Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 23, 2016

S. Matsushima
R. Wakikawa
SoftBank

March 22, 2016

Stateless user-plane architecture for virtualized EPC (vEPC)
draft-matsushima-stateless-uplane-vepc-06

Abstract

We envision a new mobile architecture for the future Evolved Packet Core (EPC). The new architecture is designed to support the virtualization scheme called NFV (Network Function Virtualization). In our architecture, the user plane of EPC is decoupled from the control-plane and uses routing information to forward packets of mobile nodes. Although the EPC control plane will run on hypervisor, our proposal does not modify the signaling of the EPC control plane. The benefits of our architecture are 1) scalability, 2) flexibility and 3) Manageability. How to run the EPC control plane on NFV is out of our focus in this document.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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

This Internet-Draft will expire on September 23, 2016.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents

carefully, as they describe your rights and restrictions with respect 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

3GPP introduces Evolved Packet Core (EPC) that is fully IP based mobile system for LTE and -advanced in their Release-8 specification and beyond. Operators are now deploying EPC for LTE services and encounter rapid LTE traffic growth. There are various activities to offload mobile traffic in 3GPP and IETF such as LIPA, SIPTO and DMM. The concept is similar that traffic of OTT (Over The Top) application is offloaded at entity that is closer to the mobile node (ex. eNodeB or closer anchor).

Likewise, overload of signaling (control plane) is also increasing day by day. Network operators expect recent innovation and trends of NFV (Network Function Virtualization) to solve this overloaded control plane. NFV is discussed at the ETSI NFV ISG and is introduced in [NFV-WHITEPAPER]. Mobile operator's network is built

with variety of proprietary hardware appliances today. If we can get rid of these physical appliances and could shift to a cloud-based service, we will have a lot of benefits explained in the next section. This document assumes that NFV will push networking functions currently run on dedicated hardware onto a cloud network. Expected network functions are Mobility Management Entity (MME), Serving Gateway (SGW) PDN Gateway(PGW), etc. With NFV, EPC can be operated onto servers/hyper-visors. We name it virtualized-EPC (vEPC) in this document.

This document uses a lot of 3GPP specific terms. These terms can be found mostly at [RFC6459].

1.1. The Benefits of NFV

This section briefly explains the benefits of NFV. The detailed benefits can be found in [NFV-WHITEPAPER]. Although today's eco-system of EPC appliances might be affected, we believe there are various approaches to enhance current eco-system and migrate to new NFV approaches. For example, operators could pay monthly recurring charges for the NFV services and operations to vendors, instead of one-time purchase and a little maintenance cost.

- [Flexible Network Operations]: The control functions of EPC are no longer in appliances deployed widely in operator's network and can be run at hypervisor (cloud). It is easier to add and/ or delete functions from the services, because no physical construction is needed. Network operations will be much simpler and easier because complications of today's network are pushed to NFV (i.e. hypervisor).

- [Flexible Resource Managements]: The EPC functions can be run on hypervisor and are now less dependent on proprietary hardware. Adding additional resources is easier in hypervisor, while adding or replacing physical appliances require installation, construction, configuration, and even migration plan without service cutoff. A hypervisor can be also shared across various functions such as PGW, SGW and MME. NFV also brings multi-tenancy and allows a single platform for different services and users. The operator can optimize resources and costs to share a NFV platform for multiple customers (ex. MVNO customers) and services (ex. multiple APNs).

- [Faster Speed of Time to Market]: When an operator wants a new function to its network and services, the operator needs to negotiate appliance vendors to implement the new functions or to find alternative equipment supporting the new function. It takes a longer time to convince the vendors, or to replace existing

hardware. However, if functions can be implemented as a software, it is much faster to implement the functions on NFV. Even the operator may implement them and try the new functions by themselves. Field trial is also getting easier because of no physical installation or replacement. You may turn on a new function in NFV and observe how the new function behaves in your network. NFV can save preparation time and tuning time of the new function.

- o [Cost Optimization]: Last but not least, Cost is the most important motivation for operators to realize NFV. Operators can remove many of proprietary appliances from its network and replace them with industry standard servers, switches and routers. In addition, it is easy to scale up and down operator's services so that resources can be always tuned to the size of services. In addition, operational costs led by any physical hardware such as power supply, maintenance, installation, construction and replacement can be minimized or even removed. The network design can be simpler, because complicated functions could be handled by NFV. That simple operation may enable automatic configurations and prevent unnecessary trouble-shooting. As a result, CAPEX and OPEX can be always optimized and lowered.

2. Motivations and Requirements, - Why IETF? -

2.1. Motivations

What is a role of IETF to realize vEPC in the future? IETF is not the right place to discuss, for instance, how to run MME on hypervisor. An important IETF activity must be to decouple the control- and user- planes of mobility protocols used in EPC.The motivation of decoupling the user and control plane is discussed in [I-D.wakikawa-req-mobile-cp-separation]. In doing so, NFV-enabled solutions can be easily designed and implemented with interoperability across multiple vendors and platforms. Otherwise, NFV solutions can be easily fragmented due to many proprietary solutions for the protocol separations. As stated in [NFV-WHITEPAPER], interoperability is highly important.

In the past, IETF has developed tunnel based mechanisms for mobile nodes such as Mobile IPv6 [RFC6275][RFC5555], Proxy Mobile IPv6 [RFC5213][RFC5844] and NEMO [RFC3963]. Similarly, 3GPP has developed tunnel protocols called GPRS Tunneling Protocol (GTP). These tunnel-based protocols establish a data path for a mobile node between the mobile node and an anchor point (s). There is a case where an access router terminates a tunnel instead of a mobile node (ex. Proxy Mobile IP). In 3GPP, a tunnel is established between SGW and PGW per a mobile node by either Proxy Mobile IPv6 or GTP. The control and

the user planes of these mobility protocols are tightly related and cannot be decoupled. The signaling like Binding Update and user's packets are routed along a same path in EPC. It might be necessary to extend these mobility protocols for the user- and control- planes separation. The protocol separation of Mobile IP is discussed in [I-D.yokota-dmm-scenario].

Alternatively, if vEPC was realized, we should have an opportunity to re-visit the basic architecture of mobility system. Instead of tunneling packets on today's EPC, why can't we just route packets to a mobile node? Since a role of the user plane is "routing", BGP and other routing protocols could be used to forward UE's traffic. This document introduces a BGP-based solution. Software Defined Networking (SDN) can be an alternative solution. Open Flow and other relevant protocols can setup the forward path dynamically according to UE's states available in the control plane.

