

Personal

A. O'Neill

G. Tsirtsis

BT

Internet Draft

S. Corson

Document: [draft-oneill-craps-handoff-00.txt](#)

Ansible Systems

Category: Informational

August 2000

Expires: February 2001

**Generalized IP Handoff**  
<[draft-oneill-craps-handoff-00.txt](#)>

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#). Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at  
<http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at  
<http://www.ietf.org/shadow.html>.

Abstract

This draft describes a generalized IP handoff model that supports edge mobility such as that encountered in cellular systems and Internet roaming. The model is mobile-controlled and network-constrained. The model is generic in that it permits both planned and unplanned handoff (i.e. proactive and reactive) and supports both make-before-break and break-before-make link layers. It also supports movement within and between technologies, as well as within and between domains. Finally, it is -equally applicable to routed mobility and to tunneled mobile IP-based mobility.

## INDEX

1. Introduction
  - 1.1 Inter-technology Handoff Scenario's. . . . .
2. The Handoff Model
  - 2.1 Handoff Preparation--the Forward protocol. . . . .
  - 2.2 Handoff Completion--the Reverse protocol . . . . .
  - 2.3 Benefits of the Approach. . . . .
3. Legacy Handoff
4. References

**1. Introduction**

This draft describes a generalized IP handoff model that supports edge mobility such as that encountered in cellular systems and Internet roaming. The handoff model is network-constrained/mobile-controlled. The model is generic in that it permits both planned and unplanned handoff (i.e. proactive and reactive) and supports both make-before-break and break-before-make link layers. It also supports movement within and between technologies, as well as within and between domains. Finally, it is equally applicable to routed mobility and to tunneled mobile IP-based mobility. This is necessary to ensure that intra-domain mobile routing can make use of Mobile IP services where appropriate and Mobile IP can utilize mobile routing to limit vertical and direct signalling to the Home Agent, Gateway Foreign Agent and CN's. It also enables the MN to be isolated from the mobility support deployment decisions of the domain operator.

Agreement is being reached within the Mobile IP working group that a 'smooth' handoff is one that occurs without packet loss. A 'fast' handoff is one that occurs with no discernable latency. Here we define a 'seamless' handoff as one that is both smooth and fast. The generalized handoff model attempts to support seamless IP handoff.

**1.1 Inter-technology Handoff Scenario's**

Intra-technology handoff is both well understood and already supported, especially in the case of cellular systems. A generalized IP handoff model needs to be both applicable and complementary to

these systems. In addition, there are a number of scenarios in which an ISP would wish to support a planned handoff between two different access technologies, especially in the situation where the two access technologies are attached to the same ISP. In the case of a change of ISP, which equates to an inter-domain handoff, the facility is driven by the need for partnerships between single access owners to compete with the ISP with multiple access services.

Here are some example scenarios to help highlight the need for the different characteristics of the handoff model advocated by this

O'Neill, Corson, Tsirtsis

2

Internet Draft

Generalized IP Handoff

August 2000

draft. The model itself, however, is not tied to any specific scenario and is instead intended to be generic and independent from specific business cases.

#### Residential Handoff

A MN is able to access a home network and associated fixed WAN Internet access infrastructure through either an RF interface (e.g.: Bluetooth) on a mobile terminal or a network connected cradle. At the same time the MN has access to cellular resources that overlap the home network area. Preference may be given to the home resources due to the improved bandwidth, cost, performance characteristics. When moving out of range of these home resources and when subsequently returning home to those resources, the user would wish to undertake a seamless handoff to maintain application performance. In the case of a change of ISP it is likely that AAA processes, middleware-based preferences and user action would all likely be required to affect this handoff. In the case of an intra-ISP handoff, service based preferences in the network and MN will likely be sufficient.

#### Office Handoff

In this scenario, the user has desktop access to 100Mbit ethernet and in addition has an 802.11b PCMCIA card which provides access to scarcer radio resources whilst in transit or away from fixed resources. When leaving or returning to the desktop facility, the user would be necessarily plugging and unplugging the fixed ethernet connection, and the user would wish again to maintain ongoing application performance without additional activity. Again, it is clear that some user action is required but the fixed link loss should be sufficient to automatically trigger the handoff to the radio layer.

