                 A Survey of Mobility Support in the Internet

Abstract

   Over the last two decades, many efforts have been devoted to
   developing solutions for mobility support over the global Internet,
   resulting in a variety of proposed solutions. We conducted a
   systematic survey of the previous efforts to gain an overall
   understanding on the solution space of mobility support. This
   document reports our findings and identifies remaining issues in
   providing ubiquitous and efficient Internet mobility support on a
   global scale.

Status of This Memo

   This document is not an Internet Standards Track specification; it is
   published for informational purposes.

   This document is a product of the Internet Engineering Task Force
   (IETF). It represents the consensus of the IETF community. It has
   received public review and has been approved for publication by the
   Internet Engineering Steering Group (IESG). Not all documents
   approved by the IESG are a candidate for any level of Internet
   Standard; see Section 2 of RFC 5741.

   Information about the current status of this document, any errata,
   and how to provide feedback on it may be obtained at
   http://www.rfc-editor.org/info/rfc6301.

   to this document. Code Components extracted from this document must
   include Simplified BSD License text as described in Section 4.e of
   the Trust Legal Provisions and are provided without warranty as
   described in the Simplified BSD License.

Table of Contents

1. Introduction

   This document reports our findings from a historical survey of the
   Internet mobility research and standardization efforts since the
   early 1990s. Our survey was motivated by two factors. First,
   supporting mobility over the Internet has been an active research
   area and has produced a variety of solutions, some of which have
   become the Internet standards. Yet, new issues continue to arise and
   new solutions continue to be developed to address them, making one
   wonder how much more we have yet to discover about the problem space
   as well as the solution space. The second factor is the rapid growth
   in Internet access via mobile devices in recent years, which will
   inevitably lead to new Internet application development in the coming
   years and further underscore the importance of Internet mobility
   support. We believe that a historical review of all the proposed
   solutions on the table can help us not only identify their
   commonalities and differences but also clarify remaining issues and
   shed insight on future efforts.

   In this document, we provide an overview of mobility support
   solutions from the early results to the most recent proposals. In
   the process, we also discuss the essential components in mobility
   support and analyze the design space. Through sharing our
   understanding of the current stage of the art, we aim to initiate an
   open discussion about the general direction for future mobility
   support.

   Note that the solutions discussed in this document are proposed
   designs. They have been implemented in many cases, but only a few
   have been widely deployed in the Internet.

2. Terminology

   This document uses the following terms to refer to the entities or
   functions that are required in mobility support. Readers are
   expected to be familiar with RFC 3753 "Mobility Related Terminology"
   [RFC3753] before reading this document.

   Identifier: A stable value that can be used to identify a mobile
           node. Any unique value can be used as an Identifier as long
           as it is topologically and geographically independent, i.e.,
           remains unchanged when the mobile node roams around.

   Locator: The IP address that indicates the mobile node's current
           attachment point to the Internet. It could be the IP address
           of the mobile node itself or the IP address of the network
           entity that is currently serving the mobile node.

Mapping: In this document, mapping specifically means the mapping
        between a mobile's Identifier and its Locator.

Rendezvous Point (RP): The place where the mapping is held. Some
        other functions such as data forwarding may also be co-
        located on the rendezvous point.

Global Mobility Management: A system that keeps track of each
        mobile's reachability when the mobile is moving, either
        geographically or topologically, on a global scale.

Local Mobility Management: A system that keeps track of each
        mobile's reachability within a topologically scoped local
        domain. It keeps the mobile's local movements transparent to
        all entities that are outside of the local scope.

Operator-Controlled Mobility Management: The mobile node itself is
        unaware of mobility management. Instead, certain network
        entities, which are controlled by the network operators,
        perform all the mobility-related signaling job on behalf of
        the mobile node.

User-Controlled Mobility Management: The mobile node participates in
        the mobility management. Typically, the mobile updates its
        reachability information after it changes locations and
        refreshes its reachability at a user-defined frequency.

3. Basic Components in Mobility Support Protocols

   The basic question in Internet mobility support is how to send data
   to a moving receiver (a mobile, in short). Note that here we do not
   distinguish between mobile nodes and mobile subnets. We call the
   host that sends data to a mobile the Correspondent Node (CN). To
   send data to a moving receiver M, the CN must have means to obtain
   M's latest IP address (solution type-1) or be able to reach M using a
   piece of stable information, where "stable" means that the
   information does not change as the mobile moves (solution type-2).

   Among the existing solutions, a few fall under type-1, and most of
   them use DNS as the means to provide the CN with the mobile's most
   current IP address information. The rest of the existing solutions
   fall under type-2, which must provide the function to reach the
   mobile's dynamically changing location by using that unchanged
   Identifier of the mobile known to the CN. We can summarize all the
   mobility support solutions as essentially involving three basic
   components:

o a stable Identifier for a mobile;
o a Locator, which is usually an IP address representing the
   mobile's current location; and
o a mapping between the two.

We show in the next section that different mobility support designs
are merely different approaches to provide mapping between the
Identifiers and the mobiles' current IP addresses. In type-1
solutions, the stable Identifier of a mobile is its DNS name, the
Locator is its current IP address, and the DNS server provides the
mapping function. In type-2 solutions, because the CN must be able
to reach the mobile using the stable Identifier, the Identifier
itself is typically an IP address; either the network can dynamically
find a path to reach the mobile or the IP address leads to the "home"
of the mobile that knows the mobile's current Locator and can thus
forward the CN's packets to the mobile. All the type-2 solutions
face two common issues. One issue is how to carry out this
forwarding task, given the original packet sent by the CN has the
mobile's "home address" as the destination; the other issue is how to
avoid triangle routing between the CN, the home location, and the
mobile.

4. Existing Mobility Support Protocols

   In this section, we review the existing mobility support protocols
   roughly in the time order, with a few exceptions where we grouped
   closely related protocols together for writing clarity. We briefly
   describe each design and point out how it implements the three basic
   mobility support components defined in the last section.

   Figure 1 shows a list of mobility support protocols and the time they
   were first proposed.

```
+---------------+-----+---------------+-----+
| Protocol Name |Year | Protocol Name |Year |
+---------------+-----+---------------+-----+
|    Columbia   |1991 |     TIMIP     |2001 |
+---------------+-----+---------------+-----+
|      VIP      |1991 |    M-SCTP     |2002 |
+---------------+-----+---------------+-----+
|      LSR      |1993 |      HIP      |2003 |
+---------------+-----+---------------+-----+
|   Mobile IP   |1996 |    MOBIKE     |2003 |
+---------------+-----+---------------+-----+
|     MSM-IP    |1997 |   Connexion   |2004 |
+---------------+-----+---------------+-----+
|   Cellular IP |1998 |    ILNPv6     |2005 |
+---------------+-----+---------------+-----+
|      HMIP     |1998 |  Global HAHA  |2006 |
+---------------+-----+---------------+-----+
|      FMIP     |1998 |     PMIP      |2006 |
+---------------+-----+---------------+-----+
|     HAWAII    |1999 |     BTMM      |2007 |
+---------------+-----+---------------+-----+
|      NEMO     |2000 |    WINMO      |2008 |
+---------------+-----+---------------+-----+
|      E2E      |2000 | LISP-Mobility |2009 |
+---------------+-----+---------------+-----+
```

Figure 1

4.1. Columbia Protocol

   This protocol [Columbia] was originally designed to provide mobility
   support on a campus. A router called Mobile Support Station (MSS) is
   set up in each wireless cell and serves as the default access router
   for all mobile nodes in that cell. The Identifier for a mobile node
   is an IP address derived from a special IP prefix, and the mobile
   node uses this IP address regardless of the cell to which it belongs.

   Each MSS keeps a tracking list of mobile nodes that are currently in
   its cell by periodically broadcasting beacons. The mobile replies to
   the MSS with a message containing its stable Identifier and its
   previous MSS when it receives the beacon from a new MSS. The new MSS
   is responsible to notify the old MSS that a mobile has left its cell.
   Each MSS also knows how to reach other MSSs (e.g., all MSSs could be
   in one multicast group, or a list of IP addresses of all MSSs could
   be statically configured for each MSS).

When a CN sends a packet to a mobile node, the packet goes to the MSS
nearest to the CN (MC), which either has the mobile node in the same
cell and can deliver directly or broadcasts a query to all other MSSs
and gets a reply from the MSS (MM) with the mobile node. If it is
the latter case, MC tunnels the packet to MM, which will finally
deliver the packet to the mobile node.

Hence, in this scheme, CN uses the Identifier to reach the mobile.
It largely avoids triangle routing because the router next to CN is
mobility-aware and can intercept CN's data destined to the mobile and
forward to destination MSS. Since a mobile keeps the same IP address
independent from its movement, mobility does not affect TCP
connections.

