance (e.g., paging latency), paging-related signaling overhead, and network complexity. A good paging strategy should allow a network operator to achieve a balance among

  * Paging-related signaling overhead,
  * Paging latency, and
  * Network complexity

Which paging strategies fit a network operator's business needs depend on the specific network environment and business needs of each specific network operator. It is therefore important for a paging protocol to support different paging strategies.

Furthermore, paging strategy, paging area design and location update strategy depend on each other closely and impact the design of the paging protocol. For example, if only blanket paging is used, each mobile has to update its location every time it moves into a new location area and the location update has to be performed reliably. On the other hand, if sequential paging is used, the network could send subsequent paging messages to enlarge its search area if it cannot find a dormant mobile in one location area. Therefore, with sequential paging, a mobile may not necessarily have to update its location every time it moves into a new location area (e.g., this is likely the case in movement-based location update strategies) and the location updates do not necessarily have to be delivered reliably.

This indicates that the range of paging strategies to be supported could impact the requirements and the design of the paging protocol. In other words, the paging protocol design could significantly impact

what paging strategies and location update strategies can be supported.

The discussion on different paging strategies is out of the scope of this document, but are discussed inter alia in [4] and [5].

6. Dormant mode location management

6.1 General description

This section addresses the location of the Tracking Agent (TA) function, tracking a host's location while being dormant. In general, having the TA function close to the paging areas a mobile terminal is roaming in allows short signaling paths for dormant mode location updating. Otherwise, coupling the TA function with a mobility agent in a mobile terminal's home domain would allow the paging protocol to benefit from mobility management specific functions. To find an appropriate solution on where to implement the TA function, more characteristics of individual approaches have to be studied. This incorporates in particular security related issues as well as signaling costs for dormant mode location updating and the paging process. In general, the location of the TA shall consider the balance of signaling load for registration and paging, L2 technology characteristics as well as the mobile terminal's roaming behavior.

6.2 Proposal evaluation

6.2.1 Mobility proprietary solutions

This proposal assumes that each individual IP mobility management proposal provides proprietary extensions for paging support. This allows taking advantage of each mobility protocol's specific functions for paging. Otherwise, this will result in a set of specific paging protocol extensions and does not meet the requirement on mobility protocol independence [2], aiming at an IP paging architecture and protocol solution, which can be deployed with various IP mobility managed systems. Evaluation of mobility proprietary paging proposals is out of the scope of this document.

6.2.2 Mobility independent solutions

This proposal assumes the TA function as well as the DMA function to be entirely de-coupled from the mobility management protocol and to be part of the generic IP paging framework. Furthermore, both functions are located in the dormant mobile terminal's visited domain.

- Disadvantage: Location info database as well as stateful

entities are to be maintained at home as well as in the visited domain. Required interactions between the two sites are to be studied thoroughly.

- Extra paging trigger packet buffer required in the visited domain for each registered (dormant) mobile terminal.
- Advantage: Keeps dormant mode, paging area location management and paging inside the visited domain (transparency).
- Extra requirements on the introduction of redundancy and on fail-over handling mechanisms (handling of stateful entities).
- Implies requirements on the dynamic establishment of security associations between a mobile terminal and the TA/DMA function.
- Allows flexible deployment of the DMA function w.r.t. its location and mode (capturing or reception of paging trigger packets).

6.2.3 Hybrid solutions

This proposal assumes partitioning the protocol into a core protocol part, which is independent of the mobility protocol, and separate mobility protocol dependent parts. Dormant mode location management is done in a mobile terminal's home domain. Placing only the DMA function in the visited domain, but keeping the TA in the home domain, would not be a reasonable approach, since paging trigger packets, either captured or received in the visited domain, would initiate requesting the TA, located again in the terminal's home domain, of the addressed dormant mobile terminal's location. The only reasonable scenario, when placing the TA function into the mobile terminal's home domain, would be to place the DMA function close to the TA function, or even to co-locate the two functions.

- Requires coordination with the mobility management approach, but would reduce the deployment of appropriate fail-over handling mechanisms to one node, when co-locating the TA/DMA function with a home mobility agent (assumes existence of a mobility agent).
- Disadvantage: Requires inter-domain signaling for paging area updates while a mobile terminal is dormant.
- Disadvantage: L2 support for location tracking is difficult with regard to the security associations between a mobile terminal, its visited domain and the home domain.
- Advantage: Implicit security association available between a

mobile terminal and functions located in its home domain.