We have to remember that there is a good reason of adapting tunneling in Mobile IP based solutions, that is global mobility and signaling. A mobile node should be able to move anywhere on the Internet and be reachable from anyone on the Internet. There were routing based global mobility solutions like Boeing global mobility [Boeing-BGP] and WINMO [RFC6301]. In these proposals, BGP was used to propagate forwarding information of mobile nodes to the Internet. Whenever a mobile node changes its point of attachment, the route must be updated. Due to scalability and stability issues of the Internet, this solution was not recommended by IETF [Boeing-BGP]. However, as Boeing showed, it is doable to support global mobility by using BGP routing update. If scalability is not your concern, a routing based approach becomes a candidate of the mobility solution.

While global mobility is important, the "reality" is that your cell phones (i.e. UE/mobile node) are moving just within an operator's network and fully controlled in your local EPC. If mobility is limited within an operator, we believe a routing based approach is feasible and practical for today's mobile system. Instead of dedicated proprietary equipment like SGW and PGW to manage a tunnel path for a mobile node, multiple industry standard routers and switches are configured in the user plane. These switches and routers receive mobile nodes' forwarding information from the control plane of vEPC by routing update.

2.2. Requirements

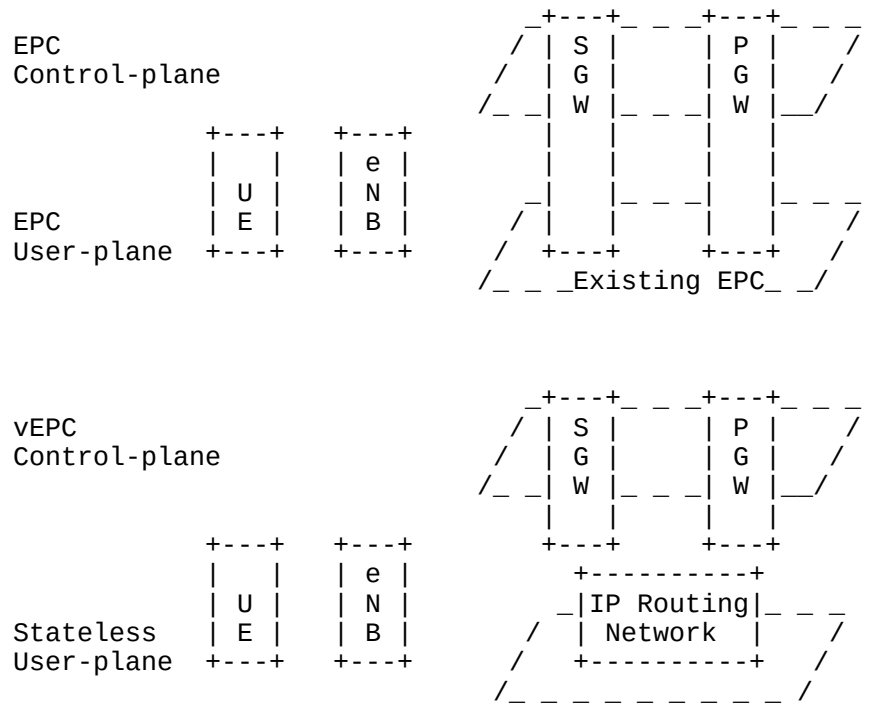
Requirements of our stateless user plane for vEPC are followings.

NFV Support

The future EPC architecture must support NFV capability. The control plane of EPC operated on NFV framework is named "virtualized EPC (vEPC)" in this document. The control plane of vEPC should keep backward compatibility with the today's EPC's control plane. It means this document doesn't modify the control plane at all. It only assumes software-based MME, SGW, and PGW run on hypervisor.

Separation of Control- and User- Planes

Due to tight relationship of the control- and user- planes in today's EPC, resource increase is always provisioned to both planes at once. It prevents flexible resource arrangement and introduces high capital investment and over-provisioned resources to one of planes. If NFV is deployed, it is expected that computing resources can be independently allocated to the control planes of the vEPC in a flexible manner.



NFV enabled EPC architecture

Figure 1

Figure 1 shows a possibility that the entities of EPC Control- plane are virtualized in generic cloud environment, however user packets won't go through those virtualized EPC nodes. Decoupling User-plane from the Control-plane entities will be made virtualized Control-plane nodes relax hyper- visor data- path capacity requirements. On the other hand, decoupled User-plane into IP routing network will be agnostic from sessions and bearers states, of which are generated and maintained in the Control-plane. In terms of IP routing, forwarding packets through the networks is based on the destination address of the packets evaluated with network reachable information in the routing table that accommodated in the routing nodes. To forward EPC User-plane packets correctly, those states must be indicated by network reachable information.

Flat Design for Distributed Operation

Today's 3GPP architecture introduces PDN gateway (PGW) as a gateway to external networks like the Internet. PGW manages all traffic from and to UEs and could be a bottleneck and single point of failure of network connectivity. In addition, due to recent rapid traffic increase, it is important to perform traffic engineering and to offload traffic to multiple locations (ex. SGW, PGW, eNodeB). For enhancements of traffic engineering capability, more flat design with multiple gateways is expected so that traffic can be distributed to all these gateways. There were proposals how to enable flat design to (Proxy) Mobile IP such as [I-D.wakikawa-mext-haha-interop2008] in IETF. Distributed Mobility Management (DMM) Working Group has also discussed how to extend Mobile IP-based solutions to support traffic distribution in an optimal way by removing centrally deployed anchors that is like a Home Agent.

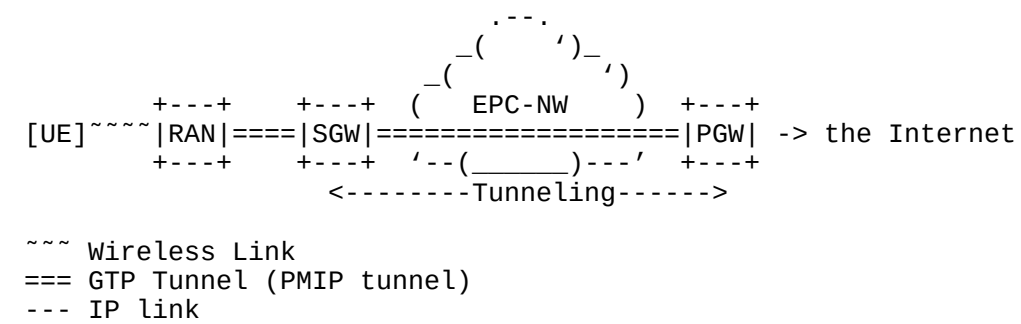
Stateless in User Plane
: Ultimate goal of vEPC is to remove all mobility specific states from the forwarding nodes in the user-plane of vEPC. If we succeed in this, industry standard routers and switches can be used to forward mobile nodes traffic in the user plane of vEPC. A mobile node's specific states are kept in both an IP header of the mobile node's packets and a routing entry of the mobile node. The detail is described in Section 3.2

3. Stateless user-plane architecture for virtualized EPC

This section explains our solution that is the stateless user-plane architecture for vEPC. This solution is basically a combination of existing protocols defined in IETF. A minor extension might be needed but it should be easily addressed in IETF. We first introduce our architecture and then protocol overview.