An illustration of the Columbia Protocol is shown in Figure 2.

```
                        +---------+
                        |         |
               .------->|   MSS   |
               |        |         |
               |        +---------+
               |
               | query
               |
       +--------+     query     +--------+
       |        | -------------->|        |
       |  MSS   | <------------- |  MSS   |
       |        |     reply      |        |
       +--------+ ============>+--------+
         /\           data          ||
         ||                          ||
         ||                          \/
       +--------+             +---------+
       |        |             |         |
       |  CN    |             |  MN     |
       |        |             |         |
       +--------+             +---------+


           ===>: data packets
           --->: signaling packets


                          Figure 2
```
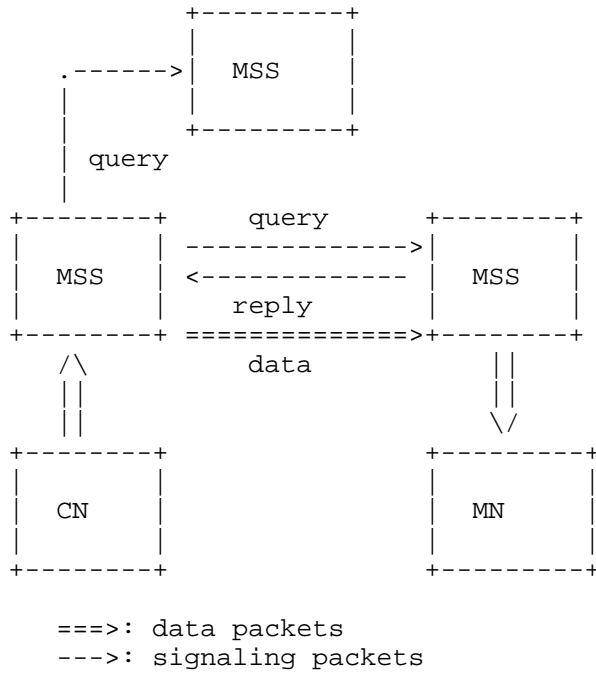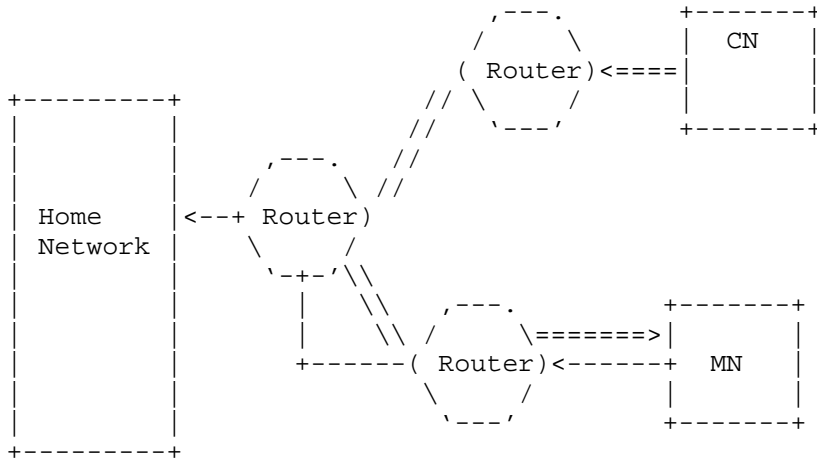
4.2. VIP

   Virtual Internet Protocol [VIP] has two basic ideas. First, a packet
   carries both Identifier and Locator; second, the Identifier is an IP
   address that leads to the home network where the mapping is kept.

   The IP header is modified to allow packets sent by a mobile to carry
   two IP addresses: a Virtual IP address (Identifier) and a regular IP
   address (Locator). Every time the mobile node changes its location,
   it notifies the home network with its latest IP address. A mobile's
   virtual address never changes and can be used to support TCP
   connections independent of mobility.

   To deliver data to a mobile, the CN first uses the mobile's Virtual
   IP address as the destination IP address, i.e., the Locator is set to
   be the same as the Identifier. As a result, the packet goes to the
   home network and the Home Agent redirects the packet to mobile's
   current location by replacing the regular IP destination address
   field with the mobile's current address.

   To reduce triangle routing, the design lets CNs and routers learn and
   cache the Identifier-Locator mapping carried in the packets from
   mobile nodes. When a CN receives a packet from the mobile, it learns
   the mobile's current location from the regular IP source address
   field. The CN keeps the mapping and uses the Locator as the
   destination in future exchanges with the mobile. Similarly, if a
   router along the data path to a mobile finds out that the mapping
   carried in the packet differs from the mapping cached by the router,
   it changes the destination IP address field to its cached value.
   This router-caching solution is expected to increase the chance that
   packets destined to the mobile get forwarded to the mobile's current
   location directly, by paying a cost of having all routers examine and
   cache all the mobile's Identifier-Locator mappings.

   Figure 3 shows how the VIP Protocol works.

```
                                  ,---.         +-------+
                                 /     \        |  CN   |
                                ( Router)<====| |       |
          +---------+           // \     /     |       |
          |         |          //   `---'      +-------+
          |         |   ,---.  //
          |         |  /     \ //
          | Home    |<--+ Router)
          | Network |   \     /
          |         |   `-+-'\\
          |         |    |   \\   ,---.         +-------+
          |         |    |    \\ /     \======>|       |
          |         |    +------( Router)<------+  MN   |
          |         |           \     /        |       |
          |         |           `---'          +-------+
          +---------+
```

```
               ===>: data packet
               --->: location update message
```

                          Figure 3

4.3. Loose Source Routing (LSR) Protocol

   In the Loose Source Routing (LSR) Protocol [LSR], each mobile has a
   designated router, called a Mobile Router, that manages its mobility.
   The Mobile Router assigns an IP address (used as an Identifier) for
   each mobile it manages and announces reachability to those IP
   addresses. Another network entity in the LSR design is Mobile Access
   Station (MAS), through which a mobile gets its connectivity to the
   Internet. The mobile node reports the IP address of its current
   serving MAS (Locator) to its Mobile Router.

   The CN uses the Identifier to reach the mobile node in the first
   place. If the CN and the mobile node are attached to the same MAS,
   the MAS simply forwards packets between the two (in this case CN is
   also mobile); otherwise, the packet from CN is routed to the Mobile
   Router of the mobile. The Mobile Router looks up the mappings to
   find the serving MAS of the mobile node and inserts the loose source
   routing (LSR) option into the IP header of the packet with the IP
   address of the MAS on it. In this way, the packet is redirected to
   the MAS, which then delivers the packet to the mobile. To this
   point, the Locator of the mobile node is already included in the LSR
   option, and the two parties can communicate directly by reversing the
   LSR option in the incoming packet. Hence, the path for the first
   packet from CN to the mobile is CN->Mobile Router->MAS->mobile node,
   and then the bidirectional path for the following packets is mobile
   node <->MAS<->CN.

Triangle routing is avoided by revealing the mobile's Locator to the
CN in the LSR option.

Figure 4 shows the basic operation of the LSR Protocol.

```
                                   +---------+
                                   |         |
               _____ |   CN    |
              |                    |         |
              |                    +---------+
              |                        /\
              V                        ||
       +-------+                       ||
       |Mobile |                       ||
       |Router |                       ||
       |       |                       ||  Reversing LSR
       +---+---+                       ||
           |                           \/
           |                    +---------+    +----------+
           |    LSR Inserted    |         |<====>|          |
           +------------------->|   MAS   |    |    MN    |
                                |         |----->|          |
                                +---------+    +----------+
```

```
              -->: first data packet
              ==>: following data packets
```
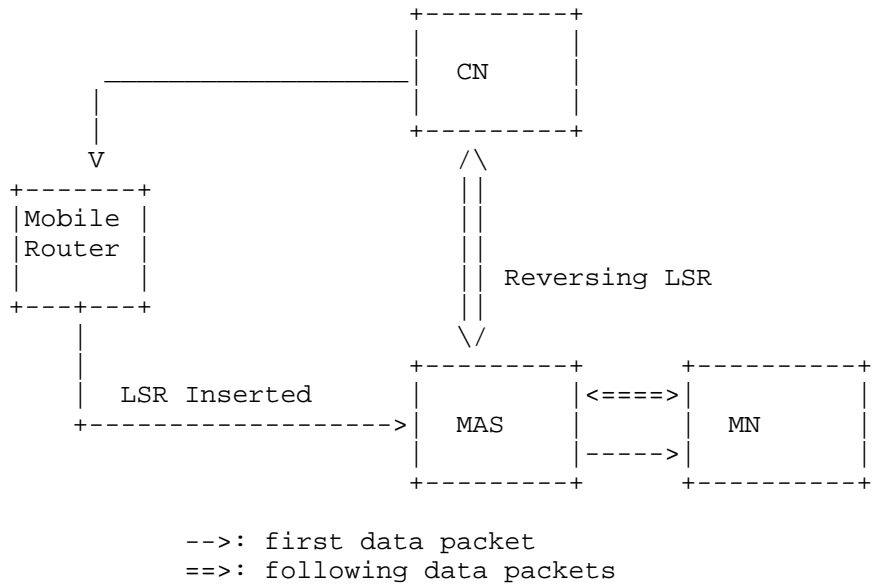
                          Figure 4

4.4. Mobile IP

   The IETF began standards development in mobility support soon after
   the above three protocols. The first version of the Mobile IP
   standard was developed in 1996. Later, the IETF developed the Mobile
   IPv4 [RFC3344] and Mobile IPv6 [RFC3775] standards in 2002 and 2004,
   respectively. In 2010, the Mobile IPv4 standard was revised
   [RFC5944]. In 2009, Dual-Stack Mobile IPv4 [RFC5454] was
   standardized to allow a dual-stack node to use IPv4 and IPv6 home
   addresses and to move between IPv4 and dual-stack network
   infrastructures.