6.2.4 Last-hop solutions

This proposal has been addressed to provide paging functionality with a minimum degree of complexity. It assumes all paging related functional entities to be placed on Access Routers. This keeps complexity of functions required for paging low, but restricts the flexibility in paging area design and deployment of enhanced paging functions.

The basic approach would rely on forwarding of paging trigger packets to the last registered IP subnet by means of standard mechanisms provided by the mobility management scheme. As an example, in case of Mobile IPv6 the Home Agent intercepts packets according to standard behavior and forwards these packets to the last registered care-of address of the dormant mobile terminal. The registered subnet's Access Router compares the destination address of the incoming packet with its routing cache entries. If the mobile terminal is active, the packets could be forwarded by means of standard mechanisms. If the mobile terminal has entered a dormant mode before, the Access Router initiates polling the IP subnet to re-activate the mobile terminal. As an option, a set of L2 Access points, covered by the IP subnet, might support multiple L2 paging areas. The latter function would then require again an appropriate protocol between a mobile terminal and its current AR to handle the different paging areas within the current IP subnet.

An argument to advocate such an approach is the reduced complexity of additional functions for paging and the fact that actually no special paging protocol is required. Only the Access Router should discover the respective mobile terminal's state and to initiate a paging function on the IP subnet, which might comprise multiple L2 access points.

This procedure implies the following restrictions:

- A paging area is restricted to the scope of an individual Access Router's IP subnet and cannot comprise multiple subnets.
- Does not provide flexibility to access providers with respect to paging area design.
- No benefit in saving signaling costs when compared to standard IP mobility management approaches.

6.3 3GPP considerations

Lack of interest in a solution for IP paging for 3GPP Release 99 is obvious, since paging is done via proprietary architecture and protocol solutions. Core components deploy the Radio Access Network

Application Part (RANAP) protocol to initiate paging at Radio Network Controllers (RNC). On the other hand, the standardization considers

more and more the integration of Internet Protocol solutions for various functions. Proposals on integrating an IP based framework for Authentication, Authorization and Accounting (AAA) as well as IP based mobility management has been considered. IP migrates more and more towards the architecture's radio access network.

Ideas on "all-IP" systems consider the integration of a set of IP based infrastructure servers for security and mobility management. Hence, in case of introducing Mobile IP approaches for mobility management and to replace the GPRS tunneling protocol (GTP) with Mobile IP tunnels, stepwise replacement of RANAP procedures with IP based approaches is possible. In addition to functions for radio access bearer management and relocation handling, IP paging is a function to be integrated and handled efficiently.

With respect to current discussion on IP RAN migration scenarios, stepwise replacement of RANAP procedures for paging could be considered, terminating the protocol more and more towards the core network and deployment of IP based approaches towards the RAN seems to be a reasonable procedure for smooth migration.

Consideration of RNC decomposition into a control node (control-plane) and a user data packet handling node (user-plane) puts additional demands on a generic IP paging architecture and protocol with respect to placement of individual functional entities and mapping of the generic IP paging protocol to access technology specific paging functions of future heterogeneous access networks.

7. DMA function

7.1 DMA location and modes

This section focuses on the flexible and efficient deployment of the Dormant Monitoring Agent (DMA) function. Basically, the DMA function should be able to work in co-located mode as well as in distributed mode, which means being located remotely from other paging related functions. This allows optimization on paging delay as well as on related signaling costs and performance, dependent on the registration mechanisms of appropriate mobility management approaches.

Two approaches are addressed in the following two sections, one is based on explicit addressing a dormant mobile terminal's DMA function, the other is based on a mode, where the DMA function has to capture data packets getting by. Reception of paging trigger packets assumes the DMA function to be co-located with at least the TA function in a separate node.

