

INTERNET-DRAFT
Intended status: Informational

S. Hu
F. Qin
Z. Li
China Mobile
T. Chua
Singapore Telecommunications Ltd
V. Lopez
Telefonica
D. Eastlake
Z. Wang
J. Song
Huawei
March 11, 2019

Expires: September 10, 2019

Architecture for Control Plane and User Plane Separated BNG
draft-cuspd-t-rtgwg-cu-separation-bng-architecture-04.txt

Abstract

This document defines an architecture for Broadband Network Gateway (BNG) devices with control plane (CP) and user plane (UP) separation. A BNG-CP is a user control management component while a BNG-UP takes responsibility as the network edge and user policy implementation component. Both BNG-CP and BNG-UP are core components for fixed broadband services and are deployed separately at different network layers.

Status of This Memo

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

Distribution of this document is unlimited. Comments should be sent to the authors or the RGTWG working group mailing list: rtgwg@ietf.org.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

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

Table of Contents

1. Introduction.....	3
1.1 Motivation.....	3
2. Terminology.....	4
3. CU Separated BNG Architecture.....	5
3.1 Internal Interfaces Between the CP and UP.....	7
4. Usage of the CU Separation BNG.....	8
5. Security Considerations.....	10
6. IANA Considerations.....	10
Normative References.....	11
Informative References.....	11
Authors' Addresses.....	12

1. Introduction

A Broadband Network Gateway (BNG) device is defined as an Ethernet-centric IP edge router, and the aggregation point for user traffic. It performs Ethernet aggregation and packet forwarding via IP/MPLS, and supports user management, access protocols termination, QoS, policy management, etc.

This document describes an architecture for BNG devices with control plane (CP) and user plane (UP) separation. A BNG-CP is a user control management component while a BNG-UP takes responsibility as the network edge and user policy implementation components. Both BNG-CP and BNG-UP are core components for fixed broadband services and are deployed separately at different network layers in the network.

1.1 Motivation

The rapid development of new services, such as 4K TV, IoT, etc., and increasing numbers of home broadband service users present some new challenges for BNGs such as:

Low resource utilization: The traditional BNG acts as both a gateway for user access authentication and accounting and an IP network's Layer 3 edge. The mutually affecting nature of the tightly coupled control plane and forwarding plane makes it difficult to achieve the maximum performance of either plane.

Complex management and maintenance: Due to the large numbers of traditional BNGs, configuring each device in a network is very tedious when deploying global service policies. As the network expands and new services are introduced, this deployment mode will cease to be feasible as it is unable to manage services effectively and rectify faults rapidly.

Slow service provisioning: The coupling of control plane and forwarding plane, in addition to a distributed network control mechanism, means that any new technology has to rely heavily on the existing network devices.

To address these challenges for fixed networks, the framework for a cloud-based BNG with CU separation conception is defined in [TR-384]. The main idea of Control-Plane and User-Plane separation is to extract and centralize the user management functions of multiple BNG devices, forming a unified and centralized control plane (CP). And the traditional router's Control Plane and Forwarding Plane are both preserved on BNG devices in the form of a user plane (UP). Note that the CU separation concept has also been introduced in the 3GPP 5G architecture [3GPP.23.501].

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

The following acronyms are used as specified below:

AAA: Authentication Authorization Accounting.

BNG: Broadband Network Gateway. A broadband remote access server (BRAS (Broadband Access Server), B-RAS or BBRAS) that routes traffic to and from broadband remote access devices such as digital subscriber line access multiplexers (DSLAM) on an Internet service provider's (ISP) network. BRAS can also be referred to as a Broadband Network Gateway (BNG).

CP: Control Plane. The CP is a user control management component which manages the UP's resources such as the user entry and user's QoS policy

DHCP: Dynamic Host Configuration Protocol.

EMS: Element Management System.

IPoE: IP over Ethernet.

MANO: Management and Orchestration.

NFV: Network Function Virtualization.

NFVI: NFV Infrastructure.

PPPoE: Point-to-Point Protocol over Ethernet.

UP: User Plane. UP is a network edge and user policy implementation component. The traditional router's Control Plane and forwarding plane are both preserved on BNG devices in the form of a user plane.

3. CU Separated BNG Architecture

The functions in a traditional BNG can be divided into two parts: one is the user access management function, the other is the router function. In a cloud-based BNG, we find that tearing these two functions apart can make a difference. The user management function can be centralized and deployed as a concentrated module or device, called the BNG-CP (Control Plane). The other functions, such as the router function and forwarding engine, can be deployed in the form of the BNG User Plane. Thus, the Cloud-based BNG architecture is made up of control plane and user plane.

The following figure describes the architecture of CU separated BNG:

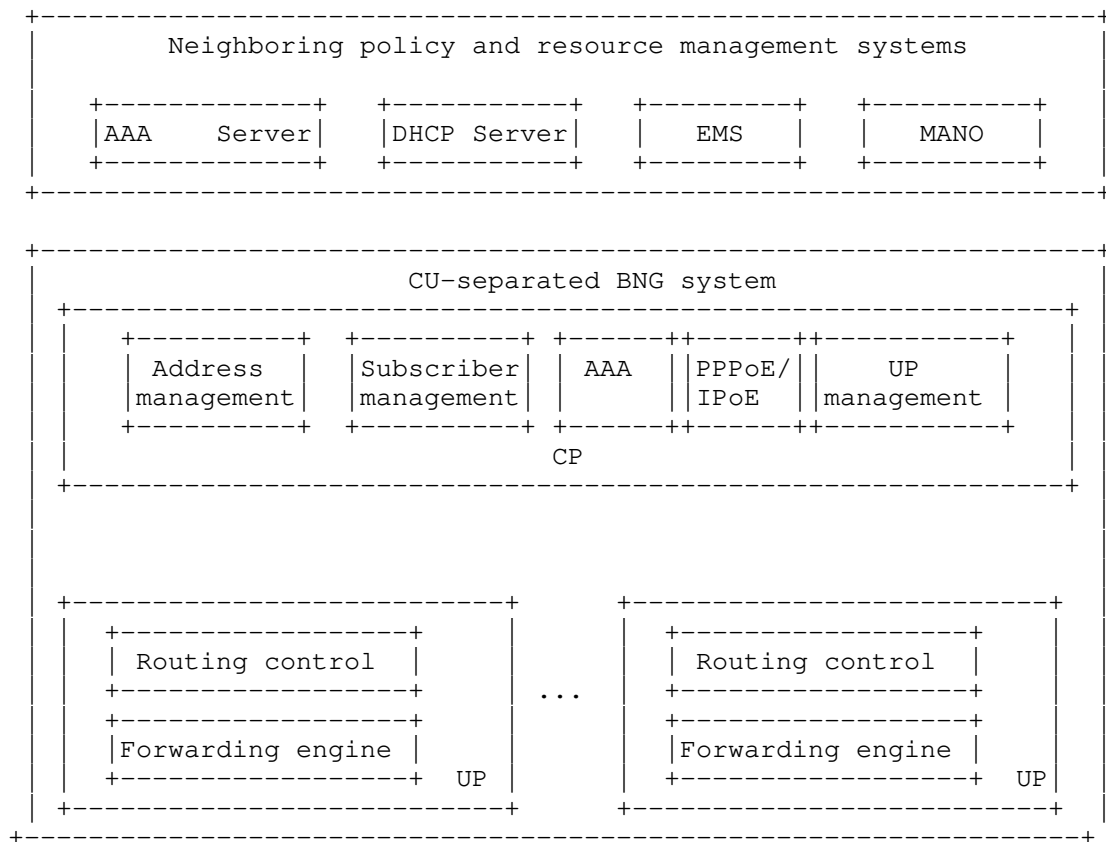


Figure 1. Architecture of CU Separated BNG

As in Figure 1, the BNG Control Plane could be virtualized and centralized, which provides significant benefits such as centralized session management, flexible address allocation, high scalability for subscriber management capacity, and cost-efficient redundancy, etc.

The functional components inside the BNG Service Control Plane can be implemented as Virtual Network Functions (VNFs) and hosted in a Network Function Virtualization Infrastructure (NFVI).

The User Plane Management module in the BNG control plane centrally manages the distributed BNG User Planes (e.g. load balancing), as well as the setup, deletion, and maintenance of channels between Control Planes and User Planes. Other modules in the BNG control plane, such as address management, AAA, etc., are responsible for the connection with outside subsystems in order to fulfill those services. Note that the User Plane SHOULD support both physical and virtual network functions. For example, BNG user plane L3 forwarding related network functions can be disaggregated and distributed across the physical infrastructure. And the other control plane and management plane functions in the CU Separation BNG can be moved into the NFVI for virtualization [TR-384].

The details of CU separated BNG's function components are as following:

The Control Plane should support:

- (1) Address management: unified address pool management.
- (2) AAA: This component performs Authentication, Authorization and Accounting, together with RADIUS/DIAMETER. The BNG communicates with the AAA server to check whether the subscriber who sent an Access-Request has network access authority. Once the subscriber goes online, this component together with the Service Control component implement accounting, data capacity limitation, and QoS enforcement policies.
- (3) Subscriber management: user entry management and forwarding policy management.
- (4) PPPoE/IPoE: process user dialup packets via PPPoE/IPoE.
- (5) UP management: management of UP interface status, and the setup, deletion, and maintenance of channels between CP and UP.

The User Plane should support:

- (1) Control plane functions including routing, multicast, and MPLS.
- (2) Forwarding plane functions including traffic forwarding, QoS and traffic statistics collection.

3.1 Internal Interfaces Between the CP and UP

To support the communication between the Control Plane and User Plane, several interfaces are involved. Figure 2 illustrates the internal interfaces of CU Separated BNG.

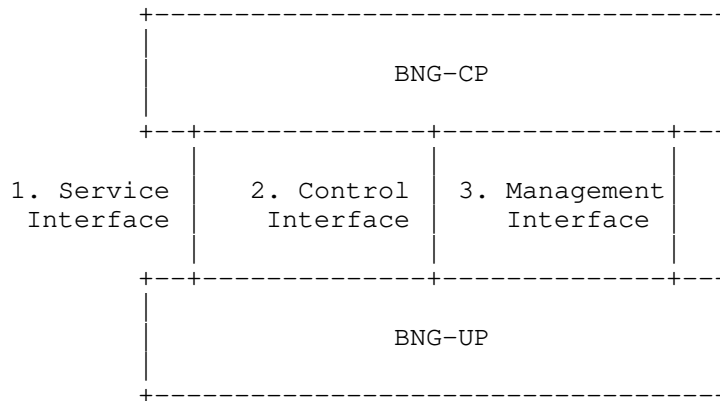


Figure 2. Internal Interfaces Between the CP and UP of the BNG

Service Interface: The CP and UP use this interface to establish tunnels with each other and transmit PPPoE and IPoE packets over those tunnels. VXLAN is commonly used for such tunnels as discussed in [hu-nvo3-vxlan-gpe-extension-for-vbng].

Control Interface: The CP uses this interface to deliver service entries, and the UP uses this interface to report service events to the CP. The requirements of this interface are introduced in [cuspdrt-gtwg-cusp-requirements], and the carrying protocol is presented in [cuspdrt-gtwg-cu-separation-bng-protocol] which specifies the Simple Control and User Plane Separation protocol (S-CUSP). The information model of this interface is presented in [cuspdrt-gtwg-cu-separation-infor-model].

Management Interface: The CP uses this interface to deliver configurations to the UP. This interface uses NETCONF [cuspdrt-gtwg-cu-separation-yang-model].

4. Usage of the CU Separation BNG

In the CU separated BNG scenario, there are several processes when a home user accesses the Internet:

- (1) User dialup packets via PPPoE or IPoE from the BNG-UP are sent to the BNG-CP through the BNG-UP's Service Interface.
- (2) BNG-CP processes the dialup packet. Confirming the user's authorization with the outside neighboring systems in the management network, the BNG-CP makes the decision to permit or deny the user access.
- (3) After that, the BNG-CP tells the UP to do perform authorized forwarding actions with appropriate QoS policies.
- (4) If the user is certificated and permitted, the UP forwards the traffic into the Internet with appropriate QoS policies such as limited bandwidth, etc. Otherwise, the user is denied to access the Internet.

In actual deployments, a CU separated BNG device is composed of a CP and one or more UPs. The CP is usually centrally deployed and takes responsibility as a user control management component managing UP's resources such as the user entry and forwarding policy. The UPs are distributed and act as a network edge and user policy implementation component.

In order to fulfill a service, neighboring policy and resource management systems are deployed outside the BNG. In the neighboring systems, different service systems such as RADIUS/DIAMETER server, DHCP server and EMS are included. If a BNG-CP is virtualized as a NFV, the NFVI management system MANO is also included here. A BNG-CP has connections with the outside neighboring systems to transmit management traffic.

The deployment scenario is shown in the following figure:

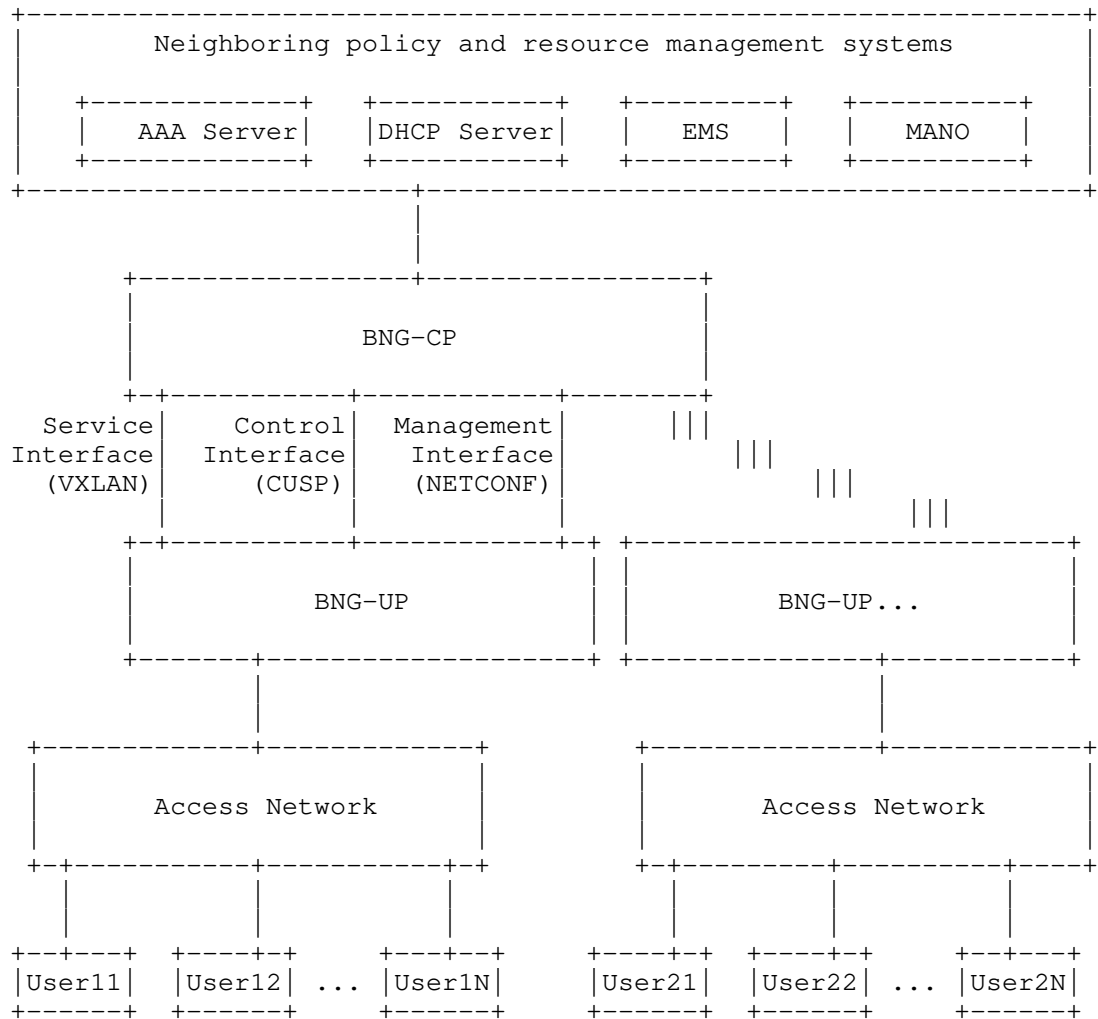


Figure 3. Deployment Example

5. Security Considerations

The Service, Control, and Management Interfaces between the CP and UP might be across the general Internet or other hostile environment. Thus, appropriate protections **MUST** be implemented to provide integrity, authenticity, and secrecy of traffic over those interfaces. For example, the implementation of IPSEC, DTLS, or TLS as appropriate. However, such security protocols need not always be used and lesser security precautions might be appropriate because, in some cases, the communication between the CP and UP might be in a more benign environment.

6. IANA Considerations

This document requires no IANA actions.

Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

Informative References

- [_3GPP.23.501] "System Architecture for the 5G System", 3GPP GPP TS 23.501 15.0.0, 2018.
- [cuspdtd-rtgwg-cu-separation-bng-deployment] Gu, R., "Deployment Model of Control Plane and User Plane Separated BNG", draft-cuspdtd-rtgwg-cu-separation-bng-deployment, work in progress, 2018.
- [cuspdtd-rtgwg-cu-separation-bng-protocol] Wang, Z., "Control-Plane and User-Plane separation BNG control channel Protocol", draft-cuspdtd-rtgwg-cu-separation-bng-protocol, work in progress, 2018.
- [cuspdtd-rtgwg-cu-separation-infor-model] Wang, Z., "Information Model of Control-Plane and User-Plane separation BNG", draft-cuspdtd-rtgwg-cu-separation-infor-model, work in progress, 2018.
- [cuspdtd-rtgwg-cusp-requirements] Hu, S., "Requirements for Control Plane and User Plane Separated BNG Protocol", draft-cuspdtd-rtgwg-cusp-requirements, work in progress, 2018.
- [cuspdtd-rtgwg-cu-separation-yang-model] Hu, F., "YANG Data Model for Configuration Interface of Control-Plane and User-Plane separation BNG", draft-cuspdtd-rtgwg-cu-separation-yang-model, work in progress, 2018.
- [hu-nov3-vxlan-gpe-extension-for-vbng] Huang, L., "VXLAN GPE Extension for Packets Exchange Between Control and User Plane of vBNG", draft-hu-nvo3-vxlan-gpe-extension-for-vbrg, work in progress, 2017.
- [TR-384] Broadband Forum, "Cloud Central Office Reference Architectural Framework", BBF TR-384, 2018.

Authors' Addresses

Shujun Hu
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: hushujun@chinamobile.com

Fengwei Qin
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: qinfengwei@chinamobile.com

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: lizhenqiang@chinamobile.com

Tee Mong Chua
Singapore Telecommunications Limited
31 Exeter Road, #05-04 Comcentre Podium Block
Singapore City 239732
Singapore

Email: teemong@singtel.com

Victor Lopez
Telefonica
Spain

Email: victor.lopezalvarez@telefonica.com

Donald Eastlake, 3rd
Huawei Technologies
1424 Pro Shop Court
Davenport, FL 33896
USA

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Zitao Wang
Huawei Technologies
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: wangzitao@huawei.com

Jun Song
Huawei Technologies
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: song.jun@huawei.com

Copyright, Disclaimer, and Additional IPR Provisions

Copyright (c) 2019 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.

INTERNET-DRAFT
Intended status: Informational

R. Gu
S. Hu
China Mobile
M. Wang
D. Eastlake
Huawei
F. Hu
ZTE
December 12, 2018

Expires: June 11, 2019

Control Plane and User Plane Separated BNG Deployment Model
draft-cuspdrt-rtgwg-cu-separation-bng-deployment-02

Abstract

This document describes the deployment model for a Broadband Network Gateway (BNG) device with Control Plane (CP) and User Plane (UP) separation. It is intended to give guidance for the deployment of CP and UP separated BNG devices in an operators' network.

Status of This Memo

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

Distribution of this document is unlimited. Comments should be sent to the authors or the BESS working group mailing list: bess@ietf.org.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

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

Gu, et al.
*

[Page 1]

INTERNET-DRAFT

Separated BNG Deployment Model

Table of Contents

1. Introduction and Overview.....	3
2. Concept and Terminology.....	5

2.1 Terminology.....	5
3. BNG with CP and UP Separation Deployment Model.....	6
3.1 CP and UP of BNG Deployment Within One District.....	6
3.2. CP and UP of BNG Deployment in Multiple Districts.....	7
4. The Process of BNG with CUPS in Home Service.....	10
5. High Availability Considerations.....	11
6. Security Considerations.....	12
7. IANA Considerations.....	12
Normative References.....	13
Informative References.....	13
Authors' Addresses.....	14

<

1. Introduction and Overview

A Broadband Network Gateway (BNG) is an Ethernet-centric IP edge router and acts as the aggregation point for the user traffic with some additional functions such as address management and cooperating with AAA (Radius/Diameter) systems and subscriber management. Because of the rapid development of new services, such as 4K, IoT, etc. and the increasing numbers of distributed home broadband service users, high resource utilization, high-efficiency management, and fast service provisioning are required. This calls for a new BNG architecture with CP and UP separation, which is also called Cloud BNG, as proposed in [BBF-CloudCO] [TR-384].

The CP and UP separation architecture of the BNG is composed of a Control Plane and a User Plane, with the concentrated CP responsible for control and management of the UP's resources and subscribers' information, and with the distributed UP taking charge of policy implementation and traffic forwarding. The obvious advantages of this new architecture are listed below.

Resource Utilization Improvement: A centralized Control Plane provides unified management capability for network resources and users information. The CP has an overview of all the resources and can distribute resources as specific users require, thus resources can be totally controlled and balanced.

Management with High Efficiency: A centralized CP provides a unified management interface to the outside systems such as EMS, DHCP Server, AAA Server, etc. In this situation, management can be easier for the centralized CP as it's the only device interfacing with the outside systems.

Dynamic and Flexible: The CP can be virtualized as a VNF with MANO management in an NFVI, while the UP can be a virtual machine or physical device as needed. A software-oriented CP can be designed with flexibility. The CP can handle all the situations dynamically over a wide range from few users accessing to large numbers of users accessing.

Fast TTM: The CP and UP can be deployed separately with the CP deployed centrally and the UP deployed in distribution closer to users. Thus, according to different situations such as session overload or extremely high throughput, the CP and UP can be extended separately. This can help shorten the time to market (TTM).

As noted, the new BNG architecture has CP and UP separation. The CP and UP are deployed with separation due to practical requirements. This document gives the CU separation BNG deployment model for actual

deployments.

2. Concept and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.1 Terminology

BNG: Broadband Network Gateway. A broadband remote access server (BRAS, B-RAS or BBRAS) routes traffic to and from broadband remote access devices such as digital subscriber line access multiplexers (DSLAM) on an Internet service provider's (ISP) network. BRAS can also be referred to as a Broadband Network Gateway (BNG).

CP: Control Plane. The CP is a user control management component which manages UP's resources such as the user entry and user's QoS policy

CUPS: Control/User Plane Separation

UP: User Plane. The UP is a network edge and user policy implementation component. The traditional router's Control Plane and forwarding plane are both preserved on BNG devices in the form of a user plane.

TTM: Time to Market. It is the length of time it takes from a product or a service being conceived until it is available for sale.

MANO: Management and Orchestration. Functions are collectively provided by NFVO, VNFM and VIM.

VNF: Virtual Network Function. Implementation of a Network Function that can be deployed on a Network Function Virtualization Infrastructure (NFVI).

PNF: Physical Network Function

DHCP: Dynamic Host Configuration Protocol

PPPoE: Point-to-Point Protocol over Ethernet

IPoE: Internet Protocol over Ethernet

3. BNG with CP and UP Separation Deployment Model

3.1 CP and UP of BNG Deployment Within One District

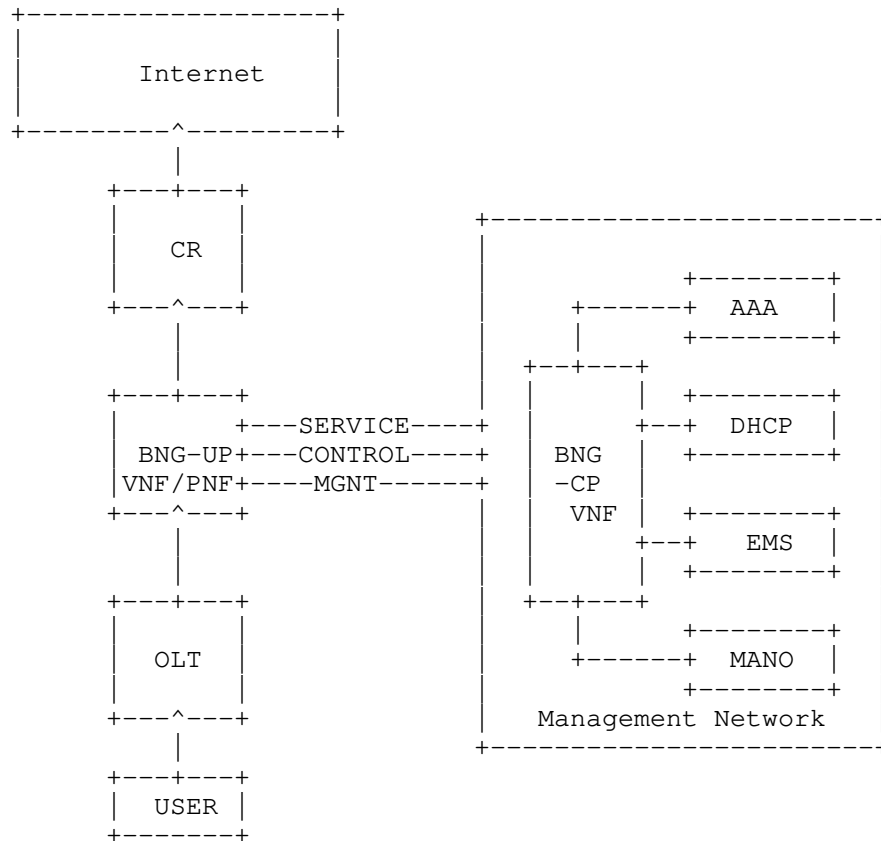


Figure 1: Cloud BNG Deployed in One District

Take a one district example as in Figure 1. Here BNG-CP and BNG-UP are separated as deployed. Since the CP is computationally intensive, a virtualized CP acting as a VNF can meet the requirements of flexibility and fast calculation. The UP is traffic intensive, which can be virtualized or stay physical depending on traffic. The virtualized UP with low expense and high flexibility can be suitable for light traffic. In high traffic, special hardware is needed with high traffic forwarding performance.

In order to fulfill the function of a BNG, the BNG-CP needs to communicate with outside systems such as a AAA (Radius/Diameter) server and many others in the management network. In addition, the

BNG-CP has three interfaces with the BNG-UP separated by their traffic categories: Service Interface, Control Interface, and Management Interface.

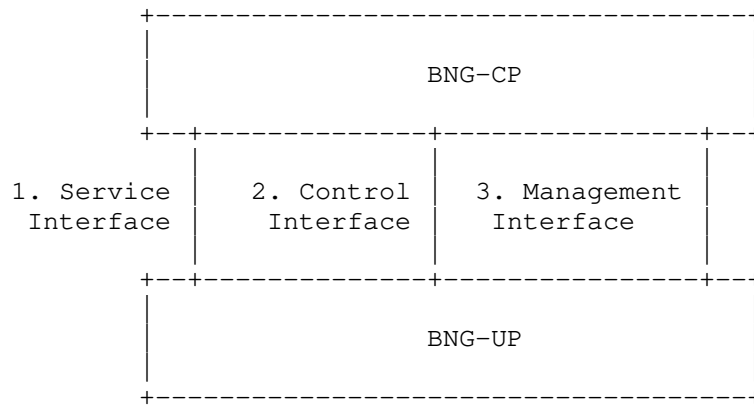


Figure 2. Internal Interfaces Between the BNG CP and UP

The functions of the three interfaces are as follows:

Service Interface: The CP and UP use this interface to establish VXLAN tunnels with each other and transmit PPPoE and IPoE packets over the VXLAN tunnels for authentication.

Control Interface: The CP uses this interface to deliver service entries to the UP, and the UP uses this interface to report service events to the CP.

Management Interface: The CP uses this interface to deliver basic configurations to the UP. This interface uses NETCONF.

Several related drafts exist describing these interfaces in detail. The VXLAN-GPE extension draft for C/U separated BNG is related to the Service Interface [huang-nov3-vxlan-gpe-extension-for-vbng]. The draft YANG data model for CU separated BNG focuses on Management Interface, seeing in [cuspdrt-gwg-cu-separation-yang-model]. Another two drafts [cuspdrt-gwg-cusp-requirements] and [cuspdrt-gwg-cu-separation-infor-model] are related to the control interface giving an information model abstraction and suitable protocol.

3.2. CP and UP of BNG Deployment in Multiple Districts

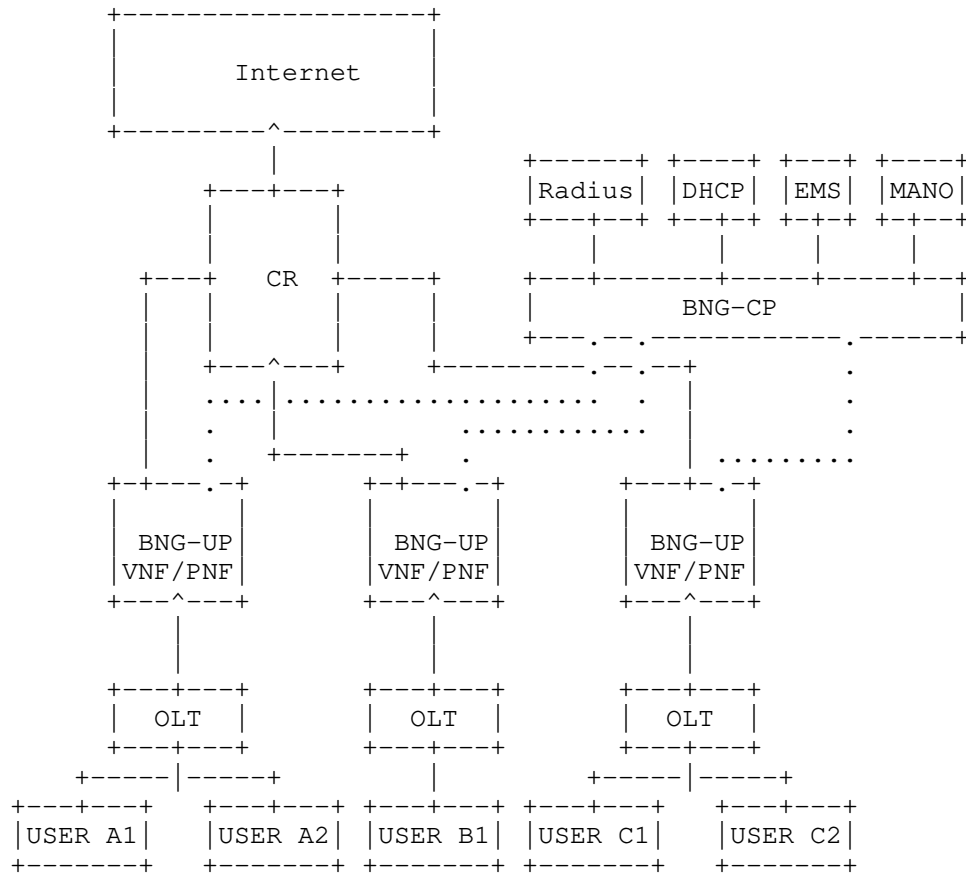


Figure 3: Cloud BNG Deployed in Several Districts

If subscribers are distributed in several districts, the CP can be deployed centrally with the UP deployed in different districts close to subscribers as shown in Figure 3. Thus the deployment model can be a bit complex.

Take three districts A, B. and C for example. Here three UPs are placed with one shared CP. The CP is usually deployed in a Core Data Center such as in a provincial datacenter with UPs in edge Data Centers such as city datacenters. In this Data Centers design, we have core data centers and edge data centers according to their location and responsibility. Core data centers are often planned in provinces for control and management, while edge data centers are in cities or towns for easy service access.

In this scenario, a centralized CP interfaces to the subsystems outside and communicate with all these UPs for control and management.

Under the CP's control, the corresponding traffic is forwarded by UP to the Internet.

4. The Process of BNG with CUPS in Home Service

Take a user Bob accessing to the Internet using Home Broadband Service as an example. The process includes the service traffic from user to the internet and signaling traffic between BNG-UP and BNG-CP. Below is the whole process.

- (1) User Bob dials up with packets of PPPoE or IPoE from BNG-UP which will be sent to the BNG-CP with the user's information. This is signaling traffic.
- (2) The BNG-CP processes the dialup packets. Confirming with the outside neighboring systems in the management network, the BNG-CP makes the decision to permit or deny of the dial access through certification. In this step, the BNG-CP manages resources and generates tables with information such as User Info, IP Info, QoS Info, etc. This is signaling traffic.
- (3) The BNG-CP sends tables to the corresponding UP or to one UP it chooses from the corresponding UPs. This is signaling traffic.
- (4) The BNG-UP receives the tables, matches rules and performs corresponding actions.
- (5) If Bob is certificated and permitted, the UP forwards their traffic into the Internet with related policies such as limited bandwidth, etc. Otherwise, Bob is denied to access the Internet. This is service traffic.

From Step 2 to Step 4, the information model defined in [cuspd-t-rtgw-g-cu-separation-infor-model] can be used.

5. High Availability Considerations

As the BNG-CP takes responsibility for control and management, such as communicating with outside systems, generating flow tables, and managing the UP's resources, high availability of this key component should be considered. Some redundancy should be adopted for reliability, such as N+N or N+K active standby BNG-CPs. N+N standby means 1:1 backup for each BNG-CP, which enables easy rapid switch of any number of BNG-CP to their backup but is expensive because it requires a large number of backup CPs. N+K means a smaller number of backup CPs, for example N2:1 backup where $N2 < N$ which is less expensive but does not handle more than 1 failure in the N2 subset of N BNG-CPs.

6. Security Considerations

TBD.

7. IANA Considerations

This document requires no IANA actions.

Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>> .in -10

Informative References

- [hu-nov3-vxlan-gpe-extension-for-vbng] Huang, L., "VXLAN GPE Extension for Packets Exchange Between Control and User Plane of vBNG", draft-hu-nvo3-vxlan-gpe-extension-for-vbrg, work in progress, 2017.
- [cuspdrt-rtgwg-cu-separation-yang-model] Hu, F., "YANG Data Model for Configuration Interface of Control-Plane and User-Plane separation BNG", draft-cuspdrt-rtgwg-cu-separation-yang-model, work in progress, 2018.
- [cuspdrt-rtgwg-cusp-requirements] Hu, S., "Requirements for Control Plane and User Plane Separated BNG Protocol", draft-cuspdrt-rtgwg-cusp-requirements, work in progress, 2018.
- [cuspdrt-rtgwg-cu-separation-infor-model] Wang, Z., "Information Model of Control-Plane and User- Plane separation BNG", draft-cuspdrt-rtgwg-cu-separation-infor-model, work in progress, 2018.
- [TR-384] BroadBand Forum, "Cloud Central Office Reference Architectural Framework", January 2018.