   Although the three documents differ in details, the high-level design
   is similar. Here we use Mobile IPv6 as an example. Each mobile node
   has a Home Agent (HA), from which it acquires its Home Address (HoA),
   the Identifier. The mobile node also obtains its Locator, a Care-of
   Address (CoA), from its current access router. Whenever the mobile
   node gets a new CoA, it sends a Binding Update message to notify the

Home Agent. Conceptually, Mobile IPv6 design looks similar to the
VIP Protocol, with the mobile's HoA corresponding to the Virtual IP
Address in VIP and the CoA corresponding to the regular IP address.

The CN uses the mobile's HoA as the destination IP address when
sending data to a mobile. The packets are forwarded to the Home
Agent, which then encapsulates the packets to mobile node's CoA
according to the mapping.

To alleviate triangle routing, the CN, if it supports Route
Optimization, also keeps the mapping between the mobile's HoA and
CoA. Thus, the CN can encapsulate packets to the mobile directly,
without going through the Home Agent. Note that in this case, the
mobile needs to update its CoA to CNs as well.

Figure 5 illustrates the data path of Mobile IPv6 without Route
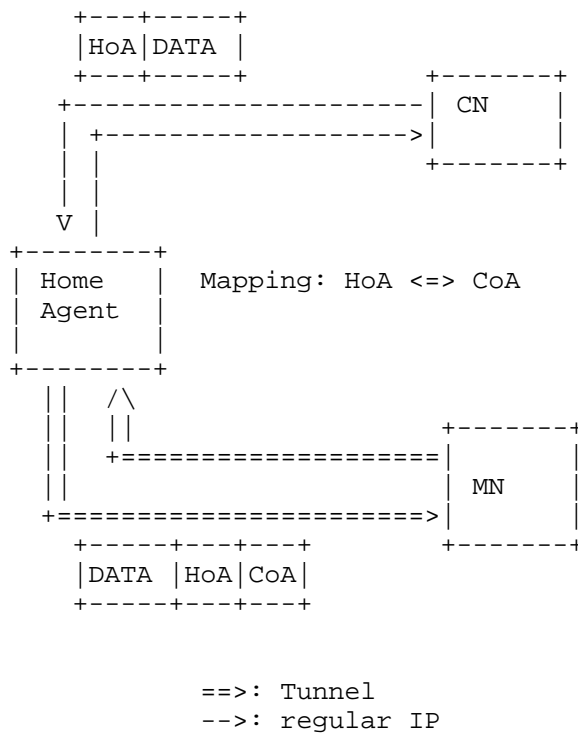Optimization.

```
                    +---+-----+
                    |HoA|DATA |
                    +---+-----+            +-------+
              +--------------------| CN     |
              | +------------------>|        |
              | |                   +-------+
              | |
              | |
              V |
        +--------+
        | Home    |    Mapping: HoA <=> CoA
        | Agent   |
        |         |
        +--------+
          ||   /\
          ||   ||                    +-------+
          ||   +==================|        |
          ||                      | MN     |
        +=====================>|        |
        +-----+---+---+            +-------+
        |DATA |HoA|CoA|
        +-----+---+---+


                ==>: Tunnel
                -->: regular IP


                Figure 5
```

4.5. HMIP

   Hierarchical Mobile IP (HMIP) [RFC5380] is a simple extension to
   Mobile IP. It aims to improves the performance of Mobile IP by
   handling mobility within a local region locally. A level of
   hierarchy is added to Mobile IP in the following way. A Mobility
   Anchor Point (MAP) is responsible for handling the movements of a
   mobile in a local region. Simply speaking, MAP is the local Home
   Agent for the mobile node. The mobile node, if it supports HMIP,
   obtains a Regional CoA (RCoA) and registers it with its Home Agent as
   its current CoA; while RCoA is the Locator for the mobile in Mobile
   IP, it is also its regional Identifier used in HMIP. At the same
   time, the mobile obtains a Local CoA (LCoA) from the subnet to which
   it attaches. When roaming within the region, a mobile only updates
   the MAP with the mapping between its RCoA and LCoA. In this way, the
   handoff performance is usually better due to the shorter round-trip
   time between the mobile and the MAP, as compared to the delay between
   the mobile and its HA. It also reduces the burden of the Home Agents
   by reducing the frequency of sending updates to Home Agents.

4.6. FMIPv6

   Fast Handover for Mobile IPv6 (FMIPv6) [RFC5568] is another extension
   to Mobile IP, which reduces the Binding Update latency as well as the
   IP connectivity latency. It is not a fully fledged mobility support
   protocol; rather, its only purpose is to optimize the performance of
   Mobile IP.

   This goal is achieved by three mechanisms. First, it enables a
   mobile node to detect that it has moved to a new subnet while it is
   still connected to the current subnet by providing the new access
   point and the corresponding subnet prefix information. Second, a
   mobile node can also formulate a prospective New Care-of Address
   (NCoA) when it is still present on the previous link so that this
   address can be used immediately after it attaches to the new subnet
   link. Third, to reduce the Binding Update interruption, FMIP
   specifies a tunnel between the Previous Care-of Address (PCoA) and
   the NCoA. The mobile node sends a Fast Binding Update to the
   previous access router (PAR) after the handoff, and PAR begins to
   tunnel packets with PCoA as the destination to NCoA. These packets
   would have been dropped if the tunnel were not established. In the
   reverse direction, the mobile node also tunnels packets to PAR until
   it finishes the Binding Update process (the mobile node can only use
   PCoA now because the binding in HA or the correspondent nodes may
   have not been updated yet).

4.7. NEMO

   It is conceivable to have a group of hosts moving together. Consider
   vehicles such as ships, trains, or airplanes that may host a network
   with multiple hosts. Because Mobile IP handles mobility per host, it
   is not efficient when handling such mobility scenarios. Network
   Mobility (NEMO) [RFC3963], as a backward-compatible extension to
   Mobile IP, was introduced in 2000 to provide efficient support for
   network mobility.

   NEMO introduces a new entity called a Mobile Router (note that this
   is different from the "Mobile Router" in the LSR Protocol). Every
   mobile network has at least one Mobile Router. A Mobile Router is
   similar to a mobile node in Mobile IP, but instead of having a single
   HoA, it has one or more IP prefixes as the Identifier. After
   establishing a bidirectional tunnel with the Home Agent, the Mobile
   Router distributes its mobile network's prefixes (namely, Mobile
   Prefixes) through the tunnel to the Home Agent. The Mobile Prefix of
   a mobile network is not leaked to its access router (i.e., the access
   router never knows that it can reach the Mobile Prefixes via the
   Mobile Router). The Home Agent in turn announces the reachability to
   the Mobile Prefix. Packets to and from the mobile network flow
   through the bidirectional tunnel between the Mobile Router and the
   Home Agent to their destinations. Note that mobility is transparent
   to the nodes in the moving network.

4.8. Mobility Support Using Multicast in IP (MSM-IP)

   MSM-IP [MSM-IP] stands for Mobility Support using Multicast in IP.
   As one can see from its name, MSM-IP leverages IP multicast routing
   for mobility support. In IP multicast, a host can join a group
   regardless of the network to which it attaches and receive packets
   sent to the group after its join. Thus, mobility is naturally
   supported in the domains where IP multicast is deployed. Note that
   MSM-IP does not address the issue of the feasibility of supporting
   mobility through IP multicast; rather, it simply shows the
   possibility of using IP multicast to provide mobility support once/if
   IP multicast is universally deployed.

   MSM-IP [MSM-IP] assigns each mobile node a unique multicast IP
   address as the Identifier. When the mobile node moves into a new
   network, it initiates a join to its own address, which makes the
   multicast router in that subnet join the multicast distribution tree.
   Whoever wants to communicate with the mobile node can just send the
   data to the mobile's multicast IP address, and the multicast routing
   will take care of the rest.

