Virtual ID: ID/locator split in a mobile IP environment for mobility, multihoming and location privacy for the next generation wireless networks

Chakchai So-In

Department of Computer Science,
Faculty of Science,
Khon Kaen University, Thailand
E-mail: chakso@kku.ac.th

Raj Jain*, Subharthi Paul and Jianli Pan

Department of Computer Science and Engineering,
Washington University in St. Louis,
St. Louis, MO 63143, USA
E-mail: jain@cse.wustl.edu
E-mail: pauls@cse.wustl.edu
E-mail: jp10@cse.wustl.edu
*Corresponding author

Abstract: Current networking protocols designed around single interface stationary end-systems clearly fail to represent the present communication context of mobile, multi-interface end-systems. Also, the convergence of wired and wireless technologies into an all-IP (AIP) next generation network will make available multiple diversified contexts that can be leveraged for better fault tolerance, higher availability and improved performance of end-to-end communications. A major reason for the weakness of current protocols may be attributed to the contextual overloading of IP addresses to serve as identifiers as well as locators. Mobile IP (MIP) and its extensions are well-known proposals developed primarily to alleviate the mobility and deployability problems. In this paper, we develop the concept of a virtual identity (ID), as an explicit ID/locator extension of mobile IPv6 and explore its applicability to address the issues of mobility, multihoming, and location privacy in the context of next generation wireless networks.

Keywords: mobile IP; virtual identity; mobility; multihoming; location privacy; ID/locator split; multi-interface selection; next generation wireless networks; NGWNs; future cellular networks; future wireless internet.

Reference to this paper should be made as follows: So-In, C., Jain, R., Paul, S. and Pan, J. (2010) 'Virtual ID: ID/locator split in a mobile IP environment for mobility, multihoming and location privacy for the next generation wireless networks', *Int. J. Internet Protocol Technology*, Vol. 5, No. 3, pp.142–153.

Biographical notes: Chakchai So-In received his BEng and MEng from Kasetsart University, Thailand in 1999 and 2001. He also received his MS and PhD from Washington University in St. Louis in 2006 and 2010. All are in Computer Engineering. In 2003, he was an Intern in a CNAP at NTU and obtained CCNP and CCDP certifications. He was Interns at Cisco Systems, WiMAX Forums and Bell Labs during summer 2006, 2008 and 2010, respectively. His research interests include architectures for future wireless networks; congestion controls; protocols to support network and transport mobility, multihoming, and privacy; and quality of service in broadband wireless access networks.

Raj Jain is a Fellow of IEEE, a Fellow of ACM, a winner of ACM SIGCOMM Test of Time award and ranks among the top 50 in Citeseer's List of Most Cited Authors in Computer Science. He is currently a Professor of Computer Science and Engineering at Washington University in St. Louis. Previously, he was one of the Co-founders of Nayna Networks, Inc. He was a Senior Consulting Engineer at Digital Equipment Corporation in Littleton, Mass and then a Professor of Computer and Information Sciences at Ohio State University in Columbus, Ohio. He is the author of *Art of Computer Systems Performance Analysis*, which won the 1991 'Best-Advanced How-to Book, Systems' award from Computer Press Association.

Subharthi Paul received his BS from University of Delhi, Delhi, India, and his Master's degree in Software Engineering from Jadavpur University, Kolkata, India. He is presently a doctoral student in Computer Science and Engineering at Washington University in St. Louis, MO, USA. His primary research interests are in the area of future internet architectures.

Jianli Pan received his BE in 2001 from Nanjing University in Posts and Telecommunications (NUPT), and his MS in 2004 from the Beijing University of Posts and Telecommunications (BUPT), China. He is currently a PhD student in the Department of Computer Science and Engineering in Washington University in Saint Louis, MO USA. His current research is on the next generation internet architecture and related issues. He is currently a student member of the IEEE.

1 Introduction

Next generation wireless networks (NGWNs) will be a convergence of different wireless technologies, such as cellular networks (2G/3G/4G), wireless broadband networks – e.g., (mobile) WiMAX and long term evolution (LTE), wireless sensor networks, and so on, and at the same time be interoperable with traditional IP-based wired networks. In NGWNs, the nodes or hosts will be mobile. In such a mobile wireless environment, the channel capacity varies over time and distance. Short-time disruptions may occur more frequently and result in a disconnection of operation. Therefore, *mobility* support is clearly one of the key requirements for NGWNs.

With the advance of networking technologies, the concept of a single host – single interface – single network will no longer be true in the context of NGWNs. A node or host may consist of many different networking interfaces incorporating different types of quality of service (QoS) controls for various applications, including voice, video, TV broadcasting, online games, medical applications, etc. This form of *multihoming* (device multihoming) offers many advantages, such as enhanced availability (fault tolerance), and traffic engineering (i.e., load balancing and load sharing).

Another feature of NGWNs is that the networks will be more user-centric. The user will be allowed to decide his/her preferred paths through *user path selection mechanisms*. The service provider should provide useful information with inherent security to aid such mechanisms. For example, with multiple networking interfaces in a single mobile device, the mobile users may choose their preferred paths for each task, probably based on the price paid and on the quality of the service offered by various service providers. Similar to a traditional cellular phone system, the users may be required to pay air-time charges.

From the user's perspective, there are always limitations and constraints. The user path selection decision may be guided by several factors, such as:

1 modes of operation of the mobile device QoS (QoS battery or power-line)

2 application requirements, e.g., throughput, delay, and completion time

3 economic viability QoS (QoS per min or flat rate, etc).

Therefore, there is no clear solution on how to choose the proper networking interfaces to make use of the resources efficiently and also meet the user requirements and constraints. Also, users may want to keep their location information private from their correspondent users. This is the so-called location privacy issue. Finally, the security of data is always of concern to users.

The issues of mobility, multihoming, location privacy, etc., discussed above, represent some of the key requirements for the design of NGWNs. Since future networks are expected to be all-IP (AIP), the question is how to make them support these features. Notice that The 3rd Generation Partnership Project (2007a, 2007b) has decided to use IP for the next generation of cellular wireless networks. System architecture evolution (SAE) is the core networking architecture, and SAE is AIP based.

IP cannot support the features required from the network layer of future networks. However, with the next generation wireless technologies – migrating to AIP architecture, it is clear that the current IP layer design has to be enhanced to suit the requirements of future wireless networks (FWNs).

One of the key design issues of the current IP layer is *the overloading of IP address semantics* (Jain, 2006; So-In et al., 2009a; Paul et al., 2009; Meyer et al., 2007; ITU-T, 2007). The IP address acts as a *host or node identifier* and *a locator* in the routing space. This contextual overloading implicitly binds a host to its point-of-attachment into the network. There is no independent namespace to represent the end host. Thus, every time the end host moves to a new network or changes its interface; and consequently obtains a new IP address, all the sessions bound to the previous IP address are broken. Such an implicit overloading makes it difficult to support full mobility, multihoming, traffic engineering, etc.