3.1. Architecture Overview

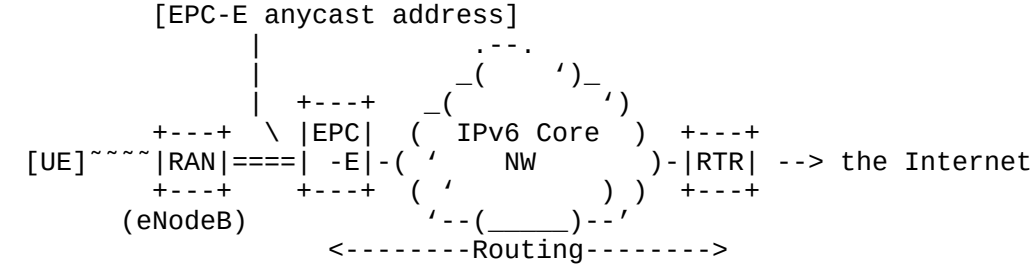
Figure 2 shows the user plane of the current EPC architecture. A tunnel is established between SGW and PGW by either Proxy Mobile IP or GTP. PGW is an anchor point of UE for incoming packets. All the packet destined to UE is routed first to PGW. The UE's packets are intercepted by PGW and tunneled to SGW. SGW then forwards the packet to UE via access points (i.e. eNodeB) over Radio Area Network (RAN).



User plane of the current EPC

Figure 2

Figure 3 is our proposed user plane of vEPC. The control plane is not shown in this figure.



User plane of vEPC

Figure 3

We introduce two new entities such as

EPC Edge Router (EPC-E)
EPC-E is located at the same place of today’s SGW and terminates GTP tunnel established with eNodeB (RAN). EPC-E supports the user plane functions of SGW and PGW. EPC-E is configured an anycast address to the network interface facing to eNodeB. The eNodeB establishes a GTP tunnel per UE with this anycast address. Thanks for anycast address, UE’s traffic forwarded by eNodeB is always routed to the closest EPC-E of UE. EPC-E is a router and

maintains routing information of every UE that is notified by the control plane. Detail of routing mechanism can be found in Section 3.4.

Router (RTR)
It is a regular IP router. The control plane of vEPC distributes routing information of every UE by a routing protocol like BGP. Therefore any additional protocols other than routing protocols are not needed for RTR. Multiple RTRs can be configured anywhere in the user plane of vEPC. RTRs announce UE's routing information to the external network (ex. The Internet).

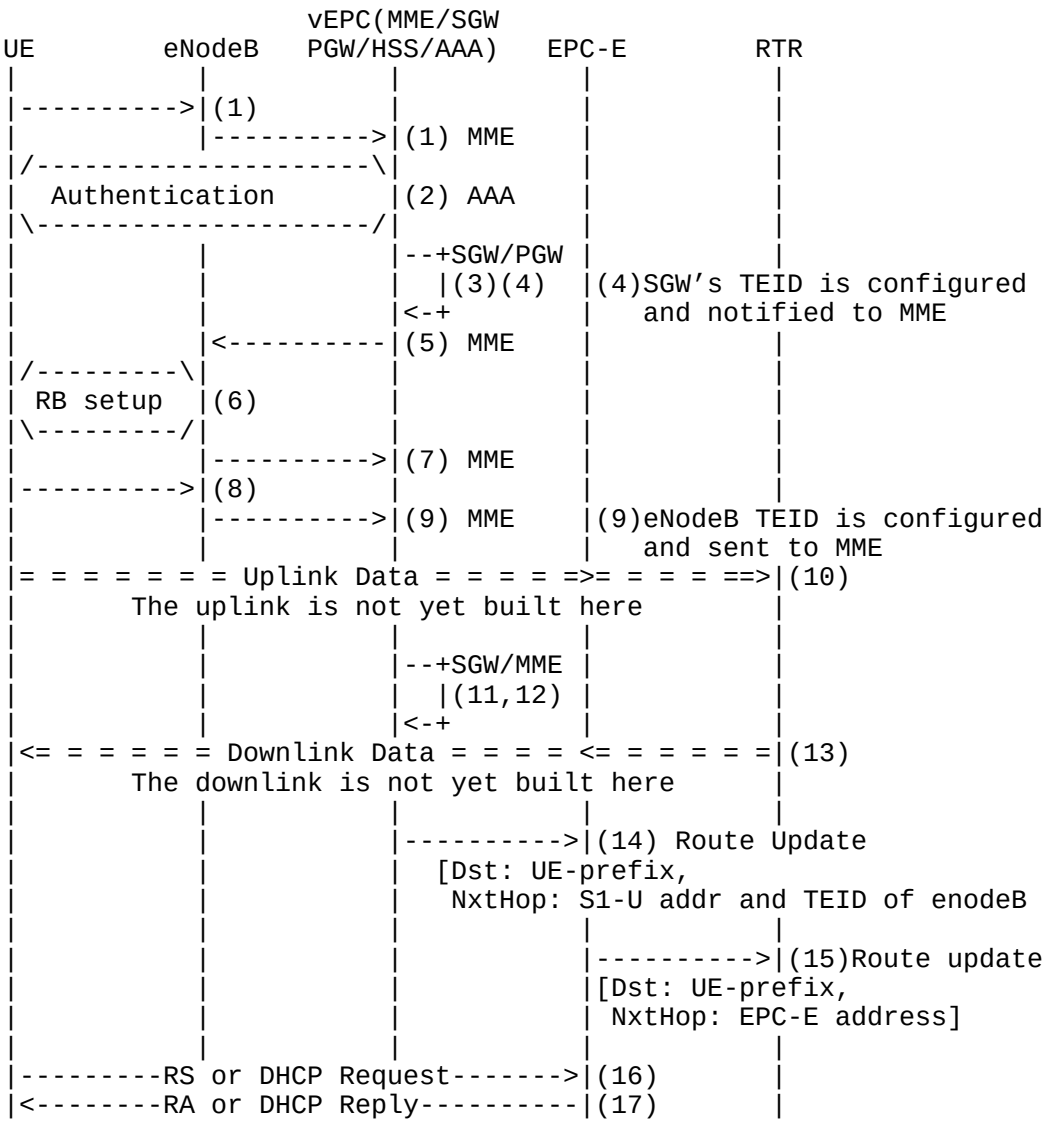
As you see in Figure 3, we omit a tunneling mechanism originally established between SGW and PGW for routing UE's packets in the user plane. By removing this tunnel, UE's packets are forwarded to and from the Internet according to routing tables on routers in the core network. Note that, although we remove the tunnel for UE's traffic in the user plane, the control-plane signaling stays same in the control plane. If Proxy Mobile IP is used for this tunnel, Proxy Binding Update and Acknowledgment are exchanged between PGW and SGW that are managed by NFV on servers/hyper-visor. Instead of a tunnel setup, states created by Proxy Mobile IP are distributed to all routing entities (EPC-E and RTR) by a routing protocol. From the user plane point of view, these states are just seen as routing entries. EPC-E and RTR are not involved in any signaling of the control plane. The control plane just injects routing information to EPC-E and RTR to setup routing paths to and from UEs.

