A Normalized Average Cost Model for ID/LOC Separation Networks

Jungha Hong and Heeyoung Jung
Internet Future Technology Research Department
Electronics and Telecommunications Research Institute (ETRI)
218 Gajeongno, Yuseong-gu, Dejeon, Korea
{jhong, hyjung}@etri.re.kr

Abstract—The current Internet has several known challenges, such as mobility, multihoming, routing scalability, traffic engineering, etc. due to the overloaded semantics of IP address, i.e. the IP address is used as a node identifier (ID) as well as a node locator (LOC). Thus, the research community has proposed the ID/LOC separation scheme. In this paper, we propose a normalized average cost model to analyze the ID/LOC separation networks assuming a local caching mechanism under a single source-destination model. We analyze the model assuming the general packet arrival process and the general sojourn time for destination, respectively to show the impact of the packet arrival process and the mobility pattern of the destination on the probability that a packet is delivered to the correct location of the destination.

Keywords—cost analysis; Future Internet; ID/LOC separation; mobility

I. INTRODUCTION

With wide popularity of smart phones and emergence of various wireless/mobile networks, the environment for Internet services has rapidly changed from fixed-based to mobile-based. It was also reported that the number of mobile users would be more than 1.6 billion in around 2014 and thus exceed the number of desktop users [1]. Such trends show that the mobile network environment is becoming a primary factor to be considered in designing the Future Internet. However, such mobile trend was not considered in the design of original Internet, which means that the current Internet has been designed for fixed hosts rather than for mobile ones. Thus, the current Internet has inherently had limitations on mobility support. For this reason, the mobility support in networks has been one of the main considerations in designing an addressing scheme and networking architecture for the Future Internet.

In the current Internet, the IP address is used as a node identifier (ID) as well as a node locator (LOC), which may make no problem in the fixed environments. However, the situation is quite different in the mobile environments since a node changes its IP address whenever it moves into another location. In order to support mobility, its LOC should be changed to be routable, whereas its ID should not be changed to keep the service continuity. The change of a moving node's ID also incurs the problem of the session continuity. Therefore, the overloaded semantics of IP address causes the difficulty in supporting mobility.

Another important problem of the current Internet is the continued growth of the BGP routing tables in the default-free zone (DFZ) [2]. A lot of solutions for this issue are based on the idea of separating the network node's identifier (ID) from its topological location (LOC).

There are several protocols to deal with the overloaded semantics issue of IP address. The Mobile IP (MIP) protocols [3, 4] use two types of IP addresses: home address (HoA) as an ID and care-of-address (CoA) as a LOC, where the mobility of a host is supported by the mobility agents such as Home Agent (HA) and Local Mobility Anchor (LMA), even when a host is away from its home domain. On the other hand, Host Identity Protocol (HIP) [5] and Locator Identifier Separation Protocol (LISP) [6] are new schemes to address the problems of IP address through separating ID and LOC. As a host-based scheme, HIP uses host identity tag for ID and IP address of host for LOC, whereas LISP is a network-based scheme that uses IP address of a host for Endpoint ID (EID) and IP address of a router for Routing LOC (RLOC).

Most ID/LOC separation schemes have been proposed mainly to make the BGP routing tables scalable. There are several cost models for the ID/LOC separation schemes focused on the aspect of the scalability issue of the current Internet. L. Iannone and O. Bonaventure [7] worked on the performance evaluation for caching of the mapping entries. They evaluated the cost of maintaining caches when the distribution mechanism was based on a pull model. Taking the LISP protocol as a reference, they based their evaluation on real Net-flow traces collected on the border router of their campus network. They also analyzed the impact of the locator/ID separation and related cost, showing that there was a trade-off between the dynamism of the mapping distribution protocol, the demand in terms of bandwidth, and the size of the caches. L. Iannone and T. Levä [8] proposed the LOC/ID Split cost model and analysis which was a useful tool in deriving the relation between important architectural parameters, incentives, and the possible deployment scenarios. The proposed analysis showed that to have a cost reduction and hence incentives to adopt LOC/ID Split, the sum of the cost of the cache, the connectivity infrastructure, and the mapping distribution system must be lower than the original cost of today's BGP Internet. They noted that the adoption process would offer more incentives if the interoperability technology greatly reduces the number of prefixes through aggregation.
Nevertheless, if the interoperability technology performs too well, late adopters can actually increase the overall cost, which may prevent complete adoption.