#### Corporate Handoff

This is an example of an inter-domain handoff and supports the

transition between public resources and corporate resources. The MN is entering the corporate facilities and is on a call, with the service being paid for by the corporate on behalf of the user. The corporate can provide the service cheaper than the public facility so would wish to trigger a handoff when its staff are in range of corporate facilities. The handoff should be seamless if possible to avoid staff simply delaying entering the handoff area whilst on a call, and hence defeating the aim of reducing cost. Non seamless handoff may be permissible for specific applications and terminal types, particularly in exchange for better performance once the handoff has occurred.

#### Handoff timing observation

In the general case handoffs should be delayed as their occurrence increases signalling overhead. Ideally they should only be undertaken when they decrease the probability of service interruption. Thus, for example, a MN coming in range of a new AR does not mean that the MN should immediately handoff to this AR. Instead, more sophisticated policies accounting for MN-AR Signal

O'Neill, Corson, Tsirtsis

3

Internet Draft

Generalized IP Handoff

August 2000

Interference Ratio (SIR) levels, MN QoS considerations and AR air interface loading issues should form the basis of handoff timing. In some cases "early" handoffs may be required for network efficiency reasons; in others handoff may be delayed. Equally QoS/service policy in the MN may require an early handoff to another technology as in the examples above.

## **2. The Handoff Model**

The handoff model consists of two principle mechanisms--proactive (or forward) and reactive (or reverse)--that work together to achieve seamless handoffs. The first mechanism is optional, preparatory in nature, and can be viewed as a performance optimization. It may be invoked when either the network (i.e. some notional network controller entity (NC)) or the mobile host has advance knowledge of a pending handoff. After this (and possibly without it), the reverse mechanism is invoked to complete handoff. This mechanism is always mobile-initiated. The resultant messaging sequence partially depends upon whether or not the forward mechanism has occurred.

The handoff model separates the horizontal signalling (dealing with handoff) from the vertical signalling (dealing with routing) because there are substantial differences in the vertical plane between Mobile IP and intra-domain mobile (host) routing solutions, whilst

there is great commonality in the horizontal plane. The horizontal plane is associated with the desire to locally support the movement of the MN between the OAR and the NAR using local authentication, temporary tunneling of packets, and the transfer of critical protocol state, to the NAR, to minimize service disruption. The vertical plane either causes a routing update which maintains the utility of the existing Care of Address (CoA) in the domain by transferring its route to the NAR, or in the case of MobileIP, the vertical plane causes a Binding Update to either the Home Agent or the Gateway Foreign Agent to bind the Home Address (HoA) to the new CoA at the NAR. Both options are generalised by the Update (UPD) and Update Acknowledge (UPDack) messages from/to the NAR.

The model is network constrained, mobile controlled such that the network constrains the MN's choices for the handoff NAR, between multiple offered NAR's, with the MN being in control of the final selection. The network is viewed as providing guidance and hints to the MN as to the timing of the handoff and choices between different handoff destinations. The timing could vary between 'immediate' in the case of rapid movement / link degradation, through to 'undefined'. These hints can be triggered autonomously by network processes that track and manage resources and policies. In addition, the hints can be triggered by MN handoff requests and actions, which the network can constrain, modify and/or supplement to enable it to police and optimize system resources.

O'Neill, Corson, Tsirtsis

4

Internet Draft

Generalized IP Handoff

August 2000

The MN must always ultimately be in control of the handoff. This is to avoid the inevitable situation in which the network cannot fully appreciate the application needs and/or personal circumstance of a user, or user process, at a given time. The user should be able to view any assistance as providing useful automation that can be overridden based on local preferences, knowledge or actions. A good example is in the case of overlapping service areas covered by multiple disconnected networks, both of which only see their own view of the situation and would wish to impose conflicting instructions to a MN. The process of constraining the timing and handoff locations should be sufficient for the network to protect and manage its resources whilst enabling the MN to recover from network control problems and to utilize facilities from multiple disconnected layer 2 networks.

The full message set is shown in Figure 1.