Although this architecture just uses IPv6 core network, it supports both IPv4 and IPv6 packets. The detailed operation of IPv4 support will be discussed in Section 3.5.

3.2. Protocol Overview

This section gives an example of protocols used for vEPC. Figure 4 is the procedure of the PDN connection setup in vEPC. This figure is copied from the section 3 of [RFC6459]. All the steps from (1) to (13) are same as the original except for NFV-based MME, SGW, PGW, HSS, and AAA.

The vEPC introduces two new steps, (14) and (15), to setup paths in the user-plane after finishing all the signaling on the control-plane. (16) and (17) are the steps to assign IP address to the mobile node.



Extended PDN Connection Setup Procedure (copied Figure 8 of RFC6459)

Figure 4

In (14), vEPC advertises a routing information of UE to EPC-Es immediately right after the control-plane signaling completion. The routing information contain UE's prefix as destination, remote

endpoint of GTP-U tunnel as next-hop which is S1-U addresses and TEIDs of serving eNodeB/EPC-E, and also QoS class applied to the UE.

In this document, the advertising entity is a BGP speaker so that the BGP speaker is required to indicate those in BGP message. To achieve that, the BGP speaker and EPC-E should be capable of (1) BGP Tunnel Encaps Attribute [I-D.ietf-idr-tunnel-encaps] which specifies the form to encode GTP-U endpoints, and (2) Dissemination of Flow Specification Rule [RFC5575] with IPv6 amendment [I-D.ietf-idr-flow-spec-v6] to indicate applied QoS class.

It is noted that the control-plane needs to expose user-plane information of UEs to BGP speaker. The means of how the control-plane and the BGP speaker deal with that is discussed in Section 3.6.

The EPC-E has a peering with the BGP speaker directly. It is thus expected that there is no additional propagation delay of traversing multiple BGP speakers between EPC-E and vEPC. Adding that kind of surplus delay affects user-plane to be interrupted so that it should be avoided as much as possible for user experience.

In step (15), the EPC-E advertises routes to upstream routers such as the RTR. For scalable routing operation, UE’s prefixes should be aggregated into more shorter length prefixes. Due to that reason, the EPC-E generates routing information and advertised it to the RTR that includes aggregated prefix instead of UE’s prefixes and EPC-E address as the next-hop.

UE requests an IPv6 prefix for its address assignment in the step (16). In our architecture, an IPv6 prefix is still assigned by vEPC in the control plane, as PDN-GW does in the legacy EPC. However, EPC-E is responsible to deliver the IPv6 prefix to UE by DHCP or Stateless address autoconfiguration (SLAAC).

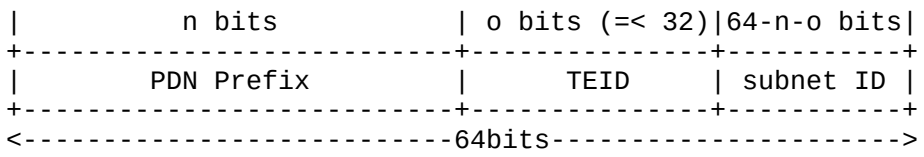
We now explain how EPC-E can know the prefix assigned to UE from vEPC for address configuration steps (16 and 17). When (1) to (15) are completed, vEPC has already advertised the UE’s prefix as route information to all the EPC-E. Therefore, when EPC-E receives a packet of either Router Solicitation or DHCPv6 request message, it just looks up the remote next-hop field of its routing information base (RIB) with the source IP address and the TEID of the received packet. A route entry matched for this search is the prefix delegated to the requesting UE. Therefore, EPC-E simply uses the prefix of the route entry as an assigned UE’s prefix.

In (17), EPC-E returns the found prefix to UE by either Router Advertisement or DHCPv6 reply message. UE now creates an address(es) from the received prefix. It is important to highlight that UE can

obtain the same prefix information from any EPC-E all the time because the same UE's route information is available on all the EPC-E.

It would be convenient to use automatic UE's prefix creation rule or algorithm for vEPC. There are various mechanisms to create UE's prefix. As an example, Stateless IPv6 Prefix Delegation [I-D.savolainen-stateless-pd] is introduced as an algorithm to create UE's prefix in vEPC below. It important to mention that our architecture of the stateless user plane does not rely on any particular prefix creation mechanisms like [I-D.savolainen-stateless-pd] and can be run with any of them.

In the case of an UE's prefix length is equal, or shorter than /64, the generated prefix is consisted as shown in Figure 5. Each PDN is assumed to have single or several prefixes (named PDN prefix) used to generate UE's address. Followed by the PDN prefix, there is TEID field assigned for a UE's session on S1-U interface of vEPC. TEID is 32 bits identifier in GTP header to distinguish each bearer. The remaining bits are filled by subnet ID.