Our interest is to analyze the ID/LOC separation schemes on the aspect of the mobility issue, which has not been done much in the literature. V. Ishakian et al. [9] performed an average cost analysis to compute the mobility/multihoming support of RINA (Recursive INternet Architecture), against that of other approaches such as LISP and Mobile-IP. In [9], multihoming was viewed as a special case of mobility and a first cost comparison of these approaches was presented. The definition of "cost" captured both the average number of packets generated by a source node to a (mobile or multihomed) destination node and the average path length from the source to the destination (as indication of delays or bandwidth usage). They computed the overall average cost for a single interface change experienced by the mobile or multihomed destination node. They also validated their analytical model using simulation.

We remark that a local caching mechanism is not considered for the LISP architecture in [9] so the lookup cost is incorporated in the delivery cost of every single data packet. The lookup cost is also incorporated in the average total cost as a term as well as the delivery cost, which may cause confusion. However, in [10], we separate the lookup cost from the delivery cost assuming a local caching mechanism and consider the source mobility as well as the destination mobility. We also provide the criterion for judgment of cost-effectiveness by comparing the average total cost of typical ID/LOC separation schemes and the one of the mobility support protocols such as MIP in [10].

In this paper, we generalize the cost model presented in [9] on the aspect of mobility support. Here, we emphasize our contributions as follows:

- We propose a normalized average cost model for ID/LOC separation networks.
- We generalize the inter-arrival times of data packets received by destination and the sojourn time of the destination in a location to show the impact of the packet arrival process and the mobility pattern of the destination on the probability that a packet is delivered to the correct location of the destination.
- We assume a local caching mechanism and take into account a Time to Live (TTL) associated with the cache entry.

The remainder of this paper is organized as follows. We briefly review the Locator Identifier Separation Protocol (LISP) and Host Identity Protocol (HIP) in Section 2. We propose a mathematical model and analyze the normalized average cost for ID/LOC separation networks in Section 3. We present some numerical results and conclude this paper with future work in Section 4 and Section 5, respectively.

II. OVERVIEW OF LISP AND HIP

The current Internet has several known challenges, such as mobility, multihoming, routing scalability, traffic engineering, etc. due to the overloaded semantics of IP address, i.e. the IP address is used as a node identifier (ID) as well as a node locator (LOC). Thus, the research community has proposed the ID/LOC separation scheme which is based on the idea of separating the network node's identifier from its topological location [11]-[13] and one of the most successful proposals is the Locator/ID Separation Protocol (LISP) at the Internet Engineering Task Force (IETF) [14].

The identity locator separation proposals are categorized as either map-and-encap or address rewriting scheme. The proposals differ in how the indication is achieved, through either the use of tunneling or address rewriting. In a map-and-encap scheme, packets are delivered using a tunneling mechanism where packets are encapsulated with an extra IP header when routed in the network. In an address rewriting scheme, the address in the IP header is rewritten instead of using of encapsulation.

In this section, we briefly provide an overview of Locator/ID Separation Protocol (LISP) and Host Identity Protocol (HIP) as examples of map-and-encap and address rewriting schemes, respectively.

A. Locator/ID Separation Protocol (LISP)

LISP [6] is designed to be a simple network-based map-and-encap protocol that implements separation of Internet addresses into Endpoint Identifiers (EIDs) and Routing Locators (RLOCs). EIDs are assigned to end-hosts independently from the network topology, used for numbering devices, and aggregated along administrative boundaries. RLOCs are topologically assigned to network attachment points (and are therefore amenable to aggregation) and used for routing and forwarding of packets through the network.