Note that, due to the nature of multicast routing, the mobile node
can have the new multicast router join the group to cache packets in
advance before it detaches the old one, resulting in smoother
handoff.

4.9. Cellular IP, HAWAII, and TIMIP

This is a group of protocols that share the common idea of setting up
a host route for each mobile in the local domain. The mobile retains
a stable IP address as long as it is within the local domain, and
this IP address is used as a regional Identifier. The gateway router
of the local domain will use this Identifier to reach the mobile
node. All three protocols are intended to work with Mobile IP as a
local mobility management protocol. By describing them together, we
can more easily show the differences by comparison.

Cellular IP [CIP] handles the local mobility in a network consisting
of Cellular IP routers. A mobile reports the IP address of the
gateway for the local network as the RCoA to its Home Agent and
retains its locally assigned IP address (the regional Identifier)
when it roams within the Cellular IP network. The routers in the
network monitor the packets originating from mobile nodes and
maintain a distributed, hop-by-hop reverse path for each mobile node.
Cellular IP utilizes the paging technique from the cellular network
to track the location of each mobile: idle mobile nodes send dummy
packets to the gateway router with a relatively low frequency to
update their reverse paths in the routers. The outdated path will
not be cleared explicitly after the mobile changes its location;
instead, it will be flushed by the routers if the paging timer
expires before the next dummy packet comes. To reduce the paging
cost, only a subset of the routers would set up a reverse path for
the idle mobile nodes.

When a packet from the CN arrives at the gateway, the gateway
initiates a controlled flooding query. If a router knows where to
forward a packet, it forwards it immediately; otherwise, it forwards
the packet to all its interfaces except the one from which the packet
came. Due to the paging technique, this will not become a broadcast.
Once the mobile receives the query, it replies with a route-update
message to the gateway, and a much more precise reverse path is then
maintained by all the routers along the data path, via which the
gateway router forwards packets from CN to the mobile. Note that the
timer value for the precise data path is much smaller than the paging
timer value, in order to avoid sending duplicate data packets to
multiple places if the mobile moves during the data communication.

Similarly, Handoff-Aware Wireless Access Internet Infrastructure
(HAWAII) [HAWAII] also aims to provide efficient local mobility
support. Unlike Cellular IP, the route between the gateway router
and the mobile is always maintained. When the mobile moves, HAWAII
dynamically modifies the route to the mobile by installing a host-
based forwarding entry on the routers located along the shortest path
between the old and new base stations of the mobile. It is possible
that a longer suboptimal routing path will be constructed (e.g.,
gateway router->old base station->new base station->mobile).
Alternatively, a new sub-path between the mobile and the cross-over
router can be established. Here, the cross-over router is the router
at the intersection of two paths, one between the gateway and the old
base station and the second between the old base station and the new
base station. In HAWAII, the mobile only periodically sends refresh
messages to the base station, and the base station along with other
routers take care of the path maintenance.

TIMIP [TIMIP], which stands for Terminal Independent Mobile IP,
integrated the design of Cellular IP and HAWAII. On one hand, it
refreshes the routing paths with dummy packets if the mobile node is
idle. On the other hand, handoff within a domain results in the
changes of routing tables in the routers. Besides, the IP layer is
coupled with layer 2 handoff mechanisms, and special nodes can work
as Mobile IP proxies for legacy mobiles that do not support Mobile
IP. Thus, as long as the mobile roams within the domain, the legacy
node has the same degree of mobility support as a Mobile-IP-capable
node.

4.10. E2E and M-SCTP

E2E (End-to-End) communication [E2E] gets its name from its end-to-
end architecture and is the first proposal that utilizes existing DNS
service to track a mobile node's current location. The stable
Identifier here is the domain name of the mobile. The mobile uses
Dynamic DNS to update its current IP address in DNS servers. To keep
the ongoing TCP connection unaffected by mobility, a TCP Migrate
option is introduced to allow both ends to replace the IP addresses
and ports in TCP 4-tuple on the fly. Thus, the CN can query DNS to
obtain the current Locator of the mobile, and after the TCP
connection is established, the mobile will be responsible for
updating its Locator for this session.

Inspired by E2E, Mobile Stream Control Transmission Protocol (M-SCTP)
[M-SCTP] was proposed in 2002. Similarly, it uses Dynamic DNS to
track the mobile nodes and allows both ends to add/delete IP
addresses used in Stream Control Transmission Protocol (SCTP)
associations during the move.

4.11. Host Identity Protocol

   The Host Identify Protocol (HIP) [RFC5201] assigns each host an
   Identifier made of cryptographic keys and adds a new Host Identity
   layer between the transport and network layers. Host Identities,
   which are essentially public keys, are used to identify the mobile
   nodes, and IP addresses are used only for routing purposes. In order
   to reuse the existing code, a Host Identity Tag (HIT), which is a
   128-bit hash value of the Host Identity, is used in transport and
   other upper-layer protocols.

   HIP can use DNS as the rendezvous point that holds the mappings
   between HITs and IP addresses. However, HIP by default uses its own
   static infrastructure Rendezvous Servers in expectation of better
   rendezvous service. Each mobile node has a designated Rendezvous
   Server (RVS), which tracks the current location of the mobile node.
   When a CN wants to communicate with mobile node, it queries DNS with
   a mobile node's HIT to obtain the IP address of the mobile node's RVS
   and sends out the first packet. After receiving this first packet,
   RVS relays it to the mobile node. The mobile node and correspondent
   node can then start communication on the direct path. If the mobile
   node moves to a new address, it notifies the CN by sending HIP UPDATE
   with LOCATOR parameter indicating its new IP address (Locator).
   Meanwhile, it also updates the mapping in RVS.

4.12. MOBIKE

   IKEv2 Mobility and Multihoming Protocol (MOBIKE) [RFC4555] is an
   extension to Internet Key Exchange (IKEv2) to support mobility and
   multihoming. The main purpose of MOBIKE is to allow roaming devices
   to keep the existing IKE and IPsec Security Associations (SAs)
   despite IP address changes. The mobility support in MOBIKE allows
   both parties to move, but it does not provide a rendezvous mechanism.
   In other words, simultaneous movement of both parties is not
   supported.

   MOBIKE allows both parties to have a set of addresses, and the party
   that initiated the IKE_SA is responsible for deciding which pair of
   addresses to use. During the communication session, if the initiator
   wishes to change the addresses due to movement, it updates the IKE_SA
   with new IP addresses and also updates the IPsec SAs associated with
   this IKE_SA. Then it sends an INFORMATIONAL request containing the
   UPDATE_SA_ADDRESSES notification to the other party. The responder
   then checks the local policy and updates the IP addresses in the
   IKE_SA with the values from the IP header. It replies to the
   initiator with an INFORMATIONAL response, initiates a return
   routability check if it wants to, and updates the IPsec SAs
   associated with this IKE_SA.

MOBIKE is not a fully fledged mobility protocol, and it does not
intend to be one. Nevertheless, through the use of IPsec tunnel
mode, MOBIKE partially supports mobility as it can dynamically update
the tunnel endpoint addresses.

4.13. Connexion and WINMO

Connexion [Boeing] was a mobility support service provided by Boeing
that uses BGP to support network mobility. Every mobile network is
assigned a /24 IP address prefix (stable Identifier), and the CN uses
this Identifier to reach the moving network, which means that the
global routing system is responsible for finding a path to the mobile
network. When an airplane moves between its access routers on the
ground, it withdraws its prefix from the previous access router and
announces the prefix via the new access point. As a result, the
location change of the plane is effectively propagated to the rest of
the world. However, if the number of moving networks becomes large,
the amount of BGP updates will also increase proportionally,
resulting in severe global routing dynamics.

WINMO [WINMO] (which stands for Wide-Area IP Network Mobility) was
introduced in 2008 to address the routing update overhead problem of
Connexion. Like Connexion, WINMO also assigns each mobile network a
stable prefix. However, through two new approaches, WINMO can reduce
the BGP updates overhead for mobile networks by orders of magnitude
lower than those of Connexion. First, WINMO uses various heuristics
to reduce the propagation scope of routing updates caused by mobile
movements. Consequently, not every router may know all the mobiles'
current locations. Handling this issue led to the second and more
fundamental approach taken by WINMO: it adopts the basic idea from
Mobile IP by assigning each mobile network a "home" in the following
way. WINMO assigns each mobile network a prefix out of a small set
of well-defined Mobile Prefixes. These Mobile Prefixes are announced
by a small set of Aggregation Routers, which also keep track of the
mobile network's current locations. Therefore, these Aggregation
Routers play a similar role to Home Agents in Mobile IP and can be
counted on as a last resort to reach mobile networks globally.