There have been many attempts to resolve some of these issues, especially in the traditional AIP based wired-networks (So-In et al., 2009a). However, no clear consensus has been reached. The problem is more serious within the mobile wireless environment. In general, the techniques to resolve these problems are based on redirection and indirection techniques, such as HIP (Moskowitz et al., 2006), SHIM6 (Nordmark and Bagnulo, 2007), LISP (Meyer, 2008), Enhanced MILSA (Pan et al., 2009), etc. The main differences among these techniques are their varying focus *on the different protocol layers, on the introduction of new naming spaces, on the required changes of a protocol stack, and on the ways to separate a host's identity from its locator* (So-In et al., 2009a). Mobile IP (MIP) (Perkins et al., 2002; Johnson et al., 2004; Gundavelli et al., 2008; Soliman, 2007; Soliman et al., 2008; Koodli et al., 2005; Campbell et al., 2002) is another well-known approach primarily designed to resolve the mobility issue (focus on network layer). The 3GPP has also adapted MIP in SAE. However, MIP and its extensions fail to fully support other features for NGWNs.

In this paper, our focus is on a network layer approach to mobility. A key advantage of this approach is that the network-layer based solutions require no change in the higher layers of the protocol stack, and so the solutions work for all applications. Then, we have applied the ID/locator split idea, a well-known approach used to resolve the mobility and location privacy issues, into a MIP environment. We have separated IP address space into identity (ID) and locator spaces (So-In et al., 2010a). The IP address from the ID space is used as the node identity, over the entire duration of the session. A home agent (HA) represents a rendezvous server or the mapping server to resolve the identity from/to the locators. These concepts make MIP fully support mobility and location privacy.

In addition, we have used multiple care-of-address (CoA) registration (Wakikawa et al., 2009) and flow binding option (Soliman et al., 2009c) features to support multiple interfaces. However, these combinations allow mobile IPv6 to achieve *per flow* multihoming in terms of flow sharing and flow balancing. Neither provides a mechanism to map or select a proper flow into each interface using path characteristic information. Therefore, to fully support multihoming, in this paper, we also introduce a policy-based multiple interface selection procedure to choose the best N interfaces and/or paths to meet the user requirements and constraints (So-In et al., 2010b).

This paper is organised as follows: In Section 2, we briefly survey proposals and/or techniques for mobility, multihoming (including user path preference and multi-interface selection), and location privacy. Then, we introduce the virtual ID concepts used to resolve key issues in NGWNs. In this section, we also discuss how to make use of multiple interfaces in an efficient way. Finally, the conclusions are drawn in Section 5.

2 Related work

In this section, we describe two main categories of approaches including the optimisation extensions used to resolve key issues in NGWNs: IP-in-IP tunnelling or encapsulation (i.e., MIP) and ID/locator split. In addition, we discuss their pros and cons in handling the three main features: mobility, multihoming, and location privacy as well as multi-interface selection problems.

2.1 MIP and its extensions

MIP (Perkins et al., 2002; Johonson et al., 2004) is a well-known technique designed to resolve the *mobility* problem in traditional wired and wireless networks. The 3GPP has adopted these concepts for SAE. Notice that most of the concepts discussed in this paper apply to both IPv4 and IPv6. However, for simplicity, we limit our discussion to IPv6 because it has sufficient address space and is preferred for public wireless networks.

2.1.1 Mobility

If nodes, hosts, or users change their networks and/or locations, then their IP addresses may also change. Consequently, their transmission control protocol (TCP) connections at the transport layer are broken. Mobile IPv6 is potentially used to maintain the connection and/or session regardless of time and location, with an IP-in-IP encapsulation technique. In other words, mobile IPv6 is used to preserve the connection.

Briefly, mobile IPv6 functions as follows: first the mobile node (MN) is assigned a home IP address (HoA). When the node moves from one network to another network, it informs its home network (HA) about its new IP address (CoA). When a correspondent node (CN) wants to contact this node, the CN sends a packet to the home network; the packet is then intercepted by the HA and forwarded to the MN's new address (CoA).

This basic technique has a *triangulation problem*: if the MN is far away from the home network but close to the CN, all packets from correspondent still have to go to the home network and be forwarded from there to the MN. Route-optimisation functionality (So-In et al., 2009a; Johonson et al., 2004) can be used to solve the triangulation problem. This, however, introduces other issues as discussed later in this section.

A second problem with basic MIP is that of ingress filtering. The MN puts its home address in the source IP address fields in packets. However, some routers may not forward packets with source addresses that are not from the local IP address space. Two optimisation techniques have been developed to address this issue. The first is to use a reverse tunnelling technique, i.e., to send packets back toward the home network but this technique introduces additional delay. The second optimisation technique is to use an IP-in-IP encapsulation with destination option in IPv6 (So-In et al., 2009a; Johonson et al., 2004) with the increase of header overhead trade-off.

Other approaches to mitigate the route-to-home network delay and/or hand-off latency are HAWAII, cellular IP, and hierarchical MIP (HMIP) (Soliman et al., 2008; Campbell et al., 2002). These approaches all deploy several HAs in a hierarchical manner, especially at the edge routers. With HMIP, the binding update is only sent to the local HA, which decreases delay latency. However, these approaches require synchronisation among HAs and additional nodes.

A fast handover in MIP (Koodli et al., 2005) is used to allow MNs to configure new CoAs before moving to a new network. When the MNs attach to a new base station, they can communicate using the new CoA. The fast handover in MIP requires that some packets be forwarded from the old to the new base station.

Proxy-MIP (Gundavelli et al., 2008) was originally introduced to improve the deployability of MIP using network mobility or NEMO (Nagami et al., 2007). The idea is to use the router or proxy agent to act on behalf of the

MN and to perform the MIP functionality. In other words, with proxy-MIP, no changes are made in the MN to support MIP.

Dual stack MIP (MIPv4 and MIPv6), or DSMIPv6 (Soliman, 2007), was developed to allow backward interoperability between IPv4 and IPv6. DSMIPv6 applies IP-in-IP encapsulation in which the MN or proxy agent does not support IPv6. Note that both proxy MIP and dual stack MIP have been selected by 3GPP for use in SAE.

2.1.2 Multihoming

MIP cannot support multihoming because each single MN is bound to only one IP address. Recently, some have suggested allowing multiple CoAs registrations (Wakikawa et al., 2009; Soliman et al., 2009c) to allow multihoming. There is no detailed discussion on the user path selection issue and how to optimally use multiple interfaces underneath.

2.1.3 Location privacy

MIP has no concept to resolve location privacy; however, MIP implicitly supports it if and only if the MN is in foreign networks because the current location is no longer bound to the home address. However, this scenario introduces a triangular routing problem. The mobile IPv6 route optimisation feature was introduced to resolve this triangular routing problem, that is, to allow the MN and the CN to communicate with each other directly. However, again, this direct communication introduces the location privacy issue for mobile users.

2.2 ID/locator split

The ID/locator split concept (Jain, 2006; So-In et al., 2009a; Paul et al., 2009; Meyer et al., 2007; ITU-T, 2007) is a well-known approach used to resolve *mobility and location privacy*. Basically, the idea is to separate the functionality of the identity from that of the locator.