Authors' Addresses

Rong Gu
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: gurong_cmcc@outlook.com

Sujun Hu
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: shujun_hu@outlook.com

Michael Wang
Huawei Technologies
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: wangzitao@huawei.com

Donald Eastlake, 3rd
Huawei Technologies
1424 Pro Shop Court
Davenport, FL 33896
USA

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Fangwei Hu
ZTE Corporation
No.889 Bibo Rd
Shanghai 201203
China

Phone: +86 21 68896273
Email: hu.fangwei@zte.com.cn

Copyright, Disclaimer, and Additional IPR Provisions

Copyright (c) 2018 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.

INTERNET-DRAFT
Intended status: Proposed Standard

S. Hu
China Mobile
D. Eastlake
Futurewei Technologies
M. Chen
Huawei Technologies
F. Qin
Z. Li
China Mobile
T. Chua
Singapore Telecommunications
D. Huang
ZTE
July 3, 2019

Expires: January 2, 2020

Control-Plane and User-Plane Separation BNG
Simple Control Channel Protocol (S-CUSP)
draft-cuspd-rtgwg-cu-separation-bng-protocol-06

Abstract

This document specifies the Simple Control Plane (CP) and User Plane (UP) Separation Broadband Network Gateway (BNG) control channel Protocol (S-CUSP) for communications between a CP and a UP. S-CUSP is designed to be flexible and extensible so as to easily allow for the addition of further messages and data items to meet future requirements.

Status of This Memo

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

Distribution of this document is unlimited. Comments should be sent to the authors or the RGTWG working group mailing list: rtgwg@ietf.org.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

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

Table of Contents

1. Introduction.....	6
2. Terminology.....	7
2.1 Implementation Requirement Keywords.....	7
2.2 Terms.....	7
3. BNG CUPS Overview.....	10
3.1 BNG CUPS Motivation.....	10
3.2 BNG CUPS Architecture Overview.....	10
3.3 BNG CUPS Interfaces.....	12
3.3.1 Service Interface.....	13
3.3.2 Control Interface.....	14
3.3.3 Management Interface.....	14
3.4 BNG CUPS Procedure Overview.....	14
4. S-CUSP Protocol Overview.....	18
4.1 Control Channel Related Procedures.....	18
4.1.1 S-CUSP Session Establishment.....	18
4.1.2 Keep Alive.....	19
4.2 Node Related Procedures.....	20
4.2.1 UP Resource Report.....	20
4.2.2 Update BAS Function on Access Interface.....	21
4.2.3 Update Network Routing.....	21
4.2.4 CGN Public IP Address Allocation.....	22
4.2.5 Data Synchronization between the CP and UP.....	23
4.3 Subscriber Session Related Procedures.....	24
4.3.1 Create Subscriber Session.....	25
4.3.2 Update Subscriber Session.....	26
4.3.3 Delete Subscriber Session.....	27
4.3.4 Subscriber Session Events Report.....	27
5. S-CUSP Call Flows.....	29
5.1 IPoE.....	29
5.1.1 DHCPv4 Access.....	29
5.1.2 DHCPv6 Access.....	30
5.1.3 IPv6 SLAAC Access.....	32
5.1.4 DHCPv6 + SLAAC Access.....	33
5.1.5 DHCP Dual Stack Access.....	35
5.1.6 L2 Static Subscriber Access.....	37
5.2 PPPoE.....	40
5.2.1 IPv4 PPPoE Access.....	40
5.2.2 IPv6 PPPoE Access.....	41
5.2.3 PPPoE Dual Stack Access.....	43
5.3 WLAN Access.....	45
5.4 L2TP.....	47
5.4.1 L2TP LAC Access.....	47
5.4.2 L2TP LNS IPv4 Access.....	49
5.4.3 L2TP LNS IPv6 Access.....	51
5.5 CGN (Carrier Grade NAT).....	54

Table of Contents (continued)

5.6 L3 Leased Line Access.....	55
5.6.1 Web Authentication.....	55
5.6.2 User Traffic Trigger.....	57
5.7 Multicast Access.....	58
6. S-CUSP Message Formats.....	60
6.1 Common Message Header.....	60
6.2 Control Messages.....	61
6.2.1 Hello Message.....	61
6.2.2 Keepalive Message.....	62
6.2.3 Sync_Request Message.....	62
6.2.4 Sync_Begin Message.....	62
6.2.5 Sync_Data Message.....	63
6.2.6 Sync_End Message.....	63
6.2.7 Update_Request Message.....	64
6.2.8 Update_Response Message.....	64
6.3 Event Message.....	65
6.4 Report Message.....	66
6.5 CGN Messages.....	66
6.5.1 Addr_Allocation_Req Message.....	66
6.5.2 Addr_Allocation_Ack Message.....	66
6.5.3 Addr_Renew_Req Message.....	67
6.5.4 Addr_Renew_Ack Message.....	67
6.5.5 Addr_Release_Req Message.....	67
6.5.6 Addr_Release_Ack Message.....	67
6.6 Vendor Message.....	67
6.7 Error Message.....	68
7. S-CUSP TLVs and Sub-TLVs.....	69
7.1 Common TLV Header.....	69
7.2 Basic Data Fields.....	70
7.3 Sub-TLV Format and Sub-TLVs.....	71
7.3.1 Name sub-TLVs.....	71
7.3.2 Ingress-CAR sub-TLV.....	72
7.3.3 Egress-CAR sub-TLV.....	72
7.3.4 If-Desc sub-TLV.....	73
7.3.5 IPv6 Address List sub-TLV.....	75
7.3.6 Vendor sub-TLV.....	75
7.4 The Hello TLV.....	77
7.5 The Keep Alive TLV.....	78
7.6 The Error Information TLV.....	79
7.7 BAS Function TLV.....	79
7.8 Routing TLVs.....	82
7.8.1 IPv4 Routing TLV.....	82
7.8.2 IPv6 Routing TLV.....	84
7.9 Subscriber TLVs.....	85
7.9.1 Basic Subscriber TLV.....	86
7.9.2 PPP Subscriber TLV.....	88
7.9.3 IPv4 Subscriber TLV.....	89

Table of Contents (continued)

7.9.4 IPv6 Subscriber TLV.....	90
7.9.5 IPv4 Static Subscriber Detect TLV.....	91
7.9.6 IPv6 Static Subscriber Detect TLV.....	93
7.9.7 L2TP-LAC Subscriber TLV.....	94
7.9.8 L2TP-LNS Subscriber TLV.....	95
7.9.9 L2TP-LAC Tunnel TLV.....	95
7.9.10 L2TP-LNS Tunnel TLV.....	96
7.9.11 Update Response TLV.....	97
7.9.12 Subscriber Policy TLV.....	98
7.9.13 Subscriber CGN Port Range TLV.....	100
7.10 Device Status TLVs.....	100
7.10.1 Interface Status TLV.....	101
7.10.2 Board Status TLV.....	101
7.11 CGN TLVs.....	102
7.11.1 Address Allocation Request TLV.....	102
7.11.2 Address Allocation Response TLV.....	103
7.11.3 Address Renewal Request TLV.....	104
7.11.4 The Address Renewal Response TLV.....	105
7.11.5 Address Release Request TLV.....	106
7.11.6 The Address Release Response TLV.....	106
7.12 Event TLVs.....	107
7.12.1. Subscriber Traffic Statistics TLV.....	108
7.12.2 Subscriber Detection Result TLV.....	109
7.13 Vendor TLV.....	110
8. Implementation Status.....	112
8.1 Implementations.....	112
8.1.1 Huawei Technologies.....	112
8.1.2 ZTE.....	113
8.1.3 H3C.....	113
8.2 Hackathon.....	113
8.3 EANTC Testing.....	114
9. IANA Considerations.....	115
9.1 Message Types.....	115
9.2 TLV Types.....	115
9.3 TLV Operation Codes.....	117
9.4 Sub-TLV Types.....	118
9.5 Error Codes.....	118
10. Security Considerations.....	120
Contributors.....	121
Normative References.....	122
Informative References.....	123
Authors' Addresses.....	125

1. Introduction

A fixed network Broadband Network Gateway (BNG) is an Ethernet-centric IP edge router, and the aggregation point for user traffic. To provide centralized session management, flexible address allocation, high scalability for subscriber management capacity, and cost-efficient redundancy, the Control/User (CU) separated BNG framework is described in [TR-384]. The CU separated service Control Plane (CP), which is responsible for user access authentication and setting forwarding entries in User Planes (UPs), can be virtualized and centralized. The routing control and forwarding plane, i.e. the BNG user plane (local), can be distributed across the infrastructure. Other structures can also be supported such as both CP and UP being virtual or both being physical.

This document specifies the Simple CU Separation BNG control channel Protocol (S-CUSP) for communications between a BNG Control Plane (CP) and a set of User Planes (UPs). S-CUSP is designed to be flexible and extensible so as to easily allow for additional messages and data items, should further requirements be expressed in the future.

2. Terminology

This section specifies implementation requirement keywords and terms used in this document. S-CUSP messages are described in this document using Routing Backus-Naur Form (RBNF) as defined in [RFC5511].

2.1 Implementation Requirement Keywords

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.2 Terms

This section specifies terms used in this document.

AAA: Authentication Authorization Accounting.

ACK: Acknowledgement message.

BAS: Broadband Access Server (BRAS, BNG).

BNG: Broadband Network Gateway. A broadband remote access server (BRAS (BRoadband Access Server), B-RAS or BBRAS) routes traffic to and from broadband remote access devices such as digital subscriber line access multiplexers (DSLAM) on an Internet Service Provider's (ISP) network. BRAS can also be referred to as a Broadband Network Gateway (BNG).

BRAS: BRoadband Access Server (BNG).

CAR: Committed Access Rate.

CBS: Committed Burst Size.

CGN: Carrier Grade NAT.

Ci: Control Interface.

CIR: Committed Information Rate.

CoA: Change of Authorization.

CP: Control Plane.

CP is a user control management component which supports the management of the UP's resources such as the user entry and forwarding policy.

CPE: Customer Premises Equipment.

CU: Control-plane / User-plane.

CUSP: Control and User plane Separation Protocol.

DEI: Drop Eligibility Indicator. A bit in a VLAN tag after the priority and before the VLAN ID. (This bit was formerly the CFI (Canonical Format Indicator).) [802.1Q]

DHCP: Dynamic Host Configuration Protocol [RFC2131].

dial-up: This refers to the initial connection messages when a new user appears. The name is left over from when users literally dialed up on a modem equipped phone line but herein is applied to other initial connection techniques. Initial connection is frequently indicated by the receipt of packets over PPPoE [RFC2516] or IPoE.

EMS: Element Management System.

IPoE: IP over Ethernet.

L2TP: Layer 2 Tunneling Protocol [RFC2661].

LAC: L2TP Access Concentrator.

LNS: L2TP Network Server.

MAC: 48-bit Media Access Control address [RFC7042].

MANO: Management and Orchestration.

Mi: Management Interface.

MSS: Maximum Segment Size.

MRU: Maximum Receive Unit.

NAT: Network Address Translation [RFC3022].

ND: Neighbor Discovery.

NFV: Network Function Virtualization.

NFVI: NFV Infrastructure

PBS: Peak Burst Size.

PD: Prefix Delegation.

PIR: Peak Information Rate.

PPP: Point to Point Protocol [RFC1661].

PPPoE: PPP over Ethernet [RFC2516].

RBNF: Routing Backus-Naur Form [RFC5511].

RG: Residential Gateway.

S-CUSP: Simple Control and User Plane Separation Protocol.

Si: Service Interface.

TLV: Type, Length, Value. See Sections 7.1 and 7.3.

UP: User Plane. UP is a network edge and user policy implementation component. The traditional router's Control Plane and Forwarding Plane are both preserved on BNG devices in the form of a user plane.

URPF: Unicast Reverse Path Forwarding.

User: Equivalent to "customer" or "subscriber".

VRF: Virtual Routing and Forwarding.

3. BNG CUPS Overview

3.1 BNG CUPS Motivation

The rapid development of new services, such as 4K TV, IoT, etc., and increasing numbers of home broadband service users present some new challenges for BNGs such as:

Low resource utilization: The traditional BNG acts as both a gateway for user access authentication and accounting and an IP network's Layer 3 edge. The mutually affecting nature of the tightly coupled control plane and forwarding plane makes it difficult to achieve the maximum performance of either plane.

Complex management and maintenance: Due to the large numbers of traditional BNGs, configuring each device in a network is very tedious when deploying global service policies. As the network expands and new services are introduced, this deployment mode will cease to be feasible as it is unable to manage services effectively and rectify faults rapidly.

Slow service provisioning: The coupling of control plane and forwarding plane, in addition to a distributed network control mechanism, means that any new technology has to rely heavily on the existing network devices.

To address these challenges for fixed networks, the framework for a cloud-based BNG with Control Plane and User Plane (CU) separation is described in [TR-384]. The main idea of CU separation is to extract and centralize the user management functions of multiple BNG devices, forming a unified and centralized Control Plane (CP). And the traditional router's Control Plane and Forwarding Plane are both preserved on BNG devices in the form of a User Plane (UP).

3.2 BNG CUPS Architecture Overview

The functions in a traditional BNG can be divided into two parts: one is the user access management function, the other is the router function. The user management function can be centralized and deployed as a concentrated module or device, called the BNG Control Plane (BNG-CP). The other functions, such as the router function and forwarding engine, can be deployed in the form of the BNG User Plane (BNG-UP).

The following figure shows the architecture of CU separated BNG:

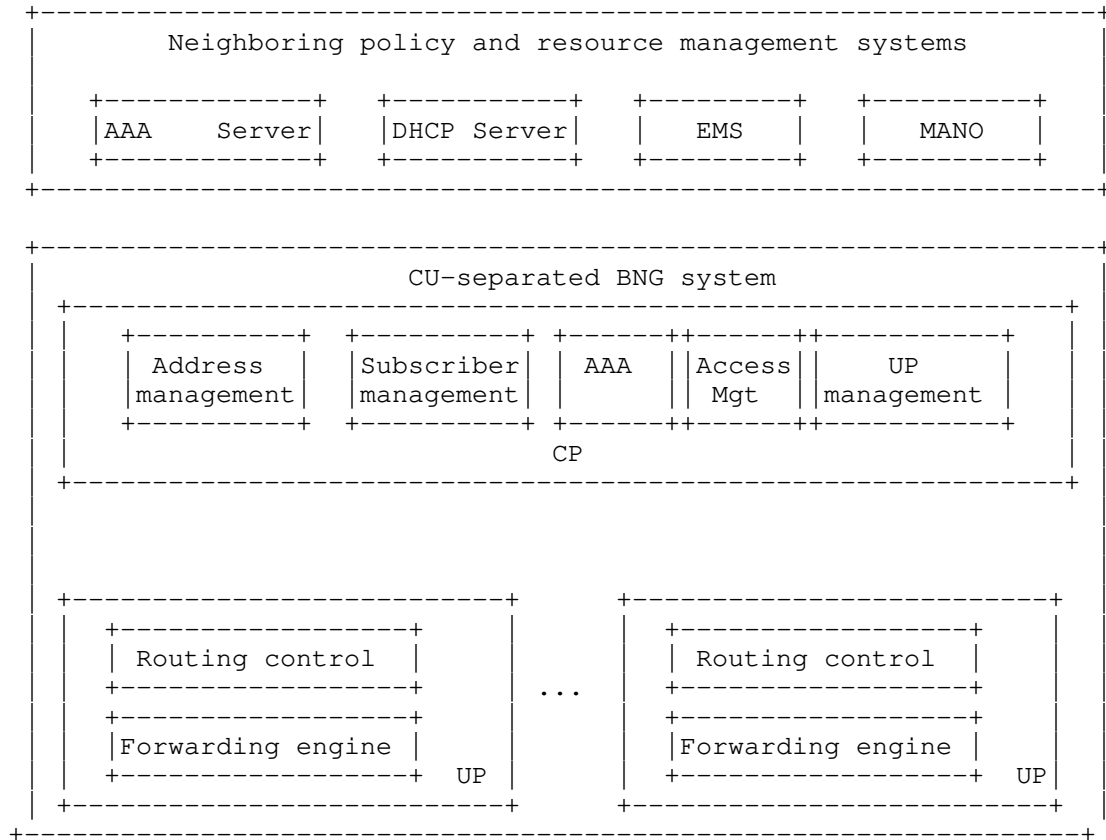


Figure 1: Architecture of CU Separated BNG

As shown in Figure 1, the BNG Control Plane could be virtualized and centralized, which provides benefits such as centralized session management, flexible address allocation, high scalability for subscriber management capacity, and cost-efficient redundancy, etc. The functional components inside the BNG Service Control Plane can be implemented as Virtual Network Functions (VNFs) and hosted in a Network Function Virtualization Infrastructure (NFVI).

The User Plane Management module in the BNG Control Plane centrally manages the distributed BNG User Planes (e.g. load balancing), as well as the setup, deletion, and maintenance of channels between Control Planes and User Planes. Other modules in the BNG control plane, such as address management, AAA, etc., are responsible for the connection with outside subsystems in order to fulfill those services. Note that the User Plane SHOULD support both physical and virtual network functions. For example, BNG user plane L3 forwarding

related network functions can be disaggregated and distributed across the physical infrastructure. And the other control plane and management plane functions in the CU Separation BNG can be moved into the NFVI for virtualization [TR-384].

The details of CU separated BNG's function components are as following:

The Control Plane is responsible for the following:

1. Address management: unified address pool management and CGN subscriber address traceability management.
2. AAA: This component performs Authentication, Authorization and Accounting, together with RADIUS/DIAMETER. The BNG communicates with the AAA server to check whether the subscriber who sent an Access-Request has network access authority. Once the subscriber goes online, this component together with the Service Control component implement accounting, data capacity limitation, and QoS enforcement policies.
3. Subscriber management: user entry management and forwarding policy management.
4. Access management: process user dial-up packets, such as PPPoE, DHCP, L2TP, etc.
5. UP management: management of UP interface status, and the setup, deletion, and maintenance of channels between CP and UP.

The User Plane is responsible for the following:

1. Routing control functions: responsible for constructing routing forwarding plane (e.g., routing, multicast, MPLS, etc.).
2. Routing and Service Forwarding plane functions: responsible including traffic forwarding, QoS and traffic statistics collection.

Subscriber detection: responsible for detecting whether a subscriber is still online.

3.3 BNG CUPS Interfaces

To support the communication between the Control Plane and User Plane, three interfaces are assumed. These are referred to as the Service Interface (Si), Control Interface (Ci), and Management Interface (Mi) as shown in Figure 2.

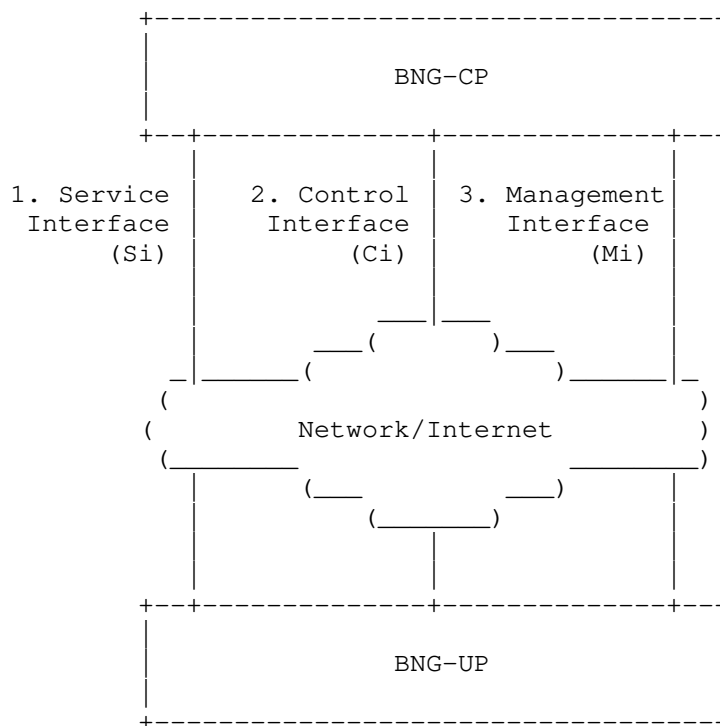


Figure 2: Interfaces Between the CP and UP of the BNG

3.3.1 Service Interface

For a traditional BNG (without CU separation), the user dial-up signals are terminated and processed by the control plane of a BNG. When the CP and UP of a BNG are separated, there needs to be a way to relay these signals between the CP and the UP.

The Service Interface (Si) is used to establish tunnels between the CP and UP. The tunnels are responsible for relaying the PPPoE, IPoE, and L2TP related control packets that are received from a Residential Gateway (RG) over those tunnels. An appropriate tunnel type is VXLAN [RFC7348].

The detailed definition of Si is out of scope for this document.

3.3.2 Control Interface

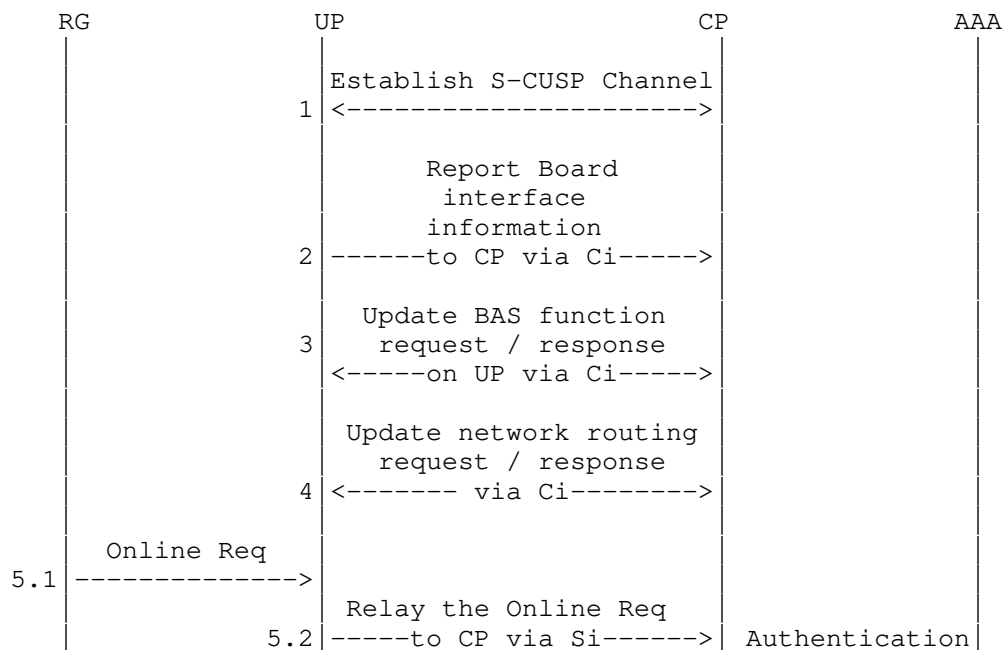
The CP uses the Control Interface to deliver subscriber session states, network routing entries, etc. to the UP (see Section 6.2.7)). The UP uses this interface to report subscriber service statistics, subscriber detection results, etc. to the CP (see Sections 6.3 and 6.4). A carrying protocol for this interface is specified in this document.

3.3.3 Management Interface

NETCONF [RFC6241] is the protocol used on the Management Interface between a CP and UP. It is used to configure the parameters of the Control Interface, Service Interface, the Access interfaces and QoS/ACL Templates. It is expected that implementations will make use of existing YANG models where possible, but that new YANG models specific to S-CUSP will need to be defined. The definitions of the parameters are out of scope for this document.

3.4 BNG CUPS Procedure Overview

The following numbered sequences (Figure 3) gives a high level view of the main BNG CUPS procedures.



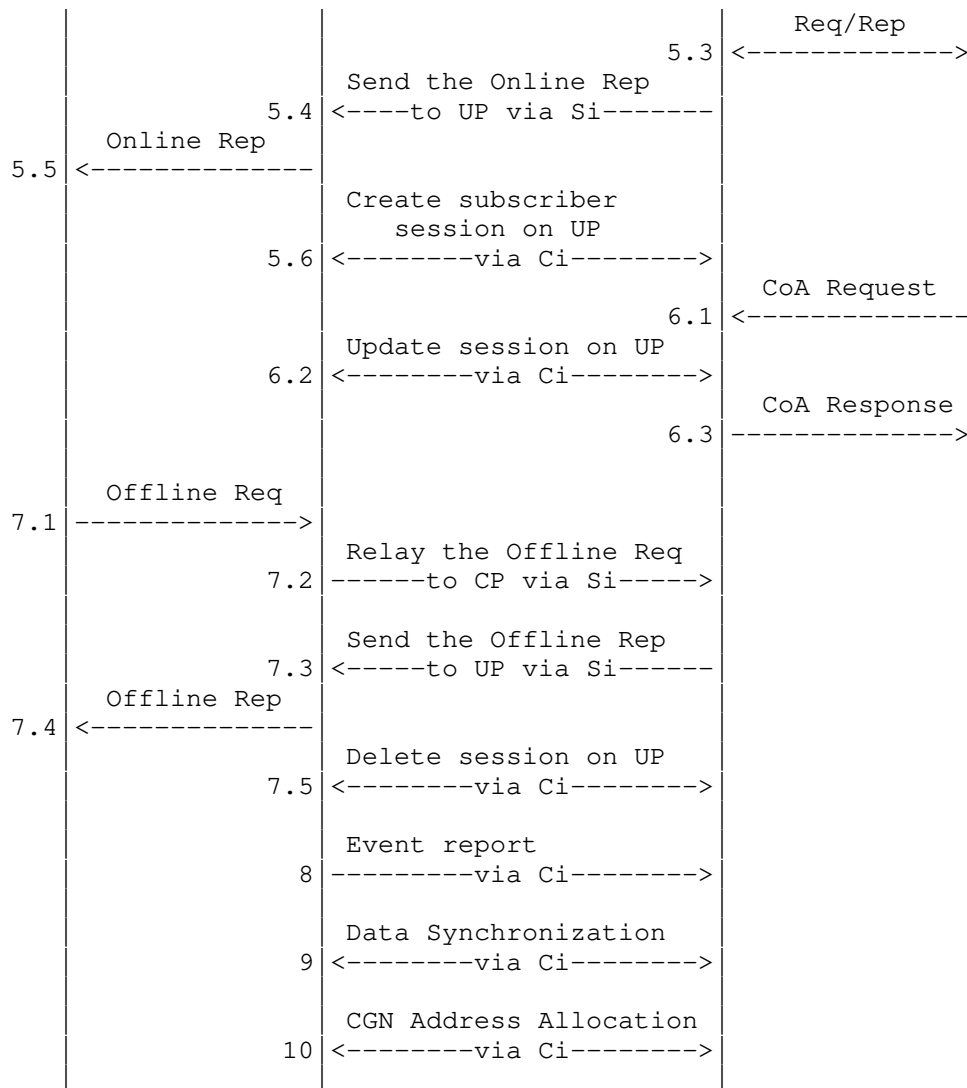


Figure 3: BNG CUPS Procedures Overview

1. S-CUSP session establishment: This is the first step of BNG CUPS procedures. Once the Control Interface parameters are configured on a UP. It will start to setup S-CUSP sessions with the specified CPs. The detailed definition of S-CUSP session establishment can be found in Section 4.1.1.
2. Board and interface report: Once the S-CUSP session is established between the UP and a CP, the UP will report status information on the boards and subscriber side interfaces of this UP to the CP. A board can also be called a Line/Service Process

Unit (LPU/SPU) card. The subscriber side interfaces refer to the interfaces that connect the Access Network nodes (e.g., OLT: Optical Line Terminal, DSLAM: Digital Subscriber Line Access Multiplexer, etc.). The CP can use this information to enable the Broadband Access Service (BAS) function (e.g., IPoE, PPPoE, etc.) on the specified interfaces. See Sections 4.2.1 and 7.10 for more details on Resource reporting.

3. BAS (Broadband Access Service) function enable: To enable the BAS function on the specified interfaces of a UP.
4. Subscriber network route advertisement: The CP will allocate one or more IP address blocks to a UP. Each address block contains a series of IP addresses. Those IP addresses will be allocated to subscribers who are dialing up from the UP. To enable other nodes in the network to learn how to reach the subscribers, the CP needs to notify the UP to advertise to the network the routes that can reach those IP addresses.
5. 5.1-5.6 is a complete call flow of a subscriber dial-up process. When a UP receives a dial-up request, it will relay the request packet to a CP through the Service Interface. The CP will parse the request. If everything is OK, it will send an authentication request to the AAA server to authenticate the subscriber. Once the subscriber passes the authentication, the AAA server will return a positive response to the CP. Then the CP will send the dial-up response packet to the UP and the UP will forward the response packet to the subscriber (RG). At the same time, the CP will create a subscriber session on the UP, which enables the subscriber to access the network. For different access types, the process may be a bit different. But the high-level process is similar. For each access type, the detail process can be found in Section 5.
6. 6.1-6.3 is the sequence when updating an existing subscriber session. The AAA server initiates a Change of Authorization (CoA) and sends the CoA to the CP. The CP will then update the session according to the CoA. See Section 4.3.2 for more detail on CP messages updating UP tables.
7. 7.1-7.5 is the sequence for deleting an existing subscriber session. When a UP receives an offline request, it will relay the request to a CP through the Service Interface. The CP will send back a response to the UP through the Service Interface. The UP will then forward the offline response to the subscriber. Then the CP will delete the session on the UP through the Control Interface.

8. Event reports include the following two parts (more detail can be found in Section 4.3.4) Both are reported using the Event message.
 - 8.1 Subscriber Traffic Statistics Report
 - 8.2 Subscriber Detection Result Report
9. Data synchronization: See Sections 4.2.5 for more detail on CP and UP Synchronization.
10. CGN address allocation: See Sections 4.2.4 for more detail on CGN Address Allocation.

4. S-CUSP Protocol Overview

4.1 Control Channel Related Procedures

4.1.1 S-CUSP Session Establishment

A UP is associated with a CP and is controlled by that CP. In the case of a hot-standby or cold-standby, a UP is associated with two CPs, one called the Master CP and the other called the Standby CP. The association between a UP and its CPs is implemented by dynamic configuration.

Once a UP knows its CPs, the UP starts to establish S-CUSP sessions with those CPs as shown in Figure 4.

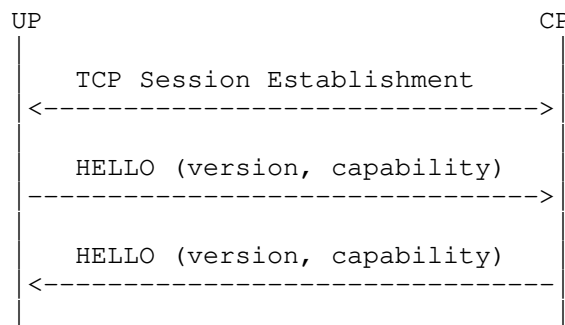


Figure 4: S-CUSP Session Establishment

The S-CUSP session establishment consists of two successive steps:

1. Establishment of a TCP [RFC793] connection (3-way handshake) between the CP and the UP using a configured port from the dynamic port range (49152-65535).
2. Establishment of a S-CUSP session over the TCP connection.

Once the TCP connection is established, the CP and the UP initialize the S-CUSP session during which the version and Keepalive timers are negotiated.

The version information (Hello TLV, see Section 7.4) is carried within Hello messages (see Section 6.2.1). A CP can support multiple versions, but a UP can only support one version. So, the version negotiation is based on whether a version can be supported by both the CP and the UP. For a CP or UP, if a Hello message is received that

does not indicate a version supported by both, a subsequent Hello message with an Error Information TLV will be sent to the peer to notify the peer of the "Version-Mismatch" error and the session establishment phase fails.

Keepalive negotiation is performed by carrying a Keepalive TLV in the Hello message. The Keepalive TLV includes a Keepalive timer and Dead Timer field. The CP and UP have to agree on the Keepalive Timer and Dead Timer. Otherwise, a subsequent Hello message with an Error Information TLV will be sent to its peer and the session establishment phase fails.

The S-CUSP session establishment phase fails if the CP or UP disagree on the version and keepalive parameters or if one of the CP or UP does not answer after the expiration of the Establishment timer. When the S-CUSP session establishment fails, the TCP connection is promptly closed. Successive retries are permitted but an implementation SHOULD make use of an exponential back-off session establishment retry procedure.

The S-CUSP session timer values that need to be configured are summarized in the table below.

Timer Name	Range in seconds	Default Value
-----	-----	-----
Establishment	1-32767	45
Keepalive	0-255	30
DeadTimer	1-32767	4 * Keepalive

4.1.2 Keep Alive

Once an S-CUSP session has been established, a UP or CP may want to know that its S-CUSP peer is still available for use.

Each end of a S-CUSP session runs a Keepalive timer. It restarts the timer every time it sends a message on the session. When the timer expires, it sends a Keepalive message.

The ends of the S-CUSP session also run DeadTimers, and they restart the timers whenever a message is received on the session. If one end of the session receives no message after the DeadTimer expires, it declares the session dead. The session will be closed.

The minimum value of the Keepalive timer is 1 second, and it is specified in units of 1 second. The RECOMMENDED default value is 30 seconds. The timer may be disabled by setting it to zero.

The recommended default for the DeadTimer is 4 times the value of the Keepalive timer used by the remote peer. This implies there is essentially no risk of TCP congestion due to excessive Keepalive messages.

The Keepalive timer and DeadTimer are initially negotiated through the Keepalive TLV carried in the Hello Message.

4.2 Node Related Procedures

4.2.1 UP Resource Report

Once an S-CUSP session has been established between a CP and an UP. The UP reports the information of the Boards and access side interfaces on this UP to the CP as shown in Figure 5. Report messages are unacknowledged and are assumed to be delivered because the session runs over TCP.

The CP can use that information to activate/enable the Broadband Access Service (BAS) functions (e.g., IPoE, PPPoE, etc.) on the specified interfaces.

In addition, the UP resource report may trigger a UP warm-standby process. In the case of warm-standby, a failure on an UP may trigger the CP to start a warm-standby process, by moving the on-line subscriber sessions to a standby UP and then direct the affected subscribers to access the Internet through the standby UP.

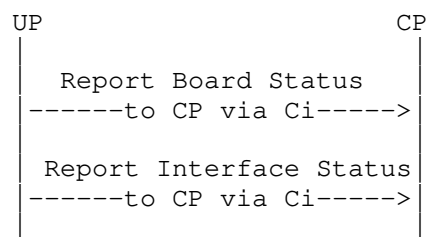


Figure 5: UP Board and Interface Report

Board status information is carried in the Board Status TLV (Section 7.10.2) and Interface status information is carried in Interface Status TLV (Section 7.10.1). Both Board and Interface Status TLVs are carried in the Report Message (Section 6.4).

4.2.2 Update BAS Function on Access Interface

Once the CP collects the interface status of a UP, it will activate/de-activate/modify the BAS functions on specified interfaces through the Update_Request and Update_Response message (Section 6.2) exchanges carrying the BAS Function TLV (Section 7.7).