7.1.1 Paging trigger packet reception

To allow explicitly addressing a mobile terminal's DMA function while

being dormant, the DMA is to be registered with the mobility management system to receive initial user-data packets on behalf of the dormant mobile terminal. To deploy appropriate registration features is according to requirement 4.8 of [2]. Initial user-data packets are received at the DMA function and are to be buffered after the function has ascertained that the packet addressed a registered dormant mobile terminal and has not been forwarded to the DMA misleadingly.
After the dormant mobile terminal has entered the active state and its current location has been ascertained after the paging process, the DMA function has to re-address and send the buffered packet(s) to the mobile terminal. The DMA is to be de-registered with the mobility management system by means of appropriate location updating mechanisms.

In this mode, the acquired and registered DMA function represents a long-term entity to be responsible for receiving paging trigger packets. The DMA should not change when a new paging area has been entered.

- o Advantage: Avoid performance degradation at individual DMA functions, since usually only packets destined to registered dormant mobile terminals are received by this functional entity. This avoids running the DMA function in promiscuous mode.

- o Advantage: Could benefit from long-term DMA registration with the platform's mobility management system while being dormant.

- o Allows to co-locate the DMA function with the TA function.

- o In case if IP-IP encapsulated packets are received at the DMA function, the actual terminal identifier might be found in the inner IP packet, which requires additional processing of received packets.

- o This approach requires that nodes, implementing the DMA function, terminate tunnels for DMHA only.

7.1.2 Paging trigger packet capturing

When no appropriate registration mechanisms are provided by the mobility management to register an individual DMA function, multiple DMA functions are necessary to be distributed and placed on relevant ingress routers or appropriate nodes, where user-data traffic packets go through on the way to a mobile terminal's location, which has been registered with the mobility management system before the terminal has entered the dormant mode. The DMA functions have to inspect all packets and to capture and process the ones addressed to mobile terminal's, which have been registered as dormant.

- Advantage: Allows to forego an explicit registration of a DMA

function with the mobility management system. This saves signaling when an individual mobile terminal decides to enter the dormant mode.
- o Requires an additional protocol and security association between the nodes implementing DMA functions and the nodes implementing other paging related functions.

7.2 Filtering for DMHA

The DMA function should allow filtering of paging trigger packets to avoid that all packets, addressing a dormant mobile terminal, trigger a paging process by default. The mobile terminal should be able to select from a list of packet types, indicating, which packets are allowed to trigger the paging process. This optimizes the dormant mode support and further allows to reduce paging specific signaling costs.
Otherwise, complex and too extensive filter configuration should be avoided. Therefore, the protocol should allow the mobile terminal to configure the filter function according to a reasonable selection of packet types.
Packet filtering at DMAs supports the dormant mode for mobile terminals as well as for stationary hosts.

8. PA function

The Paging Agent functional entity is responsible for the actual paging process. Being informed of the paging area to search for the paged mobile terminal, the PA polls the paging area by means of an appropriate paging strategy.

Thorough investigation of the PA function results in a further functional split into a generic part and into multiple attendant parts, responsible for mapping the generic protocol to access technology specific functions.

This chapter describes the PA functions and evaluation of addressing schemes for the individual PA functions and related interfaces. The latter evaluation distinguishes between addressing of core components, which includes the interface between the generic PA function and the PA attendants, and addressing of last-hop components, which is related to addressing the mobile terminal from PA attendants either on L3 or on L2. The role of multicast schemes is described and advantages and disadvantages for individual interfaces are evaluated.

8.1 Functional split of the PA function

- o The generic PA function

This part of the PA function could be placed somewhere in the

network, possibly being co-located with the TA function. Its protocol to poll the attendant functions is specified on IP layer and independent of the domain's individual IP subnet supporting access technologies.

o The paging attendant functions

The attendants are addressed from the generic PA function during a paging process. They should be located on a domain's last-hop subnets, e.g. on Access Routers. Polling a paging area, the generic PA function addresses all paging attendants comprising the paging area according to the selected paging strategy. The paging attendants are then responsible for mapping the generic IP paging protocol to L2 paging functions. Furthermore, these attendants support paging area update related signaling by means of mapping appropriate L2 signaling to IP paging area update signaling, which is to be sent to the responsible TA function.

In general, access technology specific dormant mode and paging support should be assumed to be present and is recommended to be used. This section of the DMHA framework document describes the functions to allow for this mapping, but appropriate design of paging attendant functions for individual access technologies is out of the scope of this document.