To prevent frequent Interior Border Gateway Protocol (iBGP) routing
updates due to the movement of mobile networks within an Autonomous
System (AS), WINMO also introduces a Home Agent for the Mobile
Prefixes: only a Designated BGP-speaking Router (DBR) acts as the
origin of Mobile Prefixes, and mobile networks always update the
addresses of their access routers (intra-AS Locators) with DBR, which
resembles the binding updates in Mobile IP. Thus, packets destined
to mobile networks are forwarded to DBR after they enter the border
of an AS, and DBR will tunnel them to the current locations of mobile
networks.

A new BGP community attribute, which includes the mobile network's
intra-AS Locator in each packet, is also defined to eliminate the
triangle-routing problem caused by DBR. The border routers of the AS
can tunnel packets directly to the mobile network based on the new
attribute.

4.14. ILNPv6

ILNPv6 [ILNP] stands for Identifier-Locator Network Protocol for
IPv6. The ILNPv6 packet header was deliberately made similar to the
IPv6 header. Essentially, it breaks an IPv6 address into two
components: high-order 64 bits as a Locator and low-order 64 bits as
an Identifier. The Identifier identifies a host, instead of an
interface, and is used in upper-layer protocols (e.g., TCP, FTP); on
the other hand, the Locator changes with the movement of the mobile
node, and a set of Locators can be associated with a single
Identifier. Several new DNS resource records (RRs) are required,
among which I (Identifier Record) and L (Locator Record) are most
important. As in current Internet, the CN will query the DNS about
the mobile's domain name to determine where to send the packet.
During the movement, the mobile node uses Secure Dynamic DNS update
to ensure that the Locator values stored in DNS are up to date. It
also sends Locator Update messages to the CNs that are currently
communicating with it. As an optimization, ILNPv6 supports soft-
handoff, which allows the use of multiple Locators simultaneously to
achieve smooth transition. ILNPv6 also supports mobile networks.

4.15. Global HAHA

Global Home Agent to Home Agent (HAHA) [HAHA], first proposed in 2006
as an extension to Mobile IP, aims to eliminate the triangle-routing
problem in Mobile IP and NEMO by distributing multiple Home Agents
globally. All the Home Agents join an IP anycast group and form an
overlay network. The same home prefix is announced by all the Home
Agents from different locations. Each mobile node can register with
any Home Agent that is closest to it. A Home Agent H that accepts
the binding request of a mobile node M becomes the primary Home Agent
for M and notifies all other Home Agents of the binding [M, H] so
that the binding information databases for all the mobiles in all
Home Agents are always synchronized. When a mobile moves, it may
switch its primary Home Agent to another one that becomes closest to
the mobile.

A correspondent node sends packets to a mobile's Home Address.
Because of anycast routing, the packets are delivered to the nearest
Home Agent. This Home Agent then encapsulates the packets to the IP
address of the primary Home Agent that is currently serving the
mobile node, which will finally deliver the packets to the mobile

node after striping off the encapsulation headers. In the reverse
direction, this approach works exactly the same as Mobile IP. If the
Home Agents are distributed widely, the triangle-routing problem is
naturally alleviated without Route Optimization.

The data flow in Global HAHA is shown in Figure 6.

```
        +------+             +------+      +-----+
        | HA   |-------------| HA   |      |     |
        |      |             |      |      | CN  |
        +--+---+     +------+++---+      +-----+
           |             |   ||              /\
           |             |   ||              ||
           |             |   ||              ||
           |             |   ||              ||
           |             |   ||              ||
        +--+---+------+   ||              ||
        |      |<=============+              ||
        | HA   |=============================+
        +-++---+
          || /\
          \/ ||
        +---++-+        ===>: data flow
        |      |        ----: HA overlay network
        | MN   |
        +------+
```

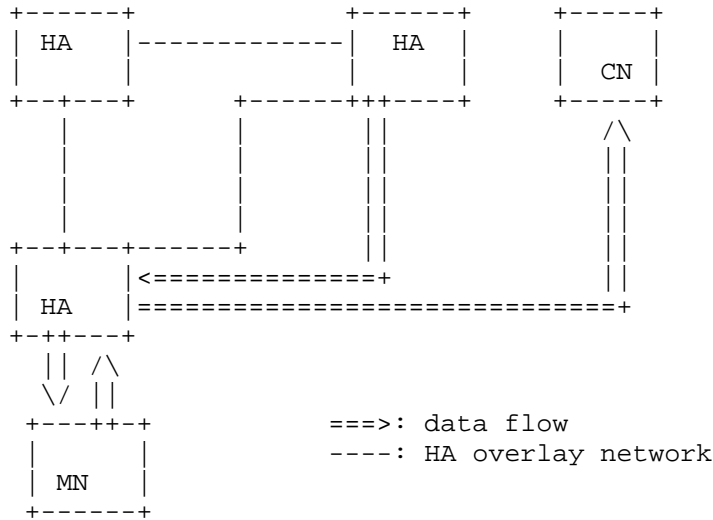   Figure 6

4.16. Proxy Mobile IP

   Proxy Mobile IP (PMIP) [RFC5213] was proposed in 2006 to meet the
   interest of mobile network operators who desire to support mobility
   in a network rather than on mobile devices and to have tighter
   control on mobility support. Mobility is completely transparent to
   the mobile devices and is provided to legacy IP devices. PMIP
   introduces two new types of network nodes, Local Mobility Anchor
   (LMA) and Mobile Access Gateway (MAG), which together can support
   mobility within an operator's network without any action taken by the
   mobile node. LMA serves as a local Home Agent and assigns a local
   Home Network Prefix for each mobile node. This prefix is the
   Identifier for the mobile node within the PMIP domain. MAGs monitor
   the attaching and detaching events of the mobile node and generate
   Proxy Binding Update to LMA on behalf of the mobile node during
   handoff. After the success of binding, LMA updates the mobile node's
   Proxy-CoA (Locator in PMIP domain) with the IP address of the MAG
   that is currently serving the mobile node. The MAG then emulates the
   mobile node's local Home Link by advertising the mobile node's local
   Home Network Prefix in Router Advertisement. When roaming in the

PMIP domain, the mobile node always obtains its local Home Prefix and
believes that it is on a local Home Link. Within the domain, the
mobile node is reached by the Identifier, and LMA tunnels packets to
the mobile node according to the mapping.

4.17. Back to My Mac

Back to My Mac (BTMM) [RFC6281] is an engineering approach to
mobility support and has been deployed since 2007 with Mac OS Leopard
release. Each user gets a MobileMe account (which includes BTMM
service), and Apple, Inc. provides DNS service for all BTMM users.
The reachability information of the user's machine is published in
DNS.

A mobile uses secure DNS update to dynamically refresh its current
location. Each host generates an IPv6 Unique Local Address (ULA)
[RFC4193] at boot time, which is stored in the DNS database as its
topologically independent Identifier. The host's current IPv4
address (which is the IPv4 address of the NAT box if the host is
behind a NAT) is stored in a SRV resource record [RFC2782] together
with a transport port number needed for NAT traversal. Every node
establishes a long-lived query (LLQ) session with the DNS server so
that the DNS server can immediately notify each node when the answer
to its query has changed. A host uses its Identifier in transport
protocols and applications and uses UDP/IPv4 encapsulation to deliver
data packets using information learned from the SRV RR. Note that
the Locator here is the IPv4 address plus the transport port number
and that the IPv6 address is only for identification purposes. In
fact, it could be any form of Identifier (e.g., domain name); BTMM
chose to use IPv6 addresses so that its implementation can reuse
existing code.

BTMM is currently used by millions of subscribers. It is simple and
easy to deploy. However, the current applications use BTMM service
in a "stop-and-reconnect" fashion. It remains to be seen how well
BTMM can support continuous communications while hosts are on the
move, for example, as needed for voice calls.