2.2.1 Mobility

Each MN has its own unique identity. When the node moves, its identity does not change, but its locator does. Consequently, the connection will not be broken because the connection will be bound to the identity, not to the locators. The locator represents the current point of attachment to the network. In other words, the locator helps decide where the packet should be routed. However, such indirection mechanisms also require new naming and name resolution mechanisms.

2.2.2 Multihoming

The ID/locator split concept implicitly supports the use of multihoming because the identity is not tied to a particular locator. However, there is no detailed discussion on user path selection, and, again, on how to efficiently use the multiple locators.

2.2.3 Location privacy

The key feature of ID/locator split concepts is that it supports location privacy since the location is invisible to the CN. CN knows only the node identity (from the node name to node identity mapping), not the actual locators.

In general, the identity can be a string of characters or digits. Currently in IPv6, each node has a name and an address. The node name is the fully qualified domain name (FQDN). The 128-bit IPv6 address is used to represent both identity and location. The domain name server (DNS) is used to convert from FQDN to the node ID (IPv6 address). Then, the same address is used as a locator for routing the packet to the node.

There are three ways to implement the ID/locator split concept: placing a split in the end host, e.g., HIP (Moskowitz et al., 2006) and SHIM6 (Nordmark and Bagnulo, 2007), in the network, e.g., LISP (Meyer, 2008), and creating a combination split, e.g., enhanced MILSA (Pan et al., 2009). The first approach requires the insertion of a new ID sub-layer usually between the transport and the network layers. Thus, the upper layers are bound to an identity instead of a locator. HIP and enhanced MILSA introduce new secure naming spaces, but SHIM6 uses one of its current locators as the identity.

The second set of splitting techniques implements the ID/locator split concept in the network. The basic idea is that there is no change to the end host. The routers take care of the split. At the edge of the network, the identities are resolved into the locators needed for communication. This requires changes to network infrastructure devices (routers). The third approach is to combine the former two, when allows splitting in both the host and the network, but with a complexity trade-off.

2.3 Interface selection and flow distribution

In the previous two sections, a multihoming support has been discussed by dynamically allowing different locators to be bound with the identity (ID/locator split) and home address (MIP); however, there is no explicit mechanism to choose or use multiple locators in an efficient way.

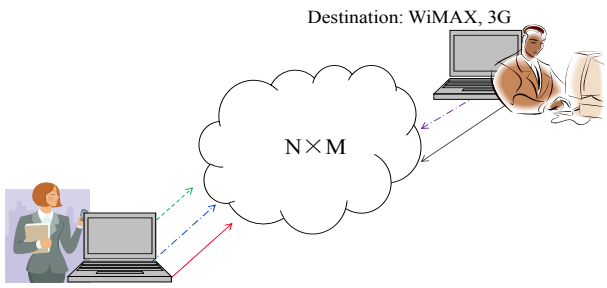
In general, the problem of selecting a networking interface (locator) has been investigated in the concept of "Always best connected" (Gustafsson and Jonsson, 2003). The basic idea is to find the *best single active* interface given the interface characteristics and network constraints such as bandwidth, power consumption, access technology, and so on. This concept is purposely used for a hard handover purpose (Wang et al., 1999), that is, only one best active interface is allowed at any moment. Mitsuya et al. (2007) also proposed a policy based on user requirements, such as air-time charges, to select the best networking interface.

Now, let us consider MIP. Again, recently a flow binding option feature (Wakikawa et al., 2009; Soliman et al., 2009) has been introduced to map a particular flow into a specific interface or CoA. MIP can use this option to uniquely identify the flow. Note that this extension is based on the use of multiple CoA registrations (Wakikawa et al., 2009) that we described briefly earlier. To meet user requirements and QoS control parameters, a mechanism similar to a policy-based model is required for MIP.

In addition, with a single best active interface, there is a fundamental limitation. The multihoming feature cannot be fully utilised using only one interface. For example, suppose a user with a device with both 3G and WiMAX interfaces wants to maximise her throughput. Also, suppose the device is power-line operated and paid for by the flat rate fee; therefore, using two interfaces simultaneously will achieve twice as much throughput available for each application. Otherwise, a per-application or a flow distribution throughput will be limited by a per-flow distribution mapped to each interface.

Figure 1 shows a simple configuration for the end-to-end multihoming. In this setup, there are three different interfaces at the source: WiMAX, 3G, and WLANs; and only two at the destination: WiMAX and 3G. Assuming there is a wired internet in between; there is at least 3 × 2 or six possible paths between these two users. The path characteristics, e.g., path throughput, congestion level, loss probability, end-to-end delay, and so on, may be different from one path to the others.

Figure 1 Example of end-to-end multihoming (N × M = 3 × 2) (see online version for colours)



Source: WiMax, 3G, WLAN

Moreover, aside from achieving throughput aggregation, enabling multiple interface transmissions simultaneously allows MIP to support soft handoff when two radio channels are used at the same time.

In fact, there have been several proposals to resolve the issue of throughput aggregation. These solutions may be classified by the aggregation layer, such as session, transport, network, or link layer. For example, SLM was introduced by Landfeldt et al. (1999) to achieve the mobility at the session layer. SLM allows multiple transport connections for each application, and results in a higher throughput. Hsieh and Sivakumar (2002) proposed a wrapper at the transport layer used to allow multiple virtual connections to aggregate, and so increases the total throughput. Adiseshu et al. (1996) applied a simple link scheduling algorithms, e.g., a deficit round robin (DRR) and a weighted fair queue (WFQ), to balance per-packet transmissions.

Each of these proposals has its pros and cons. For instance, the lower level modifications make upper layers unaware of the aggregation and multiple connections; however, they lack flow and QoS (application-based) information. The higher layer modifications do not need to change the lower protocol stack but there is no explicit mechanism to select a particular interface.

In MIP, the modification occurs at the network layer. The change can only be done in the built-in MIP agent. The upper layers do not need to be aware of this change. There are no requirements to modify other parts of the protocol stack. Generally, in a MIP communication, it is difficult to use multiple connections; especially in a scenario with a single home address and several connected CoAs. Note that for all solutions allowing multiple interface transmissions, protocol data unit re-ordering is required.

3 A new framework using the ID/locator split concepts applied to a mobile IPv6 environment

In this section, we introduce a new framework used to resolve the key issues of mobility, multihoming, and location privacy in NGWNs. For mobility, we use the mobile IPv6 concept. For multihoming including user path selection, we apply the multiple CoA registration feature. Then, an algorithm based on linear programming is used to derive the best *N* interfaces in order to set simple weights among available paths. For location privacy, we have employed the ID/locator split concept by separating the IP address space into two naming spaces: identity and locator spaces.

3.1 Virtual ID (MIP + ID/locator split)

As we discussed earlier, both MIP and ID/Locator Split have their own advantages and disadvantages. For example, MIP does not require a new naming space and is supported by the industry (3GPP SAE, Cisco Systems, etc.). However, MIP has no concept of identity and cannot distinguish itself from the actual locators. The ID/locator split concept can explicitly support location privacy; however, there is an issue of deployability. Therefore, in this section we propose the idea of virtual ID by applying the ID/locator split concept explicitly to a MIP environment to leverage the advantages from both approaches.