Some further hints on technology specific support can be found in Appendix C.

In case of not getting support from access technologies' L2, there are proposals for enhanced IP paging approaches on the access link. These approaches are based on multicast addressing or solicited node multicast addressing, as well as on slotted modes to allow a dormant mobile terminal to shut down its interface activities periodically.

L3 slotted mode approach is not evaluated in this document for the following reasons. First, L3 slotted mode might improve power saving performance of dormant hosts, but will not reduce IP signaling traffic in the last hop subnet. Technology specific paging on the access link should be utilized as much as possible when aiming at both saving power and IP signaling traffic. Second, L3 slotted mode needs more investigation on possible and required accuracy levels of synchronization, requirements on IP subnet design (including security) to realize the accuracy levels of synchronization and implementation costs. Of course, the discussed and proposed framework allows for the deployment of L3 slotted mode, but its benefit is expected to be less compared to the cost of design and implementation work for such an approach.

8.2 Addressing of core components

Tbd.

8.3 Addressing of last-hop components

Tbd.

9. Signaling transport

Since battery consumption of hosts is a common issue for both IPv4 and IPv6 nodes, it is necessary that DMHA protocol supports both IPv4 and IPv6. There are two approaches for defining DMHA transport.

First approach is to define a single protocol that is commonly used for both IPv4 and IPv6. This is equivalent to defining a single DMHA protocol above IP layer (e.g, UDP), where minimum IP version specific information such as packet filtering information might be allowed to be carried in the message.

Second approach is to define separate protocols for IPv4 and IPv6. There are following IPv4 specific candidates for DHMA protocol transport.

- Defining new ICMPv4 types
- Defining new Mobile IPv4 [6] extensions information.

There are following IPv6 specific candidates for DHMA protocol transport.

- Defining new ICMPv6 types
- Defining new Mobile IPv6 messages and/or mobility options in Mobility Header [7]
- Defining new IPv6 Destination Options. The defined IPv6 Destination Option could be piggybacked in any packets including data packets.

  There is another approach in which the first two approaches are combined, i.e., using a common transport for both IPv4 and IPv6 for DMHA message exchange among DMHA agents, while applying a separate transport for IPv4 and IPv6 for DMHA message exchange between host and DMHA agent.

  Basically, higher layer transport is suitable for commonality while lower layer transport is suitable for optimization in terms of coupling with other protocol such as Mobile IP. However, the actual

protocol design must also consider security characteristics when choosing DMHA signaling transport. DMHA security discussion is

described in section "Security support".

10. Security support

10.1 Security threat categories

DMHA security threats are described in [2,8]. There are basically two types of attacks related to DMHA. Note that some of the described attacks can be combined to organize two-party attacks [8].

o Category 1 attacks: Attacks that can be avoided or mitigated by authenticating and/or encrypting signaling packets used for DMHA. Category 1 attacks include:

- bogus paging registration messages (detailed explanation is TBD)

- bogus paging area update messages (detailed explanation is TBD)

- bogus paging request messages (detailed explanation is TBD)

- bogus paging area advertisement messages (detailed explanation is TBD)

Note that it is impossible for dormant hosts to avoid receiving and partially processing bogus paging request and bogus paging area advertisement messages, however, it is possible to mitigate the attacks by discarding received messages immediately when integrity check fails. Considering buttery power consumption of dormant hosts, the integrity check applied for those messages should be light weight as much as possible.

o Category 2 attacks: Attacks that cannot be solved by authenticating or encrypting signaling packets used for DMHA. Category 2 attacks include:

- paging trigger packets sent from malicious correspondent nodes, which would result in DoS amplification by (i) producing paging request messages to be widecast in candidate paging area(s) in which the target dormant host is expected to be located, (ii) forming a large sized queue at DMA until the target dormant host is awaken and (iii) awake the target dormant host to drain its battery. (Avoiding this attack is not possible since DMA cannot

make a distinction between malicious and legitimate correspondent nodes.)

- malicious dormant hosts that do not respond to paging request messages. This includes the case in which a host performs paging registration or paging area update with incorrect paging area information. (Avoiding this attack is not possible since PA cannot make a distinction between malicious and legitimate dormant hosts.)