Stateless-pd Prefix

Figure 5

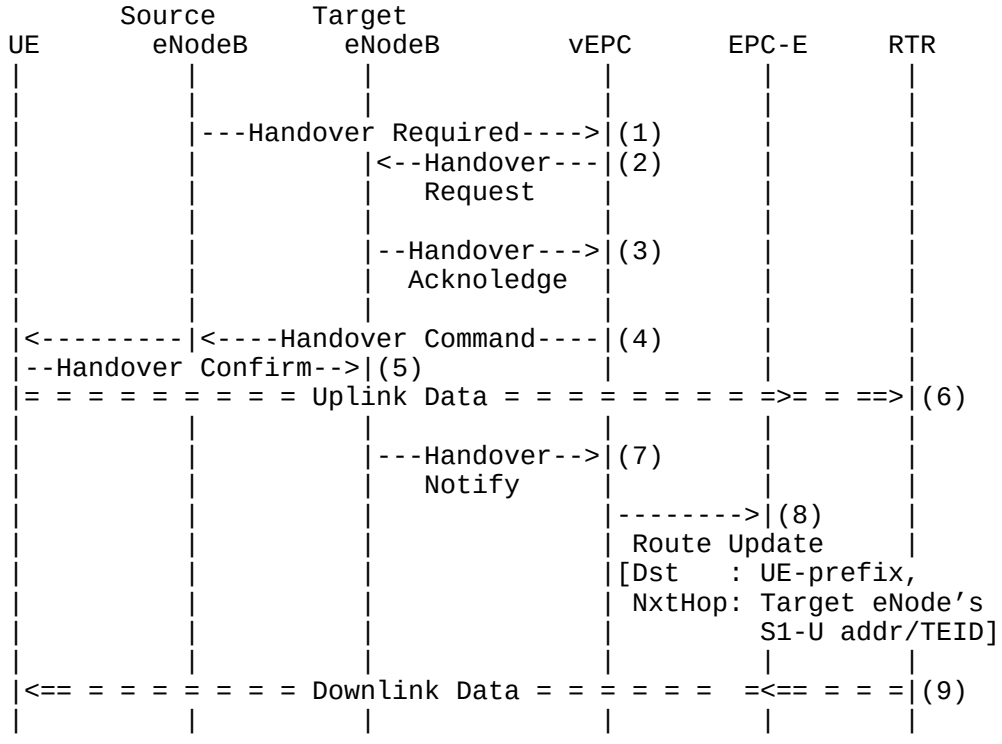
3.2.1. Hand-over

When tunnel endpoint is updated by UE hand-over between eNodeBs, vEPC must refresh the route of UE with the updated tunnel endpoint as new remote next-hop.

Figure 6 shows vEPC that advertising updated route in (8) when UE hand-over from source eNodeB to target eNodeB on simplified hand-over procedure. The updated route that points to target eNodeB's S1-U address and TEID as the next-hop should be immediately advertised to all the EPC-Es right after the procedures (1) to (7) completed.

It is noted that RTR or any upstream routers of EPC-Es do not require routing update for each of UE hand-over event. EPC-E is required to just advertise once aggregate route during at least an UE route exist

so that EPC-E does not advertise hand-over UE route in Figure 6. Operators require that their core network must be kept its routing stable. This architecture prevents routing fluctuation in the network that helps to fulfill that requirement consequently.



Simplified Hand-over Procedure

Figure 6

3.2.2. Detaching UE

In the case of UE detachment, vEPC also advertises route update that includes detached UE prefix as withdrawn route to delete the route of the detached UE from EPC-Es.

3.3. Control-plane awareness of stateless user-plane

Nodes in the control-plane in vEPC must be aware that the anycast address assigned to EPC-E is a S1-U address of vEPC. The vEPC must use the anycast address in signaling between vEPC and RAN. By doing

this, packets from RAN are correctly forwarded to an appropriate EPC-E. Due to anycast nature, it means there is no hand-off procedure between SGWs because all eNodeB in the RAN send packets to the same anycast address.

When an operator needs to increase virtualized instances to cope with just signaling overload, the operator should use the existing S1-U address (i.e. EPC-E anycast address) for the new instances. If the operator would increase the capacity of the user plane, it can add additional EPC-Es in the core network. The operator can group the new EPC-Es as a set and increase scalability and performance of the user plane. In this case, the operator uses a new anycast address to the new set of EPC-E. We will discuss operational consideration in detail in Section 4.

3.4. Routing mechanism

Figure 7 shows a packet forwarding mechanism of our stateless user plane. As an example, there are four eNodeB (illustrated as eNB-x) , three EPC-Edge routers(shown as EPC-Ex) and two routers (RTRx) in Figure 7. UE is first connected to eNB-C and then moves to eNB-D. The UE at the new location is illustrated as UE'. Routing entry for UE is also illustrated at the right side in Figure 7.

EPC-E has two interfaces facing either RAN or CORE networks. An anycast address (shown as X) is configured to the interface facing RAN of all EPC-E. EPC-E assigns an individual IPv6 address to another interface (illustrated "a" to "d" in the figure). It is important to mention that the anycast address X can be treated as the SGW's S1-U address.

Since RTRs are a gateway to the Internet, they advertise routes of an operator's prefix to the Internet. After one of RTR receives a packet of UE from the Internet, it needs to routing it to UE in the user plane. RTR has a simple routing entry for PDN prefix whose next hop points to the EPC-E. One of RTR (let's say RTR2 in this case) looks up a routing table with UE's address and matched it with a routing entry of PDN prefix. Since multiple EPC-Es advertise a route for the same PDN prefix, RTR2 should forward the packet to one of EPC-E according to the routing entry. This routing is known as hot-potato routing. In this example, the RTR2 uses EPC-E2-b as a nexthop of PDN prefix.