3.1.1 Virtual ID

In IPv6, a 128-bit address is used for both node identity and locator which introduces many disadvantages, as indicated earlier. In a mobile wireless environment, mobile IPv6 also

mixes these functionalities. Therefore, to explicitly separate the function of the identity from that of the locator in mobile IPv6, we have divided the IP address space into identity and locator spaces.

Similar to SHIM6, the 128-bit IPv6 address is used as node identity and locator. However, we do not use one of the nodes' current addresses as its identity. Instead, we use an address from a separate identity space. We call this a virtual home address or virtual ID.

In traditional MIP, when the MN is in the home network, a single IPv6 address represents both node identity and locator. However, when the MN is outside the home network, mobile IPv6 can be treated as an ID/locator split scheme because another IP address, CoA, is involved. This CoA can be treated as the node locator (the indicator of where the node is).

Virtual ID is pre-defined and randomly assigned by the service provider (again from the ID space). This ID is permanent, and thus no longer bound to the home network and/or to the location of the nodes. In other words, the virtual ID is used even when the MN resides in the home network. As in mobile IPv6, the IP-in-IP encapsulation is applied in that the nodes update their CoAs when they are in different location/networks.

We use the 128-bit IPv6 address format to represent the node's identity. This allows backward compatibility since legacy nodes (virtual ID unaware nodes) treat these identities as addresses.

3.1.2 User location privacy

The concepts of virtual ID introduced by separating the node identity from its locations helps resolve the issue of location privacy in that the CNs do not know the location of the MN, only the node identity.

Consider the ID/locator split concept. Basically there are two levels of mapping: from node name (FQDN) to node identity, and then from node identity to node location. The result of the DNS resolution is the node's identity, not its location. The other mapping level can be done at rendezvous servers. With virtual ID, an additional mapping from the virtual home address to the mobile IPv6 home address is also required. Therefore, the HA is modified to do the second level mapping.

Consider MIP with the route optimisation feature. In general, the CNs are required to send the packets through the MN's home network. Therefore, there is a triangulation issue, as discussed earlier (see related work). Route optimisation can be used, but again, lead to user location privacy problem.

To solve this problem, we propose an add-on feature to proxy mobile IPv6 (Nordmark and Bagnulo, 2007). Traditionally, in proxy mobile IPv6, a mobile access gateway or proxy node is used to provide MIP functionality on behalf of mobile-IP unaware nodes. However, with MIP-aware node (the MIP functionality is performed at the end node), the proxy is also required to optionally rewrite the address with its selected anonymity proxy address to hide the exact location.

3.1.3 Multihoming

In NGWNs, mobile users will want to exercise multi-interface availability or user path selection because they will have to pay for their choices according to bandwidth constraints and other QoS controls. For example, suppose Alice buys access services from two different service providers: one service is a 3G network accessible to her cellular phone; the other is over WLANs. When she is at home or when WLANs are available, Alice would prefer accessing the internet service through WLANs and also probably disable the 3G service, especially when air-time charges are high.

To meet these requirements of multihoming and user path selection, virtual ID also applies the multiple CoA registration concepts at the HA to support multihoming (Soliman et al., 2009c). Again, virtual ID is unique and permanent. Only the physical location or CoAs can be changed on the fly with a change of locations.

In NGWNs, users should be able to choose their own ingress and egress paths, based on price and the QoS constraints. For simplicity, we use a weight factor along with the CoA registration when the nodes update the address to the HA.

Note that in a more general case, the users could specify a set of connection rules. The HA will forward the packets to the node according to pre-selected user path rules. We will discuss the details of multi-interface selection in Section 3.2.

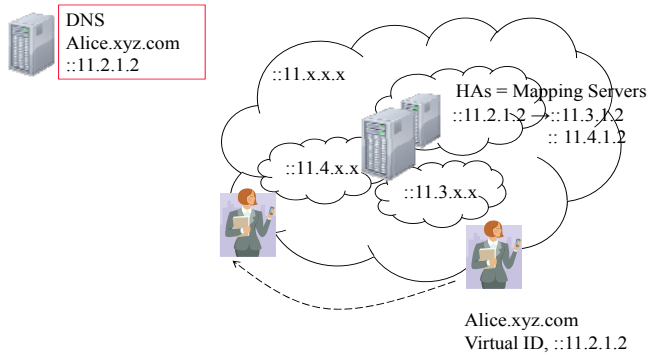
3.1.4 Multi-tiered multihoming

In the previous section, we described the concept of device multihoming. Here, we discuss multihoming in general. In the past, computers had a single networking interface. Also, most nodes stayed inside only one network with one egress. Nowadays, nodes have multiple networking interfaces resulting in *node multihoming* or *device multihoming*. In addition, each user may have multiple devices such as computers, personal digital assistant (PDA), and cellular phones. We call this *user multihoming*. Finally, the networks that users are in may contain several internet interfaces. We call this *site multihoming*. This hierarchical concept is the so-called multi-tiered multihoming phenomenon. Note that all these multihoming functionality aspects may be used to support fault-tolerance and traffic engineering (i.e., load sharing and load balancing). This explanation will be clear with detailed examples in the next section.

3.1.5 Virtual ID examples

In this section, we provide detailed examples for the virtual ID concepts.

Figure 2 Virtual ID example (see online version for colours)



Mobility example

Figure 2 shows an example of virtual ID. In this figure, Alice's node's name is *Alice.xyz.com* registered at a DNS. The service provider allocates a virtual home address (virtual ID), *::11.2.1.2*, as Alice's identity. Suppose the service provider networks are *::11.x.x.x* with *::11.3.x.x* and *::11.4.x.x* sub-networks assigned into different physical regions. The virtual home address networks *::11.2.x.x* are specifically dedicated as a virtual ID space. Only the service provider knows the mapping between virtual ID or node identity (*::11.2.1.2*) and the physical address (*::11.3.1.2*) or current IP address. This mapping can be stored at rendezvous servers or HAs. Alice's node's ID will be used regardless of the change of location for mobility purpose; when Alice moves to networks *::11.4.x.x*, Alice's node's locator is updated to *::11.4.1.2*, not her ID.

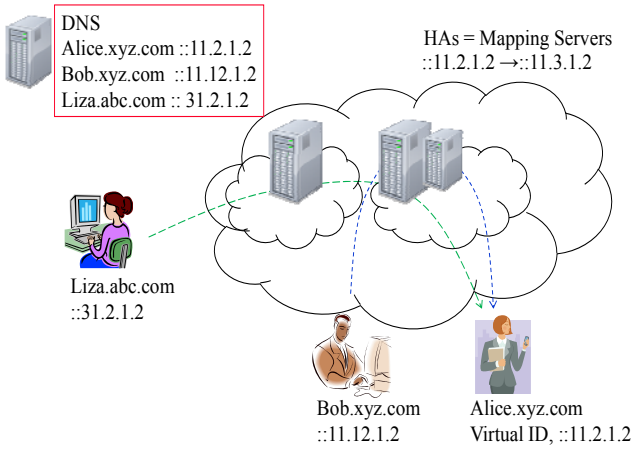
User location privacy example

This section describes two main scenarios that use virtual ID to achieve a location privacy requirement in NGWNs: when the CN resides either out of the home network or inside the home network.