- malicious nodes that make active mode hosts consume batteries and prevent from entering dormant mode, by continuously sending data packets. (Avoiding this attack is not possible since active mode hosts cannot make a distinction between malicious and legitimate senders.)

10.2 Consideration for Category 1 attacks

10.2.1 Candidate security mechanisms for DMHA

One way to secure IP paging signaling packets is to use IPsec [9]. An IPsec security association (SA) is identified based on the peers' IP addresses, security protocol (AH or ESP) and Security Parameter Index (SPI) [9]. Therefore, if a peer changes its IP address, a new IPsec SA needs to be established. With virtually any IP paging protocol, after establishing an IPsec SA with a network agent, a dormant mobile may change the IP address it uses to communicate with the agent. For example, a mobile may move into a different IP subnet where it may:

Need to receive acknowledgements to its L3 Location Updates, if a paging algorithm requires to know a mobile's precise current location area, OR

Need a new IP address to send IP packets to outside the local IP subnet (e.g., to perform location update) if the local access IP router implements packet filtering that will discard outgoing packets if the source IP address is not part of the address space in the local subnet.

Establishing a new IPsec SA requires Diffie-Hellman key exchange [10] in which intensive exponential computation is performed. This can lead to heavy battery consumption. Furthermore, establishing an IPsec SA requires at least three message round-trips that will increase signaling delay. Therefore, IPsec does not seem appropriate for securing IP paging protocols.

Otherwise, if the IPsec security association could be based on a static identifier, e.g. a common paging identifier to be unique

across the borders of the registered paging area as long as the mobile terminal is dormant, establishment of a new SA could be avoided in case of the mobile changes it's IP address. This is basically an issue of implementation and would allow IPsec mechanisms to be deployed for authentication and encryption without accessing any key distribution center or other mechanisms for the establishment of a new SA while a mobile is dormant.

When Mobile IP (v4 or v6) is used for mobility management, establishment of a new IPsec SA may be avoided if a paging protocol entity sends signaling packets destined to a mobile to the mobile's Mobile IP Home Address, which does not change as the mobile moves about. However, in order for the mobile to receive these packets in a visited IP subnet, the mobile would have to first obtain a new local Care-of Address, transition into full active mode and perform mobility binding with its Home Agent each time it moves into a new L3 Location Area. This could result in a signaling cost that is comparable to placing the TA on the Home Agent.

For other cases, it would be better to consider using a security mechanism that builds security associations based on stable identifiers that do not change while the mobile moves about, but allows mobiles to use dynamically changing local IP addresses to receive packets without having to perform mobility binding.

One choice of stable identifier is Cryptographically Generated Address (CGA) [11]. A CGA contains a part of an IP address (i.e., lower 64-bit of an IPv6 address) that is cryptographically generated based on public key cryptography. This part of the IP address will be assigned the same value by the sending entity regardless wherever the sending entity connects to the network. Using CGA, a signaling packet carries a Message Integrity Check (MIC) field calculated for the data to send as well as the sender's public key (the public key is necessary only in the first message exchange). The cryptographically generated address will be specified as the source IP address of the signaling packet so that the receiver can authenticate the sender's IP address as well as the data by using the attached public key. However, since CGA depends on a particular address structure and address assignment scheme, it cannot be applied if the address structure or address assignment scheme is different. For example, CGA is difficult to apply if the entire part of an IP address is assigned by the network.

Another choice is to use higher layer security mechanisms in which stable identifier is defined. Use of TLS [12], or defining authentication field within DMHA signaling message payload [13] matches this approach.

10.2.2 Establishing an SA

SAs that are used for avoiding or mitigating Category 1 attacks can be established either statically or dynamically. Static SAs could be applicable for small sized networks only, where scalability is not necessarily required for key management. Dynamic SAs would be necessary for other cases. There are following ways to dynamically establish a SA between host and DMHA agent, or between two DMHA agents.

o Establishing an SA based on strong authentication (i.e., by using some other pre-shared SA or using PKI)

- IKE[10] (mostly used for establishing IPsec SAs)

- AAA[14,15] with appropriate key distribution mechanisms (tied with client authentication)