Figure 7 shows the basic architecture of BTMM.

```
   DDNS update    +--------+  DDNS update
   +--------------->|        |<-------+
   |                |  DNS   |        |
   |      LLQ       |        |  LLQ   |
   |      +--------->|        |<----+  |
   |      |         |        |     |  |
   |      |         +--------+     |  |
   |      |                        |  |                 +---------+
   |      V               +---+--+----+                 |         |
   |  ++-------+          |          |       +-------|   |         |
   |  |Endhost1|  Tunnel  |   NAT    |       +------>|Endhost2 |
   |  |        |<=====================================>|         |
   |  +--------+          |          |       |      |   |         |
   |                      +----------+       |      |   |         |
                                     +----------+   +---------+
```

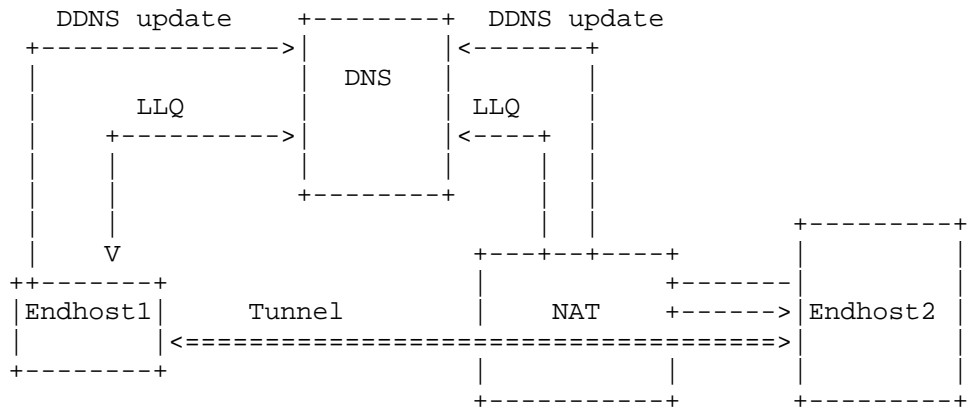   Figure 7

4.18. LISP-Mobility

   LISP-Mobility [LISP-Mobility] is a relatively new design. Its
   designers hope to utilize functions and services provided by
   Locator/ID Separation Protocol (LISP) [LISP], which is designed for
   Internet routing scalability, to support mobility as well.
   Conceptually, LISP-Mobility may seem similar to some protocols we
   have mentioned so far, such as ILNPv6 and Mobile IP. Lightweight
   Ingress Tunnel Router and Egress Tunnel Router functions are
   implemented on each mobile node, and all the packets to and from the
   mobile node are processed by the two router functions (so the mobile
   node looks like a LISP site). Each mobile node is assigned a static
   Endpoint ID, as well as a preconfigured Map-Server. When a mobile
   node roams into a network and obtains a new Routing Locator, it
   updates its Routing Locator set in the Map-Server, and it also clears
   the cached Routing Locator in the Ingress Tunnel Routers or Proxy
   Tunnel Routers of the CNs. Thus, the CN can always learn the up-to-
   date location of the mobile node by the resolution of the mobile
   node's Endpoint ID, either issued by itself or issued after receiving
   the notification from the mobile node about the staled cache. The
   data would always travel through the shortest path. Note that both
   Endpoint IDs and Routing Locators are essentially IP addresses.

5. Different Directions towards Mobility Support

   After studying various existing protocols, we identified several
   different directions for mobility support.

5.1. Routing-Based Approach versus Mapping-Based Approach

   All existing mobility support designs can be broadly classified into
   two basic approaches. The first one is to support mobility through
   dynamic routing. In such designs, a mobile keeps its IP address
   regardless of its location changes; thus, the IP address can be used
   both to identify the mobile and to deliver packets to it. As a
   result, these designs do not need an explicit mapping function.
   Rather, the routing system must continuously keep track of a mobile's
   movements and reflect its current position in the network on the
   routing table so that at any given moment packets carrying the
   (stable) receiver's IP address can be delivered to the right place.

   It is also worthwhile to identify two sub-classes in routing-based
   approaches. One is broadcast based, and the other is path based. In
   the former case, either the mobile's location information is actively
   broadcasted to the whole network or a proactive broadcast query is
   needed to obtain the location information of a mobile (e.g., Columbia
   and Connexion); in the latter case, on the other hand, a host-based
   path is maintained by the routing system instead (e.g., Cellular IP,
   HAWAII, and TIMIP).

   Supporting mobility through dynamic routing is conceptually simple;
   it can also provide robust and efficient data delivery, assuming that
   the routing system can keep up with the mobile movements. However,
   because either the whole network must be informed of every movement
   by every mobile or a host-based path must be maintained for every
   mobile host, this approach is feasible only in small-scale networks
   with a small number of mobiles; it does not scale well in large
   networks or for a large number of mobiles.

   The second approach to mobility support is to provide a mapping
   between a mobile's stable Identifier and its dynamically changing IP
   address. Instead of notifying the world on every movement, a mobile
   only needs to update a single binding location about its location
   changes. In this approach, if one level of indirection at IP layer
   is used, as in the case of Mobile IP, it has a potential side effect
   of introducing triangle routing; otherwise, if the two end nodes are
   aware of each other's movement, it means that both ends have to
   support the same mobility protocol.

   Yet, there is the third case in which the protocols combine the above
   approaches in the hope of keeping the pros and eliminating some cons
   of the two. WINMO is a typical protocol in this case.

   In Figure 8, we show the classification of the existing protocols
   according to the above analysis.

```
+--------------+-----------------------------------------+
|              | VIP, LSR, Mobile IP, HMIP, NEMO, E2E,    |
| Mapping-based| M-SCTP, ILNPv6, HIP, FMIP, PMIP,         |
|              | BTMM, GLOBAL HAHA, LISP-Mobility         |
+--------------+-----------------------------------------+
|              | Columbia, Connexion                     |
| Routing-based+-----------------------------------------+
|              | Cellular IP, HAWAII, TIMIP, MSM-IP      |
+--------------+-----------------------------------------+
| Combination  | WINMO                                   |
+--------------+-----------------------------------------+
```

Figure 8

5.2. Mobility-Aware Entities

Among the various design choices, a critical one is how many entities
are assumed to be mobility-aware. There are four parties that may be
involved during a conversation with a mobile: the mobile itself, CN,
the network, and the Home Agent or its equivalent (additional
component to the existing IP network that holds the mapping). We
mainly focus our discussion on the mapping-based approach here.

The first design choice is to hide the mobility from the CN, based on
the assumption that the CN may be the legacy node that does not
support mobility. In this approach, the IP address that is used as
the mobile's Identifier points to the Home Agent or its equivalent
that keeps track of the mobile's current location. If a
correspondent node wants to send packets to a mobile node, it sets in
the destination field of IP header an IP address, which is a mobile's
Identifier. The packets will be delivered to the location where the
mapping information of the mobile is kept, and later they will be
forwarded to the mobile's current location via either encapsulation
or destination address translation. Mobile IP and most of its
extensions, as well as several other protocols fall into this design.

The second design choice is to hide the mobility from the mobile and
CN, which is based on a more conservative assumption that both the
mobile and the CN do not support mobility. Protocols like PMIP and
TIMIP adopt this design. The protocol operations in this design
resemble those in the first category, but a significant difference is
that here the mobility-related signaling (e.g., the update Locator to
the Home Agent) is handled by the entities in the network rather than
by the mobile itself. Hence, the mobile blissfully assumes that it
is always in the same subnet.

The third one is to let both the mobile and the CN be mobility-aware.
As a result, the network is not aware of the mobility, and no
additional component is required. As an increasing number of mobile
devices are connected to the Internet (Why hide mobility from them?),
this design choice seems to be more and more appealing. One common
approach taken by this design is to use DNS to keep track of mobiles'
current locations. Mobiles use dynamic DNS updates to keep their DNS
servers updated with their current locations. This approach
re-utilizes the DNS infrastructure, which is ubiquitous and quite
reliable, and makes the mobility support protocol simple and easy to
deploy. Protocols like E2E, ILNP, and BTMM fall into this design.
Although HIP adds special-purpose rendezvous servers to the network
to replace the role of DNS, both mobile and CN are still mobility-
aware; hence, it is also classified in this category.

Figure 9 shows the three categories of protocols.

```
+-------------+--------------------------------+
| Design 1    | VIP, LSR, Mobile IP, HMIP, NEMO, |
|             | Global HAHA                    |
+-------------+--------------------------------+
| Design 2    | PMIP, TIMIP                     |
+-------------+--------------------------------+
| Design 3    | E2E, M-SCTP, ILNPv6, HIP,       |
|             | BTMM, LISP-Mobility            |
+-------------+--------------------------------+
```

                          Figure 9

5.3. Operator-Controlled Approach versus User-Controlled Approach

   At the time of this writing, cellular networks are providing the
   largest operational global mobility support, using a service model
   that bundles together device control, network access control, and
   mobility support. The tremendous success of the cellular market
   speaks loudly that the current cellular service model is a viable one
   and is likely to continue into the foreseeable future. Consequently,
   there is a strong advocate in the IETF that we continue the cellular
   way of handling mobility, i.e., instead of letting mobile devices
   participate in the mobility-related signaling themselves, the network
   entities deployed by the operators should take care of any and all
   signaling processes of mobility support. A typical example along
   this direction is Proxy Mobile IP, in which LMA works together with
   MAGs to assure the reachability to the mobile using its Home
   Prefixes, as long as the mobile roams within the same provider's
   domain.