The first scenario is when the node is in a different network. Figure 3 shows the CN or *Liza.abc.com* contacting *Alice.xyz.com,* which is in a different service provider network. First, Liza (at *::31.2.1.2*) retrieves Alice's identity, *::11.2.1.2*, from a DNS resolution process and uses that ID to route packets to Alice's home network. Since Alice's ID is used instead of her physical attached address, *::11.3.1.2*, Alice's location privacy can be maintained. Notice that if Alice is in a foreign network, her location privacy is implicitly maintained. This scenario is similar to a traditional mobile IPv6 because the permanent home address is different from her virtual ID.

The other scenario is when communication occurs within the same network, say, in *::11.x.x.x* networks. Suppose Bob's address is *::11.12.1.2* and again Alice is at *::11.3.1.2*, within her home network. With a traditional mobile IPv6, Bob knows Alice's location. However, with virtual ID, Alice's identity, *::11.2.1.2*, is used instead; therefore, Bob no longer knows Alice's location information.

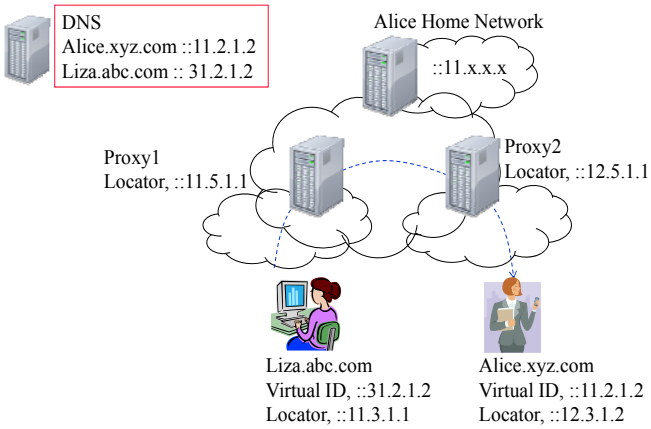
Figure 3 Virtual ID with user location privacy example (see online version for colours)



Proxy-assisted user location privacy example

Figure 4 shows an example of proxy-assisted mobile IPv6 and virtual ID providing location privacy (especially with route optimisation). In this figure, Liza, a CN, wants to contact Alice, who is not in her own home network. With the Liza node unaware of MIP, proxy mobile IPv6 can be directly applied to preserve location privacy.

Figure 4 Proxy-assisted user location privacy example (see online version for colours)



However, with a MIP enabled device, one approach is to enforce the proxy MIP functionality over the device (by treating the device as a MIP unaware node). The other approach is to enable address rewriting at the proxy and/or HA. With this functionality, the proxy hides the actual locator by rewriting the CoA with an anonymity proxy address. Note that the communication of both end proxies is maintained with routable addresses.

In this figure, suppose Liza wants to reach Alice. With route optimisation, Liza knows only her proxy not Alice's locator. Liza's proxy will rewrite Alice's locator (e.g., from *::12.3.1.2* to *::x.x.x.x*) and forward the packets to Liza. This rewriting can also be done for Liza as well. Notice that each proxy has local node location information so that the

proxy can forward the packets to the correct final destination.

Multihoming example

In this section, we provide details of how to incorporate the multihoming feature into NGWNs by pre-registering all possible CoAs regardless of the service provider. We also show a simple use of weight factor for user path selection.

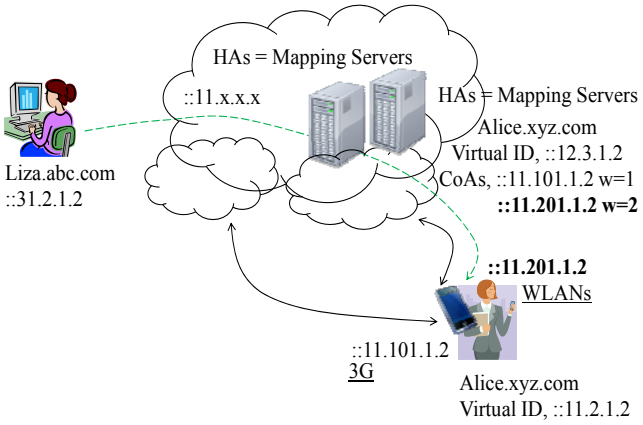
We consider two main scenarios:

1 when the multihoming interfaces are to the same service provider

2 when they are to different service providers.

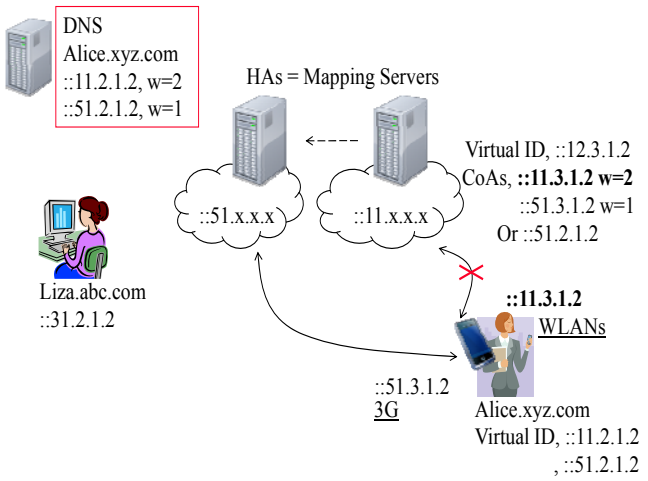
The first scenario, Figure 5 shows the process of multiple CoA registrations with the preferred path selection when both networking attachment points are with the same service provider. In this figure, Alice has two access technologies with the same service provider (*::11.x.x.x*), say, cellular networks and WLANs on her single mobile device. Alice's virtual ID is:*::11.2.1.2*, and the two physical locators or CoAs are *::11.101.1.2* (on 3G networks) and *::11.201.1.2* (on WLANs). When Alice is at home, she can send the update to her HA to set a higher priority toward the WLAN interface so that the inbound traffic can be forwarded toward this WLAN interface.

Figure 5 Multiple CoA registration example (see online version for colours)



The other scenario is when mobile users have multiple access services from different network providers. Figure 6 shows this configuration. As shown, Alice has two access services from two different service providers: cellular networks and WLANs. Since there are different service providers, Alice may acquire two different virtual IDs. Alice can send the update to the DNS server with her preferred path selections (with different weights). In this scenario, Alice is at home and she prefers the WLAN path (with a higher weight, or higher priority), which is towards *::11.x.x.x* networks.

Figure 6 Multihoming feature in mobile IPv6 with virtual ID example (see online version for colours)



Note that the CoA of Alice on the WLAN path is *::11.3.1.2*, not the virtual ID *::11.2.1.2*. In this scenario, the packets are sent only through the WLAN interface as long as Alice does not update her preferred path on her DNS. There are no requirements for cooperation and interaction between two service providers.