+-----+

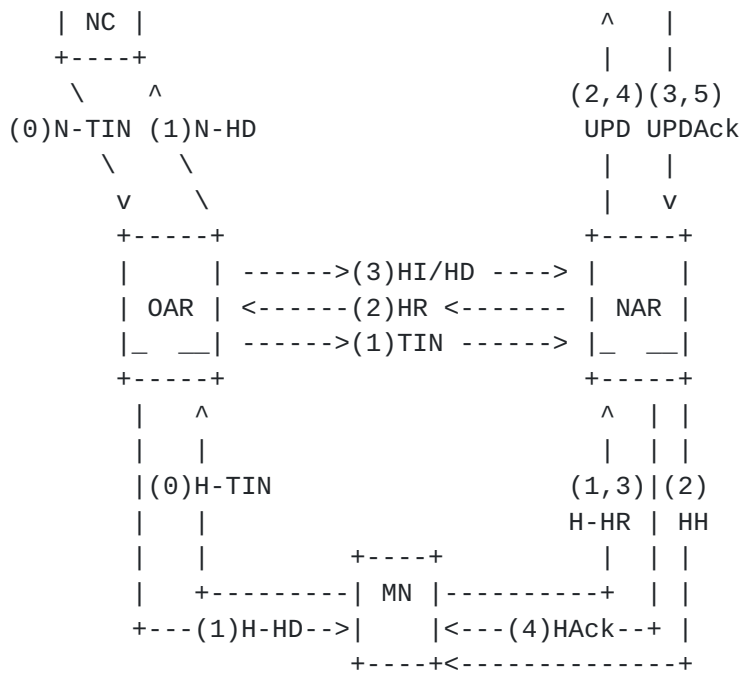


Fig1: All messages

#### Entities

NC = Network Controller  
 OAR = Old Access Router  
 NAR = New Access Router  
 MN = Mobile Node

#### Messages

TIN = Tunnel INitiation  
 H-TIN = Host TIN  
 H-HD = Host Handoff Denial  
 N-TIN = Network TIN  
 N-HD = Network Handoff Denial  
 UPD = Update

UPDAck = UPD Ack  
 HI = Handoff Initiation  
 HD = Handoff Denial  
 HH = Handoff Hint  
 HR = Handoff Request  
 H-HR = Host HR  
 HAck = Handoff Ack

## 2.1 Handoff Preparation--the Forward protocol

We assume the OAR is receiving the IP flows for delivery over the old link. The old link is used for data forwarding as long as necessary (or possible). The forward protocol is initiated by the "Tunnel INitiation" (TIN) Message and is shown in Figure 2. If mobile-initiated, the message originates at a mobile host and is called a "Host Tunnel INitiation" (H-TIN) message. If network-assisted, the message originates at a NC entity and is called a "Network Tunnel INitiation" (N-TIN) message. An H-TIN or N-TIN is sent to the OAR instructing it to prepare for possible handoff to a NAR. The message might be generated for a number of reasons, including policy, prior knowledge of movement and other parameters (outside the scope of this draft) known to the MN or NC. The OAR may deny either a H-TIN or N-TIN based on local information resulting in an optional H-HD or N-HD message back to the MN or NC, respectively. This deny is semantically equivalent to the HD message discussed later.

Arrival of a H-TIN or N-TIN message at the OAR instructs it to build a temporary tunnel to the NAR for the possible forwarding of data packets during handoff. The OAR then sends a TIN message to the NAR. The TIN effectively conveys the state necessary to achieve seamless handoff to the NAR. On TIN arrival at the NAR, the NAR establishes the necessary tunnel and optional buffer state for possible reception of packets. This tunnel is used if the old link is lost in either an unpredictable or predictable manner, and therefore provides both a safety net and an IP layer make-before-break capability to supplement break-before-make link layers or user activities. The tunnel transit time and any resultant session / flow buffering at a NAR can provide additional time for the new link to be initialized.

The OAR continues to deliver over the old link and may optionally forward a copy of the data via one or more virtual links to the potential NAR(s) to assist with seamlessness. The decision to forward may depend on the H/C-TIN request contents and may be modified by the OAR based on policy. Here we term the replicated delivery of packets to multiple NARs for possible forwarding to the MN as IP diversity. This differs from and should not be confused with physical radio layer "macro diversity" which is orthogonal to, but supplemented by IP diversity.

```
+----+
| NC |
+----+
 ^   \
```