One main reason for this approach is perhaps backward compatibility.
By not requiring the participation of mobiles in the control-
signaling process, it avoids any changes to the mobile nodes so that
the mobile nodes can stay simple and all the legacy nodes can obtain
the same level of mobility services as the latest mobile devices.
According to the claim of 3G vendors and operators, transparent
mobility support is a key aspect for success as they learn from their
deployment experience.

On the other hand, most of the mobility support protocols surveyed in
this document focus on mobility support only, assuming mobiles
already obtained network access. Mobile nodes typically update their
locations themselves to the rendezvous points chosen by the users,
and, of course, only the nodes implementing one of these solutions
can benefit from mobility support. However, this class of protocols
does offer users and mobile devices more flexibility and freedom,
e.g., they can choose whatever mobility services are available as
long as their software supports that protocol, and they can also tune
the parameters to get the services that are most suitable to them.

5.4. Local and Global-Scale Mobility

The work done on mobility management can also be divided into two
categories according to scale: local mobility management and global
mobility management.

Global mobility management is typically supposed to support mobility
of an unlimited number of nodes in a geographically as well as
topologically large area. Consequentially, it pays a lot of
attention to scalability issues. For the availability concern, it
also tries to avoid failure of single point.

Local mobility management, on the other hand, is designed to work
together with global mobility management and thus focuses more on
performance issues, such as handoff delay, handoff loss, local data
path, etc. Since it is typically used on a small scale with a not-
so-large number of mobile nodes, sometimes the designers can use some
fine-tuned mechanisms that are not scaled with a large network (such
as host route) to improvement performance. As a side effect of local
mobility management, the number of location updates sent by mobile
nodes to their global rendezvous points is substantially reduced.
Thus, the existence of local mobility management also contributes to
the scalability of global mobility management.

One problem of local mobility management is that it often requires
infrastructure support, such as MAGs in PMIP or MAPs in HMIP. These
kinds of local devices are essentially required in all small domains,
which can be a huge investment.

Nevertheless, mobility management in two scales make it possible for
designers to design protocols that fit into specific user
requirements; it also enables the gradual deployment of local
enhancement while not losing the ability of global roaming. The
coexistence of the two seems to be a right choice in the foreseeable
future.

Figure 10 shows the classification of the studied protocols according
to their serving scale.

```
+-----------+---------------------------------------+
|           | VIP, LSR, Mobile IP, NEMO, E2E, M-SCTP, |
|   Global  | HIP, ILNPv6, Connexion, WIMO, BTMM,    |
|           | MSM-IP, Global HAHA, LISP-Mobility     |
+-----------+---------------------------------------+
|   Local   | Columbia, HMIP, FMIP, Cellular IP,     |
|           | HAWAII, TIMIP, PMIP                    |
+-----------+---------------------------------------+
```

                             Figure 10

5.5. Other Mobility Support Efforts

   Despite the wide spectrum of mobility solutions covered by this
   survey, the list of mobility protocols is not exhaustive.

   The General Packet Radio Service (GPRS) Tunneling Protocol [GTP] is a
   network-based mobility support solution widely used in cellular
   networks. Its implementation only involves Gateway GPRS Support Node
   (GGSN) and Serving GPRS Support Node (SGSN). It allows end users of
   a Global System for Mobile Communications (GSM) or Universal Mobile
   Telecommunications System (UMTS) network to move from place to place
   while remaining connected to the Internet as if from on location at
   the GGSN. It does this by carrying the subscriber's data from the
   subscriber's current SGSN to the GGSN that is handling the
   subscriber's session. To some extent, it is the non-IETF variant of
   PMIP, with GGSN resembling LMA and SGSN resembling MAG, respectively.

   There is also work on application-layer mobility support, most
   notably SIP-based mobility support [ALM-SIP]. SIP was initially
   designed as an application signaling protocol for multimedia, and
   later researchers noticed its potential capability for mobility
   support. When the mobile initiates a session with CN, normal SIP-
   signaling procedure is performed to establish the session. When the
   mobile moves to a new network while the session is ongoing, it send a
   RE-INVITE message with the existing session but reveals the new IP
   address to the CN. The home SIP server is also updated with the

latest location information of the mobile after the move. However,
the SIP-based approach cannot maintain TCP connections when the
mobile's IP address changes.

A lot of enhancements to Mobile IPv6 Route Optimization have also
been developed. A comprehensive taxonomy and analysis of these
efforts can be found in [RFC4651].

6. Discussions

   In the last section, we discussed the different directions towards
   mobility support. We now turn our attention to identify both new
   opportunities and remaining open issues in providing global-scale
   mobility support for an unlimited number of online mobility devices.
   We are not trying to identify the solutions to these issues, but
   rather, the goal is to share our opinions and to initiate an open
   discussion.

6.1. Deployment Issues

   Among the various protocols we discussed in this document, few have
   been deployed in commercial networks. There are several reasons to
   explain this situation.

   First, although the research community started to develop mobility
   support protocols 20 years ago, only in recent years has the number
   of mobiles soared. Hence, operators did not see the incentive of
   deploying mobility support protocols several years back. As of
   today, the number of mobiles is still growing by leaps and bounds,
   and there is enough user demand for the operators to seriously
   consider the deployment of mobility support protocols.

   Second, the complexity of most mobility support protocols impedes the
   implementation and hence the deployment in commercial networks. The
   complexity arises from multiple aspects. One is the optimizations on
   performance. The other is the problem with the use of security
   protocols such as IPsec and IKE. The discussions regarding to these
   two problems are still ongoing in the MEXT Working Group. Some
   researchers argue that the research community should design a "barely
   work" version of a mobility support protocol first, without
   considering nice performance features and complex security
   mechanisms, roll it out in the real world, and improve it thereafter.
   However, there are different views on what the essential features are
   and which security mechanisms are better.

   Third, almost all the mobility support protocols assume that the
   mobile nodes have network connectivity anywhere, anytime. In
   reality, however, this is not always the case. Nevertheless,

wireless access is available in more and more places, and it is
foreseeable that in the near future, the coverage of wireless access
in different forms (WiFi, Wimax, 3G/4G) will be ubiquitous.

6.2. Session Continuity and Simultaneous Movements

In order for users to benefit from mobility support, it is important
to keep the TCP sessions uninterrupted by the mobility. If the
durations of the sessions are short (e.g., web browsing), the
probability is high that the TCP sessions finish before the handover
happens; even if the TCP session is interrupted by the handover, the
cost is usually low (e.g., refresh the web page). However, if the
TCP sessions are typically long (e.g., downloading large files and
voice calls), the interruptions during the handover would become
unacceptable.

It is hard to predict tomorrow's applications, but most mobility
support protocols try to keep the sessions up during movements. For
routing-based protocols, session continuity is not a problem since
the IP address of the mobile never changes. For other protocols,
either a stable IP address (e.g., HoA) or an equivalent (e.g., HIT)
is used in the transport layer so that the mobility is hidden, or TCP
is modified so that both ends can change IP addresses while keeping
the established session (e.g., E2E).

Another concern is the support of simultaneous movements. In some
scenarios, only one end is mobile, and the other end is always
static; moreover, the communication between the two is always
initiated by the mobile end. A lot of applications as of today fall
into this category. Typically, the server side is static, and the
client is mobile; usually, the client would contact the server first.
Hence, in these scenarios, the support of simultaneous movements is
not a requirement. However, in other scenarios, both ends may be
moving at the same time. For example, during a voice call, two
mobile nodes may experience handovers simultaneously. In this case,
a rendezvous point is necessary to keep the current locations of the
mobiles so that they can find each other after a simultaneous
movement. Besides, if a static server wants to push information to a
mobile client, a rendezvous point is also required.

It is clear that the number of mobile devices is rapidly growing, and
more mobiles are going to provide content in the near future. Hence,
the simultaneous-movements scenarios are considered important. In
fact, almost all the mobility support protocols are equipped with
rendezvous points, either by adding dedicated components or by
leveraging the existing DNS systems.

6.3. Trade-Offs of Design Choices on Mobility Awareness

   The mobility awareness at two communicating ends is closely related
   to the backward-compatibility problem. The Internet has been running
   for more than two decades, and the scale of the Internet gets so
   large that it is impossible to upgrade the whole system overnight.
   As a result, it is also not possible for a mobility support system
   designer to overlook this problem: how does one decide the mobility
   awareness in the protocol design, and how important is backward
   compatibility?

   In the following text, we discuss the trade-offs of the design
   choices mentioned in Section 5.2.