- Kerberos[16] (tied with client authentication)

o Establishing an SA based on weak authentication

- CGA[11]

o Purpose Built Key approach

The idea of this mechanism is to allow the establishment of a SA for a specific function (like dormant mode and paging). When this function is not required anymore (after a mobile has been paged and re-activated), the security association is deleted. In case of paging, a key could be generated before a mobile enters the dormant mode and be kept on the mobile terminal as well as on paging relevant nodes within the network as a shared secret. The association to an individual mobile terminal should be identified uniquely by means of a static identifier, which could be a proprietary paging related identifier or a mobility management related static identifier (e.g. the home address in Mobile IP). Static means, that it should not change while the respective mobile terminal is dormant. Further, it should be avoided, that validity of this identifier is based on a lifetime, which is to be refreshed while being dormant. The validity should expire after a successful paging process and after the mobile terminal has been fully re-activated.

As an option, the generation and exchange of the PBK could be coupled with the IP paging protocol procedures.

10.3 Consideration for Category 2 attacks

For paging trigger packets sent from malicious correspondent nodes, the only way to solve this problem is packet filtering. Ingress

filtering or packet filtering at access routers through which correspondent nodes are connected to the Internet is effective for blind masquerade attacks by which an attacker randomly changes the source IP address fields of packets used for the attack. For attacks from a malicious correspondent node that has a correct IP address assigned at the connected subnet, packet filtering at DMA as described in "Filtering for DMHA" would be effective for attackers who are not aware of what types of packets trigger paging.

For attacks from malicious dormant hosts that do not respond to paging request messages, adaptive paging timeout mechanism could be effective in which the paging timeout value becomes small as the number of paging failure increases. In addition, validity check for paging area information contained in paging registration and paging area update messages against the host's actual location would be necessary.

For attacks from malicious nodes that make active mode hosts consume batteries by continuously sending data packets, the target host can enter a low power mode when it receives a large amount of unimportant packets, as is described in [8].

11. Availability

Availability is an issue to be addressed and to be taken into account already before the actual design of the architecture and protocol for IP paging is done. To allow for efficient introduction of redundancy and deployment of mechanisms for fail-over handling, this should be taken into consideration when taking decisions on distributing or co-locating stateful entities.

11.1 Related nodes

With regard to IP paging, this issue is mainly related to the Tracking Agent functional entity as well as to the DMA functional entity.

o DMA co-located with TA

When having the DMA co-located with the TA, the same database could be accessed to check for individual registered mobile terminals' entries. This requires the registration of the node implementing the DMA to allow reception of paging trigger packets at the DMA, as discussed in section 7.1.1. Otherwise, such a configuration allows efficient management of redundancy, since less stateful entities are to be mirrored. Furthermore, appropriate protocols for fail-over handling could manage redundant nodes easier.

- Separated and distributed DMA(s) from TA

Such a configuration requires distribution of DMAs within the network or on a domain's last-hop subnets, as discussed in 7.1.2. A protocol is to be deployed between distributed DMAs and an individual TA function to coordinate registries. Both functions, TA as well as distributed DMAs, have to keep and manage states for individual registered dormant mobile terminals. Dependent on the network topology, multiple DMAs keep the same registry for an individual mobile terminal.

11.2 Evaluated mechanisms

To be done.

11.3 Proposed solution

To be done.

A REFERENCES

[1] J. Kempf, "Dormant Mode Host Alerting Problem Statement", RFC 3132, June 2001

[2] J. Kempf et al., "Requirements and Functional Architecture for an IP Host Alerting Protocol", RFC 3154, August 2001.

[3] J. Kempf et al., "Dormant Mode Host Alerting (DMHA) Protocol Assessment", draft-ietf-seamoby-paging-protocol-assessment-01.txt, work in progress, February 2002

[4] C. Rose / R. Yates, Ensemble Polling Strategies for Increased Paging Capacity in Mobile Communication Networks, ACM Journal of Wireless Networks, vol. 3, no. 2, pp.159--167, 1997

[5] W.-S. Wong / V.M. Leung, Location Management for Next-Generation Personal Communication Networks, IEEE Network, Sept. 2000

[6] C. Perkins, "IP Mobility Support for IPv4", RFC 3344, August 2002.