The basic idea of LISP is to tunnel the core network to deliver packets from RLOC_a (the RLOC of the source EID) to RLOC_d (the RLOC of the destination EID), as depicted in Fig. 1. When a host in a LISP-capable domain wants to send a packet, it first looks up EID_a, the destination host's EID, in the Domain Name System. It then puts its EID, EID_a, in the packet source address, and EID_d in its destination address. If the destination of the packet is in another domain, the packet traverses the source domain infrastructure to one of the domain ITRs. The ITR performs the EID-to-RLOC lookup in its local cache, or queries the mapping system if no mapping is available in the cache. Then, the ITR sets the destination RLOC in the packet using the result of the lookup or querying, and the source RLOC to RLOC_a (its RLOC). The packet is then sent over the core network to the ETR indicated in the destination RLOC, which decapsulates the packet and sends it on to the destination EID.

Indeed, LISP uses a local caching mechanism to reduce the frequency of lookup and latency so that only the first packet may trigger a query to the mapping system. In other words, an ITR maintains a table that stores some recently used EID-to-RLOC mappings. In general, each mapping has a time-to-live (TTL) value that is originally set to a cache timeout. If the mapping is used before its TTL expires, the TTL for the mapping is reset to be the given cache timeout; otherwise, the mapping is removed from the cache.
B. Host Identity Protocol (HIP)

HIP [5] separates the identity of a host from its location. The location of the host is bound to IP addresses and used for routing packets to the host in the same way as in the current Internet architecture. However, transport and application layers use host identity (HI), consisting of the public key component of a private-public key pair. Each host is responsible for creating one or more public/private key pairs to provide independent identifiers that allow a mobile host to preserve its transport layer connections upon movement. On the other hand, the host identity can be used for looking up the current location of a host because the host identity is a long-term identifier. HIP uses Domain Name System (DNS) and also a rendezvous server (RVS) which holds the latest IP address for a given HI.

In order to start communicating through HIP, two hosts must establish a HIP association which is known as the HIP Base Exchange, as depicted in Fig. 2. A CN(correspondent node) obtains the host identity of a server, the RVS IP address of the MN (mobile node), typically from the DNS. Then the CN will initiate a 4-way handshake with the MN. The first packet during the handshake is sent to the RVS which will forward it to the MN. When the MN receives this packet it will reply directly to the CN, and from now on packets will be forwarded directly between the MN and CN without passing through the RVS in order to complete the handshake and to allow data to flow. When the MN moves, it will update its RVS with its new address and it will send an Update message to the CN.

III. Cost Model and Analysis

In this section, we propose an analytical cost model to quantitatively assess the effectiveness of ID/LOC separation networks in supporting mobility.

Our model may seem to be similar to the cost model proposed in [9]. However, the model in [9] is on LISP, MIP, and RINA (Recursive Internet Architecture) in supporting multihoming, where the overall average cost for a single interface change experienced by the mobile or multihomed destination node is computed and mobility is considered a special form of multihoming. We also remark that it is assumed that the location information is not cached at the source TR (Tunnel Router) for the LISP architecture in [9], so the lookup cost is incorporated in the delivery cost of every single data packet. In our model, however, we present a normalized average cost to analyze the ID/LOC separation networks assuming a local caching mechanism.

A. Model Description

We assume a single source-destination model where the destination may move from one location to another. We also assume that the source sends data packets to the destination at a constant rate. We define the followings:

- \( X \): random variable of the inter-arrival times of data packets at destination with mean \( 1/\lambda \).
- \( Y \): random variable of the sojourn time for destination node in a location with mean \( 1/\mu \), where the sojourn time denotes the amount of time that a node spends in a location before moving into a new location.
- \( R \): random variable of the remaining time of \( Y \), i.e. the remaining time during which the destination does not change the location.
- \( F_X(x) \), \( F_Y(y) \), \( F_R(r) \): distribution functions of \( X \), \( Y \), and \( R \), respectively.
- \( f_X(x) \), \( f_Y(y) \), \( f_R(r) \): probability density functions of \( X \), \( Y \), and \( R \), respectively.
- \( F_X(s) \), \( F_Y(y) \), \( F_R(r) \): Laplace Transform (LT) of \( X \), \( Y \), and \( R \), respectively, i.e.

\[
F_X(s) = E[e^{-sx}],
\]

where \( E[\cdot] \) denotes the expected value of the argument.