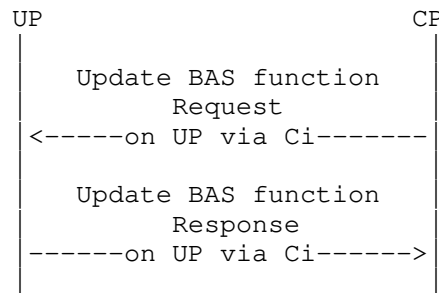


Figure 6: Update BAS Function

4.2.3 Update Network Routing

The CP will allocate one or more address blocks to a UP. Each address block contains a series of IP addresses. Those IP addresses will be allocated to subscribers who are dialing up to the UP. To enable the other nodes in the network to learn how to reach the subscribers, the CP needs to install the routes on the UP and notify the UP to advertise the routes to the network.

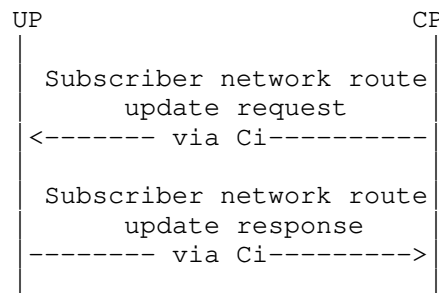


Figure 7: Update Network Routing

The subscriber network routing update request and response are achieved through the Update Request and Response Message exchanges by carrying the IPv4/IPv6 Routing Information TLVs (Section 7.8).

4.2.4 CGN Public IP Address Allocation

The following sequences describe the CGN address management related procedures. Three independent procedures are defined, one each for CGN address allocation request/response, CGN address renewal request/response, and CGN address release request/response.

CGN address allocation/renew/release procedures are designed for the case where the CGN function is running on the UP. The UP has to map the subscriber private IP addresses to a public IP addresses, and such mapping is performed by the UP locally when a subscriber dials-up. That means the UP has to ask for public IPv4 address blocks for CGN subscribers from the CP.

In addition, when a public IP address is allocated to a UP, there will be a lease time (e.g., one day). Before the lease time expires, the UP can ask for renewal of the IP address lease from the CP. It is achieved by the exchange of the Addr_Renew_Req and Addr_Renew_Ack messages.

If the public IP address will not be used anymore, the UP SHOULD release the address by sending an Addr_Release_Req message to the CP.

If the CP wishes to withdraw addresses that it has previously leased to a UP, it uses the same procedures as above. The "Oper" code in the IPv4/IPv6 Routing TLV (see Section 7.1) determines whether the request is an update or withdraw.

The relevant messages are defined in Section 6.5.

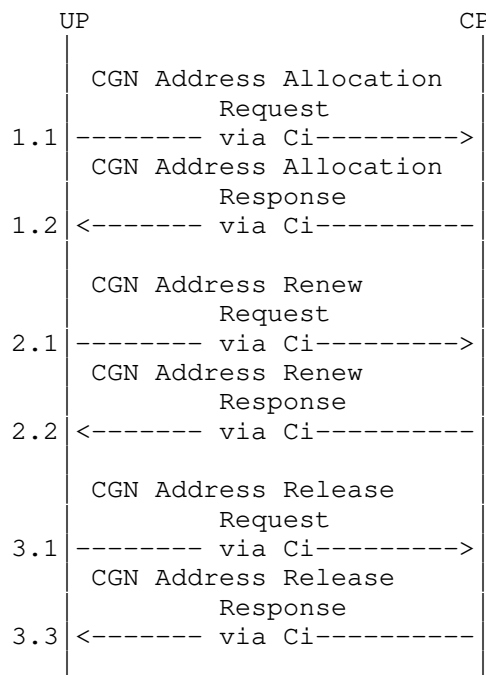


Figure 8: CGN Public IP Address Allocation

4.2.5 Data Synchronization between the CP and UP

For a CU separated BNG, the UP will continue to function using the state that has been installed in it even if the CP fails or the session between the UP and CP fails.

Under some circumstances it is necessary to synchronize state between the CP and UP, for example if a CP fails and the UP is switched to a different CP.

Synchronization includes two directions. One direction is from UP to CP; in that case, the synchronization information is mainly about the board/interface status of the UP. The other direction is from CP to UP; in that case, the subscriber sessions, subscriber network routes, L2TP tunnels, etc. will be synchronized to the UP.

The synchronization is triggered by a Sync_Request message, to which the receiver will (1) reply with a Sync_Begin message to notify the requester that synchronization will begin, and (2) then start the synchronization using the Sync_Data message. When synchronization finished, a Sync_End message will be sent.

The following figure shows the process of data synchronization between a UP and a CP.

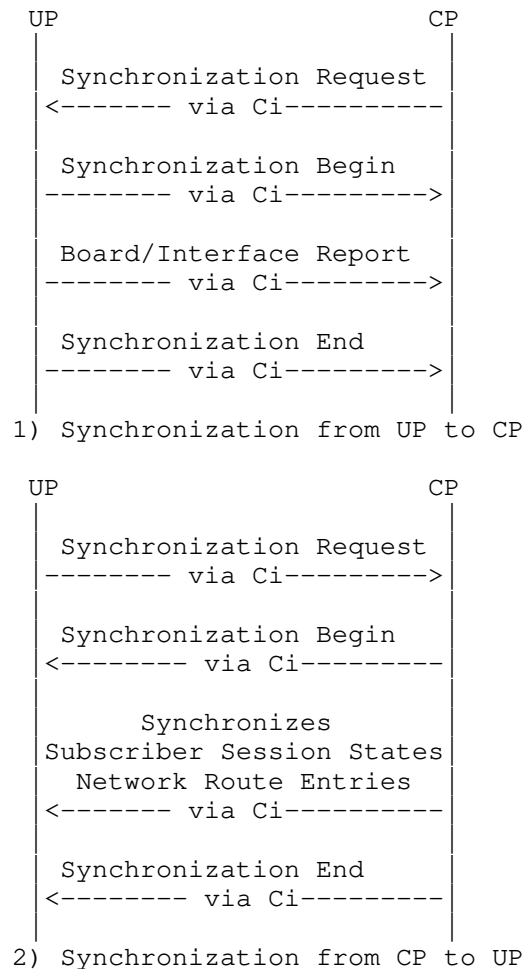


Figure 9: Data Synchronization

4.3 Subscriber Session Related Procedures

A subscriber session consists of a set of forwarding states, policies, and security rules that are applied to the subscriber. It is used for forwarding subscriber traffic in a UP. To initialize a session on a UP, a set of hardware resource have to be allocated (e.g., NP, TCAM etc.) to a session.

Subscriber session related procedures include subscriber session

create, update, delete, and statistics report. The following sub-sections give a high level view of the procedures.

4.3.1 Create Subscriber Session

The below sequence describes the DHCP IPv4 dial-up process, it is an example that shows how a subscriber session is created. (An example for IPv6 appears in Section 5.1.2.)

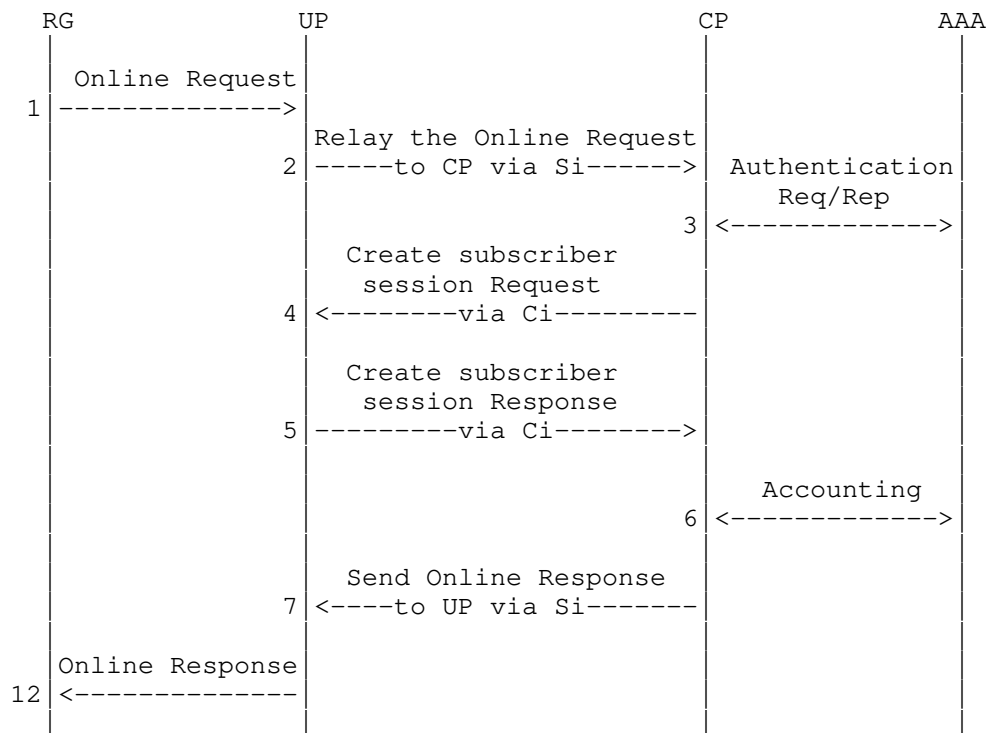


Figure 10: Subscriber Session Create

The request starts from an Online Request message (step 1) from the RG (for example, a DHCP Discovery packet). When the UP receives the Online Request from the RG, it will tunnel the Online Request to the CP through the Service Interface (Step 2). The Service Interface is implemented by a tunneling technology.

When the CP receives the Online Request from the UP, it will send an authentication request to the AAA server to authenticate and authorize the subscriber (step 3). When a positive reply is received from the AAA sever, the CP starts to create a subscriber session for the request. Relevant resources (e.g., IP address, bandwidth, etc.)

will be allocated to the subscriber, policies and security rules will be generated for the subscriber. Then the CP sends a session create request to the UP through the Control Interface (Ci) (step 4), and a response is expected from the UP to confirm the creation (step 5).

Finally, the CP will notify the AAA server to start accounting (step 6). At the same time, an Online Response message (for example, a DHCP Ack packet) will be sent to the UP through the Si (step 7). And the UP will forward the Online Response to the RG (step 8).

This completes the subscriber online process.

4.3.2 Update Subscriber Session

The following numbered sequence shows the process of subscriber session update.

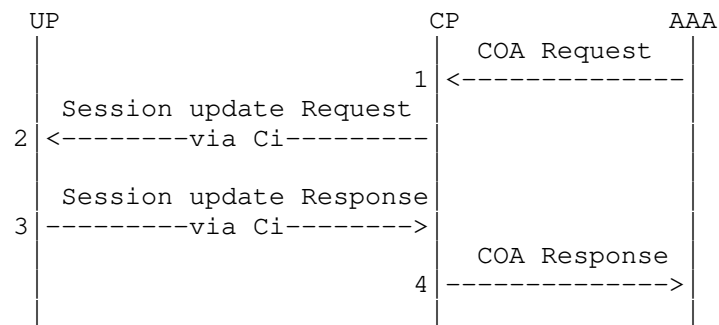


Figure 11: Subscriber Session Update

When a subscriber session has been created on a UP, there may be requirements to update the session with new parameters (e.g., Bandwidth, QoS, policies, etc.).

This procedure is triggered by a Change of Authorization (COA) request message sent by the AAA server. The CP will update the session on the UP according to the new parameters through the Control Interface.

4.3.3 Delete Subscriber Session

The below call flow shows generally how S-CUPS deals with a subscriber offline request.

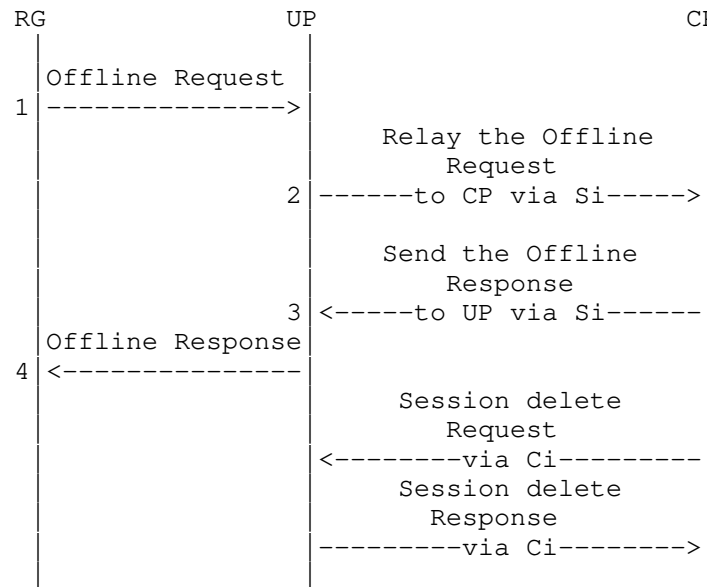


Figure 12: Subscriber Session Delete

Similar to the session creation process, when a UP receives an offline request from a RG, it will tunnel the request to a CP through the Si.

When the CP receives the offline request, it will withdraw/release the resources (e.g., IP address, bandwidth) that have been allocated to the subscriber. Then, it sends a reply to the UP through the Service Interface and the UP will forward the reply to the RG. At the same time, it will delete all the status of the session on the UP through the Ci.

4.3.4 Subscriber Session Events Report

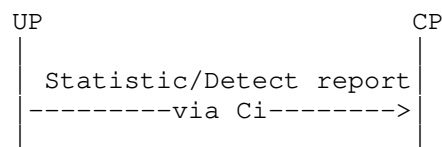


Figure 13: Events Report

When a session is created on an UP, the UP will periodically report statistics information and detect results of the session to the CP.

5. S-CUSP Call Flows

The subsections below give an overview of various "dial-up" interactions over the Service Interface followed by an overview of the setting of various information in the UP by the CP using S-CUSP over the Control Interface.

S-CUSP messages are described in this document using Routing Backus Naur Form (RBNF) as defined in [RFC5511].

5.1 IPoE

5.1.1 DHCPv4 Access

The following sequence shows detailed procedures for DHCPv4 access.

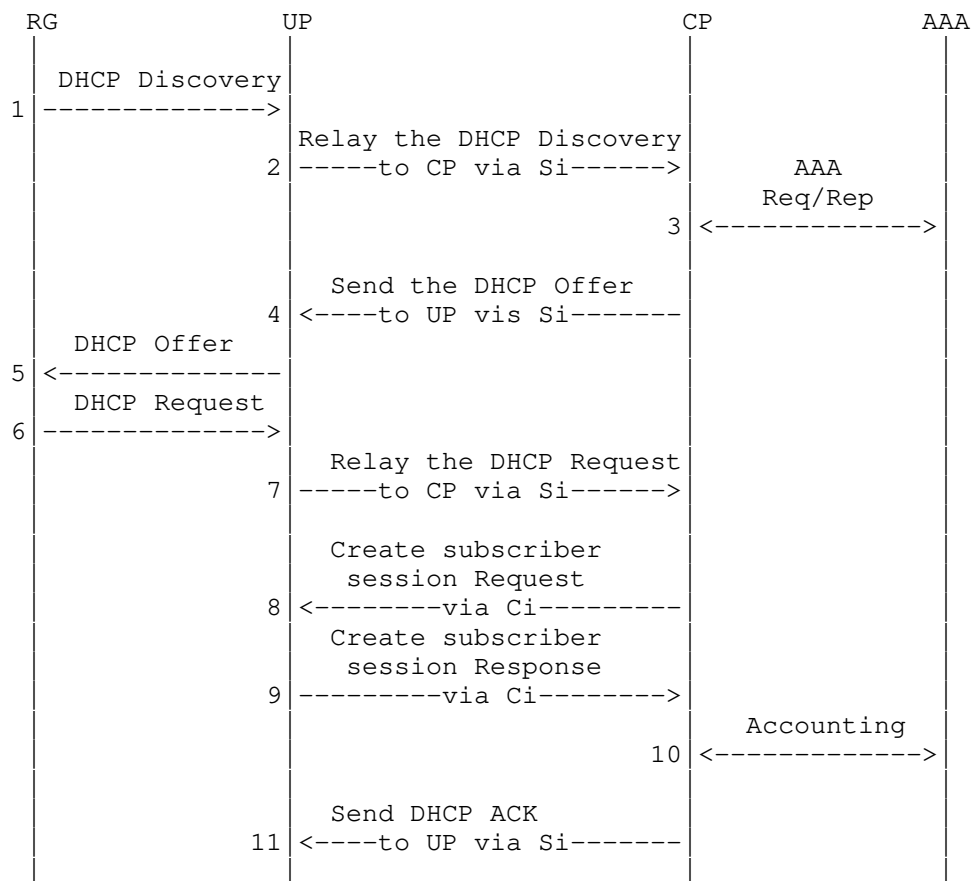




Figure 14: DHCPv4 Access

Step 8 and 9 are implemented by the S-CUSP protocol.

When a subscriber is authenticated and authorized by the AAA server, the CP will create a subscriber session on the UP. This is achieved by sending an Update_Request message to the UP.

The format of the Update_Request message is shown as follows using RBNF:

```

<Update_Request Message> ::= <Common Header>
                             <Basic Subscriber TLV>
                             <IPv4 Subscriber TLV>
                             <IPv4 Routing TLV>
                             [<Subscriber Policy TLV>]
  
```

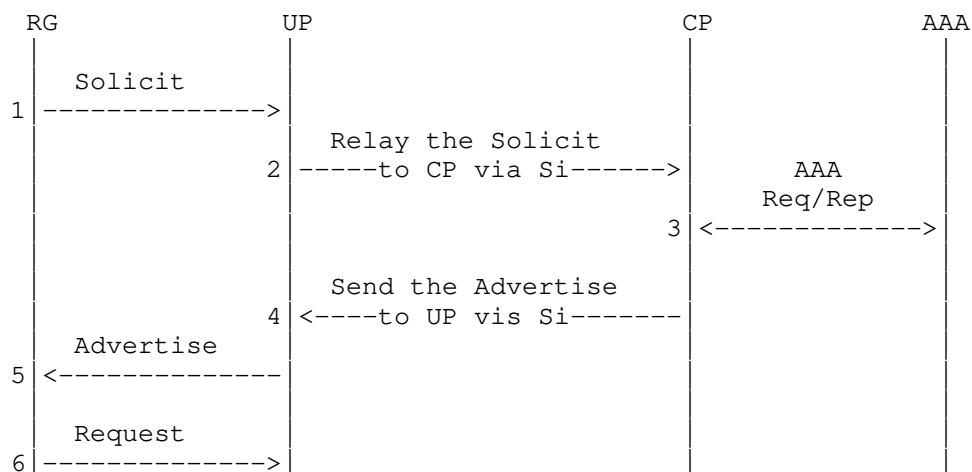
The UP will reply with an Update_Response message, the format of the Update_Response message is as follows:

```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]
  
```

5.1.2 DHCPv6 Access

The following sequence shows detailed procedures for DHCPv6 access.



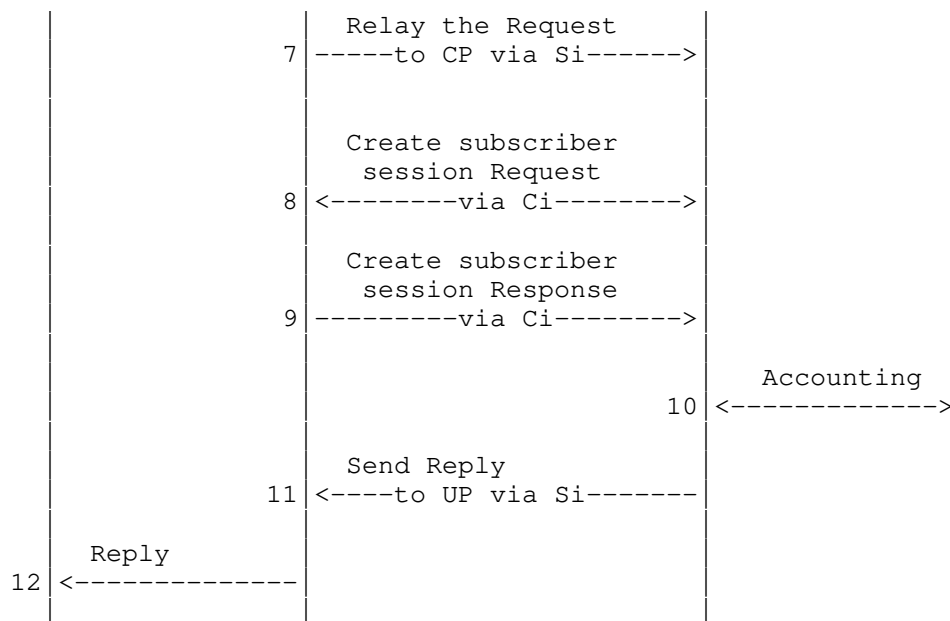


Figure 15: DHCPv6 Access

Steps 1-7 are a standard DHCP IPv6 access process. The subscriber creation is triggered by a DHCP IPv6 request message. When this message is received, it means that the subscriber has passed the AAA authentication and authorization. Then the CP will create a subscriber session on the UP. This is achieved by sending an Update_Request message to the UP (Step 8).

The format of the Update_Request message is as follows:

```

<Update_Request Message> ::= <Common Header>
                             <Basic Subscriber TLV>
                             <IPv6 Subscriber TLV>
                             <IPv6 Routing TLV>
                             [<Subscriber Policy TLV>]
  
```

The UP will reply with an Update_Response message (Step 9). The format of the Update_Response message is as follows:

```

<Update_Response Message> ::= <Common Header>
                             <Update Response TLV>
  
```

5.1.3 IPv6 SLAAC Access

The following flow shows the IPv6 SLAAC access process.

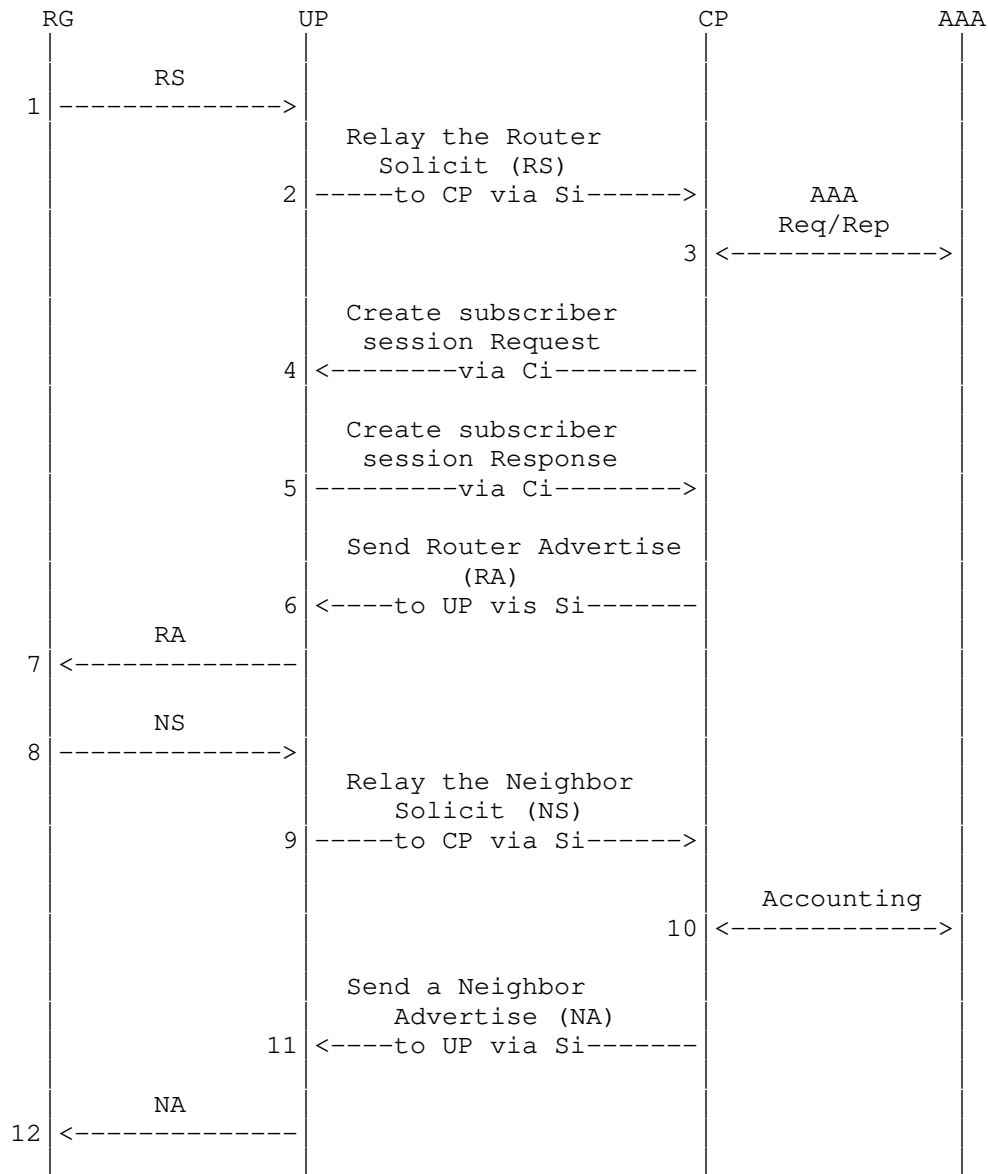


Figure 16: IPv6 SLAAC Access

It starts with a Router Solicit (RS) request from an RG that is tunneled to the CP by the UP. After the AAA authentication and authorization, the CP will create a subscriber session on the UP.

This is achieved by sending an Update_Request message to the UP (step 4).

The format of the Update_Request message is as follows:

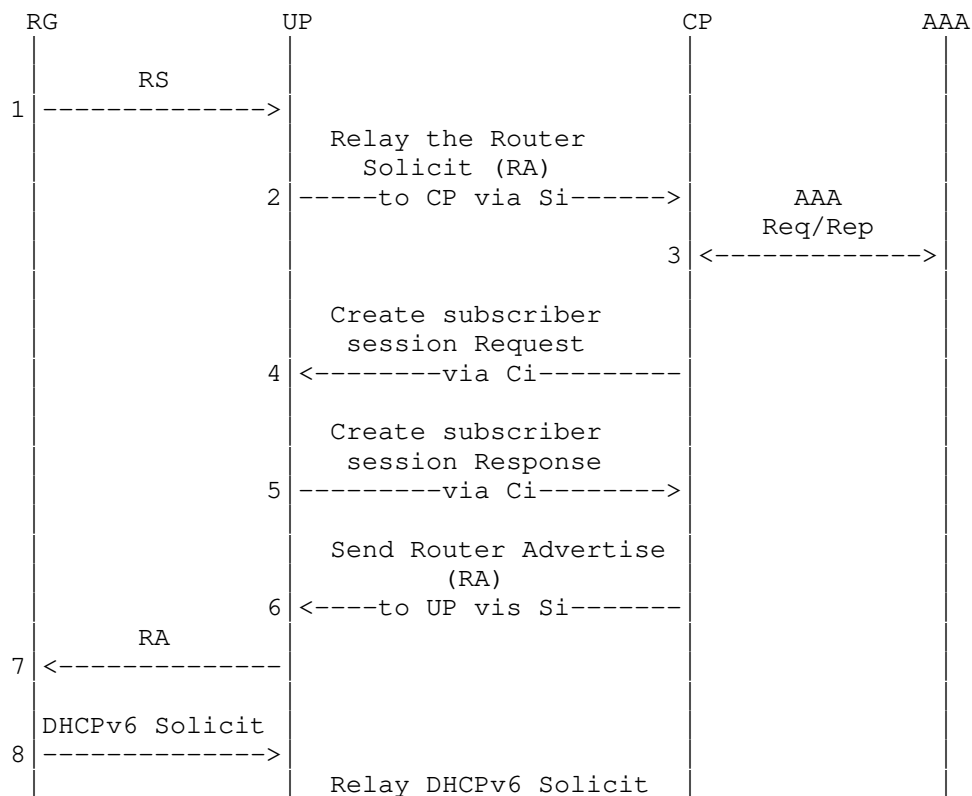
```
<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               [<Subscriber Policy TLV>]
```

The UP will reply with an Update_Response message (step 5), the format of the Update_Response message is as follows:

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

5.1.4 DHCPv6 + SLAAC Access

The following call flow shows the DHCP IPv6 and SLAAC access process.



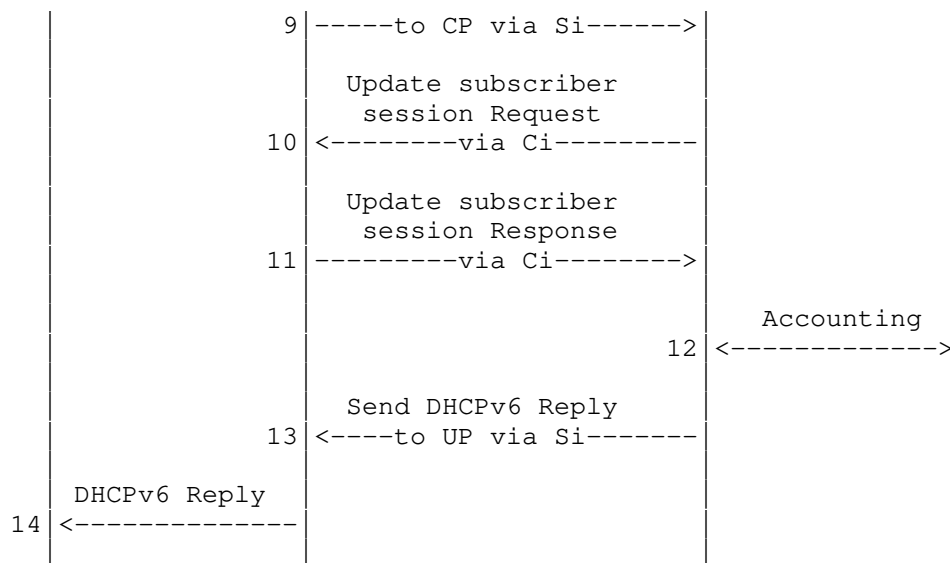


Figure 17: DHCPv6 + SLAAC Access

When a subscriber passes AAA authentication, the CP will create a subscriber session on the UP. This is achieved by sending an Update_Request message to the UP (step 4).

```

<Update_Request Message> ::= <Common Header>
                             <Basic Subscriber TLV>
                             <IPv6 Subscriber TLV>
                             <IPv6 Routing TLV>
                             [<Subscriber Policy TLV>]
  
```

The UP will reply with an Update_Response message (step 5). The format of the Update_Response is as follows:

```

<Update_Response Message> ::= <Common Header>
                              <Update Response TLV>
  
```

After receiving a DHCPv6 Solicit, the CP will update the subscriber session by sending an Update_Request message with new parameters to the UP (Step 10).

The format of the Update_Request message is as follows:

```

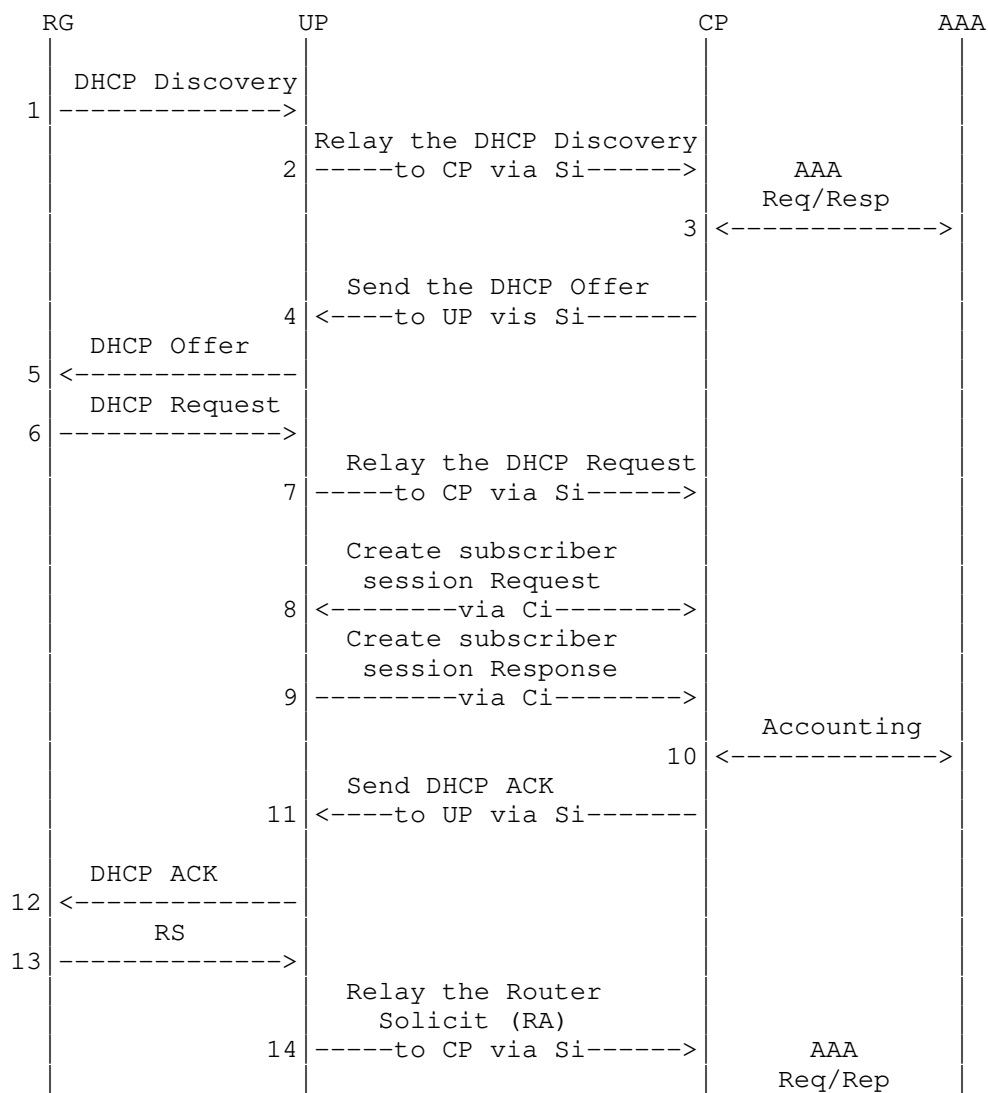
<Update_Request Message> ::= <Common Header>
                             <Basic Subscriber TLV>
                             <IPv6 Subscriber TLV>
                             <IPv6 Routing TLV>
                             [<Subscriber Policy TLV>]
  
```

The UP will reply with an Update_Response message (step 11). The format of the Update_Response is as follows:

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

5.1.5 DHCP Dual Stack Access

The following sequence is a combination of DHCP IPv4 and DHCP IPv6 access processes.