   The advantage of the first design choice is that the mobile does not
   lose the ability to communicate with legacy nodes while roaming
   around, i.e., the mobile can benefit from unilateral deployment of
   mobility support. Another potential advantage is that the static
   nodes do not need to be bothered by the mobility of the mobiles,
   which saves resources and could be desirable if the CN is a busy
   server. The disadvantage of this design is also well known: it
   introduces triangle routing, which significantly increases the delays
   in the worst cases. There are means to remedy the problem, e.g.,
   Route Optimization in Mobile IP if a CN is mobility-capable and
   distribution of Home Agents as Global HAHA does, at the expense of
   increasing complexity.

   The second design caters to the inertness of the Internet (and the
   users) by keeping everything status quo from the user's point of
   view. It is like the cellular network, with the smart network and
   dumb terminals. The advantage is that the legacy nodes can benefit
   from the mobility support without upgrade. However, the cost is also
   not trivial: the users lose the freedom of control in terms of
   mobility management, and a large number of entities in the network
   need to be upgraded.

   The third design assumes that the other end is likely also mobility-
   capable (as of today, more people are accessing the Internet via
   mobile devices than a desktop) and thus does not provide backward
   compatibility at all; however, as a trade-off, the system design
   becomes much simpler, and the data path is always the shortest one.

   We all know that backward compatibility is important in system
   design. But how important is it? How much effort should we make for
   this issue? At least for now, the answer is not yet clear.

6.4. Interconnecting Heterogeneous Mobility Support Systems

   As our survey suggests, multiple solutions for mobility support are
   already exist, and it is almost for sure that the mobility support
   systems in the future are going to be heterogeneous. However, as of
   today, the interoperation between different protocols is still
   problematic. For example, when a mobile node supporting Mobile IP
   only wants to communicate with another mobile with only HIP support,
   neither of them can benefit from mobility support.

   This situation reminds us the days before IP was adopted. In that
   time, the hosts in different networks were not able to communicate
   with each other. IP merged the networks and created the Internet,
   where each host can freely communicate with any other host. Is it
   necessary to introduce something like IP to mobility support in the
   future? Is it possible to design an architecture so that it glues
   all the mobility support systems together? We believe the answers to
   both of these questions are "yes".

   The basic idea for the solution is simple. As the famous quote says,
   "Every problem in Computer Science can be solved by adding a level of
   indirection". However, the devil is in the details, and we still
   need to figure that out.

7. Security Considerations

   Since mobility means that the location of a mobile may change at any
   time, how to secure such dynamic location updates is a very important
   consideration for all mobility support solutions. However, this
   document examines a wide range of solution proposals, so the security
   aspects also vary greatly. For example, Home-Agent-based solutions
   call for secure communications between the mobile and its Home
   Agent(s). On the other hand, for routing-based solutions, such as
   Connexion, the issue becomes one of global-routing security.
   Similarly, for those solutions that use DNS to provide mapping
   between Identifiers and Locators, the issue is essentially converted
   to how to secure DNS dynamic updates as well as queries. To keep
   this survey document both comprehensive as well as reasonably sized,
   we chose to focus the survey on describing and comparing the
   solutions to the center piece of all mobility supports -- the
   resolution between Identifiers and Locators.

8. Informative References

   [ALM-SIP] Schulzrinne, H. and E. Wedlund, "Application-Layer
              Mobility Using SIP", Mobile Computing and Communications
              Review, 2010.

   [Boeing]   Andrew, L., "A Border Gateway Protocol 4 (BGP-4)",
                Boeing White Paper, 2006.

   [CIP]   Valko, A., "Cellular IP: A New Approach to Internet Host
                Mobility", ACM SIGCOMM, 1999.

   [Columbia]   Ioannidis, J., Duchamp, D., and G. Maguire, "IP-based
                Protocols for Mobile Internetworking", ACM SIGCOMM CCR,
                1991.

   [E2E]   Snoeren, A. and H. Balakrishnan, "An End-to-End Approach
                to Host Mobility", ACM Mobicom, 2000.

   [GTP]   "GPRS Tunneling Protocol Across Gn and Gp Interface", 3G
                TS 29.060 v3.5.0.

   [HAHA]   Wakikawa, R., Valadon, G., and J. Murai, "Migrating Home
                Agents Towards Internet-scale Mobility Deployment",
                ACM CoNEXT, 2006.

   [HAWAII]   Ramjee, R., Varadhan, K., and L. Salgarelli, "HAWAII: A
                Domain-based Approach for Supporting Mobility in Wide-area
                Wireless Networks", IEEE/ACM Transcations on Networking,
                2002.

   [ILNP]   Atkinson, R., Bhatti, S., and S. Hailes, "A Proposal for
                Unifying Mobility with Multi-Homing, NAT, and Security",
                MobiWAC 2007.

   [LISP]   Farinacci, D., Fuller, V., Meyer, D., and D. Lewis,
                "Locator/ID Separation Protocol (LISP)", Work in Progress,
                March 2009.

   [LISP-Mobility]
                Farinacci, D., Lewis, D., Meyer, D., and C. White, "LISP
                Mobile Node", Work in Progress, May 2011.

   [LSR]   Bhagwat, P. and C. Perkins, "A Mobile Networking System
                Based on Internet Protocol (IP)", Mobile and Location-
                Independent Computing Symposium, 1993.

   [M-SCTP]   Xing, W., Karl, H., and A. Wolisz, "M-SCTP: Design and
                Prototypical Implementation of An End-to-End Mobility
                Concept", 5th Intl. Workshop on the Internet Challenge,
                2002.

   [MSM-IP]   Mysore, J. and V. Bharghavan, "A New Multicast-based
               Architecture for Internet Host Mobility", ACM Mobicom,
               1997.

   [RFC2782]   Gulbrandsen, A., Vixie, P., and L. Esibov, "A DNS RR for
               specifying the location of services (DNS SRV)", RFC 2782,
               February 2000.

   [RFC3344]   Perkins, C., "IP Mobility Support for IPv4", RFC 3344,
               August 2002.

   [RFC3753]   Manner, J. and M. Kojo, "Mobility Related Terminology",
               RFC 3753, June 2004.

   [RFC3775]   Johnson, D., Perkins, C., and J. Arkko, "Mobility Support
               in IPv6", RFC 3775, June 2004.

   [RFC3963]   Devarapalli, V., Wakikawa, R., Petrescu, A., and P.
               Thubert, "Network Mobility (NEMO) Basic Support Protocol",
               RFC 3963, January 2005.

   [RFC4193]   Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast
               Addresses", RFC 4193, October 2005.

   [RFC4555]   Eronen, P., "IKEv2 Mobility and Multihoming Protocol
               (MOBIKE)", RFC 4555, June 2006.

   [RFC4651]   Vogt, C. and J. Arkko, "A Taxonomy and Analysis of
               Enhancements to Mobile IPv6 Route Optimization", RFC 4651,
               February 2007.

   [RFC5201]   Moskowitz, R., Nikander, P., Jokela, P., and T. Henderson,
               "Host Identity Protocol", RFC 5201, April 2008.

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

   [RFC5380]   Soliman, H., Castelluccia, C., ElMalki, K., and L.
               Bellier, "Hierarchical Mobile IPv6 (HMIPv6) Mobility
               Management", RFC 5380, October 2008.

   [RFC5454]   Tsirtsis, G., Park, V., and H. Soliman, "Dual-Stack Mobile
               IPv4", RFC 5454, March 2009.

   [RFC5568]   Koodli, R., "Mobile IPv6 Fast Handovers", RFC 5568,
               July 2009.

   [RFC5944] Perkins, C., "IP Mobility Support for IPv4, Revised",
              RFC 5944, November 2010.

   [RFC6281] Cheshire, S., Zhu, Z., Wakikawa, R., and L. Zhang,
              "Understanding Apple's Back to My Mac (BTMM) Service", RFC
              6281, June 2011.

   [TIMIP]   Grilo, A., Estrela, P., and M. Nunes, "Terminal
              Independent Mobility For IP (TIMIP)", IEEE Communications
              Magazine, 2001.

   [VIP]   Teraoka, F., Yokote, Y., and M. Tokoro, "A Network
              Architecture Providing Host Migration Transparency",
              ACM SIGCOMM CCR, 1991.

   [WINMO]   Hu, X., Li, L., Mao, Z., and Y. Yang, "Wide-Area IP
              Network Mobility", IEEE INFOCOM, 2008.

Authors' Addresses

   Zhenkai Zhu
   UCLA
   4805 Boelter Hall, UCLA
   Los Angeles, CA 90095
   USA

   Phone: +1 310 993 7128
   EMail: zhenkai@cs.ucla.edu

   Ryuji Wakikawa
   Toyota ITC
   465 Bernardo Avenue
   Mountain View, CA 94043
   USA

   EMail: ryuji.wakikawa@gmail.com

   Lixia Zhang
   UCLA
   3713 Boelter Hall, UCLA
   Los Angeles, CA 90095
   USA

   Phone: +1 310 825 2695
   EMail: lixia@cs.ucla.edu