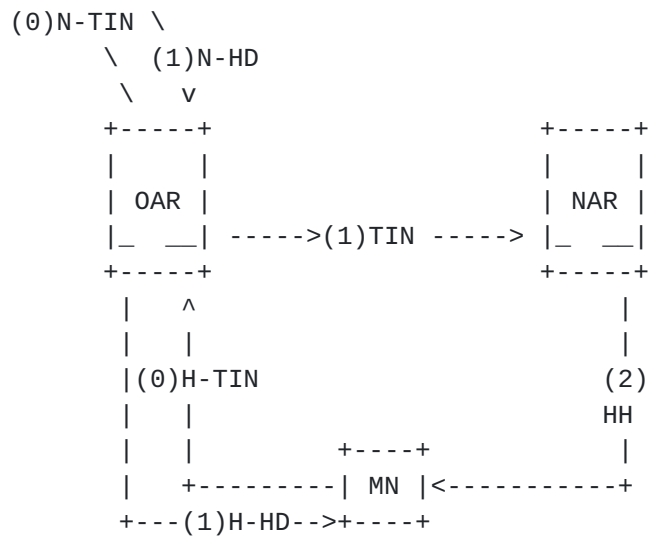


Fig2: Forward Messages

The NAR may send a "Handoff Hint" (HH) to the mobile if a layer 3 link to the mobile is present at the NAR. It is triggered by the TIN and therefore tells the MN that a TIN has succeeded. During movement it is possible that multiple tunnels may be built to multiple potential NARs. Some tunnels may be mobile-initiated. Others may be network-initiated. The mobile may receive multiple hints from these potential NARs that the network deems suitable for handoff. The HH is therefore sent from each NAR that the network is prepared to offer as a handoff point. In a dumb network, this set will be the same as the set of H-TINs sent by the MN because the network is unconstrained. In a 'clever' network, the set may be reduced or modified based on the networks needs.

The HH may report the state of any requirements and facilities requested by the MN or advertised by the NC. HHs may therefore include performance and preference information from the network perspective to assist the MN to rank NARs. If the TIN is MN initiated, the MN may in addition provide input to this process in terms of its own preferences and priorities.

A HH can indicate the status of the 'hint' that can range from mandatory/suggested/optional, and the lifetime of the handoff window during which this NAR will accept the handoff (immediate, time window). The associated processing in the MN will depend on the MN service privileges and the nature of the handoff (inter/intra domain/technology). The MN may receive mandatory immediate messages from multiple disconnected control systems with overlapping coverage which illustrates another case where the MN must be in final charge to resolve disconnects in network intelligence.



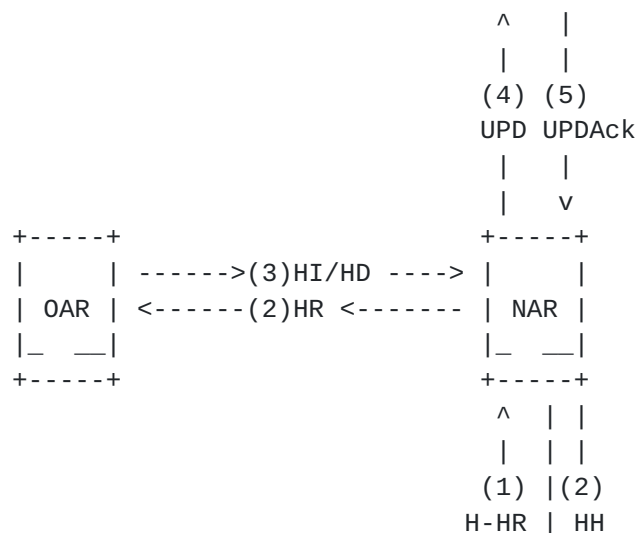
The HH also indicates whether or not replicated data is available at the NAR. If the MN requested replication and it was successful, then replicated packets may be delivered by the NAR to the MN in advance of handoff as soon as the new link is available. Alternatively it may only be forwarded after reception of the H-HR at the NAR.

A HH may never arrive due to failures, such as the loss of the TIN. Therefore the MN can request handoff independently over the new link (unplanned) on receipt of periodic or locally triggered Agent Solicitation (AS), Agent Advertisement (AA), Router Solicitation (RS) and Router Advertisement (RA) messaging.

If the N/C-TIN is rejected for administrative reasons, such as lack of resources, authorization or any other reason a "Network-Handoff Denial" (N-HD) or "Host-Handoff Denial" (H-HD) message may be send back to the NC or the MN.