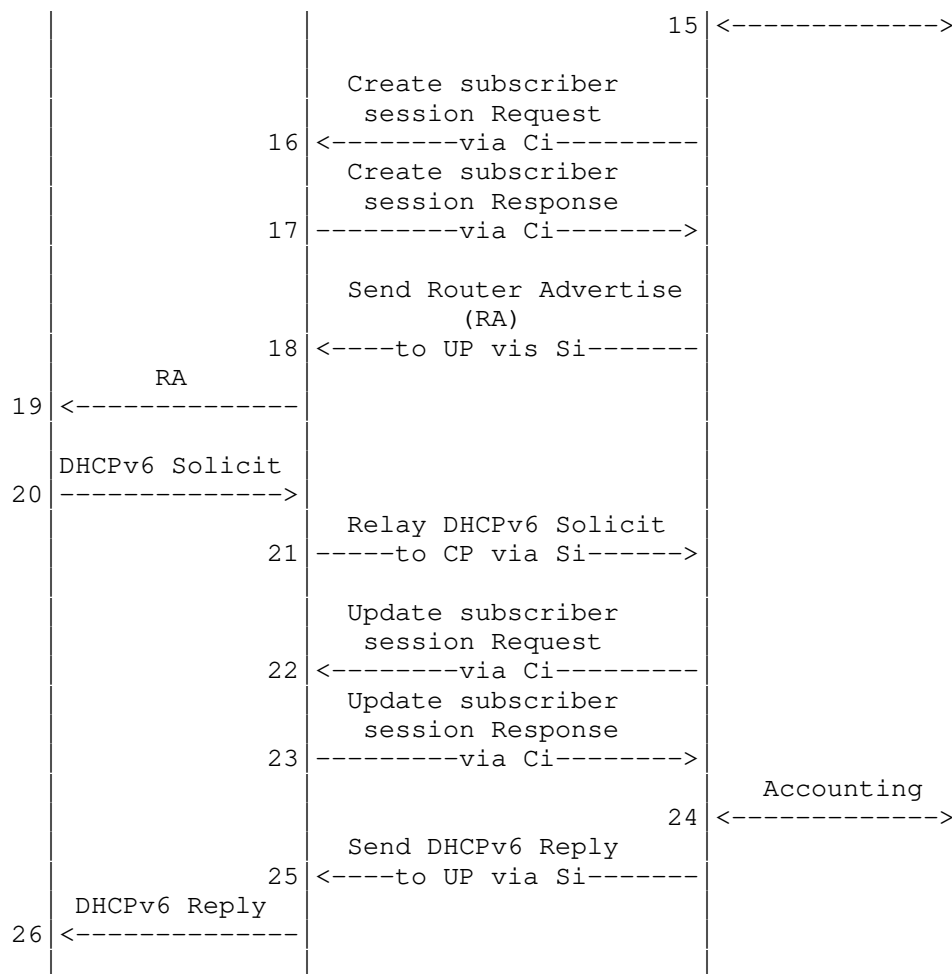


Figure 18: DHCP Dual Stack Access

The DHCP dual stack access includes three sets of Update_Request / Update_Response exchanges to create/update DHCPv4/v6 subscriber session.

1. Create DHCPv4 session (step 8 and 9)

```

<Update_Request Message> ::= <Common Header>
                             <Basic Subscriber TLV>
                             <IPv4 Subscriber TLV>
                             <IPv4 Routing TLV>
                             [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]

```

2. Create DHCPv6 session (step 16 and 17)

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               [<Subscriber Policy TLV>]

```

```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>

```

3. Update DHCPv6 session (step 22 and 23)

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               [<Subscriber Policy TLV>]

```

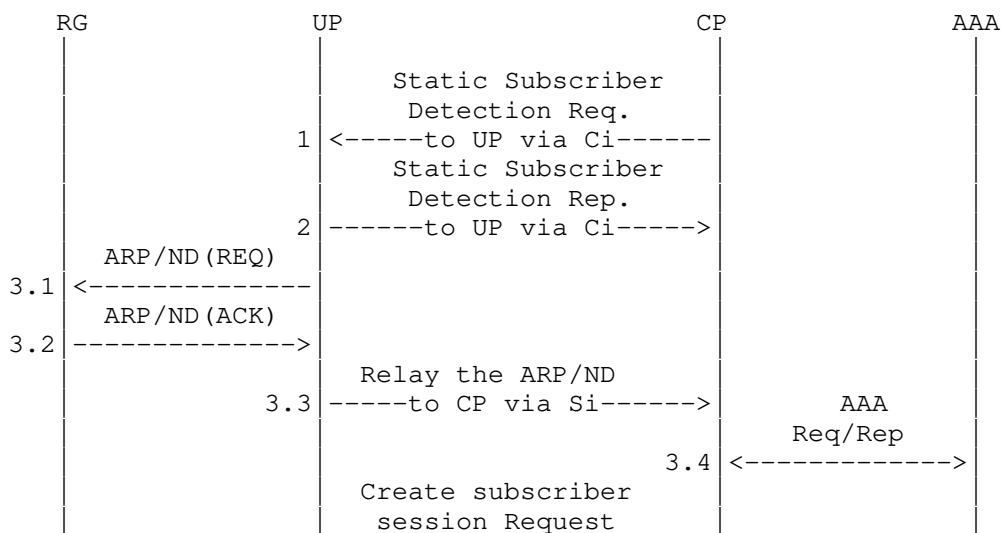
```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>

```

5.1.6 L2 Static Subscriber Access

L2 static subscriber access processes are as follows:



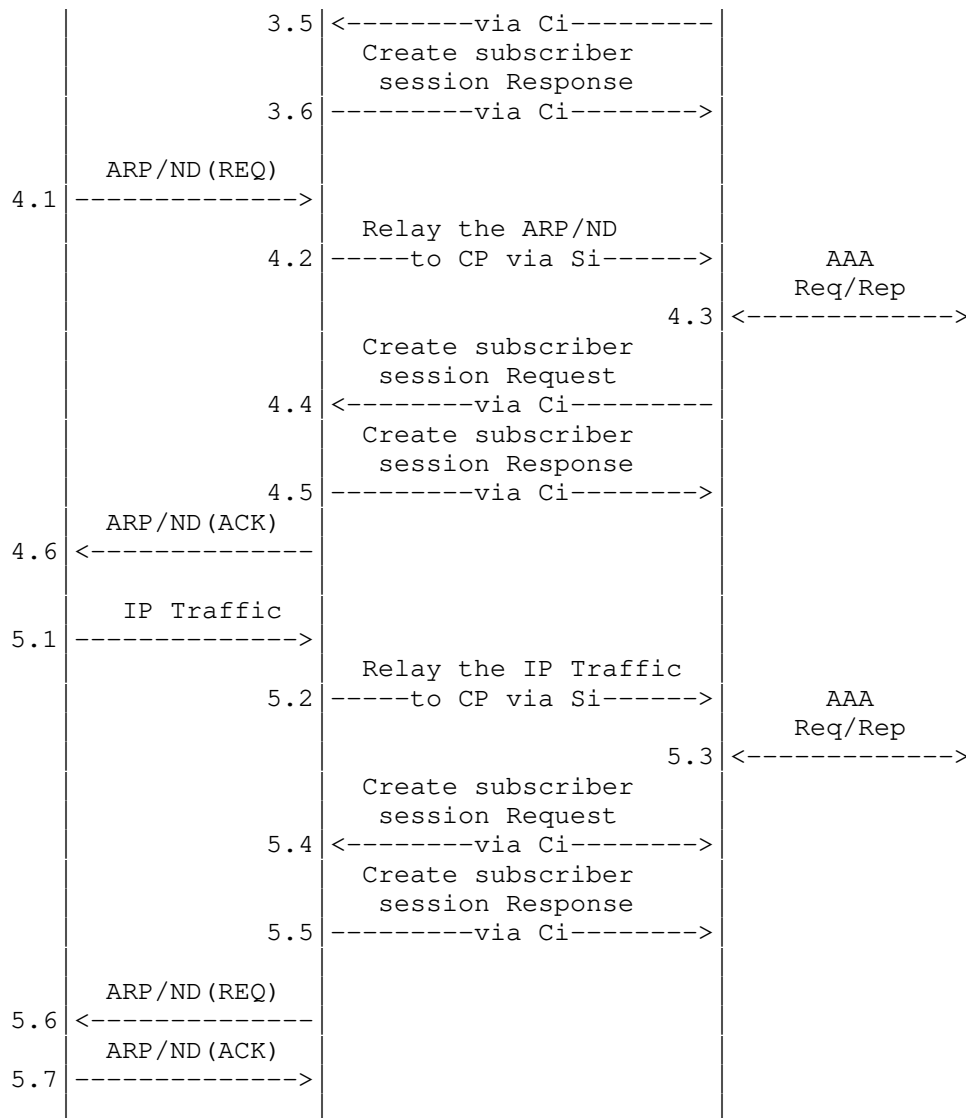


Figure 19: L2 Static Subscriber Access

For L2 static subscriber access, the process starts with a CP installing a static subscriber detection list on an UP. The list determines which subscribers will be detected. This is implemented by exchanging Update_Request and Update_Response messages between CP and UP. The format of the messages are as follows:

```

<Update_Request Message> ::= <Common Header>
                               <IPv4 Static Subscriber Detect TLVs>
                               <IPv6 Static Subscriber Detect TLVs>

```

```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>

```

For L2 Static subscriber access, there are three ways to trigger the access process:

1. Triggered by UP (3.1-3.6): This assumes that the UP knows the IP address, the access interface, and VLAN of the RG. The UP will actively trigger the access flow by sending an ARP/ND packet to the RG. If the RG is online, it will reply with an ARP/ND to the UP. The UP will tunnel the ARP/ND to the CP through the Si. The CP then triggers the authentication process. If the authentication result is positive. The CP will create a corresponding subscriber session on the UP.
2. Triggered by RG ARP/ND (4.1-4.6): Most of the process is same as option 1 (triggered by UP). The difference is that the RG will actively send the ARP/ND to trigger the process.
3. Triggered by RG IP traffic (5.1-5.7): This is for the case where the RG has the ARP/ND information, but the subscriber session on the UP is lost (e.g., due to failure on the UP, or the UP restarted). That means the RG may keep sending IP packets to the UP. The packets will trigger the UP to start a new access process.

From a subscriber session point of view, the procedures and the message formats for the above three cases are the same, as follows:

IPv4 Case:

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv4 Subscriber TLV>
                               <IPv4 Routing TLV>
                               [<Subscriber Policy TLV>]

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]

```

IPv6 Case:

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               [<Subscriber Policy TLV>]

```

<Update_Response Message> ::= <Common Header>
 <Update Response TLV>

5.2 PPPoE

5.2.1 IPv4 PPPoE Access

The following figure shows the IPv4 PPPoE access call flow.

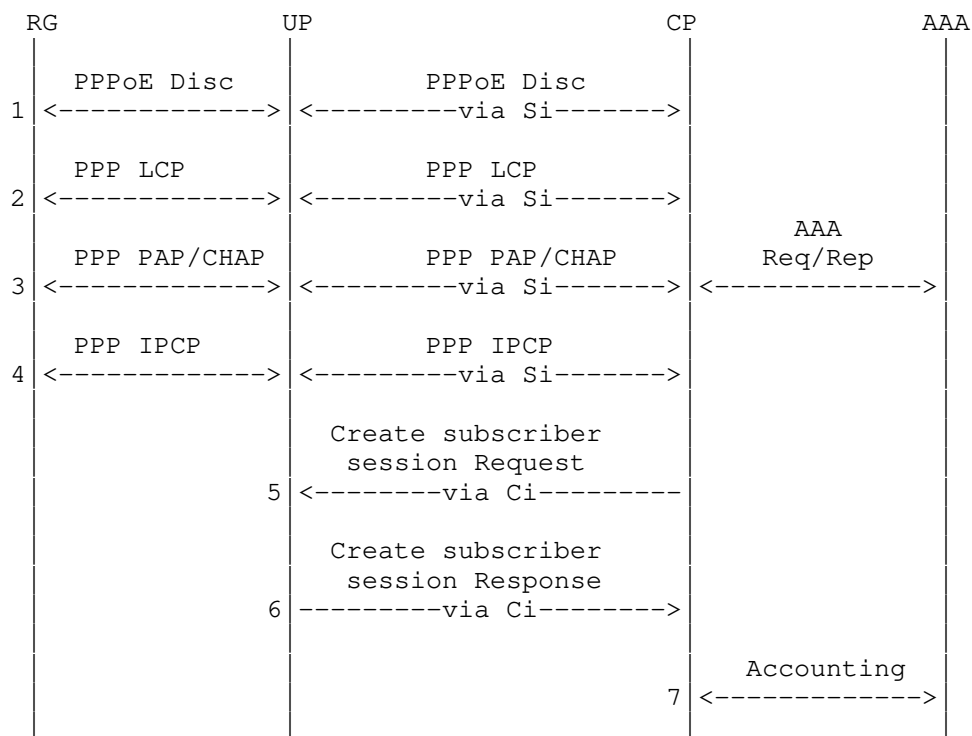


Figure 20: IPv4 PPPoE Access

From the above sequence, step 1-4 are the standard PPPoE call flow. The UP is responsible for redirecting the PPPoE control packets to the CP or RG. The PPPoE control packets are transmitted between the CP and UP through the Si.

After the PPPoE call flow, if the subscriber passed the AAA authentication and authorization, the CP will create a corresponding session on the UP through the Ci. The formats of the messages are as follows:

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv4 Subscriber TLV>
                               <IPv4 Routing TLV>
                               [<Subscriber Policy TLV>]

```

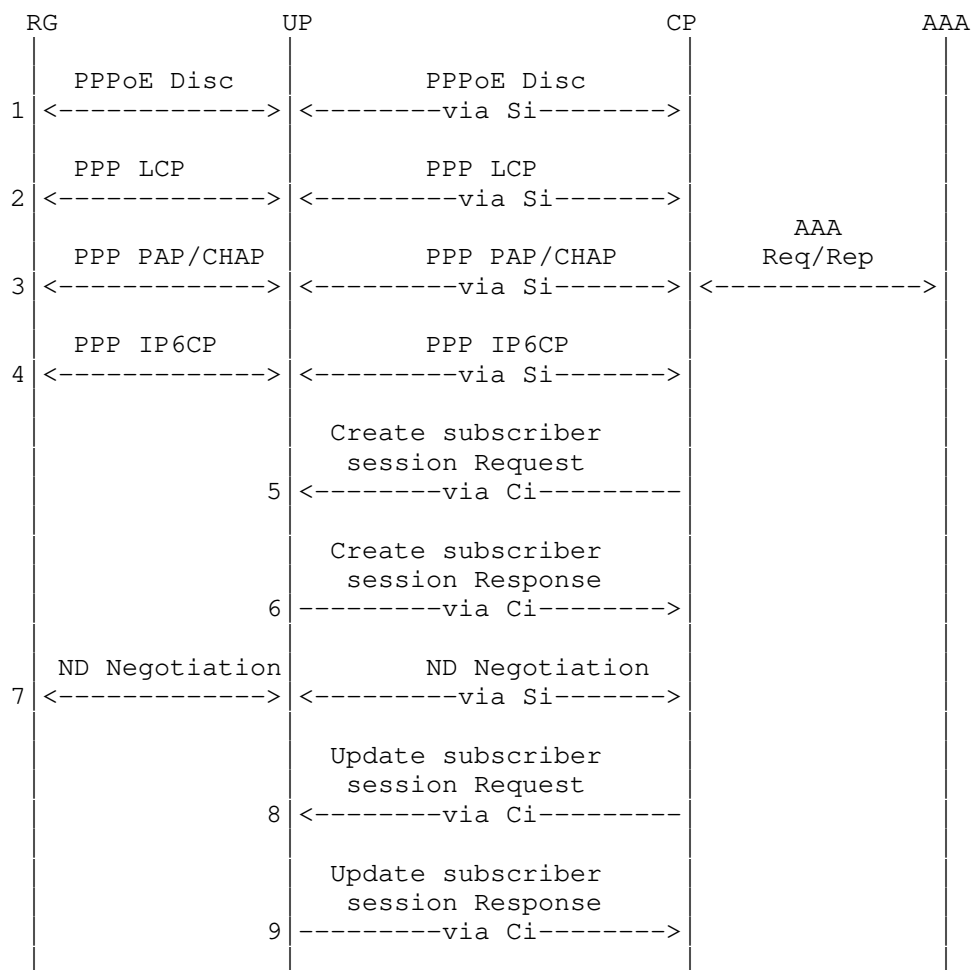
```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]

```

5.2.2 IPv6 PPPoE Access

The following figure describes the IPv6 PPPoE access call flow.



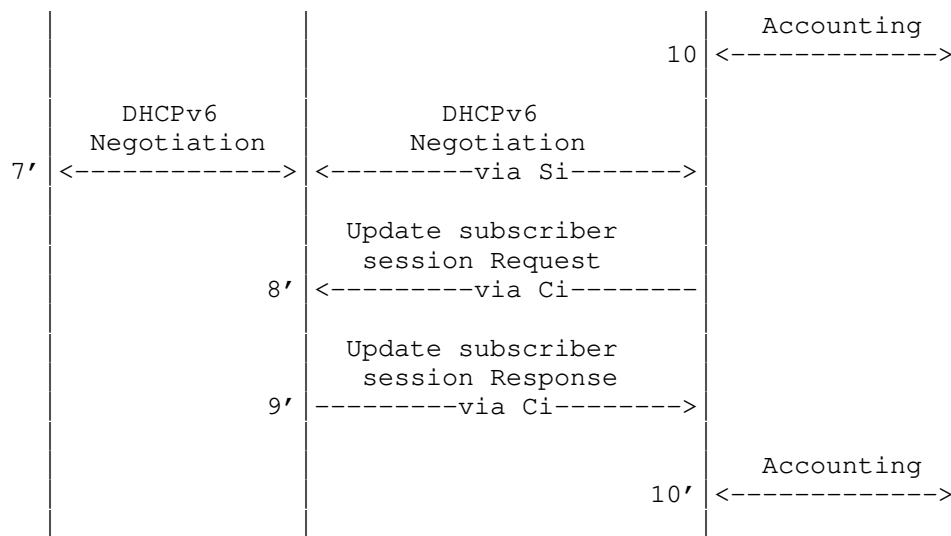


Figure 21: IPv6 PPPoE Access

From the above sequence, steps 1-4 are the standard PPPoE call flow. The UP is responsible for redirecting the PPPoE control packets to the CP or RG. The PPPoE control packets are transmitted between the CP and UP through the Si.

After the PPPoE call flow, if the subscriber passed the AAA authentication and authorization, the CP will create a corresponding session on the UP through the Ci. The formats of the messages are as follows:

```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <PPP Subscriber TLV>
    <IPv6 Subscriber TLV>
    <IPv6 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
  
```

Then, the RG will initialize a ND/DHCPv6 negotiation process with the CP (see step 7 and 7'), after that, it will trigger an update (8-9, 8'-9') to the subscriber session. The formats of the update messages are as follows:

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               [<Subscriber Policy TLV>]

```

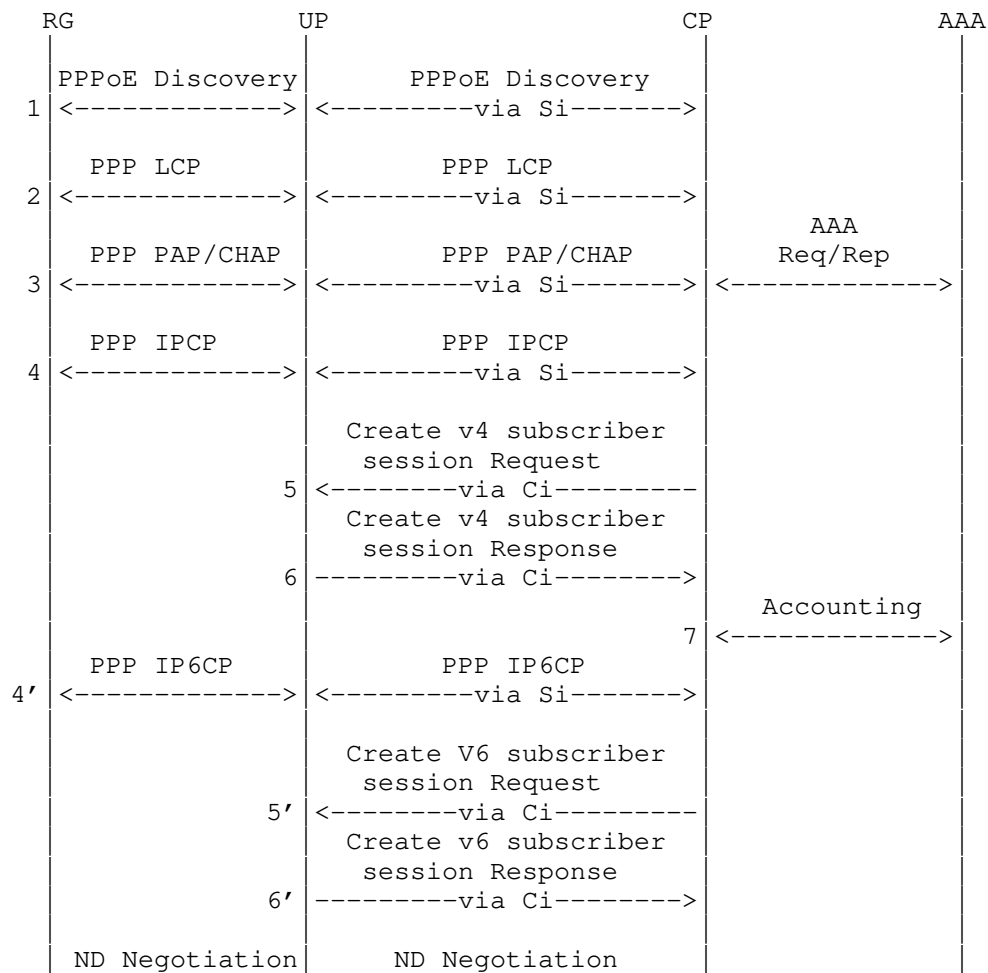
```

<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>

```

5.2.3 PPPoE Dual Stack Access

The following figure shows a combination of IPv4 and IPv6 PPPoE access call flow.



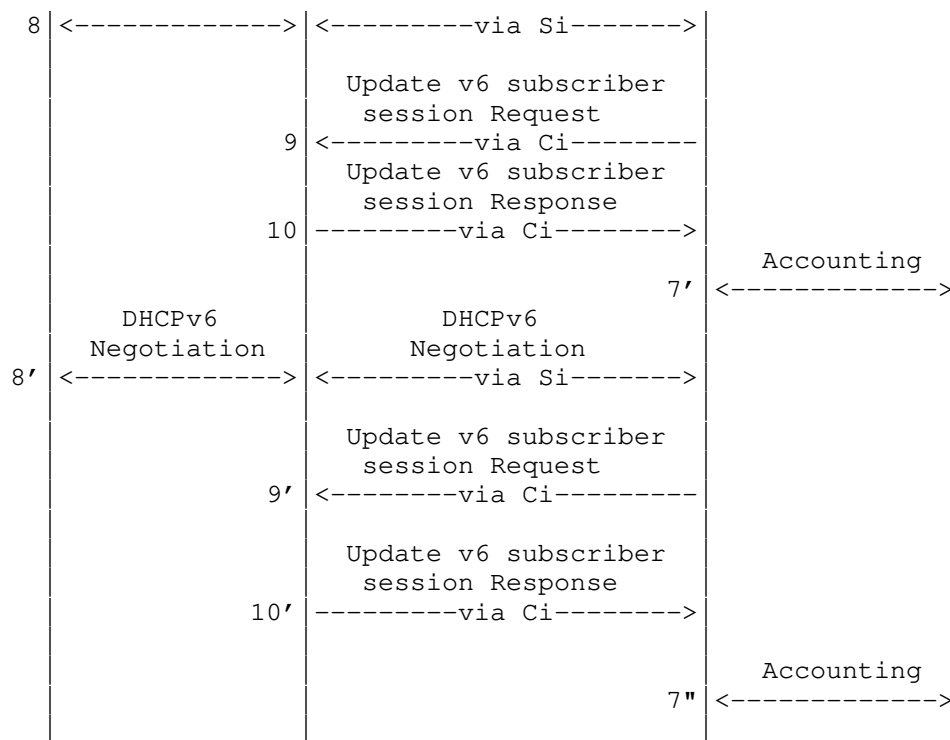


Figure 22: PPPoE Dual Stack Access

PPPoE dual stack is a combination of IPv4 PPPoE and IPv6 PPPoE access. The process is as above. The formats of the messages are as follows:

1. Create an IPv4 PPPoE subscriber session (5-6)

```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <PPP Subscriber TLV>
    <IPv4 Subscriber TLV>
    <IPv4 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
    [<Subscriber CGN Port Range TLV>]
  
```

2. Create an IPv6 PPPoE subscriber session (5'-6')

```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <PPP Subscriber TLV>
  
```

```
<IPv6 Subscriber TLV>
<IPv6 Routing TLV>
[<Subscriber Policy TLV>]
```

```
<Update_Response Message> ::= <Common Header>
                                <Update Response TLV>
```

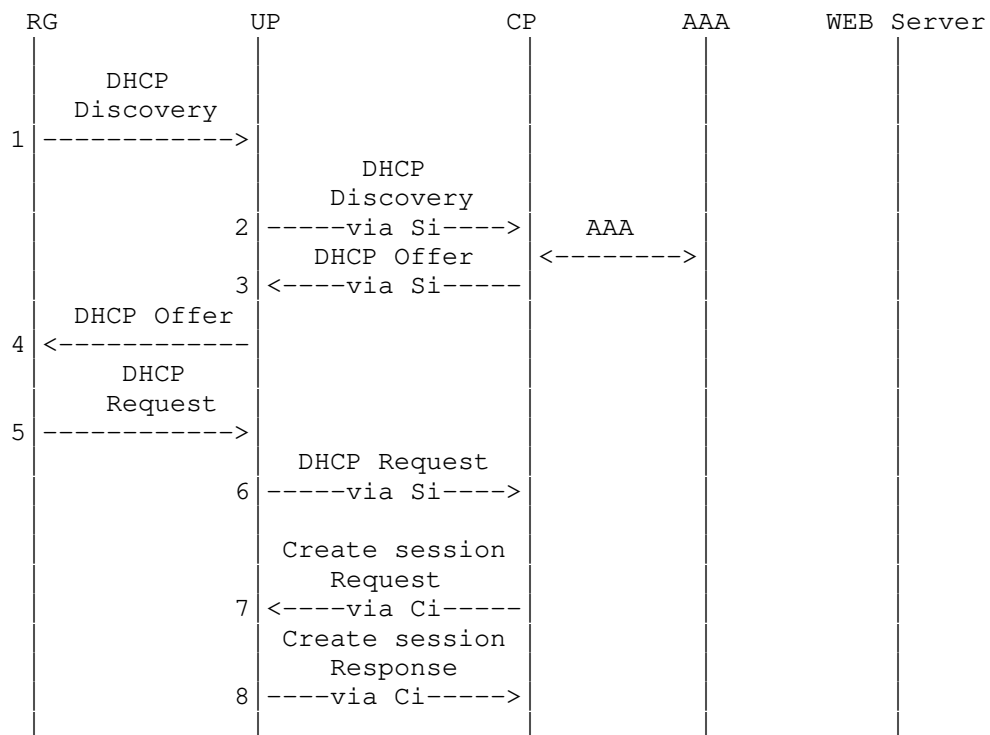
3. Update the IPv6 PPPoE subscriber session (9-10, 9'-10')

[illegible]

```
<Update_Response Message> ::= <Common Header>
                                <Update Response TLV>
```

5.3 WLAN Access

The following figure shows the WLAN access call flow.



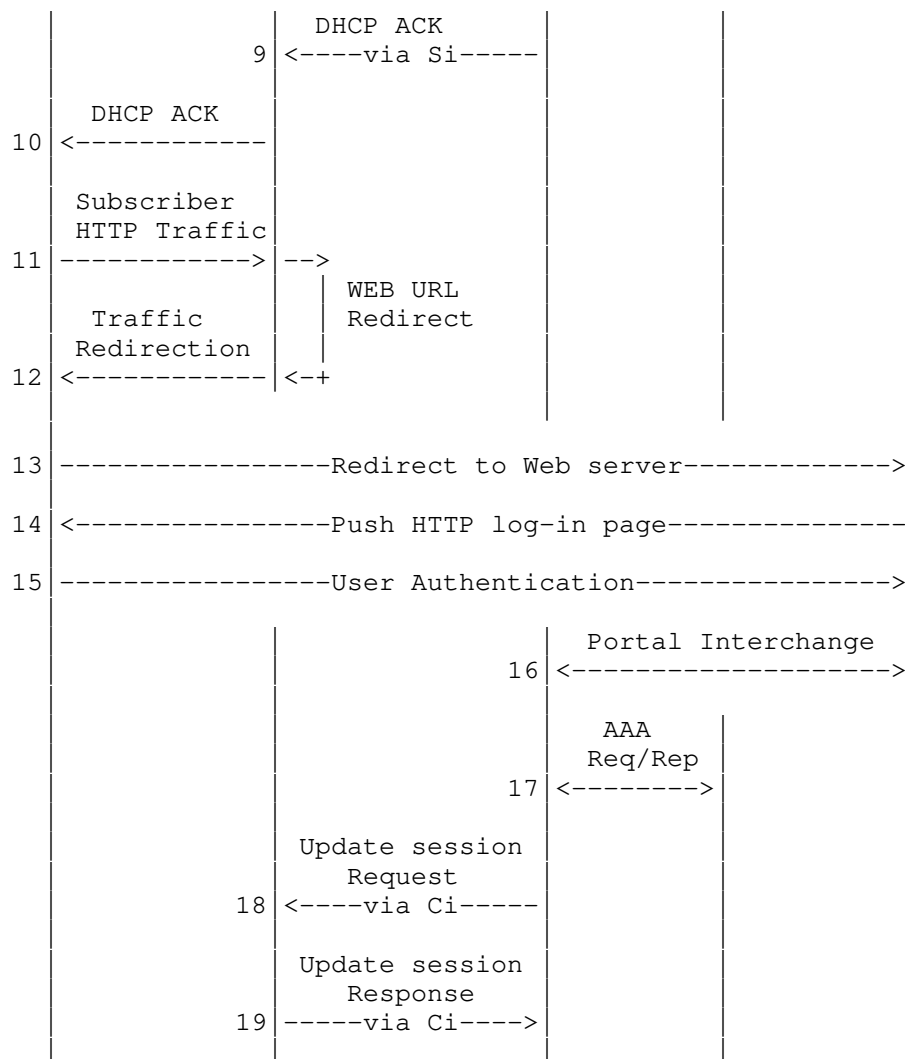


Figure 23: WLAN Access

WLAN access starts with the DHCP dial-up process (steps 1-6), after that the CP will create a subscriber session on the UP (steps 7-8). The formats of the session creation messages are as follows:

IPv4 Case:

```

<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv4 Subscriber TLV>
                               <IPv4 Routing TLV>
                               [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
                                <Update Response TLV>
                                [<Subscriber CGN Port Range TLV>]

```

IPv6 Case:

```

<Update_Request Message> ::= <Common Header>
                                <Basic Subscriber TLV>
                                <IPv6 Subscriber TLV>
                                <IPv6 Routing TLV>
                                [<Subscriber Policy TLV>]

```

```

<Update_Response Message> ::= <Common Header>
                                <Update Response TLV>

```

After step 10, the RG will be allocated an IP address and its first HTTP packet will be redirected to a WEB server for subscriber authentication (steps 11-17). After the WEB authentication, if the result is positive, the CP will update the subscriber session by using the following message exchanges:

```

IPv4 Case: <Update_Request Message> ::= <Common Header>
                                <Basic Subscriber TLV>
                                <IPv4 Subscriber TLV>
                                <IPv4 Routing TLV>
                                [<Subscriber Policy TLV>]

```

```

<Update_Response Message> ::= <Common Header>
                                <Update Response TLV>
                                [<Subscriber CGN Port Range TLV>]

```

```

IPv6 Case: <Update_Request Message> ::= <Common Header>
                                <Basic Subscriber TLV>
                                <IPv6 Subscriber TLV>
                                <IPv6 Routing TLV>
                                [<Subscriber Policy TLV>]

```

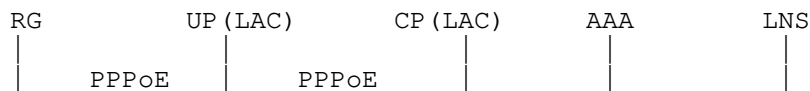
```

<Update_Response Message> ::= <Common Header>
                                <Update Response TLV>

```

5.4 L2TP

5.4.1 L2TP LAC Access



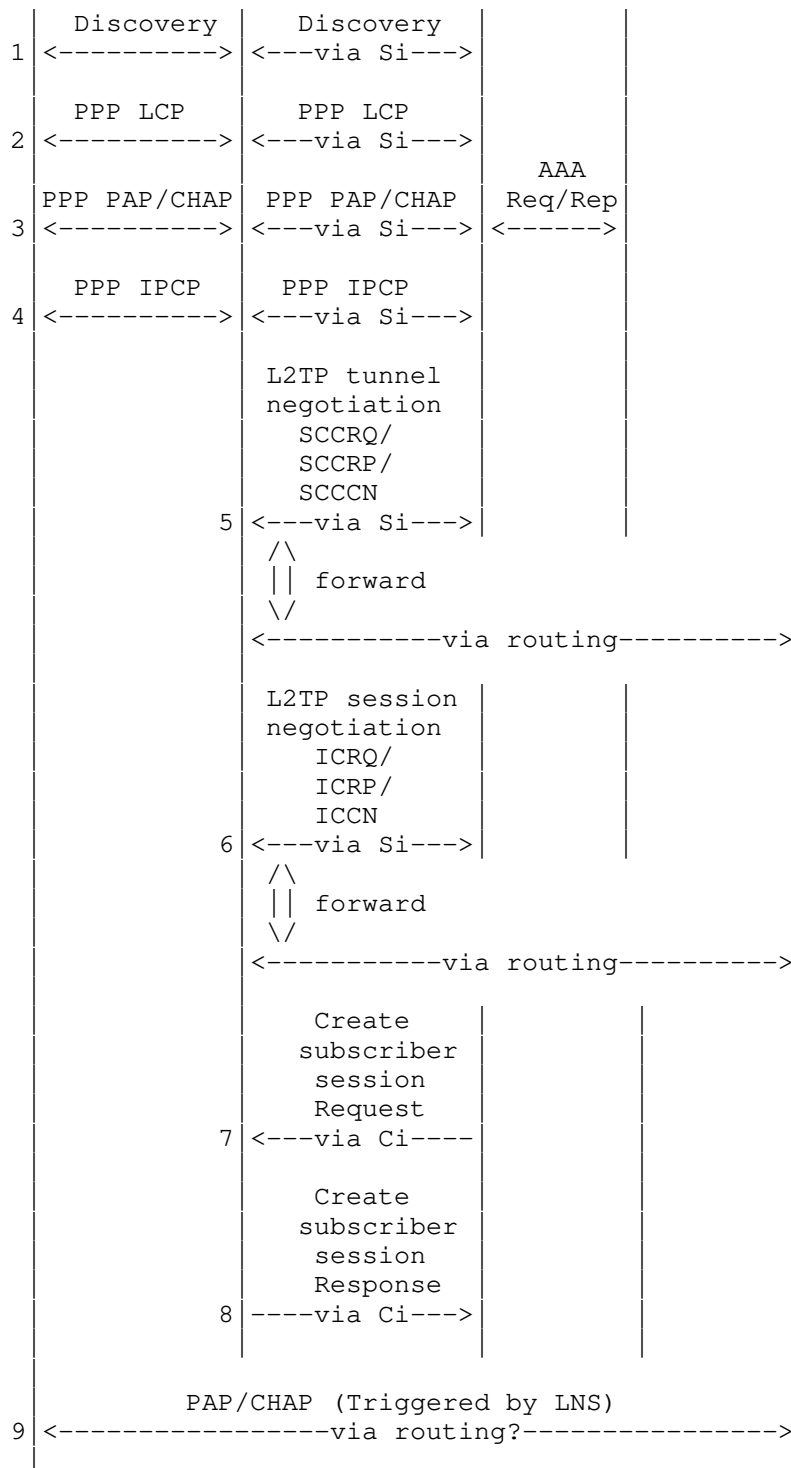


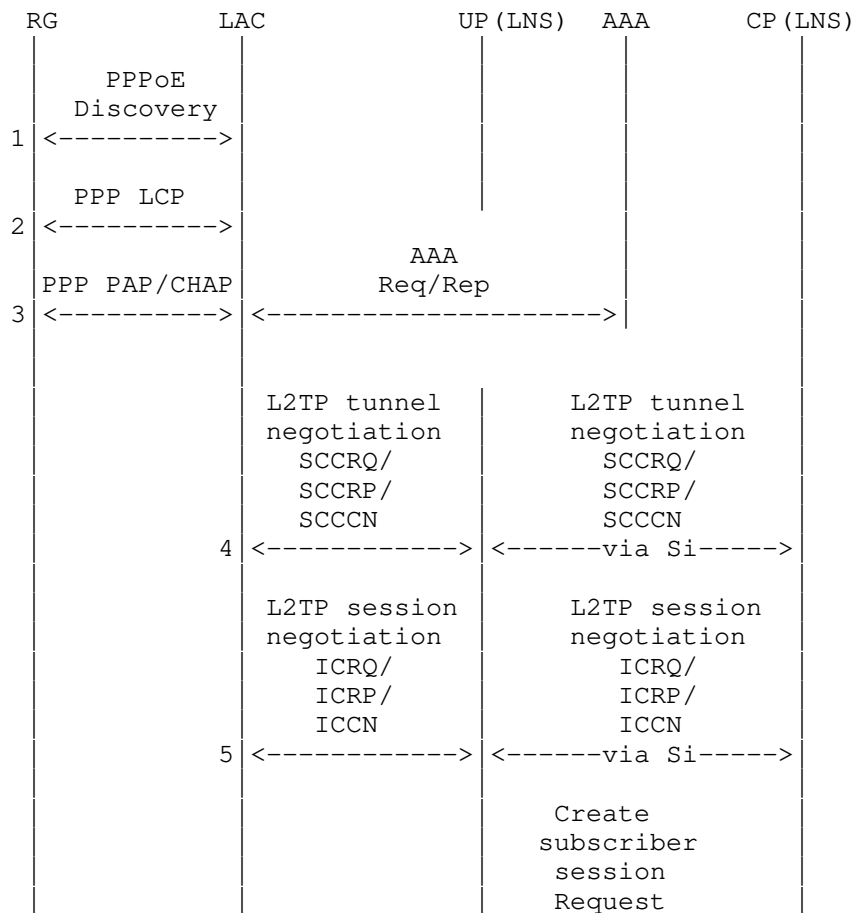
Figure 24: L2TP-LAC Access

Steps 1-4 are a standard PPPoE access process. After that the LAC-CP starts to negotiate an L2TP session and tunnel with the LNS. After the negotiation, the CP will create an L2TP LAC subscriber session on the UP through the following messages:

```
<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <L2TP-LAC Subscriber TLV>
                               <L2TP-LAC Tunnel TLV>
```