From the residual life theorem [15], we have

\[
F_R^*(s) = E[e^{-sr}] = \frac{1 - F_Y(s)}{s E[Y]}.
\]

Now, we define \( p \) as the probability that the destination has not moved into another location since the last packet delivery. In other words, \( p \) is the probability that a packet is delivered to the correct location of the destination. Then, it can be obtained as

\[
p = \text{Prob} \{ X < R \}.
\]

We also define \( q \) as the probability that TTL (Time to Live) has not been expired since the last cache hit. Then, it can be obtained as...
where \( d \) is a TTL value.

We finally define the normalized average cost \( C_N \) as the average total cost incurred during a unit time. Then, \( C_N \) is obtained in terms of the location update cost \( C_L \), the location lookup cost \( C_U \), and the packet delivery cost \( C_D \). We note that \( C_L \) is the cost of updating the location (routing) information of a node whenever it moves into another location, \( C_U \) is the cost of looking up the information of ID-to-LOC mapping for the destination by querying, and \( C_D \) is the cost of delivering a single packet from source to destination.

We can consider the case that a data packet is delivered to the wrong location due to the wrong location information of the destination, which occurs with the probability \( 1 - p \). In this case, the packet needs to look up the correct location and be delivered again. Thus, the redelivery cost of a packet is denoted by \( C_R \).

Now, we can consider all events happening during a unit time and their costs as follows:

- The destination moves \( \mu \) times so that it would need to update its location \( \mu \) times; \( \mu C_U \).
- The average number of packets arriving at destination is \( \lambda \).
- If ID-to-LOC mapping is cached with the probability \( q \), then the lookup process is not needed. Otherwise, looking up for ID-to-LOC mapping is required with the probability \( 1 - q \) and cost \( C_U \).
- Packets are delivered to the correct location with the probability \( p \); \( C_D \).
- Packets are delivered to the wrong location with the probability \( 1 - p \) so that they would need to look up the correct location and be delivered again; \( C_D^* \), where we note that
  \[
  C_D^* = C_D + C_L + C_U
  \]
Putting it all together, we therefore obtain the normalized average cost \( C_N \) as

\[
C_N = \mu C_U + \lambda [q(p C_D + (1 - p) C_D^*)] + (1 - q)(C_L + p C_D + (1 - p) C_D^*)].
\]

It can be simplified as

\[
C_N = \mu C_U + \lambda [(1 - q) C_L + p C_D + (1 - p) C_D^*].
\]

B. Mathematical Analysis

In this subsection, we compute the probability \( p \) and \( q \) given in (2) and (3) assuming the general distributions of \( X \) and \( Y \), respectively.

1) Case of the general arrival process

We first assume that the inter-arrival time \( X \) of packets has a general distribution with mean \( 1/\lambda \) and the sojourn time \( Y \) for destination is exponentially distributed with mean \( 1/\mu \). Then, the remaining time \( R \) of \( Y \) is also exponentially distributed with mean \( 1/\mu \) by memoryless property [15]. Therefore, we get

\[
p = \text{Prob}[X < R]
\]

\[
= \int_0^{\infty} \text{Prob}[X < R \mid X = x] f_X(x) \, dx
\]

\[
= \int_0^{\infty} [1 - F_R(x)] f_X(x) \, dx
\]

\[
= \int_0^{\infty} e^{-\mu x} f_X(x) \, dx
\]

\[
= F_X(\mu),
\]

\[
q = \text{Prob}[X < d] = F_X(d).
\]

2) Case of the general mobility pattern of destination

Now, we assume that the inter-arrival time \( X \) of packets is exponentially distributed with mean \( 1/\lambda \) and the sojourn time \( Y \) for destination has a general distribution with mean \( 1/\mu \). Then, we get

\[
p = \text{Prob}[X < R]
\]

\[
= \int_0^{\infty} \text{Prob}[X < R \mid R = r] f_R(r) \, dr
\]

\[
= \int_0^{\infty} F_X(r) f_R(r) \, dr
\]

\[
= \int_0^{\infty} [1 - e^{-\lambda r}] f_R(r) \, dr
\]

\[
= 1 - F_R(\lambda).
\]

From (1), we therefore get

\[
p = 1 - \frac{\mu [1 - F_V(\lambda)]}{\lambda},
\]

\[
q = \text{Prob}[X < d] = 1 - e^{-\lambda d}.
\]

In the next section, we present some numerical results to show the impact of the packet arrival process and the mobility pattern of the destination on the probability \( p \).