## **2.2 Handoff Completion--the Reverse protocol**

The mobile is ultimately responsible for completing handoff. The old link is used as long as necessary. At some time it determines it should handoff to a NAR. The timing of this decision is beyond the scope of this document but will involve both NAR and MN processes. This handoff may or may not be towards a NAR from which it has received an HH message. Clearly, the absence/suppression of an HH from a NAR may indicate that a handoff is not available to that NAR, and any handoff attempt may therefore fail. A MN should therefore prioritize NARs advertising an HH above those simply advertising asynchronous NAR advertisements. Regardless, the mobile initiates handoff by sending a "Host Handoff Request" (H-HR) to the selected NAR as shown in Figure 3.



```

+-----+          |  |  |
| MN |-----+    |  |
|      |<---(4)HAcK--+ |
+-----+<-----+

```

Fig3: Reverse Messages

On H-HR arrival at the NAR, the NAR checks to see if preparatory handoff state for the mobile is present (this may include AAA information and other handoffs-specific state sent in conjunction with tunnel establishment). If not (in the case of an unplanned handoff) a Handoff Request (HR) message is sent from the NAR to the OAR. This message likely carries the mobile's credentials to be checked at the OAR.

On HR arrival at the OAR, the OAR performs an authentication check on the mobile. If everything is in order it sends a Handoff Initiation (HI) to the NAR and creates the forwarding tunnel and buffer state in both the OAR and the NAR. Otherwise it sends a Handoff Denial (HD) packet to the NAR. The information carried in a forward TIN message as well as other information not normally sent in a TIN (e.g. authorization) comes across in an HI message.

On HI arrival at the NAR, the NAR declares the handoff request a success and initiates the vertical activity which may include either a MIP HoA/CoA binding update to the Home Agent/Gateway Foreign Agent or a routing update in the case of intra-domain mobile routing. Alternatively on HD arrival at the NAR, the NAR declares handoff failure. If the mobile has requested handoff notification in its HHR, the NAR then sends an HAcK (or HNack) accordingly.

The above unplanned handoff can be further optimized if the NAR is in a position to locally authorize and execute the handoff rather than having to query the OAR for authentication, authorization and routing/binding information. In this case the NAR can initiate the horizontal and vertical handoff activity in parallel with each other so avoiding the round trip delay to the OAR. The OAR to NAR tunnel is still useful and the completion of the horizontal phase as normal efficiently manages resources and enables the system to be resilient to vertical messaging failures and delays.

In the case of a planned handoff, where a successful forward phase has previously occurred, then the tunnel and handoff state will be in place at the NAR. The mobile is therefore immediately cleared for handoff to the NAR which triggers the vertical activity, and a HAcK

is returned to the mobile if so requested by the mobile. In addition, the reverse protocol continues to the OAR to ensure consistent behavior and state management in all elements independent of whether the handoff is planned or unplanned.

It can be seen that the de-coupling (asynchronous timing of messaging and state manipulation) of the forward and reverse phases enables the MN to recover from forward handoff messaging and network failures / delays by reverting to an unplanned handoff. The model is also resilient to vertical messaging problems as the horizontal tunnel is always installed but often never used.

### **2.3 Benefits of the Approach**

The approach permits proactively guided/constrained handoff in the forward direction. This is useful for intra-technology handoffs (e.g. cellular) where handoff planning is desired for spectrum efficiency and system load balancing. A NC element can prepare one or more NARs for handoff and signal the relative desirability of handoff to these various NARs to the mobile. Ultimately the mobile must decide to which NAR it will handoff. The network can exert additional indirect control over mobiles by denying handoffs if necessary at any time.

3G to WLAN handoff is an example of where a generalized IP handoff model must be mobile controlled. A 3G network would be unaware of the existence of the WLAN and of the mobile's desire to perform an inter-technology handoff.

The mobile-controlled reverse mechanism works on failure of (or in the absence of) any preparatory forward handoff processing. Thus it provides a way for a mobile to recover from any failure of planned handoffs, and to control handoffs across technologies. This includes movement between wired and unwired technologies.