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

5.4.2 L2TP LNS IPv4 Access



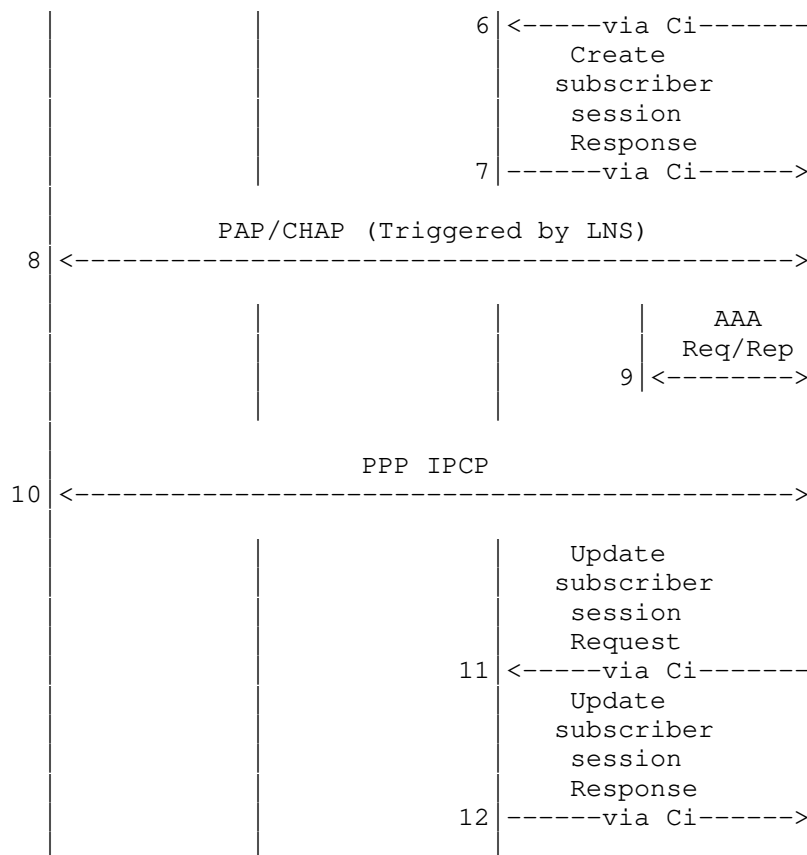


Figure 25: IPv4 L2TP-LNS Access

In this case, the BNG is running as an LNS and separated into LNS-CP and LNS-UP. Steps 1-5 finish the normal L2TP dial-up process. When the L2TP session and tunnel negotiations are finished, the LNS-CP will create an L2TP LNS subscriber session on the LNS-UP. The format of messages are as follows:

```

<Update_Request Message> ::= <Common Header>
                               <L2TP-LNS Subscriber TLV>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv4 Subscriber TLV>
                               <IPv4 Routing TLV>
                               <L2TP-LNS Tunnel TLV>
                               [<Subscriber Policy TLV>]
  
```

```

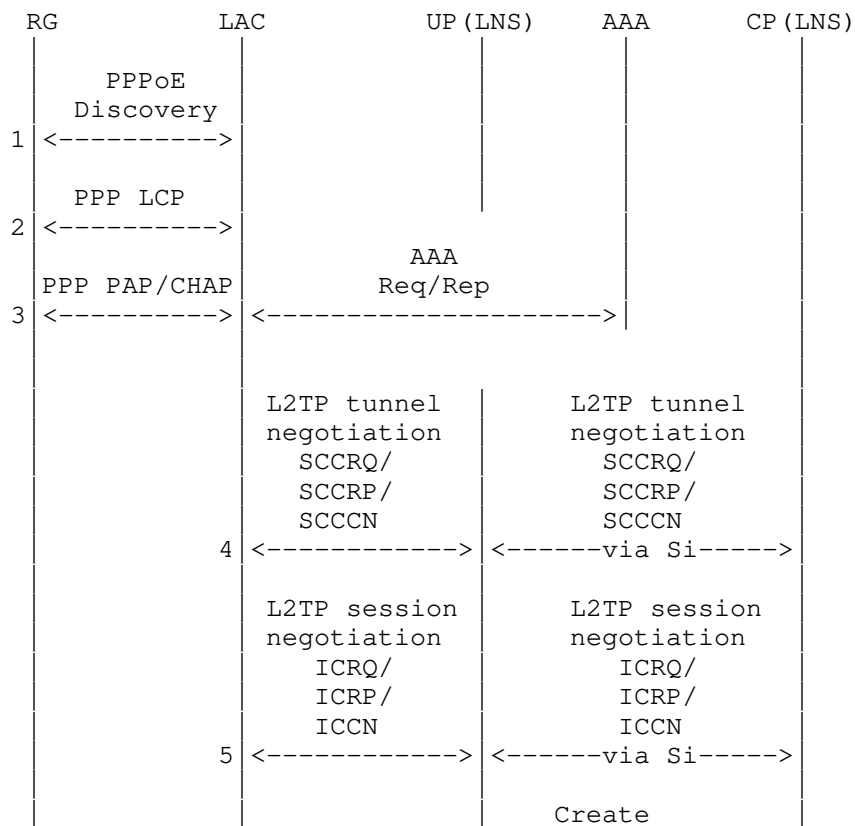
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]
  
```

After that, the LNS-CP will trigger an AAA authentication. If the authentication result is positive, a PPP IPCP process will follow, then the CP will update the session with the following message exchanges:

```
<Update_Request Message> ::= <Common Header>
                               <L2TP-LNS Subscriber TLV>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv4 Subscriber TLV>
                               <IPv4 Routing TLV>
                               <L2TP-LNS Tunnel TLV>
                               [<Subscriber Policy TLV>]
```

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]
```

5.4.3 L2TP LNS IPv6 Access



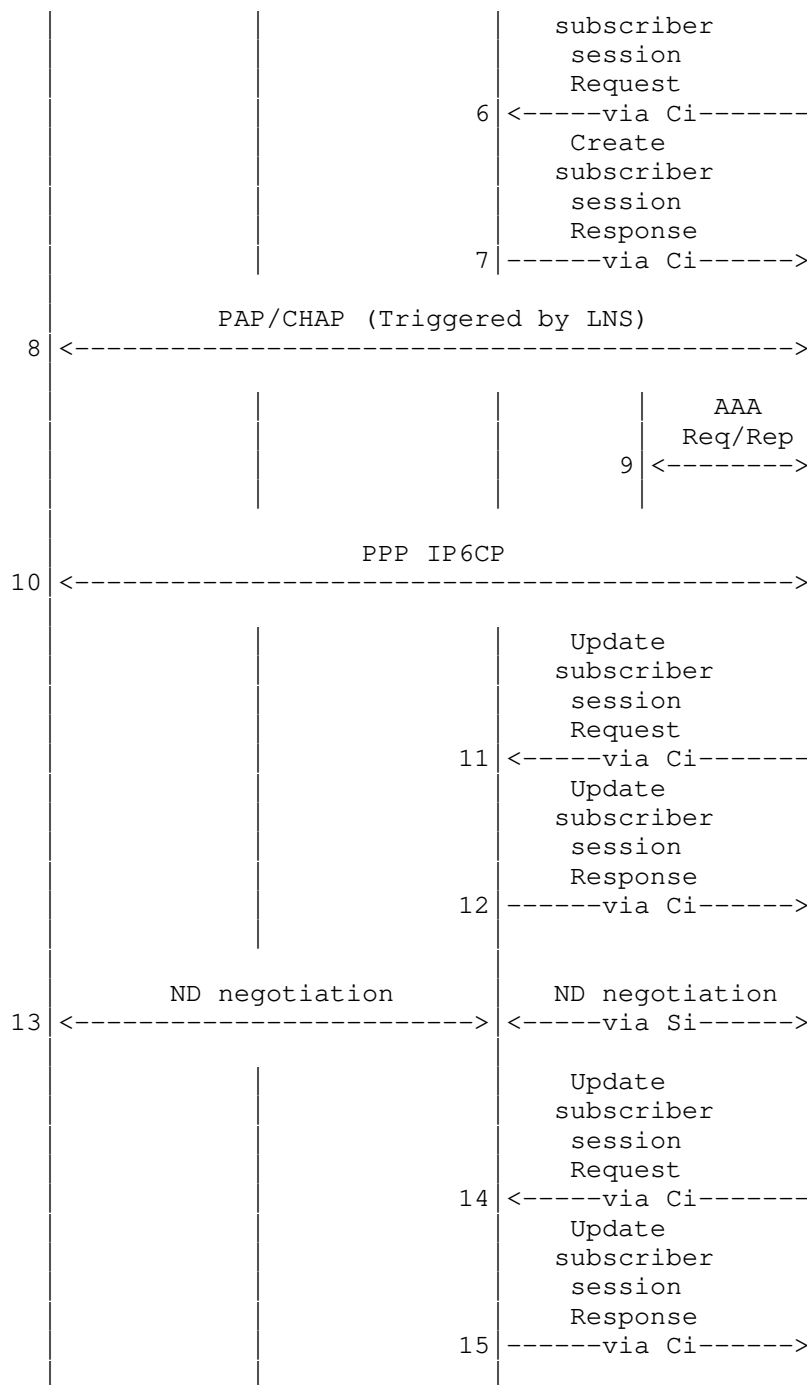


Figure 26: L2TP-LNS IPv6 Access

Steps 1-12 are the same as L2TP and LNS IPv4 Access. Steps 1-5 finish the normal L2TP dial-up process. When the L2TP session and tunnel negotiations are finished, the LNS-CP will create an L2TP LNS subscriber session on the LNS-UP. The format of the messages is as follows:

```
<Update_Request Message> ::= <Common Header>
                               <L2TP-LNS Subscriber TLV>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               <L2TP-LNS Tunnel TLV>
                               [<Subscriber Policy TLV>]
```

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

After that, the LNS-CP will trigger a AAA authentication. If the authentication result is positive, a PPP IP6CP process will follow, then the CP will update the session with the following message exchanges:

```
<Update_Request Message> ::= <Common Header>
                               <L2TP-LNS Subscriber TLV>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               <L2TP-LNS Tunnel TLV>
                               [<Subscriber Policy TLV>]
```

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

Then, an ND negotiation will be triggered by the RG. After the ND negotiation, the CP will update the session with the following message exchanges:

```
<Update_Request Message> ::= <Common Header>
                               <L2TP-LAC Subscriber TLV>
                               <Basic Subscriber TLV>
                               <PPP Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               <L2TP-LNS Tunnel TLV>
                               [<Subscriber Policy TLV>]
```

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

5.5 CGN (Carrier Grade NAT)

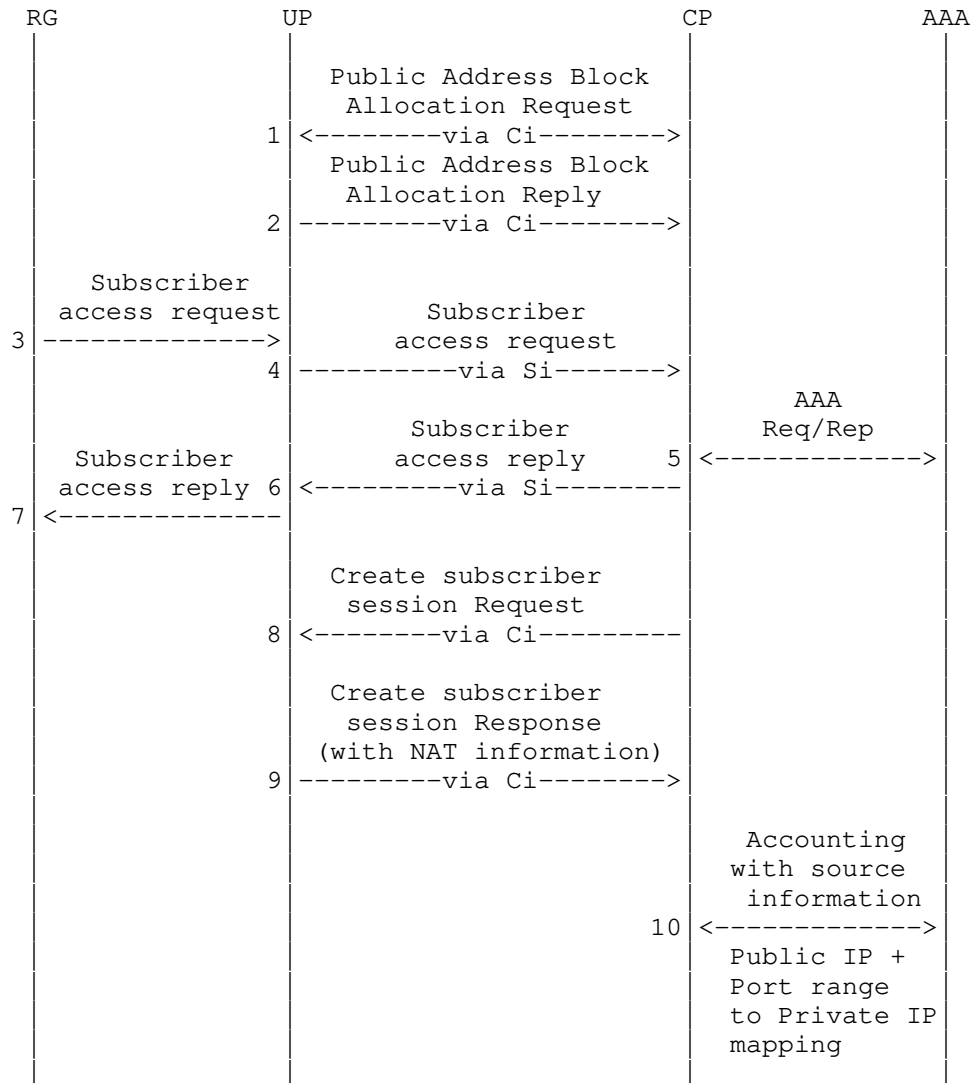


Figure 27: CGN Access

The first steps allocate one or more CGN address blocks to the UP (steps 1-2). This is achieved by the following message exchanges between CP and UP.

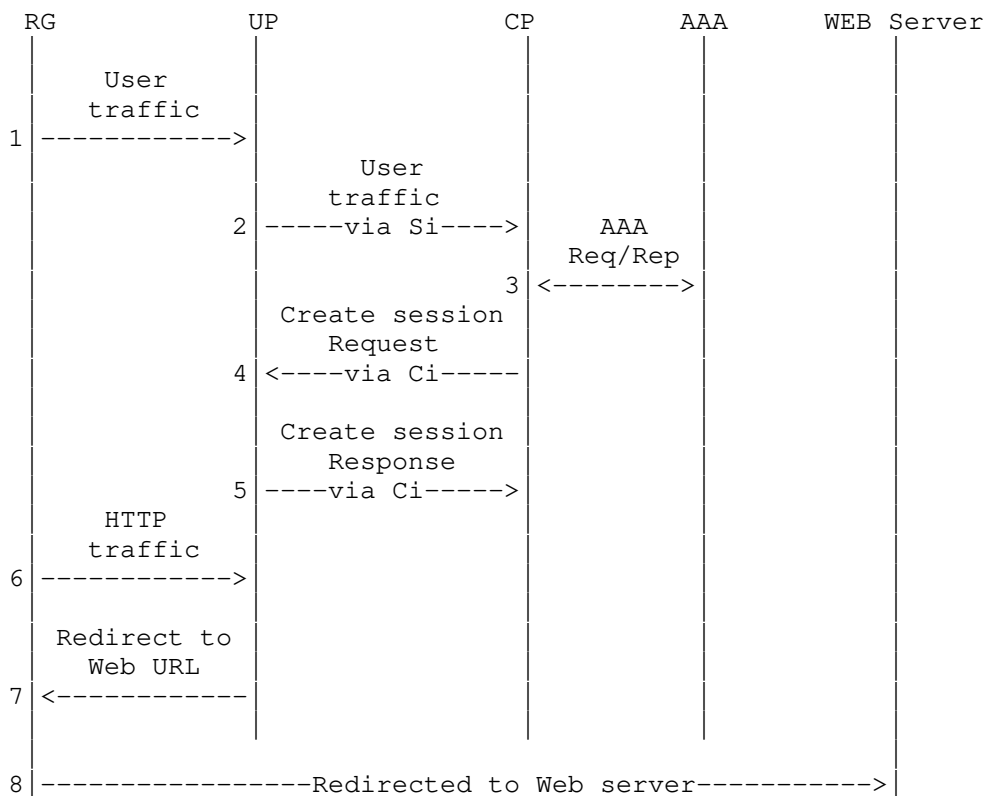
[illegible][illegible]

Steps 3-9 show the general dial-up process in the case of CGN mode. The specific processes (e.g., IPoE, PPPoE, L2TP, etc.) are defined in above sections.

If a subscriber is a CGN subscriber, once the subscriber session is created/updated, the UP will report the NAT information to the CP. This is achieved by carrying the "Subscriber CGN Port Range TLV" in the Update_Response message.

5.6 L3 Leased Line Access

5.6.1 Web Authentication



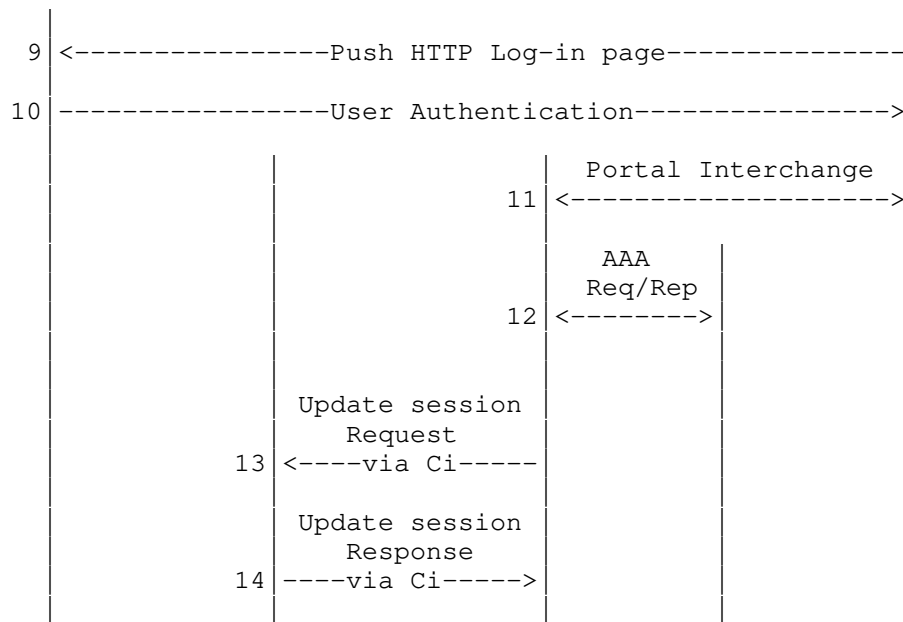


Figure 28: Web Authentication based L3 Leased Line Access

In this case, IP traffic from the RG will trigger the CP to authenticate the RG by checking the source IP and the exchanges with the AAA server. Once the RG passed the authentication, the CP will create a corresponding subscriber session on the UP through the following message exchanges:

IPv4 Case:

```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <IPv4 Subscriber TLV>
    <IPv4 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
    [<Subscriber CGN Port Range TLV>]
  
```

IPv6 Case:

```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <IPv6 Subscriber TLV>
    <IPv6 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
  
```

Then, the HTTP traffic from the RG will be redirected to a WEB server to finish the WEB authentication. Once the WEB authentication is passed, the CP will trigger another AAA authentication. After the AAA authentication, the CP will update the session with the following message exchanges:

IPv4 Case:

```
<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv4 Subscriber TLV>
                               <IPv4 Routing TLV>
                               [<Subscriber Policy TLV>]
```

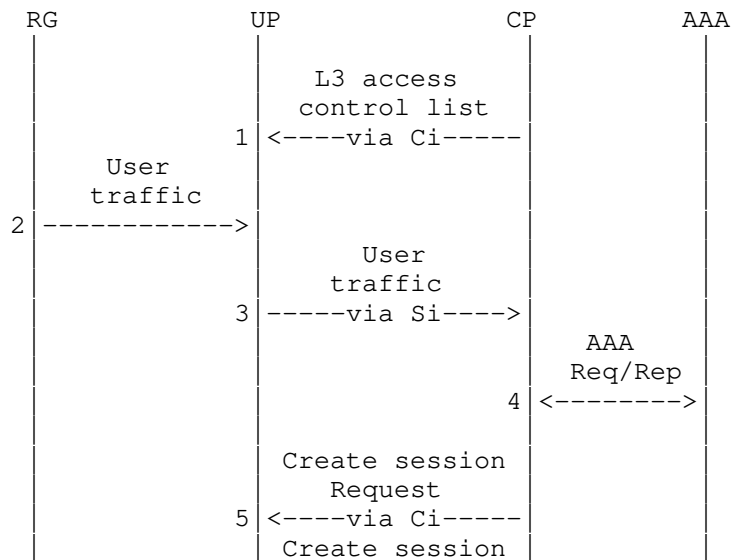
```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
                               [<Subscriber CGN Port Range TLV>]
```

IPv6 Case:

```
<Update_Request Message> ::= <Common Header>
                               <Basic Subscriber TLV>
                               <IPv6 Subscriber TLV>
                               <IPv6 Routing TLV>
                               [<Subscriber Policy TLV>]
```

```
<Update_Response Message> ::= <Common Header>
                               <Update Response TLV>
```

5.6.2 User Traffic Trigger



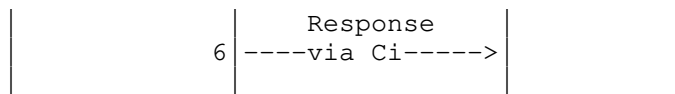


Figure 29: User Traffic Triggered L3 Leased Line Access

In this user traffic triggered case, the CP must install an access control list on the UP, which is used by the UP to determine whether an RG is legal or not. If the traffic is from a legal RG, it will be redirected to the CP through the Si. The CP will trigger a AAA interchange with the AAA server. After that, the CP will create a corresponding subscriber session on the UP with the following message exchanges:

IPv4 Case:

```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <IPv4 Subscriber TLV>
    <IPv4 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
    [<Subscriber CGN Port Range TLV>]
  
```

IPv6 Case:

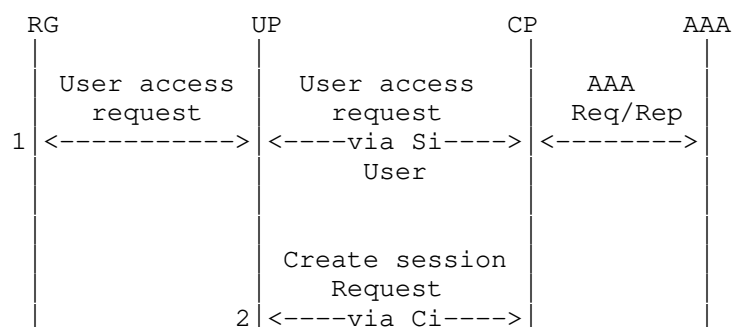
```

<Update_Request Message> ::= <Common Header>
    <Basic Subscriber TLV>
    <IPv6 Subscriber TLV>
    <IPv6 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
  
```

5.7 Multicast Access



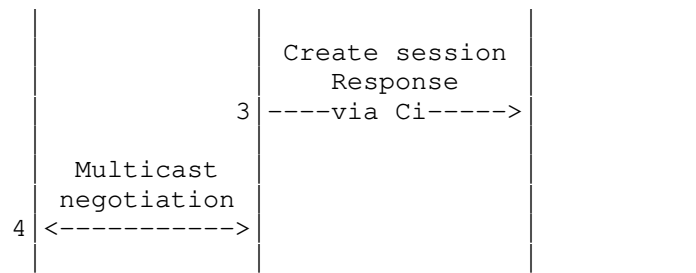


Figure 30: Multicast Access

Multicast access starts with an user access request from the RG. The request will be redirected to the CP by the Si. A follow-up AAA interchange between the CP and the AAA server will be triggered. After the authentication, the CP will create a multicast subscriber session on the UP through the following messages:

IPv4 Case:

```

<Update_Request Message> ::= <Common Header>
    <Multicast Group Information TLV>
    <Basic Subscriber TLV>
    <IPv4 Subscriber TLV>
    <IPv4 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
    [<Subscriber CGN Port Range TLV>]
  
```

IPv6 Case:

```

<Update_Request Message> ::= <Common Header>
    <Multicast Group Information TLV>
    <Basic Subscriber TLV>
    <IPv6 Subscriber TLV>
    <IPv6 Routing TLV>
    [<Subscriber Policy TLV>]
  
```

```

<Update_Response Message> ::= <Common Header>
    <Update Response TLV>
  
```

6. S-CUSP Message Formats

An S-CUSP message consists of a common header followed by a variable-length body consisting entirely of TLVs. Receiving an S-CUSP message with an unknown message type or missing mandatory TLV MUST trigger an Error message (see Section 6.7) or a response message with an Error Information TLV (see Section 7.6).

Conversely, if a TLV is optional, the TLV may or may not be present. Optional TLVs are indicated in the message formats shown in this document by being enclosed in square brackets.

This section specifies the format of the common S-CUSP message header and lists the defined messages.

Network byte order is used for all multi-byte fields.

6.1 Common Message Header

S-CUSP Common Message Header:

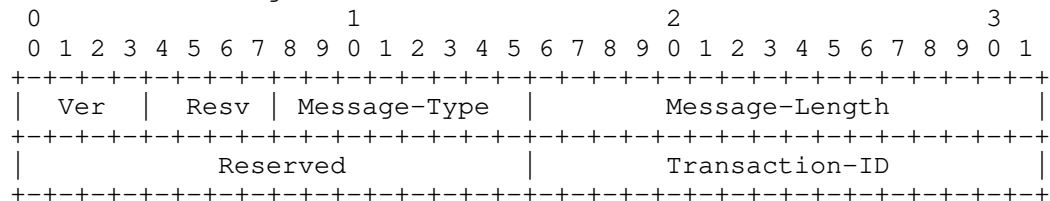


Figure 6.1: S-CUSP Message Common Header

- o Ver (4 bits): The major version of the protocol. This document specifies version 1. Different major versions of the protocol may have significantly different message structure and format except that the Ver field will always be in the same place at the beginning of each message. A successful S-CUSP session depends on the CP and the UP both using the same major version of the protocol.
- o Resv (4 bits): Reserved. MUST be sent as zero and ignored on receipt.
- o Message-Type (8 bits): The set of message types specified in this document is listed in Section 9.1.
- o Message-Length (16 bits): Total length of the S-CUSP message including the common header, expressed in number of bytes as an unsigned integer.

- o Transaction ID (16 bits): This field is used to identify requests. It is echoed back in any corresponding ACK / response / Error message. It is RECOMMENDED that a monotonically increasing value be used in successive message and that value wrap back to zero after 0xFFFF. The contents of this field is an opaque value that the receiver MUST NOT use for any purpose except to echo back in a corresponding response and, optionally, for logging.

6.2 Control Messages

This document defines the following control messages:

Type	Name	Notes and TLVs that can be carried
-----	----	-----
1	Hello	Hello TLV, Keep-Alive TLV.
2	Keepalive	A common header with the Keepalive message type.
3	Sync_Request	Synchronization request.
4	Sync_Begin	Synchronization starts.
5	Sync_Data	Synchronization data: TLVs specified in Section 5.
6	Sync_End	End synchronization.
7	Update_Request	TLVs specified in Sections 7.6-7.9.
8	Update_Response	TLVs specified in Sections 7.6-7.9.

6.2.1 Hello Message

Hello message is used for S-CUSP session establishment and version negotiation. The detail of S-CUSP session establishment and version negotiation can be found in Section 4.1.1.

The format of Hello message is as follows:

```
<Hello Message> ::= <Common Header>
                    <Hello TLV>
                    <Keepalive TLV>
                    [<Error Information TLV>]
```

The return code and negotiation result will be carried in the Error Information TLV. They are listed as follows:

- 0: Success, version negotiation success.
- 1: Failure, malformed message received.

2: One or more of the TLVs was not understood.

1001: The version negotiation fails. The S-CUSP session establishment phase fails.

1002: The keepalive negotiation fails. The S-CUSP session establishment phase fails.

```
1003: The establishment timer expires.  session establishment
phase fails.
```

6.2.2 Keepalive Message

The Keepalive message is periodically sent by each end of an S-CUSP session. It is used to detect whether the peer end is still alive. The Keepalive procedures are defined Section 4.1.2.

The format of the Keepalive message is as follows:

$$\langle \text{Keepalive Message} \rangle ::= \langle \text{Common Header} \rangle$$

6.2.3 Sync_Request Message

The Sync_Request message is used to request synchronization from an S-CUSP peer. Both CP and UP can request their peer to synchronize data.

The format of the Sync_Request message is as follows:

```
<Sync_Request Message> ::= <Common Header>
```

A `Sync_Request` message may result in a `Sync_Begin` message from its peer. The `Sync_Begin` message is defined in Section 6.2.4.

6.2.4 Sync_Begin Message

The Sync_Begin message is a reply to a Sync_Request message. It is used to notify the synchronization requester whether the synchronization can be started.

The format of Sync_Begin message is as follows:

```
<Sync_Begin Message> ::= <Common Header>
                             <Error Information TLV>
```

The return codes are carried in the Error Information TLV. The codes are listed below:

- 0: Success, be ready to synchronize.
- 1: Failure, malformed message received.
- 2: One or more of the TLVs was not understood.
- 2001: Synch-NoReady. The data to be synchronized is not ready.
- 2002: Synch-Unsupport. The data synchronization is not supported.

6.2.5 Sync_Data Message

The Sync_Data message is used to send data being synchronized between the CP and UP. The Sync_Data message has the same function and format as the Update_Request message. The difference is that there is no ACK for a Sync_Data message. An error caused by the Sync_Data message will result in a Sync_End message.

There are two scenarios:

Synchronization from UP to CP: Synchronize the resource data to CP.

```
<Sync_Data Message> ::= <Common Header>
                        [<Resource Reporting TLVs>]
```

Synchronization from CP to UP: Synchronize all subscriber sessions to UP. As for which TLVs should be carried, it depends on the specific session data to be synchronized. This is equivalent to create the specific session. Refer to Section 5 to see more details.