When WLANs are not working, either Alice or Alice's home network can detect the disconnection. Without the interaction between the service providers, the packets can continue to flow towards WLANs until Alice sends the update to the DNS server. Therefore, an additional operation is required. There are two possible solutions here: Alice can register either other CoAs or other virtual IDs acquired from different networks. Whenever the link is broken, the corresponding HA (here, at WLANs), will send all packets to alternate registered CoAs. These operations may require agreement among the involved service providers.

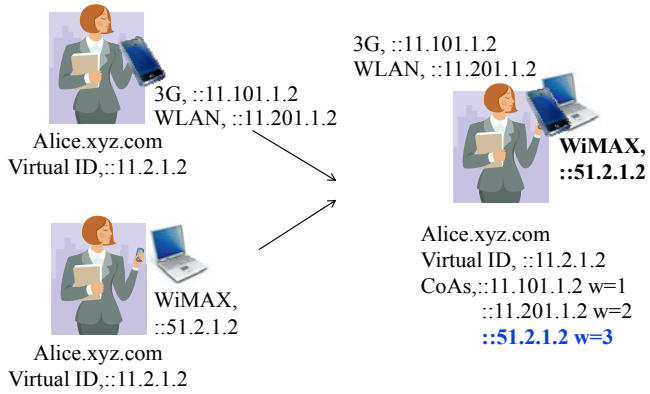
During disconnection, the steps in Figure 6 are as follows: Liza, *::31.2.1.2*, originally sends her packets to Alice through WLANs (*::11.1.3.2*). Due to a link failure, the WLAN interface of Alice is unreachable. After the link failure detection, the HA at WLANs redirects all packets to the alternate registered CoAs (either her actual locations or her other virtual IDs) in order to reach Alice. Again, this redirection is based on a roaming policy. Whenever Alice sends the update to the DNS server to withdraw the disconnected path and/or to set a lower preference, this redirection will be terminated.

Notice that in this example, we allow more than one virtual ID because there is no single network organisation to deal with diverse networks. These virtual IDs are permanent regardless of location, and are acquired during the setup during the DNS look-up process until the end of communication. However, if service providers agree on the identity naming space, only one virtual ID is required, and then CoAs among different service providers will be mapped to that virtual ID.

Multi-tiered multihoming example

Figure 7 shows an example of a combination of device and user multihoming. For simplicity, this scenario assumes that there is only one service provider.

Figure 7 Multi-tiered multihoming example (see online version for colours)



Note: Two-tiered multihoming: user and device multihoming

For device multihoming, the device may have several networking interfaces, for example, a mobile phone with 3G and WLANs with CoAs *::11.101.1.2* and *::11.201.1.2,* respectively. For user multihoming, the user may have many networking components such as a PDA (with cellular network and WLAN 802.11b functionalities) and a laptop (with WiMAX 802.16e functionality). The combination of these two layers of multihoming is called a two-tiered multihoming. In this figure, Alice has a single virtual ID, *::11.2.1.2*. However, there are three locators: a cell phone with 3G and WLAN locators and a laptop with a WiMAX locator. Alice can indicate her preference by updating the CoAs. In this example, Alice prefers a WiMAX connection, which has a CoA of *::51.2.1.2* with a weight of three (the highest priority). Now the same concept of multihoming that we described in the previous section can be applied.

3.2 Policy oriented multi-interface selection

In the previous section, we introduced a simple weight factor used to choose the interface by advertising the weight value when updating the CoAs. In this section, we describe the policy-oriented multi-interface selection model to make efficient use of multiple interfaces simultaneously.

3.2.1 A policy-based QoS model

We formulate the policy-based QoS model using linear programming (Cormen et al., 2001). The target requirement is the expected bandwidth required for each application with several constraints such as total available power, user budget, and completion time.

Figure 8 shows the minimisation problem using linear programming. The objective function is to minimise the weighted power consumption, charges, and completion time. These parameters are normalised to per time unit (t). Users can arbitrarily set the weights according to their requirements and perspectives. The weights are summed to 1. In this formulation, there are only three constraints, and so there are three weight values. In general, the number of weight factors depends on the number of user constraints.

Figure 8 Policy-based minimisation formulation

Requirements: expected throughput (B_{total})

Optional constraints: available power (P_{total}), total budget (C_{total}), and completion time (T_{total})

Each interface associated with per interface Throughput_Mbps (B_i), Charge_dollar/sec (C_i), and Power_ Joule/sec (P_i)

Turn on power charge (suppose constant charge): P_c

Objective:

$$\min\{w_1 \times (t_1P_1 + t_2P_2 + \cdots + t_iP_i + i \times P_c) + w_2 \times (t_1C_1 + t_2C_2 + \cdots + t_iC_i) + w_3 \times (t_1 + t_2 + \cdots + t_i)\}$$

$$\text{where } \sum_{i=1}^{N} w_i = 1, \qquad 1 << i << n$$

Requirements: $t_1B_1 + t_2B_2 + \cdots + t_iB_i >> B_{total}$.

Optional constraints: $t_1P_1 + t_2P_2 + \cdots + t_iP_i + i \times P_c << P_{total}$

$$t_1C_1 + t_2C_2 + \cdots + t_iC_i << C_{total}$$

$$t_1 + t_2 + \cdots + t_i << T_{total}$$

$$t_1, t_2, \ldots, t_i >> 0$$

Here, the throughput B, used to meet the expected throughput (B_{total}), is the path throughput not the access bandwidth. The connection manager, described in the next section, functions as a capacity estimator for different paths to CoAs. The resulting path throughput is used in the calculation. In addition, users can always override the rules. For example, for delay/jitter sensitive applications, users may choose the interface that meets their end-to-end jitter/delay expectations. Obviously, there is always a trade-off between user requirements and available resources.

In this formulation, we assume static linear programming. Whenever the resource changes, or the number of constraints is modified, the linear solver (Cormen et al., 2001) will need to be readjusted. It is also possible to use dynamic linear solver techniques. However, it is not in our scope here. In addition, in our linear programming formulation, the linear relationship among constraints is used. However, it is possible that in the real world, the relationship is nonlinear; then other optimisation techniques to solve nonlinear objectives can be used.

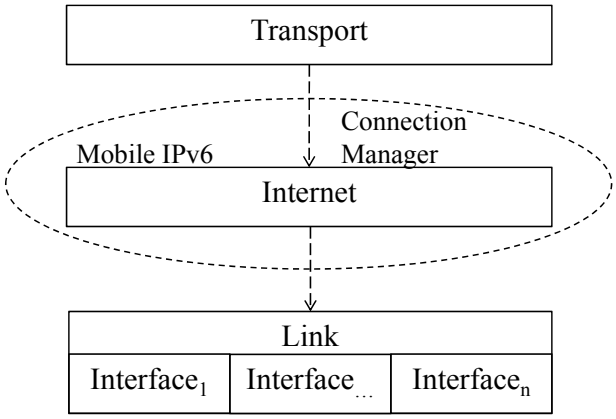
Example: Consider a device equipped with two networking interfaces: WiMAX and 3G. The total battery power is 10 Joules. For each interface: capacity, usage charge, and power consumption are as follows: 1 Mbps, 1 cent/sec, 1 Joule/sec for WiMAX and 0.5 Mbps, 3 cent/sec, and 0.01 Joule/sec for 3G. The user wants to send a 10 Mb file.