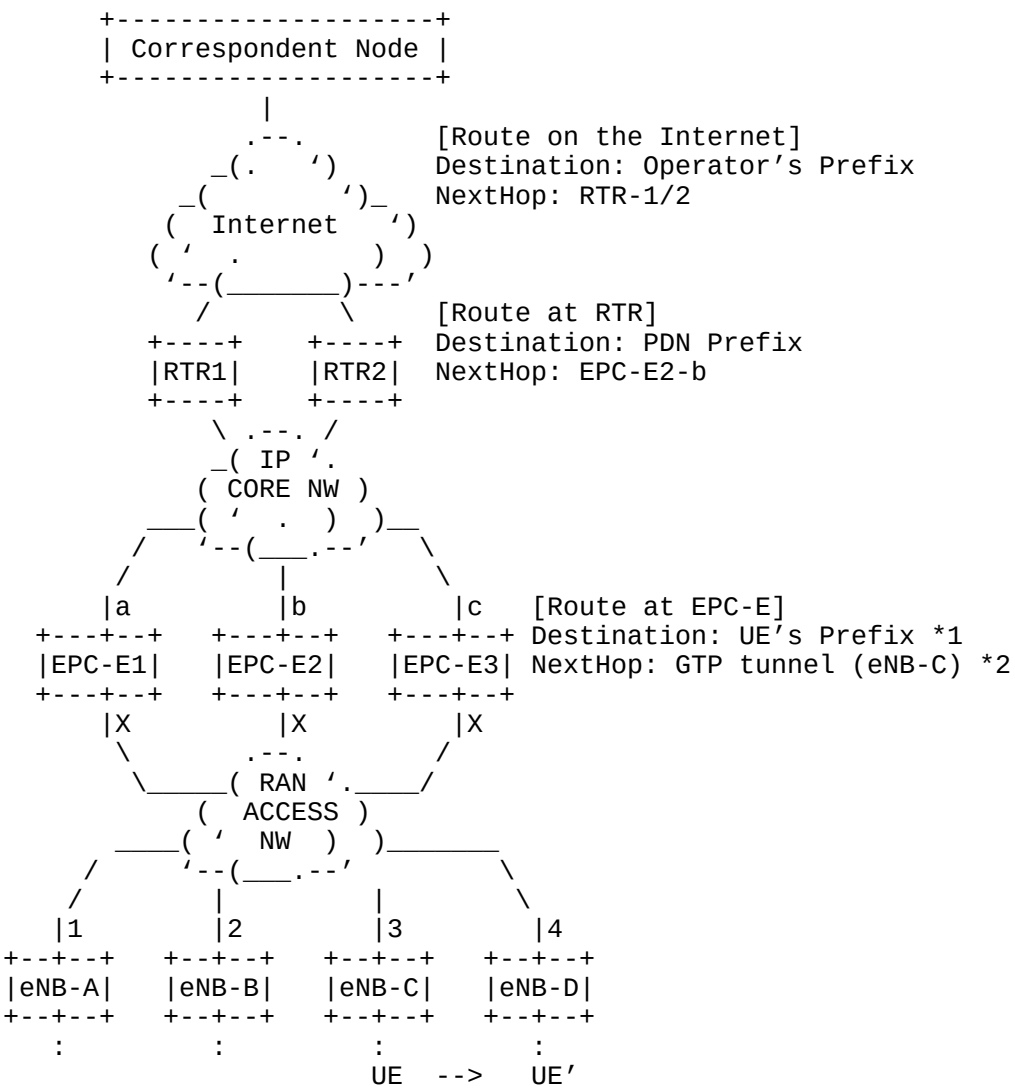
When the UE's packet is arrived at EPC-E2, EPC-E2 needs to forwards them to UE via eNodeB to which UE is connecting by using GTP tunnel. For this operation, EPC-E2 has a routing entry that destination is UE's prefix and that next hop points to GTP tunnel between eNB-C and the EPC-Es. In order to identify the GTP tunnel for UE, EPC-E needs

S1-U address and Tunnel Endpoint ID (TEID) of eNB-C that is eNB-C-3 in Figure 7. The eNB-C TEID for UE is illustrated as TEID[eNB-C]. The SGW assigned TEID is utilized to generate the UE's prefix as we explained in Section 3.2. These TEID are assigned per UE. The TEID and S1-U address of eNodeB are retrieved from the next hop field of the routing entry of the mobile node. By using the GTP information, every EPC-E can now forward the UE's packets to right eNodeB.

Routing outgoing packets from UE is much simpler. The packets from UE are always routed to the closest EPC-E to UE because of anycast routing. In Figure 7, when UE sends a packet to a destination, the packet is reached to eNB-C and tunneled to EPC-E's anycast address. The GTP-tunneled packet is routed to the closest EPC-E that is EPC-E2 in this case. The packet is decapsulated by EPC-E2 and then forwarded to one of RTR according to the routing table. Since the decapsulated packet is regular IPv6 packet, no extra control other than routing is necessary.

When UE moves to a new location (UE'), it updates its location on the control plane. After signaling completion for location update, vEPC needs to update the UE's routing entry of all EPC-E so that vEPC advertises updated route with new location to all EPC-Es by a routing protocol. The routing entry should be updated with the new eNodeB's address that is eNB-D-4. During handover, there might be some traffic arriving to the older eNodeB (eNB-C). These packets can be re-routed to the new eNodeB (eNB-D) via X2-U interface in RAN.

The UE's address isn't changed when UE changes its attachment. In our scenario, SGW run on hypervisor and is independent from network topology. Therefore, logically we don't have handover across different SGWs. UE can stay connected with the same SGW all the time and can keep using the same TEID after handover. Thus, UE's address is unchanged even after handover.



*1 TEID used at EPC-E for the UE is included in this UE's prefix. see Figure 4.

*2 GTP tunnel state is stored in the next hop field. The state information is the combination of eNB-C S1 address that is eNB-C-3 and TEID(eNB-C) assigned for the UE.

Routing Mechanism Overview

Figure 7

3.5. IPv4 Support

Recent IPv6 transition mechanisms enable IPv6-only network to forward IPv4 packet with encapsulation or translation techniques. By using one of mechanisms, we can use IPv6 for our stateless user-plane network for transporting both IPv4 and IPv6 packets. Figure 8 shows available solutions of IPv4 support for each bearer type to deal with that requirement.

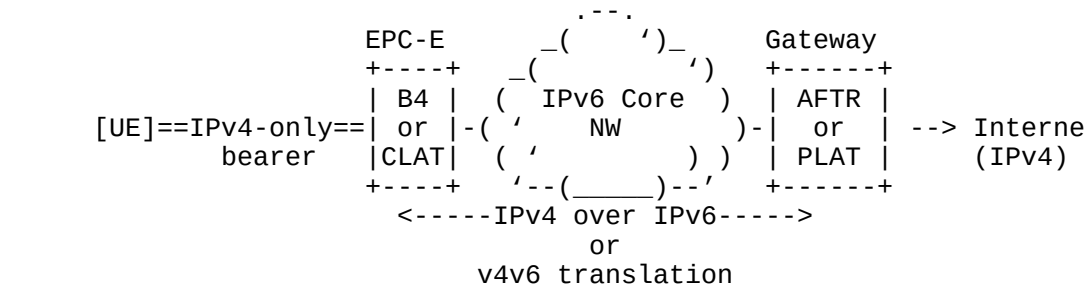
Bearer type	UE function	EPC-E function	Gateway function
IPv4	-	B4	AFTR
IPv4	-	CLAT	PLAT
IPv6	MAP-CE	-	MAP-BR
IPv6	B4	-	AFTR
IPv6	CLAT	-	PLAT

Solutions and functions for IPv4 support

Figure 8

In the case of a UE only support IPv4 bearer, B4 function of DS-Lite [RFC6333] or CLAT function of 464XLAT [RFC6877] may be implemented in a EPC-E. Both functions are stateless therefore EPC-E isn't required to maintain any tunneling or translation state.