```
<Sync_Data Message> ::= <Common Header>
                        [<User Routing TLVs>]
                        [<User Information TLVs>]
                        [<L2TP Subscriber TLVs>]
                        [<Subscriber CGN Port Range TLV>]
                        [<Subscriber Policy TLV>]
```

6.2.6 Sync_End Message

The Sync_End message is used to indicate the end of a synchronization process. The format of a Sync_End message is as follows:

```
<Sync_End Message> ::= <Common Header>
                        <Error Information TLV>
```

The return/error codes are listed as follows:

- 0: Success, synchronization finished.
- 1: Failure, malformed message received.
- 2: One or more of the TLVs was not understood.

6.2.7 Update_Request Message

The Update_Request message is a multi-task message, it can be used to create, update, and delete subscriber sessions on a UP.

For session operations, the specific operation is controlled by the "Oper" field of the carried TLVs. As defined in Section 7.1, the "Oper" can be set to either "update" or "delete" when a TLV is carried in an Update_Request message.

When the "Oper" set to update, it means to create or update a subscriber session, if the "Oper" set to delete, it indicates to delete a corresponding session on an UP.

The format of Update_Request message is as follows:

```
<Update_Request Message> ::= <Common Header>
                              [<User Routing TLVs>]
                              [<User Information TLVs>]
                              [<L2TP Subscriber TLVs>]
                              [<Subscriber CGN Port Range TLV>]
                              [<Subscriber Policy TLV>]
```

Each Update_Request message will result in an Update_Response message that is defined in Section 6.2.8.

6.2.8 Update_Response Message

The Update_Response message is a response to an Update_Request message. It is used to confirm the update request (or reject it in the case of an error). The format of an Update_Response message is as follows:

```
<Update_Response Message> ::= <Common Header>
                               [<Subscriber CGN Port Range TLV>]
                               <Error Information TLV>
```

The return/error codes are carried in the Error Information TLV. They are listed as follows:

- 0: Success.
- 1: Failure, malformed message received.
- 2: One or more of the TLVs was not understood.
- 3001 (Pool-Mismatch): The corresponding address pool cannot be found.
- 3002 (Pool-Full): The address pool is fully allocated and no address segment is available.
- 3003 (Subnet-Mismatch): The address pool subnet cannot be found.
- 3004 (Subnet-Conflict): Subnets in the address pool have been classified into other clients.
- 4001 (Update-Fail-No-Res): The forwarding table fails to be delivered because the forwarding resources are insufficient.
- 4002 (QoS-Update-Success): The QoS policy takes effect.
- 4003 (QoS-Update-Sq-Fail): Failed to process the queue in the QoS policy.
- 4004 (QoS-Update-CAR-Fail): Processing of the CAR in the QoS policy fails.
- 4005 (Statistic-Fail-No-Res): Statistics processing failed due to insufficient statistics resources.

6.3 Event Message

The Event message is used to report subscriber session traffic statistics and detection information. The format of Event message is as follows:

```
<Event Message> ::= <Common Header>
                    [<User Traffic Statistics Report TLV>]
                    [<User Detection Result Report TLV>]
```


6.4 Report Message

The Report message is used to report board and interface status on a UP. The format of Report message is as follows:

```
<Report Message> ::= <Common Header>
                        [<Board Status TLVs>]
                        [<Interface Status TLVs>]
```

6.5 CGN Messages

This document defines the following resource allocation messages:

Type	Message Name	TLV that is carried
200	Addr_Allocation_Req	Address Allocation Request
201	Addr_Allocation_Ack	Address Allocation Response
202	Addr_Renew_Req	Address Renewal Request
203	Addr_Renew_Ack	Address Renewal Response
204	Addr_Release_Req	Address Release Request
205	Addr_Release_Ack	Address Release Response

6.5.1 Addr_Allocation_Req Message

The Addr_Allocation_Req message is used to request CGN address allocation. The format of Addr_Allocation_Req message is as follows:

```
<Addr_Allocation_Req Message> ::= <Common Header>
                                   <Address Allocation Request TLV>
```

6.5.2 Addr_Allocation_Ack Message

The Addr_Allocation_Ack message is a response to an Addr_Allocation_Req message. The format of Addr_Allocation_Ack message is as follows:

```
<Addr_Allocation_Ack Message> ::= <Common Header>
                                   <Address Allocation Response TLV>
```

6.5.3 Addr_Renew_Req Message

The Addr_Renew_Req message is used to request address renewal. The format of Addr_Renew_Req message is as follows:

```
<Addr_Renew_Req Message> ::= <Common Header>
                               <Address Renewal Request TLV>
```

6.5.4 Addr_Renew_Ack Message

The Addr_Renew_Ack message is a response to an Addr_Renew_Req message. The format of Addr_Renew_Ack message is as follows:

```
<Addr_Renew_Ack Message> ::= <Common Header>
                               <Address Renewal Response TLV>
```

6.5.5 Addr_Release_Req Message

The Addr_Release_Req message is used to request address release. The format of Addr_Release_Req message is as follows:

```
<Addr_Release_Req Message> ::= <Common Header>
                               <Address Release Request TLV>
```

6.5.6 Addr_Release_Ack Message

The Addr_Release_Ack message is a response to an Addr_Release_Req message. The format of Addr_Release_Ack message is as follows:

```
<Addr_Release_Ack Message> ::= <Common Header>
                               <Address Release Response TLV>
```

6.6 Vendor Message

The Vendor message is, in conjunction with the vendor TLV and vendor sub-TLV, can be used by vendors to extend the S-CUSP protocol. It's message type is 11. If the receiver does not recognize the message, an Error message will be returned to the sender.

The format of the Vendor message is as follows:

```
<Vendor Message> ::= <Common Header>
                        <Vendor TLV>
                        [<any other TLVs as specified by the vendor>]
```

6.7 Error Message

The Error message is defined to return some critical error information to the sender. If a receiver does not know the message type of a received message, it MUST return an Error message to the sender.

The format of the Error message is as below:

```
<Error Message> ::= <Common Header>
                    <Error Information TLV>
```

7. S-CUSP TLVs and Sub-TLVs

This section specifies the following:

- o the format of the TLVs that appear in S-CUSP messages,
- o the format of the sub-TLVs that appear within the values of some TLVs, and
- o the format of some basic data fields that appear within TLVs or sub-TLVs.

See Section 9 for a list of all defined TLVs and sub-TLVs.

7.1 Common TLV Header

S-CUSP messages consist of the common header specified in Section 6.1 followed by TLVs formatted as specified in this section.

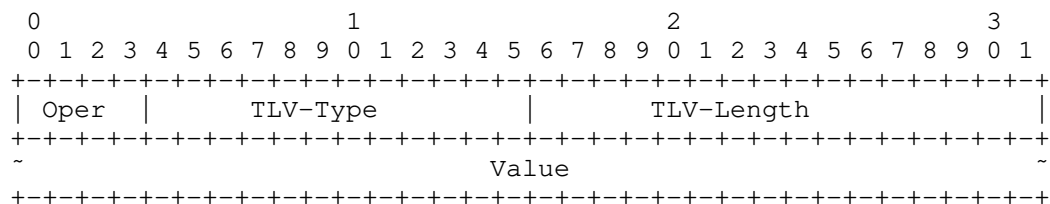


Figure 32: Common TLV Header

- o **Oper (4 bits):** For Message-Types that indicate an operation on a data set, the Oper field is interpreted as Update, Delete, or Reserved as specified in Section 9.3. For all other Message-Types, the Oper field MUST be sent as zero and ignored on receipt.
- o **TLV-Type (12 bits):** The Type of a TLV, that is the meaning and format of the Value part, are determined by the TLV-Type of the TLV. See Section 9.2.
- o **TLV-Length (2 bytes):** The length of the Value portion of the TLV in bytes as an unsigned integer.
- o **Value (variable length):** This is the value portion of the TLV whose size is given by TLV-Length. The value portion consists of fields, frequently using one of the basic data field types (see Section 7.2) and sub-TLVs (see Section 7.3).

7.2 Basic Data Fields

This section specifies the binary format of several standard basic data fields that are used within other data structures in this specification.

- o **STRING:** 0 to 255 octets. Will be encoded as a sub-TLV (see Section 7.3) to provide the length. The use of this data type in S-CUSP is to provide convenient labels for use by network operators in configuring and debugging their networks and interpreting S-CUSP messages. These labels will not normally be seen by subscribers. They are normally interpreted as ASCII [RFC20].
- o **MAC-Addr:** 6 octets. Ethernet MAC Address [RFC7042].
- o **IPv4-Address:** 8 octets. 4 octets of the IPv4 address value followed by a 4 octet address mask in the format XXX.XXX.XXX.XXX.
- o **IPv6-Address:** 20 octets. 16 octets of IPv6 address followed by a 4 octet integer n in the range of 0 to 128 which gives the address mask as the one's complement of $2^{(128-n)} - 1$.
- o **VLAN ID:** 2 octets. As follows [802.1Q]:

```

      0                               1
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+---+---+---+---+---+---+---+---+---+
| PRI |D|           VLAN-ID           |
+---+---+---+---+---+---+---+---+---+

```

PRI: Priority. Default value 7.

D: Drop Eligibility Indicator (DEI). Default value 0.

VLAN-ID: Unsigned integer in the range 1-4094. (0 and 4095 are not valid VLAN IDs [802.1Q].)

7.3 Sub-TLV Format and Sub-TLVs

In some cases, the Value portion of a TLV, as specified above, can contain one or more Sub-TLVs formatted as follows:

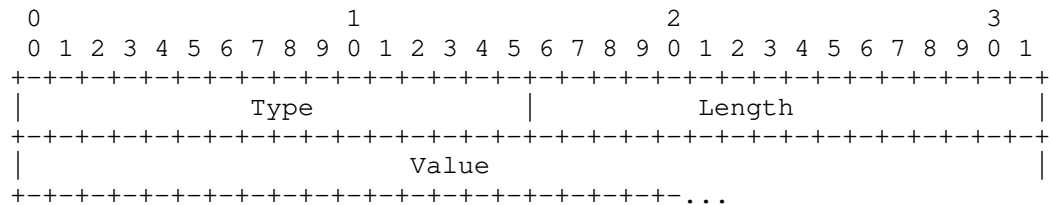


Figure 33: Sub-TLV Header

- o **Type (2 bytes):** The Type of a Sub-TLV, that is the meaning and format of the Value part, are determined by the Type of the TLV. Sub-TLV Types numbers have the same meaning regardless of the TLV Type of the TLV within which the sub-TLV occurs. See Section 9.4.
- o **Length (2 bytes):** The length of the Value portion of the sub-TLV in bytes as an unsigned integer.
- o **Value (variable length):** This is the value portion of the sub-TLV whose size is given by Length.

The sub-TLVs currently specified are defined in the following subsections.

7.3.1 Name sub-TLVs

This document defines the following name sub-TLVs that are used to carry the name of the corresponding object. The length of each of these sub-TLV is variable from 1 to 255 octets. The value is of type STRING padded with zeros octets to 4-octet alignment.

Type	Sub-TLV Name	Meaning
1	VRF-Name	The name of a VRF
2	Ingress-QoS-Profile	The name of an ingress QoS profile
3	Egress-QoS-Profile	The name of an egress QoS profile
4	User-ACL-Policy	The name of an ACL policy
5	Multicast-ProfileV4	The name of an IPv4 multicast profile
6	Multicast-ProfileV6	The name of an IPv6 multicast profile
7	NAT-Instance	The name of a NAT instance
8	Pool-Name	The name of an address pool

7.3.2 Ingress-CAR sub-TLV

The Ingress-CAR sub-TLV indicates the authorized upstream Committed Access Rate (CAR) parameters. The sub-TLV type of the Ingress-CAR sub-TLV is 9 and the sub-TLV length is 16. The format is as shown in Figure 34.

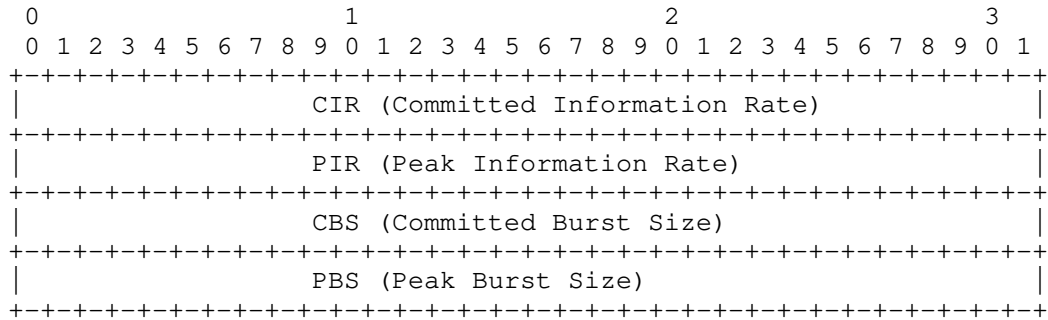


Figure 34: Ingress-CAR sub-TLV

Where:

CIR (4 bytes): Guaranteed rate in bits/second.

PIR (4 bytes): Burst rate in bits/second.

CBS (4 bytes): The token bucket in bytes.

PBS (4 bytes): Burst token bucket in bytes.

These fields are unsigned integers. More details about CIR, PIR, CBS, and PBS can be found in [RFC2698].

7.3.3 Egress-CAR sub-TLV

The Egress-CAR sub-TLV indicates the authorized downstream Committed Access Rate (CAR) parameters. The sub-TLV type of the Egress-CAR sub-TLV is 10 and its sub-TLV length is 16 octets. The format of the value part is as defined below.

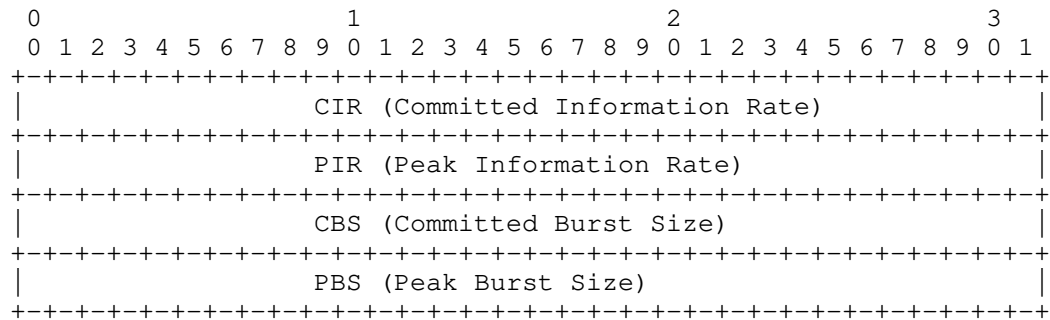


Figure 35: Egress-CAR sub-TLV

Where:

CIR (4 bytes): Guaranteed rate in bits/second.

PIR (4 bytes): Burst rate in bits/second.

CBS (4 bytes): The token bucket in bytes.

PBS (4 bytes): Burst token bucket in bytes.

These fields are unsigned integers. More details about CIR, PIR, CBS, and PBS can be found in [RFC2698].

7.3.4 If-Desc sub-TLV

The If-Desc sub-TLV is defined to designate an interface. It is an optional sub-TLV that may be carried in those TLVs that have an "if-index" or "out-if-index" field. The If-Desc sub-TLV is used as a local unique identifier within a BNG.

The sub-TLV type is 11 and the sub-TLV length is 12 octets. The format depends on the If-Type. The format of the value part is as follows:

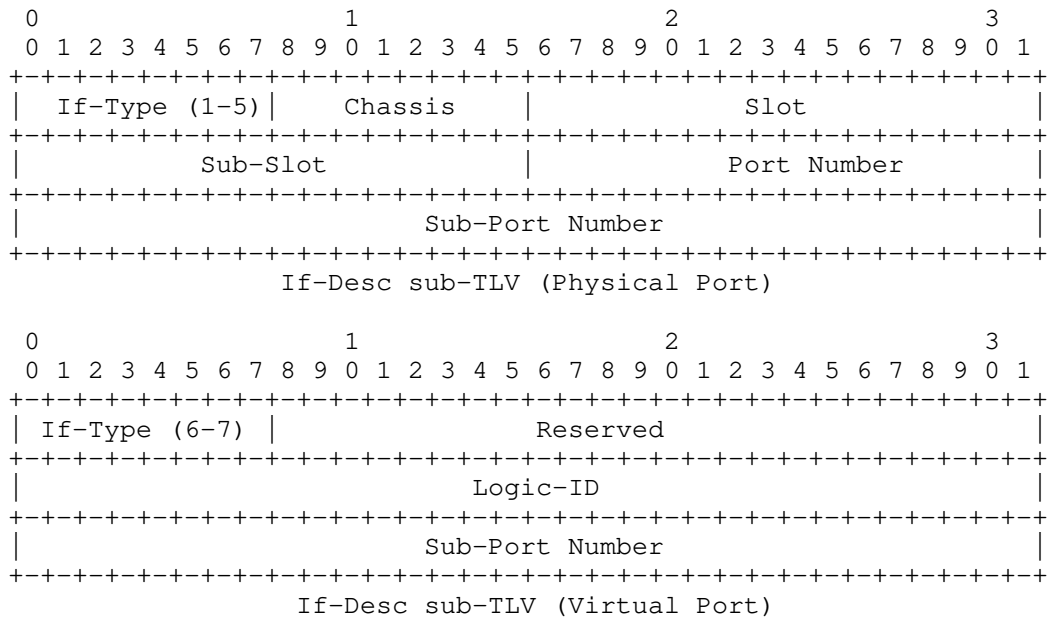


Figure 36: If-Desc sub-TLV Formats

Where:

If-Type: 8 bits in length, indicates the type of an interface. Following types are defined in this document:

- 0: Reserved
- 1: Fast Ethernet (FE)
- 2: GE
- 3: 10GE
- 4: 100GE
- 5: Eth-Trunk
- 6: Tunnel
- 7: VE
- 8-255: Reserved.

Chassis (8 bits): Identifies the chassis that the interface belongs to.

Slot (16 bits): Identifies the slot that the interface belongs to.

Sub-slot (16 bits): Identifies the sub-slot the interface belongs to.

Port Number (16 bits): An identifier of a physical port/interface (e.g., If-Type: 1-5). It is locally significant within the slot/sub-slot.

Sub-port Number (32 bits): An identifier of the sub-port. Locally significant within its "parent" port (physical or virtual).

Logic-ID (32 bits): An identifier of a virtual interface (e.g., If-Type: 6-7)

7.3.5 IPv6 Address List sub-TLV

The IPv6 Address List sub-TLV is used to convey one or more IPv6 addresses. It is carried in the IPv6 Subscriber TLV. The sub-TLV type is 12, the sub-TLV length is variable.

The format of IPv6 Addresses sub-TLV is as follows:

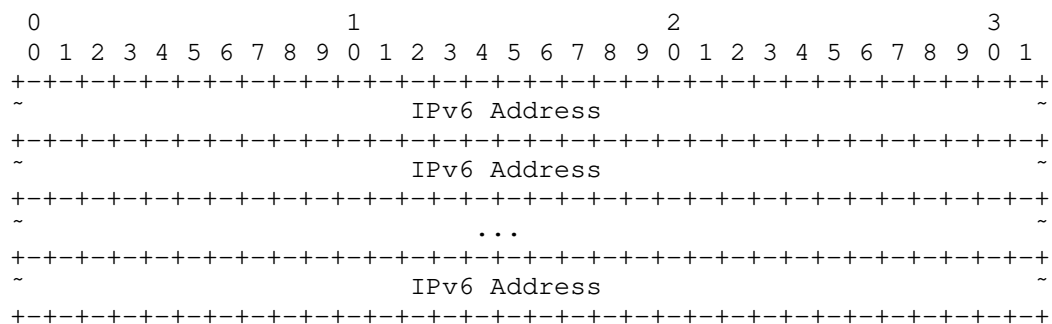


Figure 37: IPv6 Address List sub-TLV

Where:

IP Address (IPv6-Address): Each IP Address is an IP-Address type, carries an IPv6 address.

7.3.6 Vendor sub-TLV

The Vendor sub-TLV is intended to be used inside the value portion of the Vendor TLV (Section 7.13). It provides a Sub-Type that effectively extends the sub-TLV type in the sub-TLV header and provides for versioning of vendor sub-TLVs.

The value part of the Vendor sub-TLV is formatted as follows:

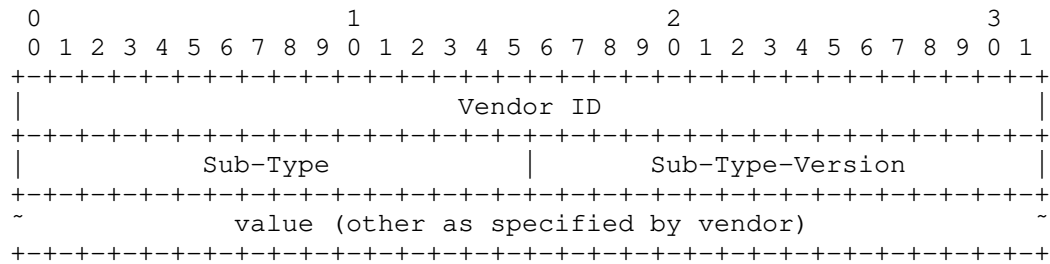


Figure 38: Vendor sub-TLV

Where:

The sub-TLV type: 13.

The sub-TLV length: variable.

Vendor-ID (4 bytes): Vendor ID as defined in RADIUS [RFC2865].

Sub-Type (2 bytes): Used by the Vendor to distinguish multiple different sub-TLVs.

Sub-Type-Version (2 bytes): Used by the Vendor to distinguish different versions of a Vendor Defined sub-TLV sub-Type.

value: as specified by the vendor.

Since Vendor code will be handling the sub-TLV after the Vendor ID field is recognized, the remainder of the sub-TLV can be organized however the vendor wants. But it is desirable for a vendor to be able to define multiple different vendor sub-TLVs and to keep track of different versions of its vendor defined sub-TLVs. Thus, it is RECOMMENDED that the vendor assign a Sub-Type value for each of that vendor's sub-TLVs that is different from other Sub-Type values that vendor has used. Also, when modifying a vendor defined sub-TLV in a way potentially incompatible with a previous definition, the vendor SHOULD increase the value it is using in the Sub-Type-Version field.

7.4 The Hello TLV

The Hello TLV is defined to be carried in the Hello message for version and capabilities negotiation. It indicates the S-CUSP sub-version and capabilities supported. The format of Hello TLV is as follows:

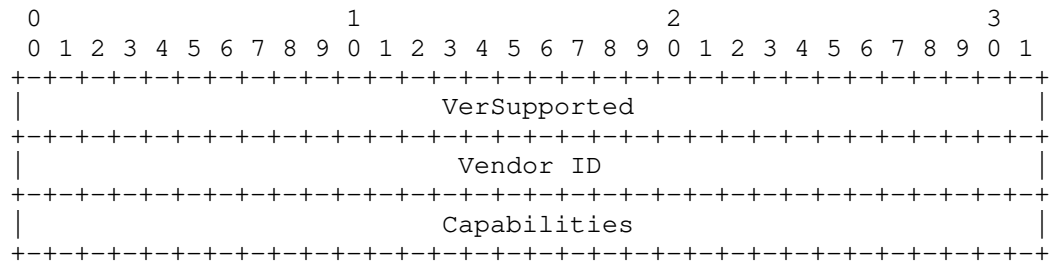


Figure 39: Hello TLV

Where:

The TLV type is 100.

The TLV length is 12 octets.

The value field consists of three parts:

VerSupported: 32 bits in length. This is a bit map of the Sub-Versions of the S-CUSP protocol that the sender supports. This document specifies Sub-Version zero of Major Version 1, that is, Version 1.0. The VerSupported field **MUST** be non-zero. The VerSupported bits are numbered from 0 as the most significant bit. Bit 0 indicates support of Sub-Version zero, bit 1 indicates support of Sub-Version one, etc.

Vendor-ID: 4 bytes in length. Vendor ID, as defined in RADIUS [RFC2865].

Capabilities: 32 bits in length. Flags that indicate the support of particular capabilities by the sender of the Hello. No Capabilities are defined in this document and so implementations will set this field to zero. The Capabilities field of the Hello TLV **MUST** be checked before any other TLVs in the Hello because capabilities defined in the future might extend existing TLVs or permit new TLVs.

After the exchange of Hello messages, the CP and UP each perform a logical AND of the Sub-Version supported by the CP and the UP and separately perform a logical AND of the Capabilities bits fields for the CP and the UP.

If the result of the AND of the Sub-Versions supported is zero, then no session can be established and the connection is torn down. If the result of the AND of the Sub-Versions supported is non-zero, then the session uses the highest Sub-Version supported by both the CP and UP.

For example, if one side supports Sub-Versions 1, 3, 4, and 5 (VerSupported = 0x5C000000) and the other side supports 2, 3, and 4 (VerSupported = 0x38000000) then 3 and 4 are the Sub-Versions in common and 4 is the highest Sub-Version supported by both sides. So Sub-Version 4 is used for the session that has been negotiated.

The result of the logical AND of the Capabilities bits will show what additional capabilities both sides support. If this result is zero, there are no such capabilities so none can be used during the session. If this result is non-zero, it shows the additional capabilities that can be used during the session. The CP and the UP MUST NOT use a capability unless both advertise support.

7.5 The Keep Alive TLV

The Keep Alive TLV is defined to be carried in the Hello message. It provides timing information for the keep alive feature. The format of Hello TLV is as follows:

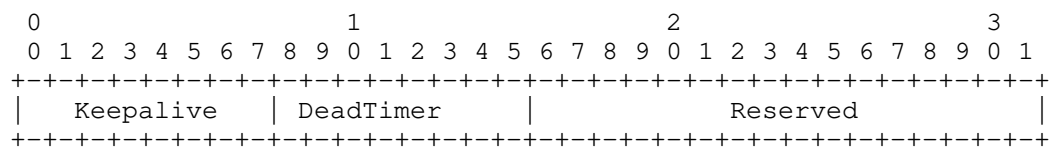


Figure 40: Keep Alive TLV

Where:

The TLV type: 102.

The value of length: 4 octets.

Keepalive (8 bits): Indicates the maximum period of time (in seconds) between two consecutive S-CUSP messages sent by the sender of the message containing this TLV as an unsigned integer. The minimum value for the Keepalive is 1 second. When set to 0, once the session is established, no further Keepalive messages are sent to the remote peer. A RECOMMENDED value for the Keepalive frequency is 30 seconds.

DeadTimer (8 bits in length): Specifies the amount of time as an unsigned integer number of seconds after the expiration of which

the S-CUSP peer can declare the session with the sender of the Hello message to be down if no S-CUSP message has been received. The DeadTimer SHOULD be set to 0 and MUST be ignored if the Keepalive is set to 0. A RECOMMENDED value for the DeadTimer is 4 times the value of the Keepalive.

The Reserved bits MUST be sent as zero and ignored on receipt.

7.6 The Error Information TLV

The Error Information TLV is a common TLV that can be used in many Response (e.g., Update_Response message) and ACK messages (e.g., Addr_Allocation_Ack message, etc.). It is used to convey the information about an error in the received S-CUSP message. The format of the Error Information TLV is as follows:

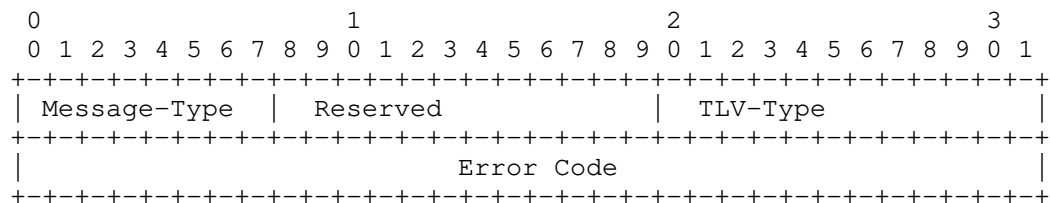


Figure 41: Error Information TLV

Where:

The TLV type: 101.

The value of length: 8 octets.

Message-Type(1 byte): This parameter is the message type of the message containing an error.

Reserved (1 byte): MUST be sent as zero and ignored on receipt.

TLV-Type (2 bytes): Indicates which TLV caused the error.

Error Code: 4 bytes in length. Indicate the specific Error Code (see Section 9.5).

7.7 BAS Function TLV

The BAS Function TLV is used by a CP to control the access mode, authentication methods, and other related functions of an interface

on a UP.

The format of the BAS Function TLV value part is as follows:

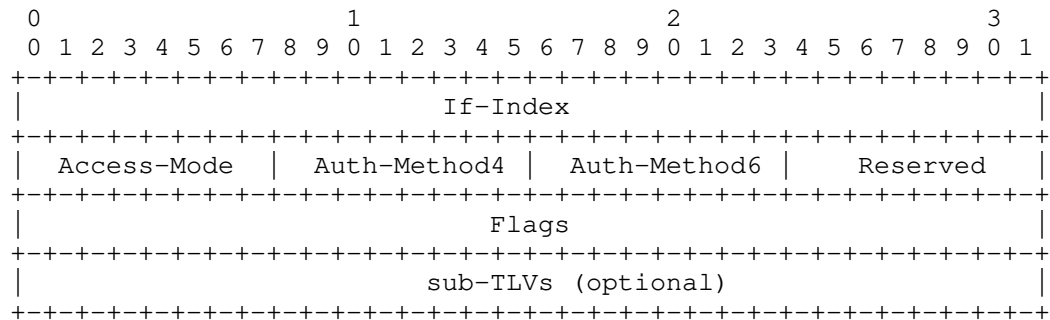


Figure 42: BAS Function TLV

Where:

The TLV type: 1.

The value of length: variable.

If-Index: 4 bytes in length, a unique identifier of an interface of a BNG.

Access-Mode: 1 byte in length, indicates the access mode of the interface. This document defines following values:

- 0: Layer 2 subscriber;
- 1: Layer 3 subscriber;
- 2: Layer 2 leased line;
- 3: Layer 3 leased line;
- 4-255: Reserved.

Auth-Method4: 1 byte in length, for IPv4 scenario, it indicates the authentication on this interface; this field is defined as a bitmap, this document defines following values (other bits are reserved and MUST be sent as zero and ignored on receipt):

- 0x1: PPPoE authentication;
- 0x2: DOT1X authentication;
- 0x4: Web authentication;
- 0x8: Web fast authentication;
- 0x10: Binding authentication.

Auth-Method6: 1 byte in length, for IPv6 scenario, it indicates the authentication on this interface; this field is defined as a bitmap, this document defines following values (other bits are

reserved and MUST be sent as zero and ignored on receipt):

0x1: PPPoE authentication;
 0x2: DOT1X authentication;
 0x4: Web authentication;
 0x8: Web fast authentication;
 0x10: Binding authentication;

sub-TLVs:

The IF-Desc sub-TLV can be carried.
 If-Desc sub-TLV: carries the interface information.

The flags field is defined as follows:

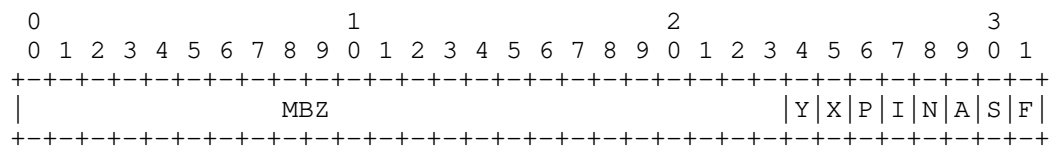


Figure 43: Interface Flags

Where:

F (IPv4 Trigger) bit: Indicates whether IPv4 packets can trigger a subscriber to go online. 1: enabled, 0: disabled.

S (IPv6 Trigger) bit: Indicates whether IPv6 packets can trigger a subscriber to go online. 1: enabled, 0: disabled.

A (ARP Trigger) bit: Indicates whether ARP packets can trigger a subscriber to go online. 1: enabled, 0: disabled.

N (ND Trigger) bit: Indicates whether ND packets can trigger a subscriber to go online. 1: enabled, 0: disabled.

I (IPoE-Flow-Check): Used for UP detection. IPoE 1: Enable traffic detection. 0: Disable traffic detection.

P (PPP-Flow-Check) bit: Used for UP detection. PPP 1: Enable traffic detection. 0: Disable traffic detection.

X (ARP-Proxy) bit: 1: The interface is enabled with ARP proxy and can process ARP requests across different Port+VLANs. 0: The ARP proxy is not enabled on the interface, and only the ARP requests of the same Port+VLAN are processed.

Y (ND-Proxy) bit: 1: The interface is enabled with ND proxy and can process ND requests across different Port+VLANs. 0: The ND proxy is not enabled on the interface, and only the ND requests of the same Port+VLAN are processed.

MBZ: Reserved bits that MUST be sent as zero and ignored on receipt.

7.8 Routing TLVs

Normally, after an S-CUSP session is established between a UP and a CP, the CP will allocate one or more blocks of IP addresses to the UP. Those IP addresses will be allocated to subscribers who will dial-up to the UP. In order to make sure that other nodes within the network learn how to reach those IP addresses, the CP needs to install one or more routes that can reach those IP addresses on the UP and notify the UP to advertise the routes to the network.

The Routing TLVs are used by a CP to notify a UP of the network routing. They can be carried in the Update_Request message and Sync_Data message.

7.8.1 IPv4 Routing TLV

The IPv4 Routing TLV is used to carry IPv4 network routing related information.

The format of the TLV value part is as below:

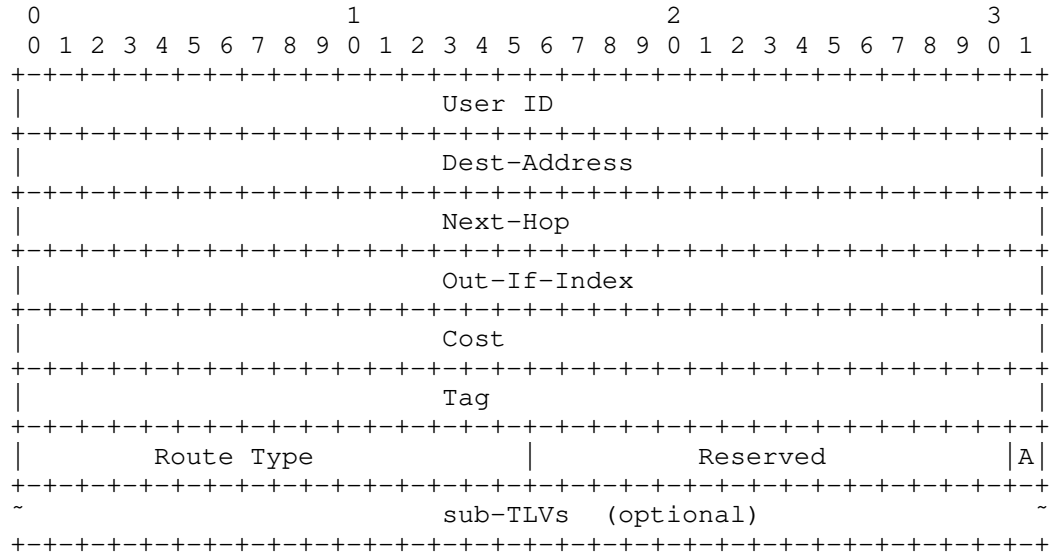


Figure 44: IPv4 Routing TLV

Where:

The TLV Type: 7

The TLV Length: Variable

User-ID: 4 bytes in length. Carries the user identifier. This field is filled with all Fs when a non-user route is delivered to the UP.

Dest-Address (IPv4-Address type): Identifies the destination address.

Next-Hop: (IPv4-Address type): Identifies the next hop address.

Out-If-Index (4 bytes): Indicates the interface index.

Cost (4 bytes): The cost value of the route.

Tag (4 bytes): The tag value of the route.

Route-Type (2 bytes): Indicates the route type. The options are as follows:

- 0: User host route
- 1: Radius authorization FrameRoute
- 2: Network segment route
- 3: Gateway route
- 4: Radius authorized IP route
- 5: L2TP LNS side user route

Advertise-Flag: 1 bit. Indicates whether the route should be advertised by the UP. Following flags are defined:

- 0: Not advertised,
- 1: advertised.

sub-TLVs: The VRF-Name and/or If-Desc sub-TLVs can be carried.

VRF-Name sub-TLV: indicates which VRF the route belongs to.

If-Desc sub-TLV: carries the interface information.

The Reserved field MUST be sent as zero and ignored on receipt.

7.8.2 IPv6 Routing TLV

The IPv6 Routing TLV is used to carry IPv6 network routing information.

The format of this TLV is as follows:

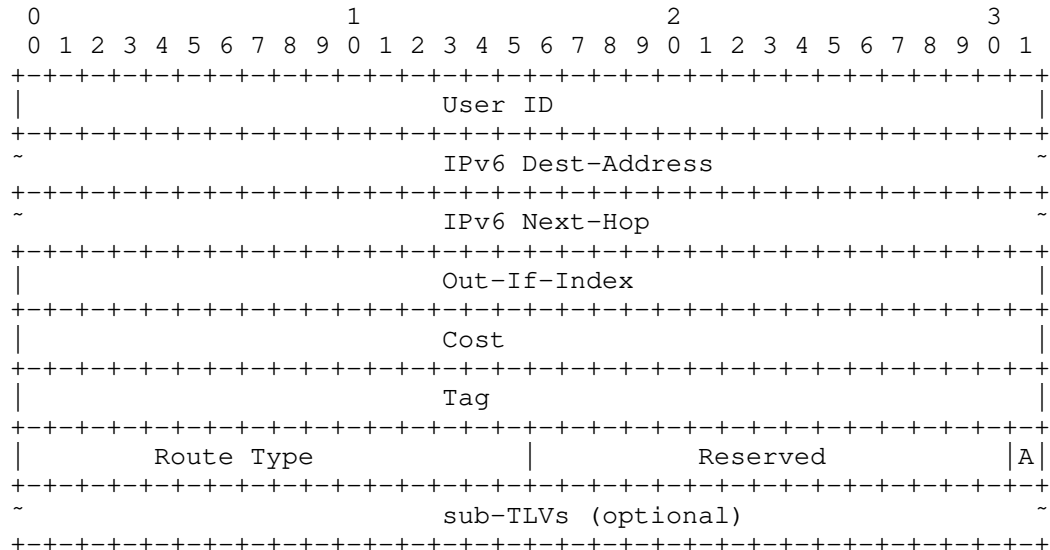


Figure 45: IPv6 Routing TLV

Where:

The TLV Type: 7

The TLV Length: Variable

User-ID: 4 bytes in length. Carries the user identifier. This field is filled with all Fs when a non-user route is delivered to the UP.

IPv6 Dest-Address (IPv6-Address type): Identifies the destination address.

IPv6 Next-Hop: (IPv6-Address type): Identifies the next hop address.

Out-If-Index (4 bytes): Indicates the interface index.

Cost (4 bytes): The cost value of the route.

Tag (4 bytes): The tag value of the route.

Route-Type: (2 bytes): Indicates the route type. The options are as follows:

- 0: User host route
- 1: Radius authorization FrameRoute
- 2: Network segment route
- 3: Gateway route
- 4: Radius authorized IP route
- 5: L2TP LNS side user route

Advertise-Flag: 1 bit. Indicates whether the route should be advertised by the UP. Following flags are defined:

- 0: Not advertised,
- 1: advertised.

sub-TLVs: If-Desc and VRF-Name sub-TLVs can be carried.

VRF-Name sub-TLV: Indicates the VRF to which the subscriber belongs.

If-Desc sub TLV: carries the interface information.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9 Subscriber TLVs

The Subscriber TLVs are defined for a CP to send the basic information about a user to a UP.

7.9.1 Basic Subscriber TLV

The Basic Subscriber TLV is used to carry the basic common information for all kinds of access subscribers. It is carried in an Update_Request message.

The format of the Basic Subscriber TLV value part is as follows:

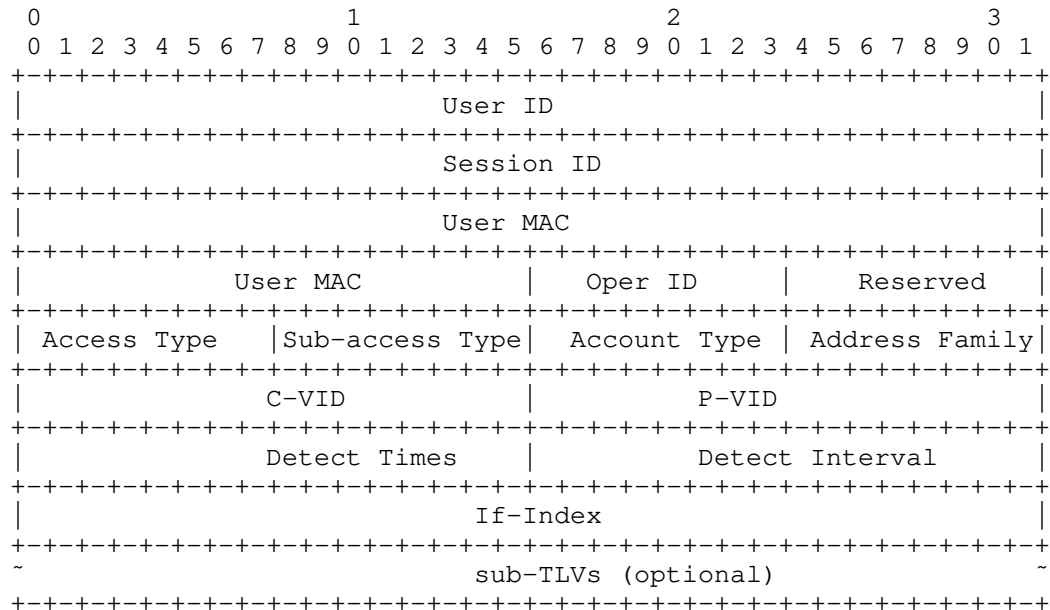


Figure 46: Basic Subscriber TLV

Where:

The TLV Type: 2.

The TLV Length: variable in length.

User-ID (4 bytes): The identifier of a subscriber.

Session-ID (4 bytes): Session ID of a PPPoE subscriber. Zero means non-PPPoE subscriber.

User-Mac (MAC-Addr type): The MAC Address of a subscriber.

Oper-ID (1 byte): Indicates the ID of an operation performed by a user. This field is carried in the response from the UP.

Reserved (1 byte): MUST be sent as zero and ignored on receipt.

Access-Type (1 byte):

- 1: PPP access (PPP [RFC1661])
- 2: PPP over Ethernet over ATM access (PPPoEoA)
- 3: PPP over ATM access (PPPoA [RFC3336])
- 4: PPP over Ethernet access (PPPoE [RFC2516])
- 5: PPPoE over VLAN access (PPPoEoVLAN [RFC2516])
- 6: PPP over LNS access (PPPoLNS)
- 7: IP over Ethernet DHCP access (IPoE_DHCP)
- 8: IP over Ethernet EAP authentication access (IPoE_EAP)
- 9: IP over Ethernet Layer 3 access (IPoE_L3)
- 10: IP over Ethernet Layer 2 Static access (IPoE_L2_STATIC)
- 11: Layer 2 Leased Line access (L2_Leased_Line)
- 12: Layer 2 VPN Leased Line access (L2VPN_Leased_Line)
- 13: Layer 3 Leased Line access (L3_Leased_Line)
- 14: Layer 2 Leased line Sub-User access
(L2_Leased_Line_SUB_USER)
- 15: L2TP LAC tunnel access (L2TP_LAC)
- 16: L2TP LNS tunnel access (L2TP_LNS)

Sub-Access-Type (1 byte): Indicates whether PPP termination or PPP relay is used.

- 0: Reserved
- 1: PPP Relay (for LAC)
- 2: PPP termination (for LNS)

Account-Type(1 byte):

- 0: Collects statistics on IPv4 and IPv6 traffic of terminals independently.
- 1: Collects statistics on IPv4 and IPv6 traffic of terminals.

Address Family (1 byte)

- 1: IPv4
- 2: IPv6
- 3: dual stack

C-VID (VLAN-ID): Indicates the inner VLAN ID. The value 0 indicates that the VLAN ID is invalid. The default value of PRI is 7, the value of DEI is 0, and the value of VID is 1~4094. The PRI value can also be obtained by parsing terminal packets.

P-VID (VLAN-ID): Indicates the outer VLAN ID. The value 0 indicates that the VLAN ID is invalid. The format is the same as that for C-VID.

Detect-Times (2 bytes): Number of detection timeout times. The value 0 indicates that no detection is performed.

Detect-Interval (2 bytes): Detection interval in seconds.

If-Index (4 bytes): Interface index.

Sub-TLVs: VRF-Name sub-TLV and If-Desc sub-TLV can be carried.

VRF-Name sub-TLV: Indicates the VRF to which the subscriber belongs.

If-Desc sub-TLV: carries the interface information.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.2 PPP Subscriber TLV

The PPP Subscriber TLV is defined to carry PPP information of a User from a CP to a UP. It will be carried in an Update_Request message when PPPoE or L2TP access is used.

The format of the TLV value part is as follows:

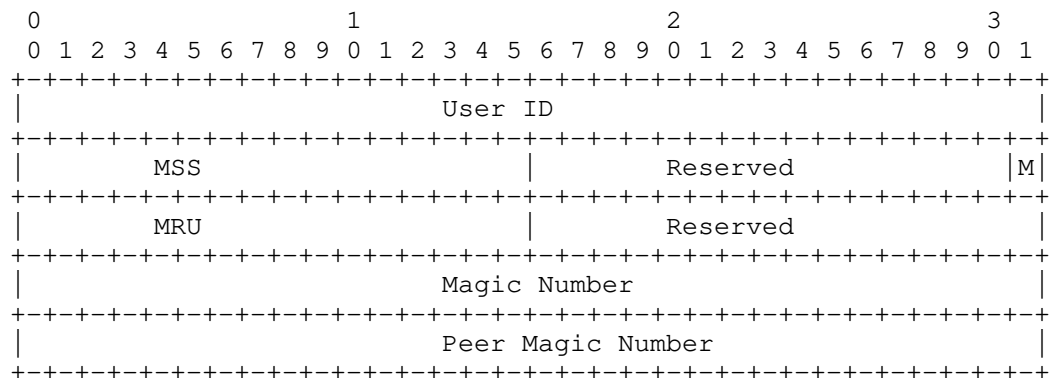


Figure 47: PPP Subscriber TLV

Where:

The TLV type: 3.

The TLV length: 12 octets.

User-ID (4 bytes): The identifier of a subscriber.

MSS-Value (2 bytes): Indicates the MSS value (in bytes).

MSS-Enable (M) (1 bit): Indicates whether the MSS is enabled.

0: Disabled.

1: Enabled.

MRU (2 bytes): PPPoE local MRU (in bytes).

Magic-Number (4 bytes): Local magic number in PPP negotiation packets.

Peer-Magic-Number (4 bytes): Remote peer magic number.

The Reserved fields MUST be sent as zero and ignored on receipt.

7.9.3 IPv4 Subscriber TLV

The IPv4 Subscriber TLV is defined to carry IPv4 related information for a BNG user. It will be carried in an Update_Request message when IPv4 IPE, or PPPoE access is used.

The format of the TLV value part is as follows:

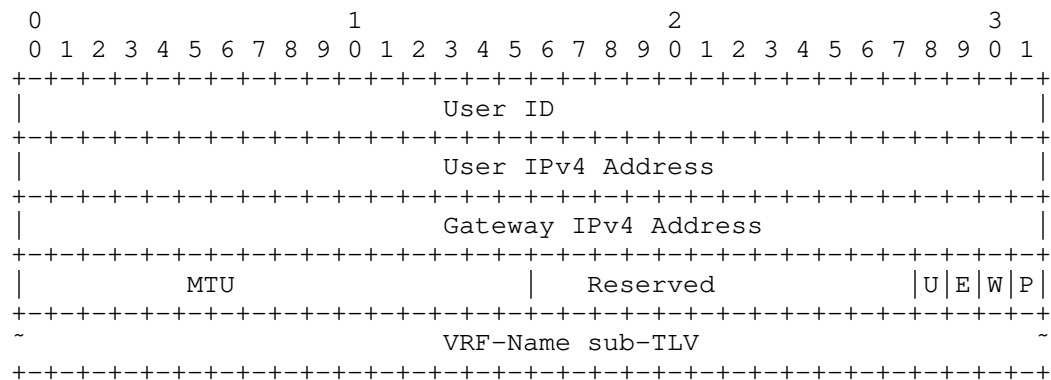


Figure 48: IPv4 Subscriber TLV

Where:

The TLV type: 4.

The TLV length: variable.

User-ID (4 bytes): The identifier of a subscriber.

User-IPv4 (IPv4-Address): The IPv4 address of the subscriber.

Gateway-IPv4 (IPv4-Address): The gateway address of the subscriber.

Portal Force (P) (1 bit): Push advertisement, 0: off, 1: on.

Web-Force (W) (1 bit): Push IPv4 Web. 0: off, 1: on.

Echo-Enable (E) (1 bit): UP returns ARP Req or PPP Echo. 0: off, 1: on.

IPv4-URPF (U) (1 bit): User Unicast Reverse Path Forwarding (URPF) flag. 0: off, 1: on.

MTU 2 bytes MTU value. The default value is 1500.

VRF-Name Sub-TLV: Indicates the subscriber belongs to which VRF.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.4 IPv6 Subscriber TLV

The IPv6 Subscriber TLV is defined to carry IPv6 related information for a BNG user. It will be carried in an Update_Request message when IPv6 IPE, or PPPoE access is used.

The format of the TLV value part is as follows:

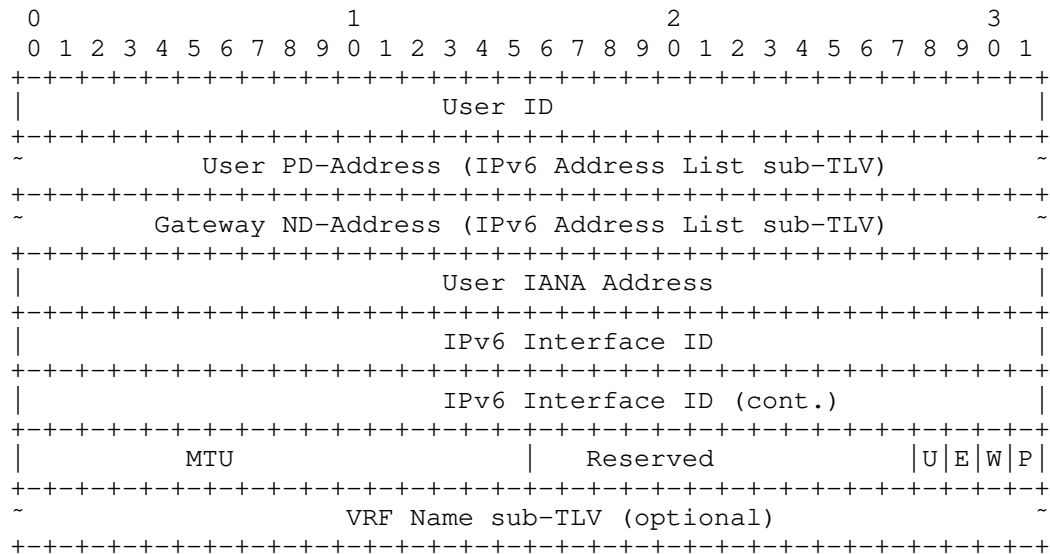


Figure 49: IPv6 Subscriber TLV

Where:

The TLV type: 5.

The TLV length: variable.

User-ID (4 bytes): The identifier of a subscriber.

User PD-Addresses (IPv6 Address List): Carries a list of Prefix Delegation (PD) addresses of the subscriber.

User ND-Addresses (IPv6 Address List): Carries a list of Neighbor Discovery (ND) addresses of the subscriber.

User IANA-Address (IPv6-Address): The IANA address of the subscriber.

IPv6 Interface ID (8 bytes): The identifier of an IPv6 interface.

Portal Force 1 bit (P): Push advertisement, 0: off, 1: on.

Web-Force 1 bit (W): Push IPv6 Web, 0: off, 1: on.

Echo-Enable 1 bit (E): The UP returns ARP Req or PPP Echo. 0: off; 1: on.

IPv6-URPF 1 bit (U): User Reverse Path Forwarding (URPF) flag, 0: off; 1: on.

MTU (2 bytes): The MTU value. The default value is 1500.

VRF-Name Sub-TLV: Indicates the VRF to which the subscriber belongs.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.5 IPv4 Static Subscriber Detect TLV

The IPv4 Static Subscriber Detect TLV is defined to carry IPv4 related information for a static access subscriber. It will be carried in an Update_Request message when IPv4 static access on a UP needs to be enabled.

The format of the TLV value part is as follows:

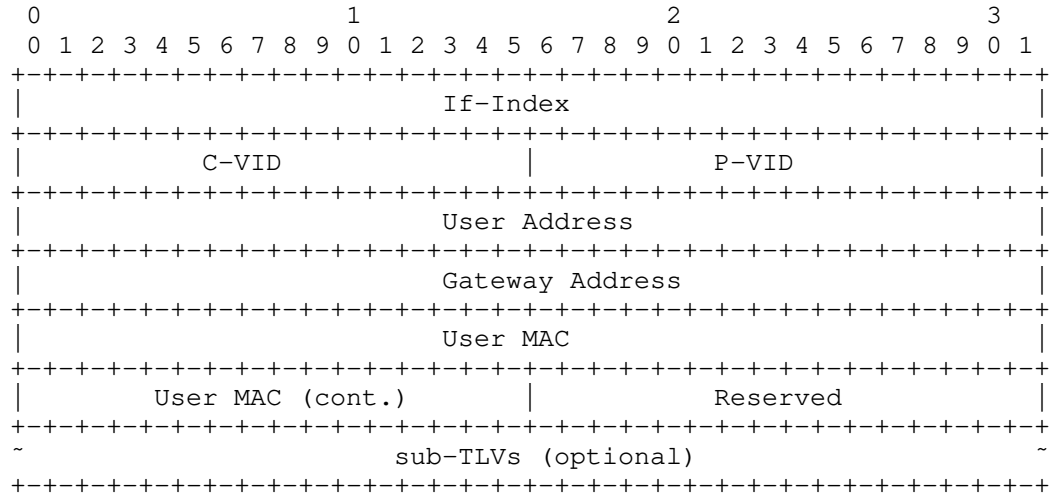


Figure 50: IPv4 Static Subscriber TLV

Where:

The TLV type: 6.

The TLV length: variable.

If-Index (4 bytes): The interface index of the interface from which the subscriber will dial-up.

C-VID (VLAN-ID): Indicates the inner VLAN ID. The value 0 indicates that the VLAN ID is invalid. A valid value is 1~4094.

P-VID (VLAN-ID): Indicates the outer VLAN ID. The value 0 indicates that the VLAN ID is invalid. The format is the same as that of the C-VID. A valid value is 1~4094. For a single-layer VLAN, set this parameter to PeVid.

User Address (IPv4-Addr): The user's IPv4 address.

Gateway Address (IPv4-Addr): The gateway's IPv4 Address.

User-MAC (MAC-Addr type): The MAC address of the subscriber.

Sub-TLVs: VRF-Name and If-Desc sub-TLVs may be carried.

VRF-Name sub-TLV: Indicates the VEF to which the subscriber belongs.

If-Desc sub-TLV: Carries the interface information.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.6 IPv6 Static Subscriber Detect TLV

The IPv6 Static Subscriber Detect TLV is defined to carry IPv6 related information for a static access subscriber. It will be carried in an Update_Request message when needed to enable IPv6 static subscriber detection on a UP.

The format of the TLV value part is as follows:

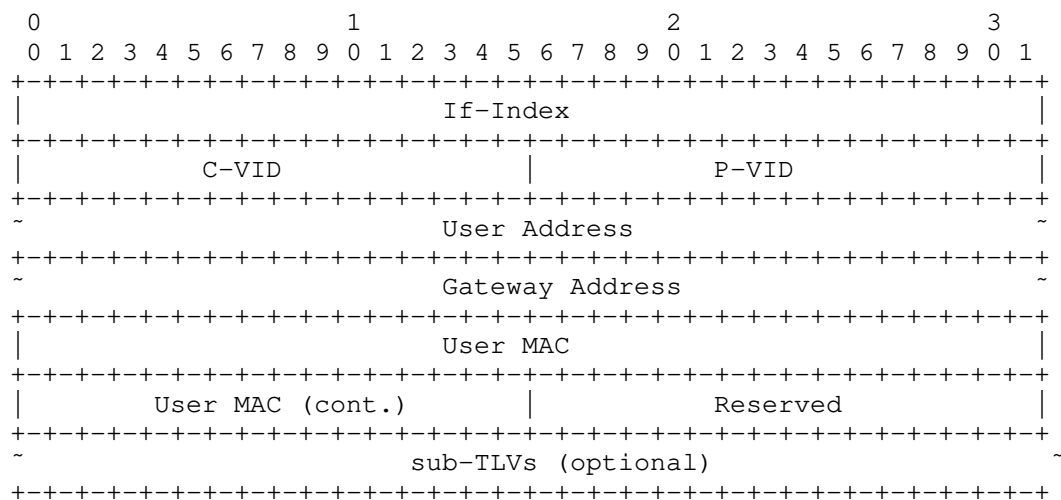


Figure 51: IPv6 Static Subscriber Detect TLV

Where:

The TLV type: 6.

The TLV length: variable.

If-Index (4 bytes): The interface index of the interface from which the subscriber will dial-up.

C-VID (VLAN-ID): Indicates the inner VLAN ID. The value 0 indicates that the VLAN ID is invalid. A valid value is 1~4094.

P-VID (VLAN-ID): Indicates the outer VLAN ID. The value 0 indicates that the VLAN ID is invalid. The format is the same as that the of C-VID. A valid value is 1~4094. For a single-layer VLAN, set this parameter to PeVid.

User Address (IPv6-Address type): The subscriber's IPv6 address.

Gateway Address (IPv6-Address type): The gateway's IPv6 Address.

User-MAC (MAC-Addr type): The MAC address of the subscriber.

sub-TLVs: VRF-Name and If-Desc sub-TLVs may be carried

VRF-Name Sub-TLV: Indicates the VRF to which the subscriber belongs.

If-Desc sub-TLV: Carries the interface information.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.7 L2TP-LAC Subscriber TLV

The L2TP-LAC Subscriber TLV is defined to carry the related information for a L2TP LAC access subscriber. It will be carried in an Update_Request message when L2TP LAC access is used.

The format of the TLV value part is as follows:

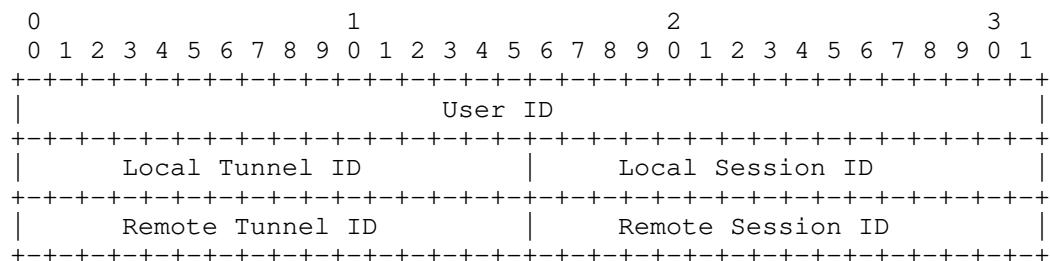


Figure 52: L2TP-LAC Subscriber TLV

Where:

The TLV type: 11.

The TLV length: 12 octets.

User-ID (4 bytes): The identifier of a user/subscriber.

Local-Tunnel-ID (2 bytes): The local ID of the L2TP tunnel.

Local-Session-ID (2 bytes): The local session ID with the L2TP tunnel.

Remote-Tunnel-ID (2 bytes): The remote ID of the L2TP tunnel.

Remote-Session-ID (2 bytes): The remote session ID of the L2TP tunnel.

7.9.8 L2TP-LNS Subscriber TLV

The L2TP-LNS Subscriber TLV is defined to carry the related information for a L2TP LNS access subscriber. It will be carried in an Update_Request message when L2TP LNS access is used.

The format of the TLV value part is as follows:

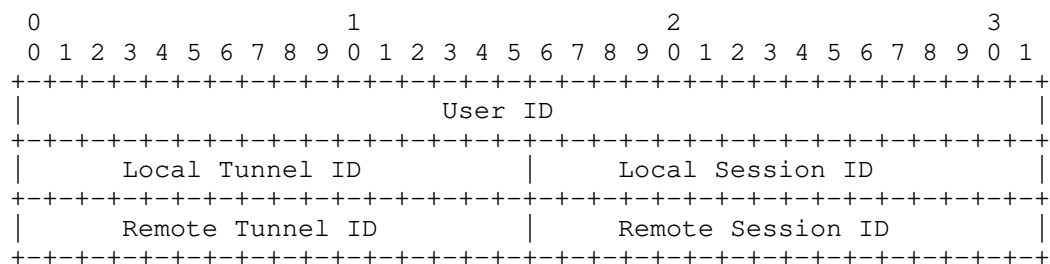


Figure 53: L2TP-LNS Subscriber TLV

Where:

The TLV type: 12.

The TLV length: 12 octets.

User-ID (4 bytes): The identifier of a user/subscriber.

Local-Tunnel-ID (2 bytes): The local ID of the L2TP tunnel.

Local-Session-ID (2 bytes): The local session ID with the L2TP tunnel.

Remote-Tunnel-ID (2 bytes): The remote ID of the L2TP tunnel.

Remote-Session-ID (2 bytes): The remote session ID of the L2TP tunnel.

7.9.9 L2TP-LAC Tunnel TLV

The L2TP-LAC Tunnel TLV is defined to carry the L2TP LAC tunnel related information. It will be carried in the Update_Request message when L2TP LAC access is used.

The format of the TLV value part is as follows:

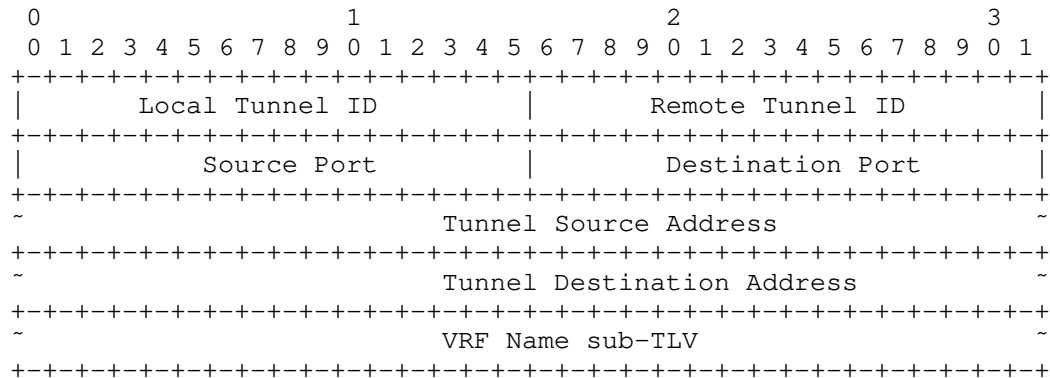


Figure 54: L2TP-LAC Tunnel TLV

Where:

The TLV type: 13.

The TLV length: variable.