If the mobile user considers only the power consumption, and there are no other constraints, such as total budget or

completion time, the mobile user can set w_1 to 1 and other weights to 0. Suppose $P_c = 0$, then the objective function here is *Min* (t_1 + 0.01 t_2). As a result, with any well-known linear programming resolvers (Cormen et al., 2001), the 3G interface is chosen for a 20-second transmission.

Mobile users can set the weights according to their requirements. For example, if w_1 and w_2 are set to 0.5, and suppose the completion time is required to be less than eight seconds. In this case, the objective function is *Min* $\{0.5(t_1 + 0.01\ t_2) + 0.5(t_1 + 3t_2)\}$. As a result, the WiMAX interface is chosen for eight seconds, and 3G for four seconds.

Figure 9 Interface of connection manager



3.2.2 Connection manager

In this section, we discuss the functionality of the connection manager. The connection manager is an extra module wrapped around the network layer or built-in the MIP agent as shown in Figure 9. The main purpose of this manager is to distribute the flow among interfaces corresponding to a particular virtual ID in an efficient way using the policy-based QoS model. There are four tasks: flow striping, interface/path characteristic estimation, TCP freeze function, and automatic setup default configuration.

Flow striping option

Recently, a MIP flow binding option has been suggested to distinguish a flow mapped to a particular CoA or interface. This allows the MN to exercise the multihoming feature of MIP; however, this option uses only one active CoA. To truly benefit from the multihoming feature, we allow a flow to be striped over multiple interfaces. We call this the flow striping option in mobile IPv6.

As shown in Figures 10 and 11, the flow striping option (registration) is included in the binding update. The type field specifies the flow striping option. The length includes a flow identifier (FID) and IPv6 CoAs, and is specified in 8-octet units. FID, or flow identifier, is a 16-bit unsigned integer that identifies the flow binding (Wakikawa et al., 2009; Soliman et al., 2009). The set of N active interfaces or CoAs for this particular flow is specified in the IPv6 CoA field.

Figure 10 Flow striping mobility option (registration)

Type (1 byte)	Length (1 byte)	FID (2 bytes)
IPv6s (CoAs)		
.		
.		

Figure 11 Flow striping mobility option (interface/path characteristic)

Type (1 byte)	Length (1 byte)	FID (2 bytes)
Interface/path characteristics e.g., buffer, timestamp, #byte count, access link technology		

Any suitable distribution algorithm, e.g., weighted deficit round robin algorithm (WDRR) or a modified version of DRR the so-called surplus round robin (Adiseshu et al., 1996), can be used to distribute packets among various paths at the network layer. Due to different round trip times (RTTs) along different paths, buffers are required in the receiver. The buffer size depends upon the difference between the maximum RTT and minimum RTT along any paths. Obviously there is a trade-off between received buffers vs. the total throughput gained from multiple interface transmissions.

To limit the variation of packet arrival times, the scheduler may select different size blocks or quantums along different paths. In addition, to mitigate out-of-order packets, a packet sequence number is used as one of the destination options to help the ordered delivery process at the receiving end. Adiseshu et al. (1996) showed that if the load sharing derived from a casual fair queuing algorithm A is used for channel striping, and A is used as the re-sequencing algorithm, then the sequence of packets output by the receiver would be the same as the sequence of packet input unless there is a lost packet.

Instead of distributing an entire flow among multiple interfaces, it is also possible to dynamically move the flow among interfaces while using only one interface. This results in less complexity but in less load balancing than multi-interface distribution.

Although per-packet based or per-bit based (fair queuing) distribution can achieve ideal load balancing and load sharing, and can make more efficient use of multihoming feature than per flow-based distribution, it may require more receiver buffers or may result in fragmentation overhead, especially if the variation of RTTs is high. As a result, the connection manager allows the user to override the rules. The connection manager allows only per flow distribution for those applications where ordered low-latency delivery is required, while it does flow striping for applications whose goals are to achieve high throughput and that ordered delivery is not critical.

Again, the interface selection is based on the user decision. The connection manager only provides the

information about interfaces and path characteristics (we will describe these in the next subsection) so the mobile user will finally select *N* interfaces for each particular flow.

Interface/path characteristic estimation

To help compute the proper weights for flow distribution or to provide useful information regarding the link characteristics, end-to-end parameters, such as path bandwidth, delay, congestion level, loss probability, end node buffer, etc., are required. Note that the connection manager can request the link layer characteristics, such as modulation and coding schemes.

To estimate the path throughput, the connection manager does heuristic path bandwidth estimation in that the number of bytes received is periodically counted. The achievable throughput is calculated by that number over time provided by a timestamp mechanism. Figure 11 shows a flow striping mobility feedback option that indicates interface/path characteristics including receiver buffers used for the flow (group of CoAs), timestamp, and byte count.

If a path is symmetric, the results from the bandwidth estimation can be used, otherwise this information can be sent by piggybacking, or in a separate control channel to the other end. Timestamps can also be used to approximately estimate the end-to-end path delay.

Path congestion and/or loss indication, e.g., obtained by the explicit congestion notification (Ramakrishnan et al., 2001), can also be used in striping decisions.

TCP freeze function

We propose a TCP freeze option (So-In et al., 2009b) in which routers send a zero window advertisement to the transport layer to freeze the TCP when there is a link failure. This allows the TCP connections to continue after the link comes back up. We believe that connection manager can make use of a modified version of the TCP freeze option and help in riding over path disconnection. However, full details still need to be worked out.

Note that link-level recovery mechanisms; for example, (hybrid) automatic repeat request (H-ARQ) (Jeffrey et al., 2007) and ARQ can be used to aid this mechanism.

Automatic setup default configuration

Due to multihoming, the device may receive many default configurations, e.g., default gateways and DNS servers for multiple interfaces or networks. The connection manager may also use a ping-like probing mechanism to check the path reachability/availability. These processes help avoid dead DNSs and routes. They help the connection manager setup proper default configurations for feasible end-to-end communications.

4 Conclusions

Next generation wireless networks, will be a ubiquitous AIP network. Users will be using devices with a variety of networking technologies and will be highly mobile. Some of the applications may need high throughput and strict delay constraints, e.g., video streaming, online gaming, and medical applications. Therefore, these networks should support key features, such as full support for user mobility, multihoming, location privacy, and so on.

MIP and its variants have been introduced to resolve some of these issues. These proposals focus on a network layer technique. Some of these techniques have been selected by 3GPP for the system evolution architecture standard. However, these techniques have several limitations, especially due to the problem of identity and locator overloading of IP addresses. The ID/locator split concept was introduced to overcome many problems; however, it requires a new naming space, and lacks detailed implementations.

In this paper, we discussed MIP and its variants, and also pointed out several drawbacks (achieved by the ID/locator split concept). Then, we introduced a new technique called virtual ID, to make the MIP fully support mobility, multihoming, and location privacy. These add-ons are based on the standard mobile IPv6 and its extensions and are, therefore, easy to deploy along with mobile IPv6.