[7] D. Johnson, et al, "Mobility Support in IPv6", draft-ietf-mobileip-ipv6-17.txt, work in progress.

[8] P. Mutaf and C. Castelluccia, "IP Paging Threat Analysis", draft-mutaf-paging-threats-00.txt, work in progress, February 2002.

[9] S. Kent and R. Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[10] D. Harkins and D. Carrel et. al., "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

[11] G. Montenegro and C. Castelluccia, "SUCV Identifiers and Addresses", Internet-Draft, draft-montenegro-sucv-02.txt, work in progress, November 2001.

[12] T. Dierks, et al., "The TLS Protocol Version 1.0", RFC 2246, January 1999.

[13] T. Zhang, et al., "A Flexible and Scalable IP Paging Protocol", to appear in IEEE Globecom 2002.

[14] C. Rigney, et al., "Remote Authentication Dial In User Service (RADIUS)", RFC 2865, June 2000.

[15] P. Calhoun, et al., "Diameter Base Protocol", Internet-Draft, work in progress, June 2002.

[16] J. Kohl, et al., "The Kerberos Network Authentication Service (V5)", RFC 1510, Semptember 1993.

[17] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications", IEEE Std 802.11, June 1997.

B AUTHOR'S ADDRESS

Marco Liebsch
NEC Network Laboratories Europe
Adenauerplatz 6, 69115 Heidelberg
Germany

Phone: +49 (0)6221 13708-11
Fax: +49 (0)6221 13708-28

E-mail: marco.liebsch@ccrle.nec.de

Yoshihiro Ohba
Toshiba America Research, Inc.
P.O. Box 136
Convent Station, NJ 07961-0136
USA

Phone: +1 973 829 5174
Fax: +1 973 829 5601

E-mail: yohba@tari.toshiba.com

Tao Zhang
Telcordia Technologies
445 South Street, Room 1J-214B
Morristown, New Jersey 07960
USA

Phone: +1 973 829 4539

E-mail: tao@research.telcordia.com

Appendix C Overview on access technologies

Appendix C.1 802.11 Power Management

The power management capability of IEEE 802.11 [17] for an infrastructure network can be summarized as follows. A station changing Power Management mode informs the Access Point (AP) of this fact using the Power Management bits in the Frame Control field of the transmitted MAC frames. An AP periodically broadcasts beacon signals to provide time synchronization information and inform stations in Power Save mode of arriving frames. A station uses the time synchronization information received from the AP to determine when it should wake up periodically from Power Save mode. The AP buffers the MAC frames destined to the station in Power Save mode and transmit them at designated times.

Unicast frames destined to a host in Power Save mode are transmitted by the AP and received by the station in different ways from broadcast/multicast frames. With every beacon transmission, the AP informs each station in Power Save mode of the unicast frames buffered by the AP for the station and whether these frames are to be sent to the station during a content-free or a contention time period. If a unicast frame is to be sent in a contention period, the station will poll the AP to receive the unicast frame. If a frame is to be sent during a contention-free period, the station will not poll the AP but will instead remain active until the frame is received or the contention-free period ends.

The AP notifies the stations of the existence of broadcast/multicast frames only via selected beacons periodically and the broadcast/multicast frames are sent immediately after these beacons.

By utilizing the timing differences for multicast/broadcast frames and unicast frames, three different dormant modes can be realized in the single Power Save mode.

o All Mode: Both unicast and multicast/broadcast frames are received.

o Unicast Only Mode: Only unicast frames are received.

o Multicast Only Mode: Only multicast/broadcast frames are received.

The dormancy levels of Unicast Only Mode and Multicast Only Mode are higher than that of All Mode. Unicast Only Mode is effective in terms of battery saving especially for a host connecting to the network where broadcast/multicast traffic is high and most of the broadcast/multicast traffic is not important to the dormant station. On the other hand, there are also important broadcast/multicast frames that need to be received by the dormant Host in order to to

receive incoming SIP calls. One example is ARP REQUEST packets which are broadcast by a node in the last hop subnet in order to obtain the

MAC address for an IP address of the host. Thus, a mechanism would be necessary for converting a particular broadcast/multicast MAC frames to a unicast MAC frame when the dormant host is operating in Unicast Only Mode.