IV. NUMERICAL RESULTS

In this section, we apply different distributions to the packet inter-arrival time and the sojourn time for destination in order to show the impact of those on the probability \( p \).

We first define \( \alpha \) as the average number of movement of destination during the average inter-arrival time, i.e.,

\[
\alpha = \frac{\mu}{\lambda}.
\]

We remark that \( \alpha \) denotes the ratio of the destination mobility rate to inter-arrival time rate.

We have used a different squared coefficient of variation (\( c^2 \)), variance divided by mean squared, for a different distribution of the packet inter-arrival time, \( X \). For \( c^2 = 0.5 \), Erlang-2 distribution has been used, i.e.

\[
F_X(s) = \left( \frac{2\lambda}{s + 2\lambda} \right)^2.
\]

Thus, we get
For $c^2 \geq 1$, we have used hyper-exponential distribution with parameters $p_1, p_2 \ (= 1 - p_1), \lambda_1, \lambda_2$ and assumed that

$$p = \left(\frac{2}{\alpha + 2}\right)^2.$$  

(i.e., the hyper-exponential distribution has "balanced mean") leaving us with two degrees of freedom for the determination of its parameters. As long as $c^2 \geq 1$, we can use arbitrary mean and variance (satisfying the obvious moment inequalities) to find its parameters. Since the mean is assumed as $1/\lambda$, we have

$$F_x(s) = \frac{1}{2} \left( \frac{1 - \frac{c^2 - 1}{\sqrt{c^2 + 1}}}{\lambda_1 s + \lambda_1} + \frac{1 - \frac{c^2 - 1}{\sqrt{c^2 + 1}}}{\lambda_2 s + \lambda_2} \right),$$

where $p_1 = \frac{1}{2} \left( 1 + \frac{c^2 - 1}{\sqrt{c^2 + 1}} \right)$, $p_2 = 1 - p_1$, $\lambda_1 = \frac{\lambda}{1 + \sqrt{c^2 + 1}}$, and $\lambda_2 = \frac{\lambda}{1 - \sqrt{c^2 + 1}}$. Thus, we get

$$p = \frac{2(1 + \alpha c^2)}{2[1 + \alpha (1 + c^2)] + \alpha^2 (1 + c^2)}.$$  

For $c^2 = 1$, which is exponential distribution, we get

$$p = \frac{1}{1 + \alpha}.$$  

Fig. 3 shows the impact of the packet arrival process on the probability $p$ with respect to $\alpha$, the ratio of the destination mobility rate to inter-arrival time rate. It shows that as $\alpha$ increases, the value $p$ decreases due to the higher mobility. It presents the probability $p$ versus the ratio $\alpha$ for various $c^2$ with the same mean $1/\lambda$. As $c^2$ increases, the value $p$ decreases, which means for higher variability of the packet arrival process, we get the higher normalized average cost ($C_B$) since the factor $(1 - p)c^2$ in (4) increases and $C_B^\alpha$ includes $C_B$. We note that the similar results have been observed for general sojourn time for the destination with Poisson arrival process.

V. CONCLUSIONS

We proposed a normalized average cost model to analyze the ID/LOC separation networks assuming a local caching mechanism under a single source-destination model. We analyzed the model assuming the general packet arrival process and the general sojourn time for destination, respectively to show the impact of the packet arrival process and the mobility pattern of the destination on the probability $p$ that a packet is delivered to the correct location of the destination.

We plan to work on getting some experimental results by applying specific ID/LOC separation schemes to the proposed model.

ACKNOWLEDGMENT

This research was supported by the KCC (Korea Communications Commission), Korea, under the IT R&D support program supervised by the KCA (Korea Communications Agency). (KCA-2011-10913-05004)

REFERENCES