In addition, to exercise the use of multiple interfaces (for multihoming purpose), we showed a simple case using weight factors to select the proper path. These weights are used along with multiple CoA registration and flow binding options. We also proposed a framework to optimally select multiple interfaces given power consumption, user budget, and completion time constraints.

The multi-interface selection is done by linear programming minimisation using a policy-based QoS model based on user expectations and constraints. We relaxed the problem of a per-packet distribution into a per-application or a per-flow distribution whereby the mobile users can override the rules. For example, RTT sensitive applications can use only per-flow distribution and allow a flow striping mechanism for others. Our formalisation can achieve the throughput aggregation goal with the minimisation of several constraints.

References

Adiseshu, H., Parlkar, G. and Varghese, G. (1996) 'A reliable and scalable striping protocol', *ACM SIGCOMM Computer Communication Review*, Vol. 26, No. 4.

Campbell, A.T. et al. (2002) 'Comparison of IP micromobility protocols', *IEEE Wireless Communication Magazine*, Vol. 9, No. 1, pp.72–82.

Cormen, T.H., Leiserson, C.E., Rivest, R.L. and Stein, C. (2001) *Introduction to Algorithms*, MIT Press, p.1180.

Gundavelli, S. et al. (2008) 'Proxy mobile IPv6', RFC 5213, Internet Engineering Task Force.

Gustafsson, E. and Jonsson, A. (2003) 'Always best connected', *IEEE Wireless Communication*, Vol. 10, No. 1, pp.49–55.

Hsieh, H. and Sivakumar, R. (2002) 'A transport layer approach for achieving aggregate bandwidths on multi-homed mobile hosts', *Proceedings of International Conference on Mobile Computing and Networking*, GA, USA, pp.83–94.

ITU-T (2007) 'General requirements for ID/locator separation in NGN', Y.2015, p.18.

Jain, R. (2006) 'Internet 3.0: ten problems with current internet architecture and solutions for the next generation', *Proceedings of IEEE Military Communication Conference*, WAHINGTON DC, USA, pp.1–9.

Jeffrey, G., Andrews, J., Arunabha-Ghosh, A. and Muhamed, R. (2007) *Fundamentals of WiMAX Understanding Broadband Wireless Networking*, Prentice Hall PTR, p.496.

Johonson, D., Perkins, C. and Arkko, J. (2004) 'Mobility support in IPv6', RFC 3775, Internet Engineering Task Force.

Koodli, R. et al. (2005) 'Fast handovers for mobile IPv6', RFC 4068, Internet Engineering Task Force.

Landfeldt, B., Larsson, T., Ismailov, Y. and Seneviratne, A. (1999) 'SLM, a framework for session layer mobility management', *Proceedings of Computer Communication and Networks*, MA, USA, pp.452–456.

Meyer, D., Zhang, L. and Fall, K. (2007) 'Report from the IAB workshop on routing and addressing', RFC 4984, Internet Engineering Task Force.

Meyer, D. (2008) 'The locator identifier separaton protocol (LISP)', *Cisco Systems: The Internet Protocol Journal*, Vol. 11, No. 1.

Mitsuya, K. et al. (2007) 'A policy management framework for flow distribution on multihomed end nodes', *Proceedings of International Workshop on Mobility in the Evolving Internet Architecture*, Kyoto, Japan.

Moskowitz, R., Nikander, P. and Jokela, P. (2006) 'Host identity protocol (HIP) architecture', RFC 4423, Internet Engineering Task Force.

Nagami, K. et al. (2007) 'Multihoming for small-scale fixed networks using mobile IP and network mobility (NEMO)', RFC 4908, Internet Engineering Task Force.

Nordmark, E. and Bagnulo, M. (2007) 'Internet draft: Shim6: level 3 multihoming Shim protocol for IPv6', Internet-Draft, draft-ietf-shim6-proto-09, Internet Engineering Task Force.

Pan, J. et al. (2009) 'Enhanced MILSA architecture for naming, addressing, routing and security issues in the next generation internet', *Proceedings of IEEE Global Communication Conference*, HAWAII, USA, pp.1–6.

Paul, S., Pan, J. and Jain R. (2009) 'A survey of naming systems: classification and analysis of the current schemes using a new naming reference model', WUSTL Technical Report, available at http://www.cse.wustl.edu/~jain/papers/naming.htm.

Perkins, C. et al. (2002) 'IP mobility support for IPv4', RFC 3220, Internet Engineering Task Force.

Ramakrishnan, K.K., Floyd, S. and Black, D. (2001) 'The addition of explicit congestion notification (ECN) to IP', RFC 2481, Internet Engineering Task Force.

So-In, C., Jain, R., Paul, S. and Pan, J. (2009a) 'Future wireless networks: key issues and a survey (ID/locator split perspective)', WUSTL Technical Report, available at http://www.cse.wustl.edu/~jain/papers/fwns.htm.

So-In, C., Jain, R. and Dommety, G. (2009b) 'PETS: persistent TCP using simple freeze', *Proceedings of the 1st International Conference on Future Information Networks*, Beijing, China, pp.97–102.

So-In, C., Jain, R., Paul, S. and Pan, J. (2010a) 'Virtual ID: a technique for mobility, multihoming, and location privacy in next generation wireless networks', *Proceedings of IEEE Consumer Communication and Networking Conference,* NV, USA.

So-In, C., Jain, R., Paul, S. and Pan, J. (2010b) 'A policy oriented multi-interface selection framework for mobile IPv6 using the ID/locator split concepts in the next generation wireless networks', *Proceedings of the 2nd International Conference on Computer and Automation Engineering*, Singapore, pp.580–584.

Soliman, H. (2007) 'Mobile IPv6 support for dual stack hosts and routers (DSMIPv6)', Internet-Draft*,* draft-ietf-mip6-nemo-v4traversal-06.txt, Internet Engineering Task Force.

Soliman, H., Castelluccia, C., Malki, K.E. and Bellier, L. (2008) 'Hierarchical mobile IPv6 (HMIPv6) mobility management', RFC 5380, Internet Engineering Task Force.

Soliman, H. et al. (2009c) 'Flow bindings in mobile IPv6 and NEMO basic support', Internet-Draft, draft-ietf-mext-flow-binding-03.txt, Internet Engineering Task Force.

The 3rd Generation Partnership Project (2007a) 'The 3GPP technical specification group service and system aspects; general packet radio service (GPRS) enhancements for evolved universal terrestrial radio access network (E-UTRAN) access', 3GPP TS 23.401 V8.0.0, p.167.

The 3rd Generation Partnership Project (2007b) 'The 3GPP technical specification group service and system aspects; architecture enhancements for non-3GPP accesses', 3GPP TS 23.402 V8.0.0, p.131.

Wakikawa, R. et al. (2009) 'Multiple care-of addresses registration', Internet-Draft, draft-ietf-monami6-multiplecoa-14.txt, Internet Engineering Task Force.

Wang, H.J., Katz, R.H. and Giese, J. (1999) 'Policy-enabled handoffs across heterogeneous wireless networks', *Proceedings of IEEE Workshops on Mobile Computing and Applications,* LA, USA.