Figure 9 shows how to support IPv4 on IPv6 core network in our vEPC. Instead of using RTR as a gateway to the Internet, DS-LITE AFTR or 464XLAT PLAT is installed as a gateway to the IPv4 Internet.



IPv4 User plane of vEPC

Figure 9

If UE supports IPv6 capable bearer, IPv6 transition function may be implemented in the UE such as MAP-CE [I-D.ietf-softwire-map], B4 or CLAT. That means an EPC-E receives IPv6 packets from UE in this case so that the EPC-E does not need to be involved in the part of IPv4 support functions.

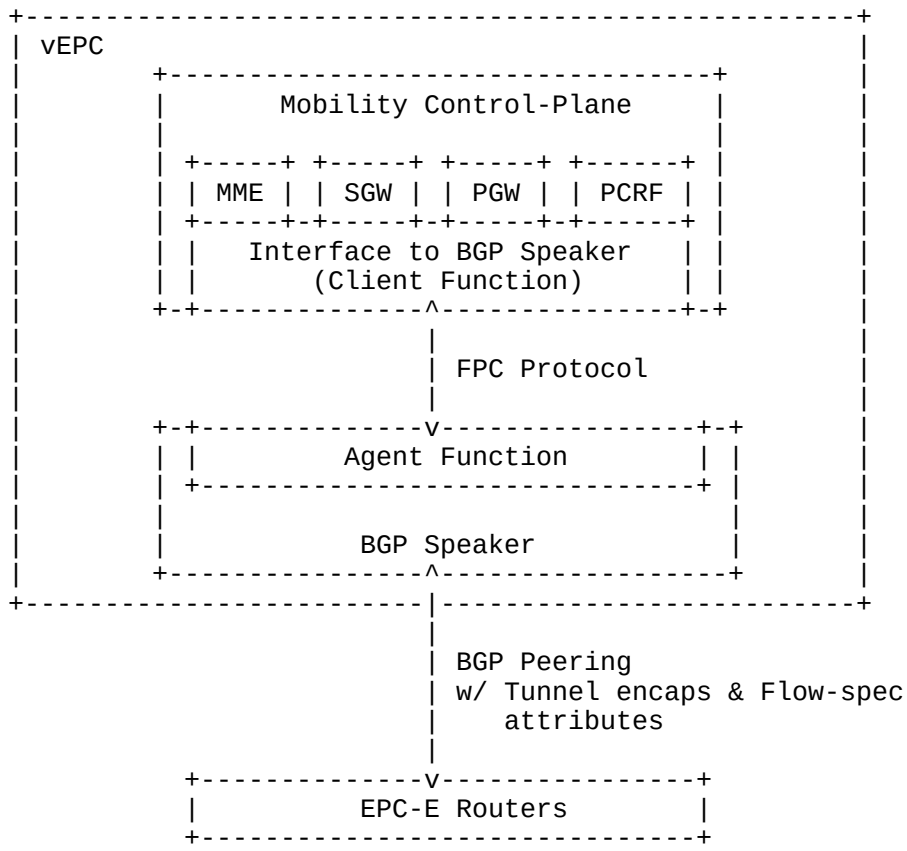
3.6. Interface between Control-plane and BGP Speaker

In Section 3.2 described, mobility control-plane and BGP speakers within a vEPC need an interface to export user-plane information from the control-plane to the BGP speakers. Perhaps many solutions would be developed proprietarily. However, adopting standardized interface will be much appropriate.

Forwarding Policy Configuration Protocol (FPCP) [I-D.ietf-dmm-fpc-cpdp] has been standardized in IETF for that purpose. That provices client function to the mobility control-plane to export user-plane information, and agent function which enables the BGP speakers to receive the user-plane information when it is implemented into them.

User-plane information contains UE’s IP prefix, GTP-U tunnel endpoints of serving eNodeB/EPC-E and applied QoS class. When those information come into the BGP speaker, the agent renders it into BGP attributes which are UE’s IP prefix, GTP-U tunnel endpoints and QoS class are indicated in NLRI, [I-D.ietf-idr-tunnel-encaps] and [RFC5575] with [I-D.ietf-idr-flow-spec-v6] respectively.

The BGP speaker generates BGP UPDATE messages based on that and then advertises it to EPC-E routers. Figure 10 depicts FPCP enabled vEPC in which mobility control-plane and BGP speaker are interfaced through FPCP client and agent functions.



FPCP enabled vEPC

Figure 10

4. Operational Considerations

4.1. Scalability and Reliability

Virtualization allows vEPC to be elastic for steep demand of requests to create and update for sessions. In our architecture, that makes routing update fluctuation from vEPC to EPC-E. This is the reason why we select BGP as a protocol between vEPC and EPC-E. BGP is scalable and stable routing protocol today.

BGP is an incremental update protocol so that once BGP peer established, millions of routes can be easily updated in stable

manners. Operators can appropriately design BGP peering between vEPC and ECP-E to secure convergence time within appropriate period.

Granularity of the peering should be aware EPC-E capacity because it is assumed that EPC-E has upper limit of routing entries. BGP peering design should makes sure that total number of routes does not exceed EPC-E capacity.

During the network planning, operators must understand EPC-E's capacity such as # of routes, bandwidth, etc. An example of estimation, if a EPC-E has 1Gbps throughput and each UE's bandwidth consumption is 10Kbps in average, the EPC-E should have 100K routes capacity.

This is an operational approach to minimize the risk of routing update fluctuation. If it is hard to support all the UEs by a EPC-E in an operators network, another EPC-E can be introduced and configured as a set of EPC-Es. The UEs are distributed and handled by the EPC-Es within the set. We don't need to support millions of UEs by a single EPC-E.

EPC-E set is also useful to have EPC-E redundancy for reliable operation. The nature of BGP makes easy to replicate UE routes to multiple EPC-Es within a EPC-E set. In that EPC-E set, when an EPC-E fall down to a failure, another EPC-E come out with same UE routes that the fall-down EPC had and immediately re-converge to core routing. That helps user-plane to minimize disruption during EPC-E failure recovery.

These are another advantage of using routing mechanism in the user plane. We already explain how to handle multiple EPC-Es and EPC-E sets in our scheme in Section 3.3.

The notion of multiple EPC-E sets is easily fitted into our today's' network. The operator's network is often separated into several regional network for geographical scalability. Therefore, the operator can assign different EPC-E set to different region for better scalability.