Establishment of a temporary tunnel from the OAR to NAR effectively establishes a virtual link to the NAR for data forwarding. On unexpected loss of the link between the OAR and the mobile, any data packets hitting the OAR can be tunneled to the NAR for the mobile (and buffered if necessary while awaiting the mobile's arrival). This effectively permits layer 3 "make-before-break" operation over systems that operate at layer 2 in a "break-before-make" mode. This is accomplished by putting the necessary intelligence only at the edge routers, and does not rely on intelligent network elements deeper (or higher) in the network.

The handoff model is orthogonal to the means by which data is routed to the mobile. Mobile IP [[MobileIP](#)], natively-routed micro-mobility [[EMA-MIP](#), [CellularIPdraft](#), [HAWAIIIdraft](#)] or some combination of these schemes may be in use.

## 2.4 Mapping the Generalized Model to MobileIP

The full message set suggested in this draft is shown in Figure 1 and is mapped to some existing Mobile IP draft messages in table 1.

O'Neill, Corson, Tsirtsis

10

Internet Draft

Generalized IP Handoff

August 2000

Messages/Draft	Proactive	PFANE	Anchor	Smooth	EMA:MIP
TIN	FBU			SHPSH	FBUack
H-TIN					FBU
N-TIN					FBU
HH	HORQ				FBUack
UPD	RREQ	RREQ		BU	UDU
UPDAck	RREP	RREP		BUAck	UDUack
HR		BU	RREQ	SHREQ	RBUack
HI			RREP	SHREP	RBUack
HD				SHREP	RBUack
H-HR	RREQ	RREQ	RREQ	SHIN/BU	RBU
HAck	RREP	RREP	RREP	SHACK	RBUack

Table 1: Generalized IP Handoff Mappings

The original EMA handoff work, covered most recently but incompletely in the [[EMA-MIP](#)] draft reflects all the messages in the generalized handoff model. In [[EMA-MIP](#)] the MIP signalling messages are used for the horizontal tunnel, state management, authentication/authorization and error control. This is supplemented by additional piggybacked messages to convey the EMA routing update information for the mobile routing update in the vertical plane as a replacement for the vertical MIP HA/GFA binding updates and registrations. The EMA:MIP horizontal MIP messaging model uses existing MIP features in a novel way to provide a comprehensive

handoff model with similarities to other MIP work. For example, the proactive [[Proactive](#)] and PFANE [[PFANE](#)] drafts have forward messaging, but lack a reverse mechanism. The anchor draft [[Anchor](#)] has reverse messaging, but lacks a forward component. The recent smooth draft [[Smooth](#)] has adopted both components, but still lacks explicit forward messages from the host or NC to initiate handoff, as well as a message acting as a hint for the mobile to handoff.

The smooth draft also has a different method of signaling interaction with the NAR, preferring IPv6 tunnels over the IPv6 Routing Option approach used in [EMA:MIP]. Further exploration of this difference is required to establish the preferred method given ALL requirements. Table 1 only reflects recent Mobile IP handoff drafts similar to the generalized model. Others diverge somewhat from the generalized model. For instance, the fast handoffs draft [[FastHandoffs](#)] has the MN register with the NAR via the OAR which is acknowledged via the OAR as opposed to the NAR in our model. This may be a useful optimization in break before make link layers which do not have a L2 make before break signaling plane (which is not the case in most cellular systems). Other drafts diverge more significantly from this generalized model but the reasons for this divergence is unclear.

### **[3. Legacy Handoff](#)**

Legacy (cellular) networks support a network controlled handoff model. This model assumes a relatively "dumb" MN that has no control over the handoff process and which may, at most, assist in the handoff procedure. The NC always knows which is the next appropriate NAR for the MN.

In the model described here, to support such mobiles, the NC can send a N-TIN to the OAR as shown in Figure 4. The OAR then sends the TIN to the NAR and sets-up the temporary tunnel for the MN. The TIN also carries AAA info for the MN so the NAR does not have to authenticate the MN. The NAR sends an HH to the MN indicating a mandatory handoff and effectively changes the state of the MN so that it is now attached to the NAR. In some cases, IP connectivity between the MN and the NAR has not yet established, in which case the HH might be send from the OAR over the existing link. At the same time the NAR can send the UPD in the network to appropriately

reconfigure the routing. This approach may be useful for very small and unintelligent devices and is supported by the generalized handoff model.