In that network, when an UE hands over between two regions, the session of the UE might be disconnected if the serving EPC-E doesn't have reachability for those region access networks. For example, in the case of regional access networks have duplicated IPv4 private address space. To enable inter-region hand-over, it is recommended that all of the access network, such as RAN, are IPv6 networks and reachable each other.

In addition, routers and EPC-E in the IPv6 core network are required to process just "route", they naturally aggregate those routing entries. It helps limiting the total number of routing entries in our core network.

4.2. Backward Compatibility

vEPC should be able to fall back to the legacy EPC based packet forwarding to secure backward compatibility which is required to connect existing system, or to connect roaming partners through legacy S5/S8 interfaces. When fallback happened, all the packets are not routed on our stateless user plane, but forwarded to vEPC (i.e. SGW and PGW instances on hypervisor). vEPC must use a S1-U address that is different from anycast address assigned to EPC-Es. This address is assigned to SGW instances in vEPC and used to terminate tunnels in vEPC servers (i.e. hypervisor).

5. IANA Considerations

This memo includes no request to IANA.

6. Security Considerations

There are no security considerations specific to this document at this moment.

7. References

7.1. Normative References

[I-D.ietf-idr-flow-spec-v6]
McPherson, D., Raszuk, R., Pithawala, B., Andy, A., and S. Hares, "Dissemination of Flow Specification Rules for IPv6", draft-ietf-idr-flow-spec-v6-07 (work in progress), March 2016.

[I-D.ietf-idr-tunnel-encaps]
Rosen, E., Patel, K., and G. Velde, "The BGP Tunnel Encapsulation Attribute", draft-ietf-idr-tunnel-encaps-01 (work in progress), December 2015.

[RFC5575] Marques, P., Sheth, N., Raszuk, R., Greene, B., Mauch, J., and D. McPherson, "Dissemination of Flow Specification Rules", RFC 5575, DOI 10.17487/RFC5575, August 2009, <http://www.rfc-editor.org/info/rfc5575>.

7.2. Informative References

[Boeing-BGP]
Andrew, , "Global IP Network Mobility using Border Gateway Protocol (BGP)", IAB Plenary IAB Plenary of IETF 62nd, March 2005.

[I-D.ietf-dmm-fpc-cpdp]
Liebsch, M., Matsushima, S., Gundavelli, S., and D. Moses, "Protocol for Forwarding Policy Configuration (FPC) in DMM", draft-ietf-dmm-fpc-cpdp-01 (work in progress), July 2015.

[I-D.ietf-softwire-map]
Troan, O., Dec, W., Li, X., Bao, C., Matsushima, S., Murakami, T., and T. Taylor, "Mapping of Address and Port with Encapsulation (MAP)", draft-ietf-softwire-map-13 (work in progress), March 2015.

[I-D.savolainen-stateless-pd]
Savolainen, T. and J. Korhonen, "Stateless IPv6 Prefix Delegation for IPv6 enabled networks", draft-savolainen-stateless-pd-01 (work in progress), February 2010.

[I-D.wakikawa-mext-haha-interop2008]
Wakikawa, R., Shima, K., and N. Shigechika, "The Global HAHA Operation at the Interop Tokyo 2008", draft-wakikawa-mext-haha-interop2008-00 (work in progress), July 2008.

[I-D.wakikawa-req-mobile-cp-separation]
Wakikawa, R., Matsushima, S., Patil, B., Chen, B., DJ, D., and H. Deng, "Requirements and use cases for separating control and user planes in mobile network architectures", draft-wakikawa-req-mobile-cp-separation-00 (work in progress), November 2013.

[I-D.yokota-dmm-scenario]
Yokota, H., Seite, P., Demaria, E., and Z. Cao, "Use case scenarios for Distributed Mobility Management", draft-yokota-dmm-scenario-00 (work in progress), October 2010.

[NFV-WHITEPAPER]
Network Operators, , "Network Functions Virtualization, An Introduction, Benefits, Enablers, Challenges and Call for Action", SDN and OpenFlow SDN and OpenFlow World Congress, October 2012.

[RFC3963] Devarapalli, V., Wakikawa, R., Petrescu, A., and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol", RFC 3963, DOI 10.17487/RFC3963, January 2005, <http://www.rfc-editor.org/info/rfc3963>.

[RFC5213] Gundavelli, S., Ed., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, DOI 10.17487/RFC5213, August 2008, <http://www.rfc-editor.org/info/rfc5213>.

[RFC5555] Soliman, H., Ed., "Mobile IPv6 Support for Dual Stack Hosts and Routers", RFC 5555, DOI 10.17487/RFC5555, June 2009, <http://www.rfc-editor.org/info/rfc5555>.

[RFC5844] Wakikawa, R. and S. Gundavelli, "IPv4 Support for Proxy Mobile IPv6", RFC 5844, DOI 10.17487/RFC5844, May 2010, <http://www.rfc-editor.org/info/rfc5844>.

[RFC6275] Perkins, C., Ed., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, DOI 10.17487/RFC6275, July 2011, <http://www.rfc-editor.org/info/rfc6275>.

[RFC6301] Zhu, Z., Wakikawa, R., and L. Zhang, "A Survey of Mobility Support in the Internet", RFC 6301, DOI 10.17487/RFC6301, July 2011, <http://www.rfc-editor.org/info/rfc6301>.

[RFC6333] Durand, A., Droms, R., Woodyatt, J., and Y. Lee, "Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion", RFC 6333, DOI 10.17487/RFC6333, August 2011, <http://www.rfc-editor.org/info/rfc6333>.

[RFC6459] Korhonen, J., Ed., Soininen, J., Patil, B., Savolainen, T., Bajko, G., and K. Iisakkila, "IPv6 in 3rd Generation Partnership Project (3GPP) Evolved Packet System (EPS)", RFC 6459, DOI 10.17487/RFC6459, January 2012, <http://www.rfc-editor.org/info/rfc6459>.

[RFC6877] Mawatari, M., Kawashima, M., and C. Byrne, "464XLAT: Combination of Stateful and Stateless Translation", RFC 6877, DOI 10.17487/RFC6877, April 2013, <http://www.rfc-editor.org/info/rfc6877>.

Authors’ Addresses

Satoru Matsushima
SoftBank
1-9-1,Higashi-Shimbashi,Minato-Ku
Tokyo 105-7323
Japan

Email: satoru.matsushima@g.softbank.co.jp

Ryuji Wakikawa
SoftBank
1-9-1,Higashi-Shimbashi,Minato-Ku
Tokyo 105-7323
Japan

Email: ryuji.wakikawa@gmail.com