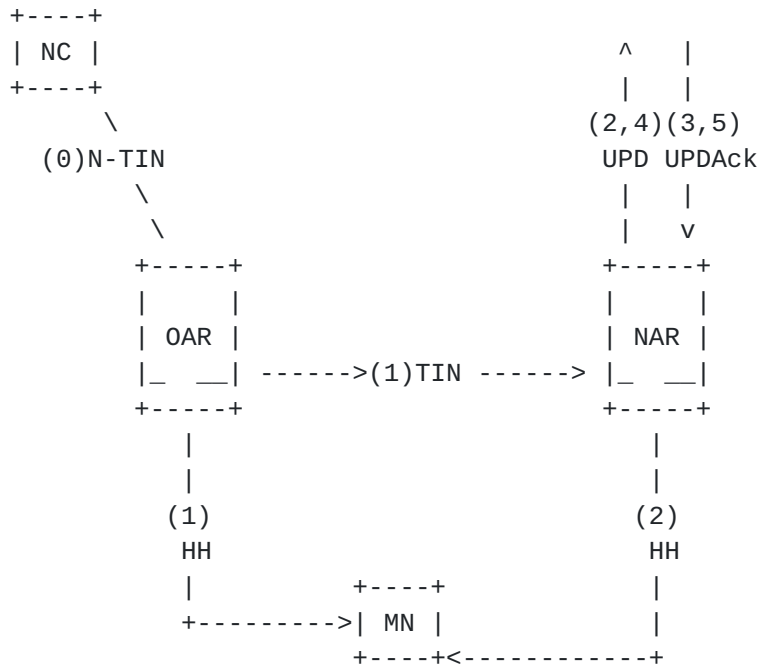


Fig4: Legacy handoff support

## 4. References

- [Anchor] G. Dommety, T. Ye, "Local and Indirect Registration for Anchoring Handoffs", [draft-dommety-mobileip-anchor-handoff-01.txt](#), July 2000.
- [EMA-MIP] A. O'Neill, S. Corson, G. Tsirtsis, "EMA Enhanced Mobile IPv6/IPv4", Internet-Draft, [draft-oneill-ema-mip-00.txt](#), July 2000.
- [CellularIPdraft] A. Campbell, J. Gomez, C-Y. Wan, Z. Turanyi and

A.Valko, "Cellular IP", Internet-Draft (work in progress), [draft-valko-cellularip-01.txt](#), October 1999.

[FastHandoffs] K. El Malki, H. Soliman, "Fast Handoffs in Mobile IPv4", [draft-elmalki-mobileip-fast-handoffs-02.txt](#), July 2000.

[HAWAIIIdraft] R. Ramjee, T. La Porta, S. Thuel, K. Varadhan and L.Salgarelli, ``IP micro-mobility support using HAWAII," Internet-Draft (work in progress), [draft-ietf-mobileip-hawaii-00.txt](#), June 1999.

[MobileIP] C.E. Perkins, ``IP Mobility Support," Internet [RFC 2002](#), Oct 1996.

[PFANE] C. Perkins, D. Johnson, "Route Optimization in Mobile IP", [draft-ietf-mobileip-optim-09.txt](#), February 2000.

[Proactive] J. Kempf, P. Calhoun, "Foreign Agent Assisted Hand-off", [draft-calhoun-mobileip-proactive-fa-00.txt](#), January 2000.

[Smooth] R. Koodli, C.E. Perkins, " A Framework for Smooth Handovers with Mobile IPv6", [draft-ietf-koodli-smoothv6-00.txt](#), July 2000.

#### Author's Addresses

Alan O'Neill  
BT  
(+44) 1473-646459  
[alan.w.oneill@bt.com](mailto:alan.w.oneill@bt.com)

M. Scott Corson  
University of Maryland  
Ansible Systems  
(+1) 301-405-6630  
[corson@isr.umd.edu](mailto:corson@isr.umd.edu)

George Tsirtsis  
BT  
(+44) 020 88260073

O'Neill, Corson, Tsirtsis

13

Internet Draft

Generalized IP Handoff

August 2000

[george.tsirtsis@bt.com](mailto:george.tsirtsis@bt.com)  
[gtsirt@hotmail.com](mailto:gtsirt@hotmail.com)





Placeholder for ISOC copyright.