Local-Tunnel-ID (2 bytes): The local ID of the L2TP tunnel.

Remote-Tunnel-ID (2 bytes): The remote ID of the L2TP tunnel.

Source-Port (2 bytes): The source UDP port number of an L2TP subscriber.

Dest-Port (2 bytes): The destination UDP port number of an L2TP subscriber.

Source-IP (IPv4/v6): The source IP address of the tunnel.

Dest-IP (IPv4/v6): The destination IP address of the tunnel.

VRF-Name Sub-TLV: The VRF name to which the L2TP subscriber tunnel belongs.

7.9.10 L2TP-LNS Tunnel TLV

The L2TP-LNS Tunnel TLV is defined to carry the L2TP LNS tunnel related information. It will be carried in the Update_Request message when L2TP LNS access is used.

The format of the TLV value part is as follows:

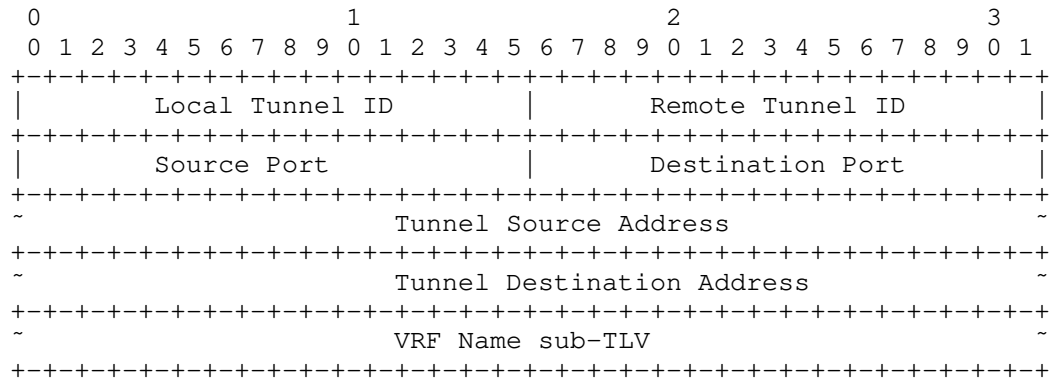


Figure 55: L2TP-LNS Tunnel TLV

Where:

The TLV type: 14.

The TLV length: variable.

Local-Tunnel-ID (2 bytes): The local ID of the L2TP tunnel.

Remote-Tunnel-ID (2 bytes): The remote ID of the L2TP tunnel.

Source-Port (2 bytes): The source UDP port number of an L2TP subscriber.

Dest-Port (2 bytes): The destination UDP port number of an L2TP subscriber.

Source-IP (IPv4/v6): The source IP address of the tunnel.

Dest-IP (IPv4/v6): The destination IP address of the tunnel.

VRF-Name Sub-TLV: The VRF name to which the L2TP subscriber tunnel belongs.

7.9.11 Update Response TLV

The Update Response TLV is used to return the operation result of an update request. It is carried in the Update_Response message as a response to the Update_Request message.

The format of Update Response TLV is as follows:

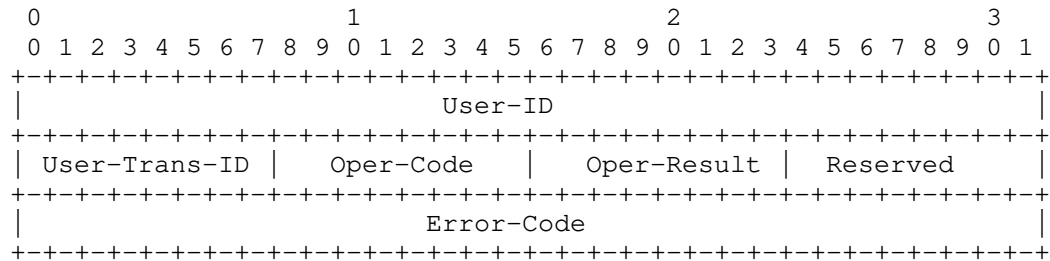


Figure 56: Update Response TLV

Where:

The TLV type: 302.

The TLV length: 12.

User-ID (4 bytes): A unique identifier of an user/subscriber.

User-Trans-ID (1 byte): In the case of dual-stack access or when modifying a session, User-Trans-ID is used to identify a user operation transaction. It is used by the CP to correlate a response to a specific request.

Oper-Code (1 byte): Operation code. 1: update, 2: delete.

Oper-Result (1 byte): Operation Result. 0: Success, Others: Failure.

Error-Code (4 bytes): Operation failure cause code. for details, see Section 9.5.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.12 Subscriber Policy TLV

The Subscriber Policy TLV is used to carry the policies that will be applied to a subscriber. It is carried in the Update_Request message.

The format of the TLV value part is as follows:

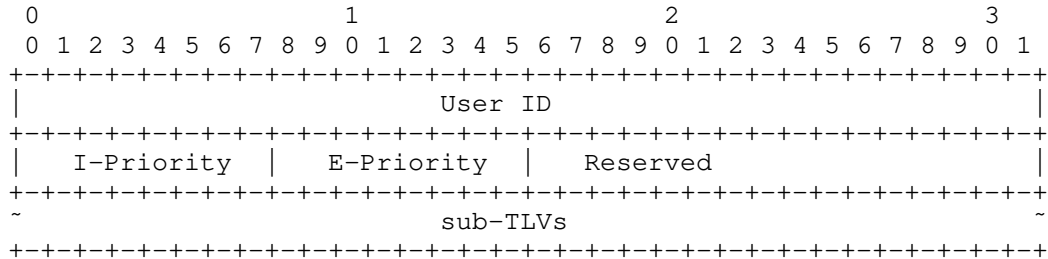


Figure 57: User QoS Policy Information TLV

Where:

The TLV type: 6.

The TLV length: variable.

User-ID (4 bytes): The identifier of a user/subscriber.

Ingress-Priority (1 byte): Indicates the upstream priority. The value range is 0~7.

Egress-Priority (1 byte): Indicates the downstream priority. The value range is 0~7.

sub-TLVs: The sub-TLVs that are present can occur in any order.

Ingress-CAR sub-TLV: Upstream CAR.

Egress-CAR sub-TLV: Downstream CAR.

Ingress-QoS-Profile sub-TLV: Indicates the name of the QoS-Profile profile in the upstream direction.

Egress-QoS-Profile Sub-TLV: Indicates the name of the QoS-Profile profile in the downstream direction.

User-ACL-Policy Sub-TLV: All ACL user policies, including v4ACLIN, v4ACLOUT, v6ACLIN, v6ACLOUT, v4WEBACL, v6WEBACL, v4SpecialACL, and v6SpecialACL.

Multicast-Profile4 Sub-TLV: IPv4 multicast policy name.

Multicast-Profile6 Sub-TLV: IPv6 multicast policy name.

NAT-Instance Sub-TLV: Indicates the instance ID of a NAT user.

The Reserved field MUST be sent as zero and ignored on receipt.

7.9.13 Subscriber CGN Port Range TLV

The Subscriber CGN Port Range TLV is used to carry the NAT public address and port range. It will be carried in the Update_Response message when CGN is used.

The format of this TLV is as follows:

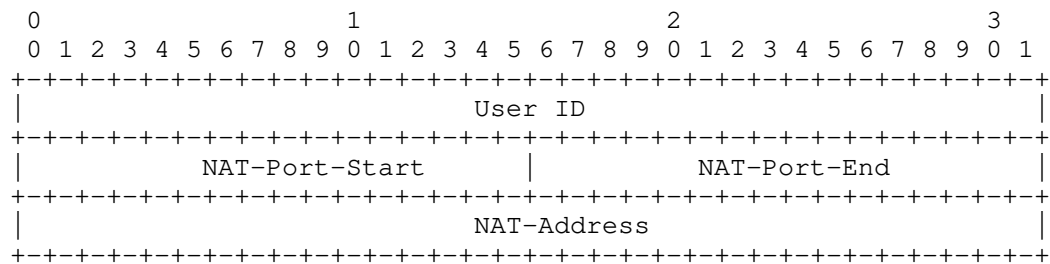


Figure 58: Subscriber CGN Port Range TLV

Where:

The TLV type: 15.

The TLV length: 12 octets.

User-ID (4 bytes): The identifier of a user/subscriber.

NAT-Port-Start (2 bytes): The start port number.

NAT-Port-End (2 bytes): The end port number.

NAT-Address (4 bytes): The NAT public network address.

7.10 Device Status TLVs

The TLVs in this section are for reporting Interface and Board level information from the UP to the CP.

7.10.1 Interface Status TLV

The Interface Status TLV is used to carry the status information of an interface on a UP. It is carried in a Report message.

The format of the value part of this TLV is as follows:

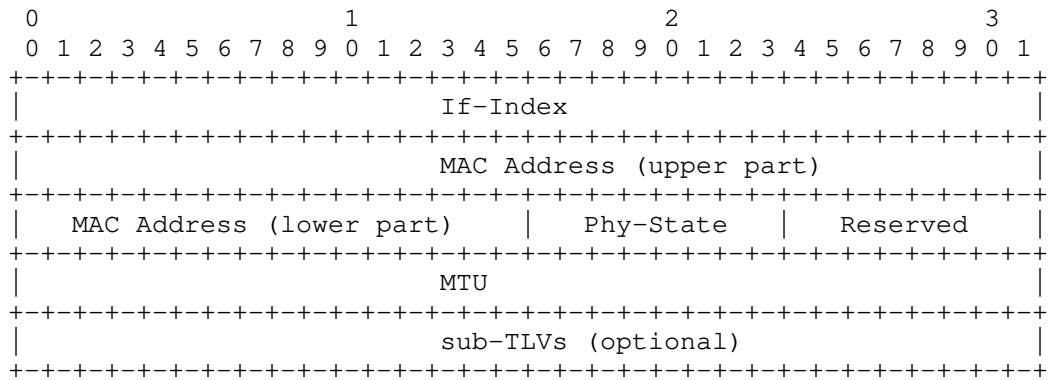


Figure 59: Interface Status TLV

Where:

The TLV type: 200.

The TLV length: variable.

If-Index (4 bytes): Indicates the interface index.

MAC-Address (MAC-Addr type): Interface MAC address.

Phy-State (1 byte): Physical status of the interface. 0: down, 1: Up

MTU (4 bytes): Interface MTU value.

sub-TLVs: The If-Desc and VRF-Name sub-TLVs can be carried.

The Reserved field MUST be sent as zero and ignored on receipt.

7.10.2 Board Status TLV

The Board Status TLV is used to carry the status information of a board on an UP. It is carried in a Report message.

The format of Board Status TLV is as follows:

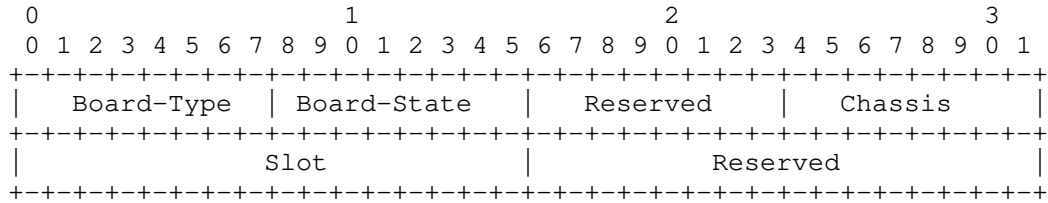


Figure 60: Interface Resource TLV

Where:

The TLV type: 201.

The TLV length: 8 octets.

Chassis (1 byte): The chassis number of the Board.

Slot (1 byte): The slot number of the Board.

Sub-Slot (1 byte): The sub-slot number of the Board.

Board-Type (1 byte):

1: CGN Service Process Unit (SPU) board.

2: Line Process Unit (LPU) Board.

Board-State (1 byte):

0: Normal.

1: Abnormal.

The reserved field MUST be sent as zero and ignored on receipt.

7.11 CGN TLVs

7.11.1 Address Allocation Request TLV

The Address Allocation Request TLV is used to request address allocation from CP. An address Pool-Name sub-TLV is carried to indicate from which address pool to allocate addresses. The Address Allocation Request TLV is carried in the Addr_Allocation_Req message.

The format of the value part of this TLV is as follows:

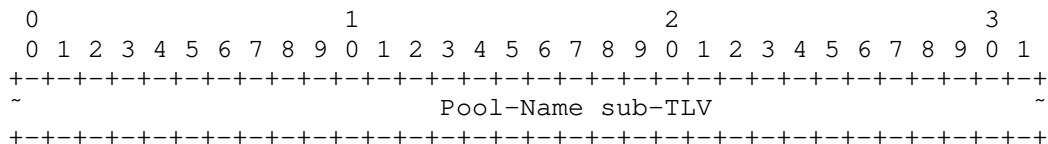


Figure 61: Address Allocation Request TLV

Where:

The TLV type: 400.

The TLV length: variable.

Pool-Name sub-TLV: Indicates from which Address pool to allocate address.

7.11.2 Address Allocation Response TLV

The Address Allocation Response TLV is used to return the address allocation result, it is carried the Addr_Allocation_Ack message.

The value part of the Address Allocation Response TLV is formatted as follows:

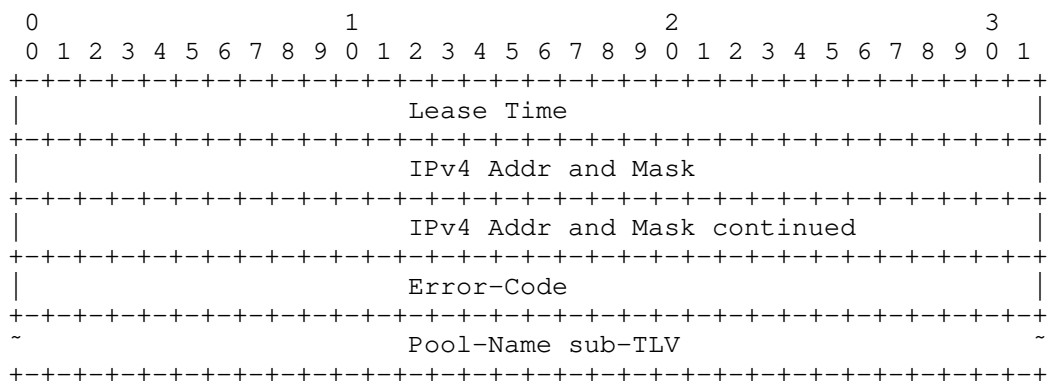


Figure 62: Address Assignment Response TLV

Where:

The TLV type: 401.

The TLV length: variable.

Lease Time (4 bytes): Duration of address lease in seconds.

IPv4 Addr and Mask (IPv4-Address type): The allocated IPv4 address.

Error-Code (4 bytes): Indicates success or an error.

0: Success.

1: Failure.

3001 (Pool-Mismatch): The corresponding address pool cannot be found.

3002 (Pool-Full): The address pool is fully allocated and no address segment is available.

Pool-Name sub-TLV: A Pool-Name sub-TLV to indicate from which Address pool the address is allocated.

7.11.3 Address Renewal Request TLV

The Address Renewal Request TLV is used to request address renewal from the CP. It is carried the Addr_Renew_Req message.

The format of this TLV value is as follows:

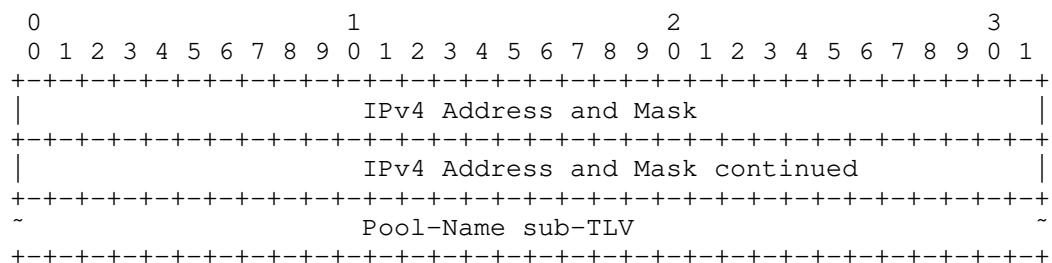


Figure 63: Address Renewal Request TLV

Where:

The TLV type: 402.

The TLV length: variable.

IPv4 Addr and Mask (IPv4-Addr): The IPv4 address to be renewed.

Pool Name sub-TLV: A Pool-Name sub-TLV to indicate from which

Address pool to renew the address.

7.11.4 The Address Renewal Response TLV

The Address Renewal Response TLV is used to return the address renewal result. It is carried in the Addr_Renew_Ack message.

The format of this TLV value is as follows:

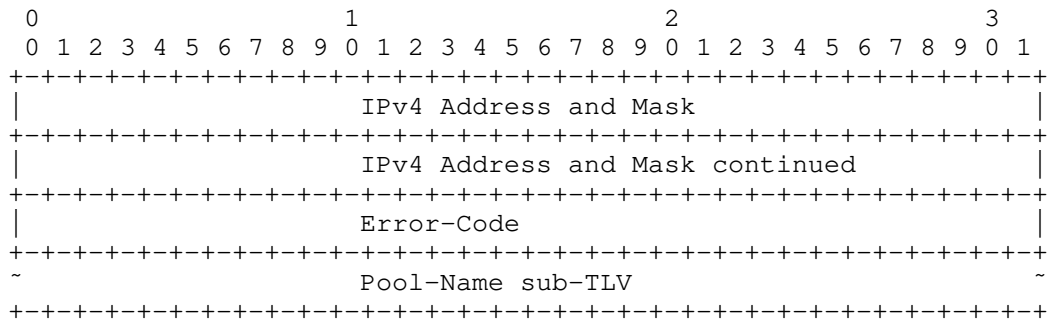


Figure 64: Address Renewal Response TLV

Where:

The TLV type: 403.

The TLV length: variable.

Client-IP (IPv4-Address type): The renewed IPv4 address.

Error Code (4 bytes): Indicates success or an error:

0: Renew success.

1: Renew failed.

3001 (Pool-Mismatch): The corresponding address pool cannot be found.

3002 (Pool-Full): The address pool is fully allocated and no address segment is available.

3003 (Subnet-Mismatch): The address pool subnet cannot be found.

3004 (Subnet-Conflict): Subnets in the address pool have been assigned to other clients.

Pool Name sub-TLV: A Pool-Name Sub-TLV to indicate from which Address pool to renew the address.

7.11.5 Address Release Request TLV

The Address Release Request TLV is used to release an IPv4 address. It is carried in the Addr_Release_Req message.

The value part of this TLV is formatted as follows:

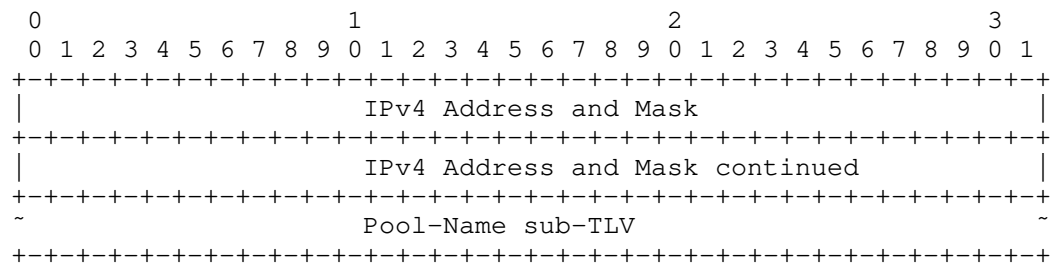


Figure 65: Address Release Request TLV

Where:

The TLV type: 404.

The TLV length: variable.

IPv4 Address and Mask (IPv4-Address type): The IPv4 address be released.

Pool-Name sub-TLV: A Pool-Name Sub-TLV to indicate from which Address pool to release the address.

7.11.6 The Address Release Response TLV

The Address Release Response TLV is used to return the address release result. It is carried in the Addr_Release_Ack message.

The format of this TLV is as follows:

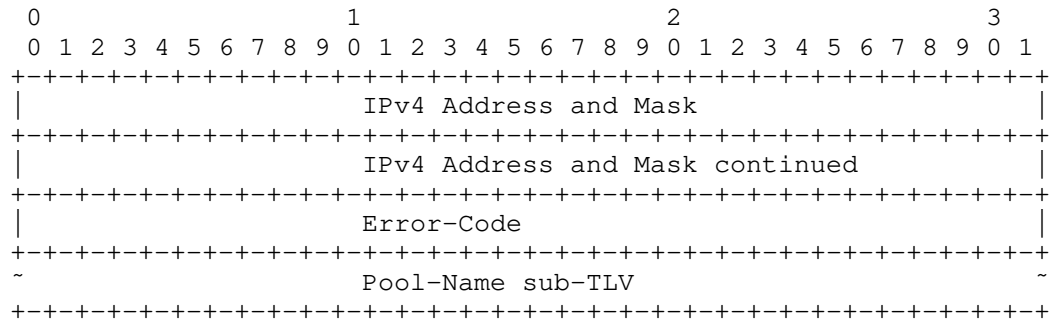


Figure 66: Address Renewal Response TLV

Where:

The TLV type: 405.

The TLV length: variable.

Client-IP (IPv4-Address type): The released IPv4 address.

Error-Code (4 bytes): Indicates success or an error.

0: Address release success.

1: Address release failed.

3001 (Pool-Mismatch): The corresponding address pool cannot be found.

3003 (Subnet-Mismatch): The address cannot be found.

3004 (Subnet-Conflict): The address has been allocated to another subscriber.

Pool-Name sub-TLV: A Pool-Name Sub-TLV to indicate from which address pool to release the address.

7.12 Event TLVs

7.12.1. Subscriber Traffic Statistics TLV

The Subscriber Traffic Statistics TLV is used to return the traffic statistics of a user/subscriber. The format of this TLV is as follows:

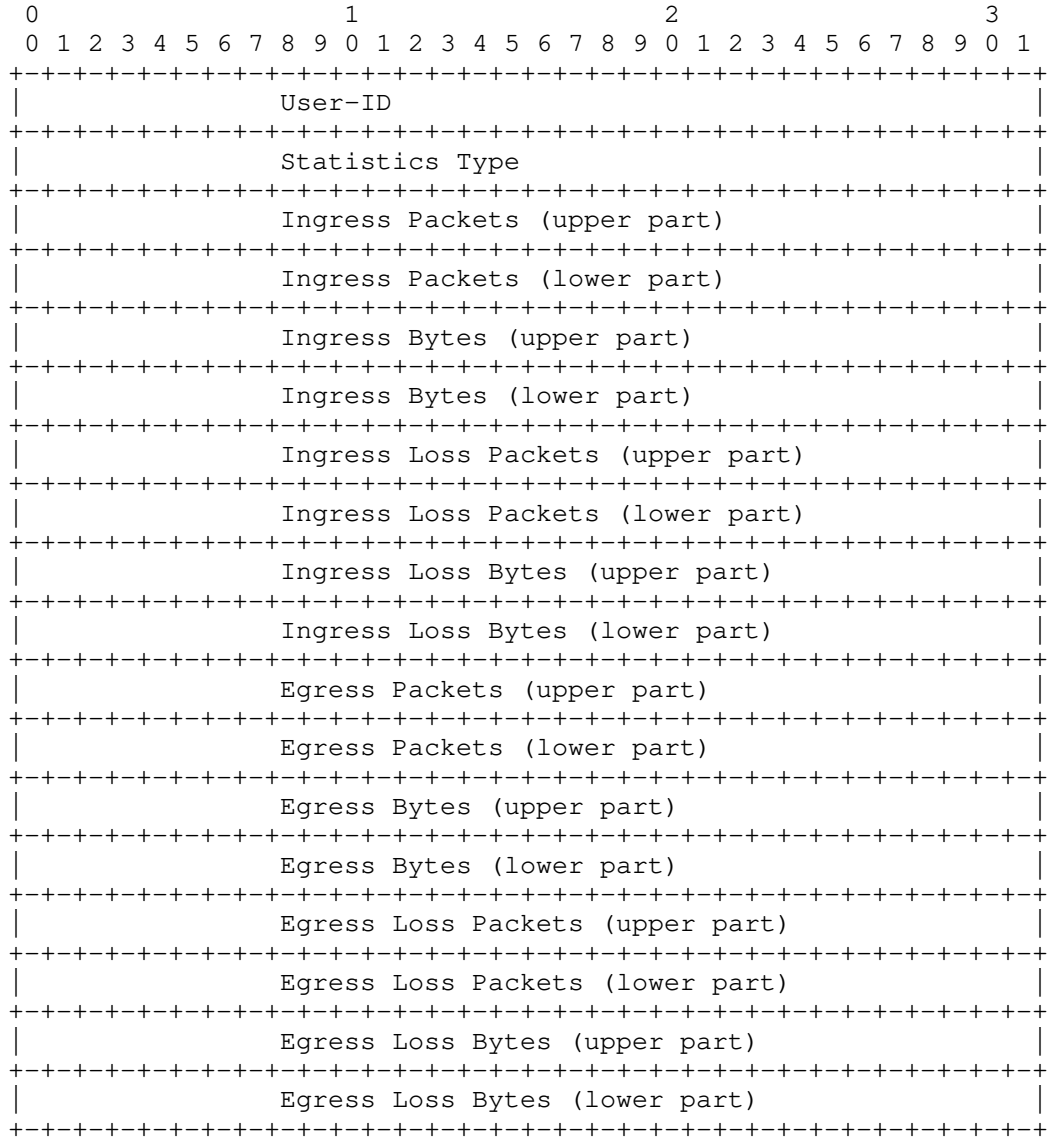


Figure 67: Subscriber Traffic Statistics TLV

Where:

The TLV type: 300.

The TLV length: 72 octets.

User-ID (4 bytes): The user/subscriber identifier.

Statistics-Type (4 bytes): Traffic type. It can be one of the following options:

- 0: IPv4 traffic.
- 1: IPv6 traffic.
- 2: Dual stack traffic.

Ingress Packets (8 bytes): The number of the packets in upstream direction.

Ingress Bytes (8 bytes): The bytes of the upstream traffic.

Ingress Loss Packets (8 bytes): The number of the lost packets in upstream direction.

Ingress Loss Bytes (8 bytes): The bytes of the lost upstream packets.

Egress Packets (8 bytes): The number of the packets in downstream direction.

Egress Bytes (8 bytes): The bytes of the downstream traffic.

Egress Loss Packets (8 bytes): The number of the lost packets in downstream direction.

Egress Loss Bytes (8 bytes): The bytes of the lost downstream packets.

7.12.2 Subscriber Detection Result TLV

The Subscriber Detection Result TLV is used to return the detection result of a subscriber. Subscriber detection is a function to detect whether a subscriber is online or not. The result can be used by the CP to determine how to deal with the subscriber session. (e.g., delete the session if detection failed).

The format of this TLV value part is as follows:

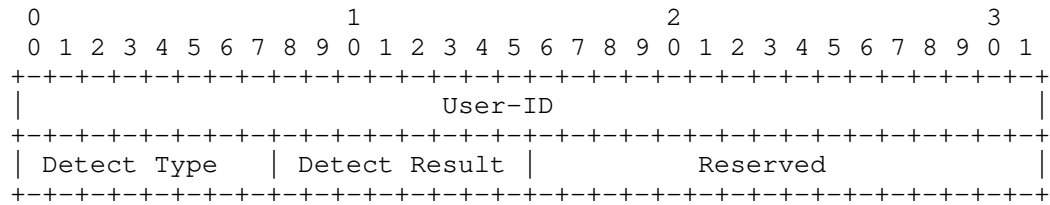


Figure 68: Subscriber Detection Result TLV

Where:

The TLV type: 301.

The TLV length: 8 octets.

User-ID (4 bytes): A user/subscriber identifier.

Detect-Type (1 byte):

0: IPv4 detection.

1: IPv6 detection.

2: PPP detection.

Detect-Result (1 bytes):

0: Indicates that the detection is successful.

1: Detection failure. The UP needs report only when the detection fails.

The Reserved field MUST be sent as zero and ignored on receipt.

7.13 Vendor TLV

The Vendor ID TLV occurs as the first TLV in the Vendor message (Section 6.6). It provides a Sub-Type that effectively extends the message type in the message header, provides for versioning of vendor TLVs, and can accommodate sub-TLVs.

The value part of the Vendor TLV is formatted as follows:

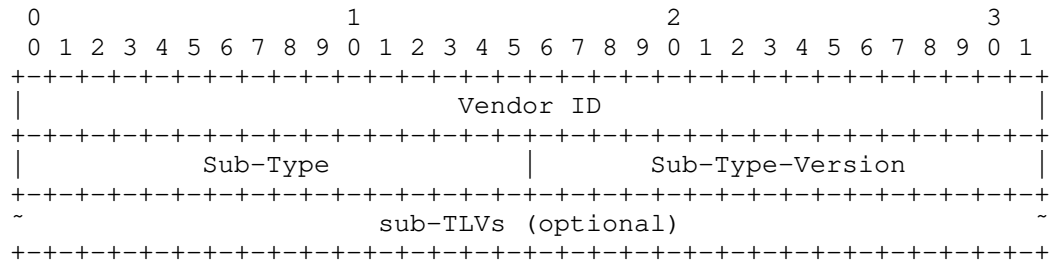


Figure 69: Vendor TLV

Where:

The TLV type: 1024.

The TLV length: variable.

Vendor-ID (4 bytes): Vendor ID as defined in RADIUS [RFC2865].

Sub-Type (2 bytes): Used by the Vendor to distinguish multiple different vendor messages.

Sub-Type-Version (2 bytes): Used by the Vendor to distinguish different versions of a Vendor Defined message sub-type.

Sub-TLVs (variable): Sub-TLVs as specified by the vendor.

Since Vendor code will be handling the TLV after the Vendor ID field is recognized, the remainder of the TLV value can be organized however the vendor wants. But it is desirable for a vendor to be able to define multiple different vendor messages and to keep track of different versions of its vendor defined messages. Thus, it is RECOMMENDED that the vendor assign a Sub-Type value for each vendor message that it defines different from other Sub-Type values that vendor has used. Also, when modifying a vendor defined message in a way potentially incompatible with a previous definition, the vendor SHOULD increase the value it is using in the Sub-Type-Version field.

8. Implementation Status

RFC Editor: Please remove this section before publication.

This section discusses the status of implementations that have provided information and the testing of this protocol at the time of posting of this Internet-Draft, and is based on the proposal in [RFC7942] ("Improving Awareness of Running Code: The Implementation Status Section"). The description of implementations in this section is intended to assist in processing drafts to RFCs.

Please note that the listing of any individual implementation or test results here does not imply endorsement by the RFC Editor or the IETF. Furthermore, no effort has been spent to verify the information presented here that was supplied by contributors. This is not intended as, and must not be construed to be, a catalog of available implementations or their testing or features. Readers are advised to note that other implementations may exist.

According to [RFC7942], "this will allow reviewers ... to assign due consideration to documents that have the benefit of running code, which may serve as evidence of valuable experimentation and feedback that have made the implemented protocols more mature."

8.1 Implementations

Information on three S-CUSP implementations appears below.

8.1.1 Huawei Technologies

Name: Cloud-based BNG.

Maturity: Production.

Coverage: According to S-CUSP protocol.

Contact information:

Zhouyi Yu: yuzhouyi@huawei.com

Date: 2018-11-01

8.1.2 ZTE

Name: ZXR10 V6000 vBRAS

Maturity: Production

Coverage: According to S-CUSP protocol.

Contact information:

Yong Chen: 10056167@zte.com.cn

Huaibin Wang: 10008729@zte.com.cn

Date: 2018-12-01

8.1.3 H3C

Name: CUSP protocol for BRAS Control Plane and User Plan Separation

Maturity: Research

Coverage: According to S-CUSP protocol

Contact information: mengdan@h3c.com; liuhanlei@h3c.com

Date: 2019-1-30

8.2 Hackathon

Successful use of the protocol at the IETF-102 Hackathon, Montreal, Quebec, in 2018.

Hackathon Project: Control Plane and User Plane Separation BNG control channel Protocol (CUSP)

Champions: Zhenqiang Li, Michael Wang

Report: See

github.com/IETF-Hackathon/ietf102-project-presentations/blob/master/IETF102-hackathon-presentation-CUSP.pptx

8.3 EANTC Testing

EANTC (European Advanced Networking Test Center (www.eantc.de)) tested the Huawei implementation. Their summary was as follows:
"EANTC tested advanced aspects of the Cloud-based Broadband Network Gateway (vBNG) with a focus on performance, scalability and high availability with up to 20 Million emulated subscribers. The solution performed very well across all test scenarios."

See report at
www.eantc.de/fileadmin/eantc/downloads/News/2018/EANTC-vBRAS-phase2.pdf

9. IANA Considerations

IANA is requested to create an "S-CUSP Parameters" web page and include thereon the registries set up in the Sub-Sections below.

9.1 Message Types

IANA is requested to create an S-CUSP Message Types registry on the S-CUSP Parameters Web Page as follows:

Registry Name: S-CUSP Message Types
 Registration Procedure: Expert Review
 Reference: [this document]

Type	Name	Section of [this document]
0	- Reserved	
1	Hello	6.2.1.
2	Keepalive	6.2.2.
3	Sync_Request	6.2.3.
4	Sync_Begin	6.2.4.
5	Sync_Data	6.2.5.
6	Sync_End	6.2.6.
7	Update_Request	6.2.7.
8	Update_Response	6.2.8.
9	Report	6.4.
10	Event	6.3.
11	Vendor	6.6.
12	Error	6.7.
13-199	- Unassigned	
200	Addr_Allocation_Req	6.5.1.
201	Addr_Allocation_Ack	6.5.2.
202	Addr_Renew_Req	6.5.3.
203	Addr_Renew_Ack	6.5.4.
204	Addr_Release_Req	6.5.5.
205	Addr_Release_Ack	6.5.6.
206-254	- Unassigned	
255	- Reserved	

9.2 TLV Types

IANA is requested to create an S-CUSP TLV Types registry on the S-CUSP Parameters Web Page as follows:

Registry Name: S-CUSP TLV Types
 Registration Procedure: Expert Review
 Reference: [this document]

Type	Name	Usage Description
0	reserved	-
1	BAS Function	Carries the BNG access functions to be enabled or disabled on specified interfaces.
2	Basic Subscriber	Carries the basic information about a BNG subscriber.
3	PPP Subscriber	Carries the PPP information about a BNG subscriber.
4	IPv4 Subscriber	Carries the IPv4 address of a BNG subscriber.
5	IPv6 Subscriber	Carries the IPv6 address of a BNG subscriber.
6	Subscriber Policy	Carries the policy information applied to a BNG subscriber.
7	IPv4 Routing	Carries the IPv4 network routing information.
8	IPv6 Routing	Carries the IPv6 network routing information.
9	IPv4 Static Subscriber Detect	Carries the IPv4 static subscriber detect information.
10	IPv6 Static Subscriber Detect	Carries the IPv6 static subscriber detect information.
11	L2TP-LAC Subscriber	Carries the L2TP LAC subscriber information.
12	L2TP-LNS Subscriber	Carries the L2TP LNS subscriber information.
13	L2TP-LAC-Tunnel	Carries the L2TP LAC tunnel subscriber information.
14	L2TP-LNS-Tunnel	Carries the L2TP LNS tunnel subscriber information.
15	Subscriber CGN Port Range	Carries the public IPv4 address and related port range of a BNG subscriber.
16-99	unassigned	-
100	Hello	Used for version and Keep Alive timers negotiation.
101	Error Information	Carried in the Ack of the control message. Carries the processing result, success, or error.
102	Keep Alive	Carried in the Hello message for Keep Alive timers negotiation.

103-199	unassigned	-
200	Interface Status	Interfaces status reported by the UP including physical interfaces, sub-interfaces, trunk interfaces, and tunnel interfaces.
201	Board Status	Board information reported by the UP including the board type and in-position status.
202-299	unassigned	-
300	Subscriber Traffic Statistics	User traffic statistics.
301	Subscriber Detection Results	User detection information.
302	Update Response	The processing result of a subscriber session update.
303-299	unassigned	-
400	Address Allocation Request	Request address allocation.
401	Address Allocation Response	Address allocation response.
402	Address Renewal Request	Request for address lease renewal.
403	Address Renewal Response	Response to a request for extending an IP address lease.
404	Address Release Request	Request to release the address.
405	Address Release Response	Ack of a message releasing an IP address.
406-1023	unassigned	-
1024	Vendor	As implemented by vendor.
1039-4095	unassigned	-

9.3 TLV Operation Codes

IANA is requested to create an S-CUSP TLV Operation Codes registry on the S-CUSP Parameters Web Page as follows:

Registry Name: S-CUSP TLV Operation Codes
 Registration Procedure: Expert Review
 Reference: [this document]

Code	Operation
----	-----
0	- reserved
1	Update
2	Delete
3-15	- unassigned

9.4 Sub-TLV Types

IANA is requested to create an S-CUSP Sub-TLV Types registry on the S-CUSP Parameters Web Page as follows:

Registry Name: S-CUSP Sub-TLV Types
 Registration Procedure: Expert Review
 Reference: [this document]

Type	Name	Section of [this document]
0	- reserved	
1	VRF Name	7.3.1.
2	Ingress-QoS-Profile	7.3.1.
3	Egress-QoS-Profile	7.3.1.
4	User-ACL-Policy	7.3.1.
5	Multicast-ProfileV4	7.3.1.
6	Multicast-ProfileV6	7.3.1.
7	Ingress-CAR	7.3.2.
8	Egress-CAR	7.3.3.
9	NAT-Instance	7.3.1.
10	Pool-Name	7.3.1.
11	If-Desc	7.3.4.
12	IPv6-Address List	7.3.5.
13	Vendor	7.3.6.
12-64534	- unassigned	
65535	- reserved	

9.5 Error Codes

IANA is requested to create an S-CUSP ERRID Codes registry on the S-CUSP Parameters Web Page as follows:

Registry Name: S-CUSP ERRID Codes
 Registration Procedure: Expert Review
 Reference: [this document]

Value	Name	Remarks
0	Success	Success
1	Fail	Malformed message received.
2	TLV-Unknown	One or more of the TLVs was not understood.
3	TLV-Length	The TLV length is abnormal.
4-999	- unassigned	Unassigned basic error codes.
1000	- reserved	
1001	Version-Mismatch	The version negotiation fails. Terminate.

		The subsequent service processes corresponding to the UP are suspended.
1002	Keepalive Error	The keepalive negotiation fails.
1003	Timer Expires	The establishment timer expired.
1004-1999	- unassigned	Unassigned error codes for version negotiation.
2000	- reserved	
2001	Synch-NoReady	The data to be smoothed is not ready.
2002	Synch-Unsupport	The request for smooth data is not supported.
2003-2999	- unassigned	Unassigned data synchronization error codes.
3000	- reserved	
3001	Pool-Mismatch	The corresponding address pool cannot be found.
3002	Pool-Full	The address pool is fully allocated and no address segment is available.
3003	Subnet-Mismatch	The address pool subnet cannot be found.
3004	Subnet-Conflict	Subnets in the address pool have been classified into other clients.
3005-3999	- unassigned	Unassigned error codes for address allocation.
4000	- reserved	
4001	Update-Fail-No-Res	The forwarding table fails to be delivered because the forwarding resources are insufficient.
4002	QoS-Update-Success	The QoS policy takes effect.
4003	QoS-Update-Sq-Fail	Failed to process the queue in the QoS policy.
4004	QoS-Update-CAR-Fail	Processing of the CAR in the QoS policy fails.
4005	Statistic-Fail-No-Res	Statistics processing failed due to insufficient statistics resources.
4006-4999	- unassigned	forwarding table delivery error codes.
5000-4294967295	- reserved	

10. Security Considerations

The Service, Control, and Management Interfaces between the CP and UP might be across the general Internet or other hostile environment. The ability of an adversary to block or corrupt messages or introduce spurious messages on any one or more of these interfaces would give the adversary the ability to stop subscribers from accessing network services, disrupt existing subscriber sessions, divert traffic, mess up accounting statistics, and generally cause havoc. Damage would not necessarily be limited to one or a few subscribers but could disrupt routing or deny service to one or more instances of the CP or otherwise cause extensive interference. If the adversary knows the details of the UP equipment and its forwarding rule capabilities, the adversary may be able to cause a copy of most or all user data to be sent to an address of the adversary's choosing, thus enabling eavesdropping.

Thus, appropriate protections **MUST** be implemented to provide integrity, authenticity, and secrecy of traffic over those interfaces. Whether such protection is used is a network operator decision. See [RFC6241] for Management Interface / NETCONF security. Security on the Service Interface is dependent on the tunneling protocol used which is out of scope for this document. Security for the Control Interface, over which the S-CUSP protocol flows, is further discussed below.

S-CUSP messages do not provide security. Thus, if these messages are exchanged in an environment where security is a concern, that security **MUST** be provided by another protocol such as TLS 1.3 [RFC8446] or IPSEC. TLS 1.3 is the mandatory to implement protocol for interoperability. The use of a particular security protocol on the Control Interface is determined by configuration. Such security protocols need not always be used and lesser security precautions might be appropriate because, in some cases, the communication between the CP and UP is in a benign environment.

Contributors

Zhouyi Yu
Huawei Technologies

Email: yuzhouyi@huawei.com

Chengguang Niu
Huawei Technologies

Email: chengguang.niu@huawei.com

Zitao Wang
Huawei Technologies

Email: wangzitao@huawei.com

Jun Song
Huawei Technologies

Email: song.jun@huawei.com

Dan Meng
H3C Technologies
No.1 Lixing Center
No.8 guangxun south street, Chaoyang District,
Beijing, 100102
China

Email: mengdan@h3c.com

Hanlei Liu
H3C Technologies
No.1 Lixing Center
No.8 guangxun south street, Chaoyang District,
Beijing, 100102
China

Email: hanlei_liu@h3c.com

Normative References

- [RFC20] Cerf, V., "ASCII format for network interchange", STD 80, RFC 20, DOI 10.17487/RFC0020, October 1969, <<https://www.rfc-editor.org/info/rfc20>>.
- [RFC793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <<https://www.rfc-editor.org/info/rfc793>>.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC2661] Townsley, W., Valencia, A., Rubens, A., Pall, G., Zorn, G., and B. Palter, "Layer Two Tunneling Protocol "L2TP"", RFC 2661, DOI 10.17487/RFC2661, August 1999, <<https://www.rfc-editor.org/info/rfc2661>>.
- [RFC2865] Rigney, C., Willens, S., Rubens, A., and W. Simpson, "Remote Authentication Dial In User Service (RADIUS)", RFC 2865, DOI 10.17487/RFC2865, June 2000, <<https://www.rfc-editor.org/info/rfc2865>>.
- [RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <<https://www.rfc-editor.org/info/rfc6241>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

Informative References

- [802.1Q] IEEE, "IEEE Standard for Local and metropolitan area networks / Bridges and Bridged Networks", IEEE Std 802.1Q-2014, 3 November 2013.
- [RFC1661] Simpson, W., Ed., "The Point-to-Point Protocol (PPP)", STD 51, RFC 1661, DOI 10.17487/RFC1661, July 1994, <<https://www.rfc-editor.org/info/rfc1661>>.
- [RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <<https://www.rfc-editor.org/info/rfc2131>>.
- [RFC2516] Mamakos, L., Lidl, K., Evarts, J., Carrel, D., Simone, D., and R. Wheeler, "A Method for Transmitting PPP Over Ethernet (PPPoE)", RFC 2516, DOI 10.17487/RFC2516, February 1999, <<https://www.rfc-editor.org/info/rfc2516>>.
- [RFC2698] Heinanen, J. and R. Guerin, "A TwoRate Three Color Marker", RFC 2698, DOI 10.17487/RFC2698, September 1999, <<https://www.rfc-editor.org/info/rfc2698>>.
- [RFC3022] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, DOI 10.17487/RFC3022, January 2001, <<https://www.rfc-editor.org/info/rfc3022>>.
- [RFC3336] Thompson, B., Koren, T., and B. Buffam, "PPP Over Asynchronous Transfer Mode Adaptation Layer 2 (AAL2)", RFC 3336, DOI 10.17487/RFC3336, December 2002, <<https://www.rfc-editor.org/info/rfc3336>>.
- [RFC5511] Farrel, A., "Routing Backus-Naur Form (RBNF): A Syntax Used to Form Encoding Rules in Various Routing Protocol Specifications", RFC 5511, DOI 10.17487/RFC5511, April 2009, <<https://www.rfc-editor.org/info/rfc5511>>.
- [RFC7042] Eastlake 3rd, D. and J. Abley, "IANA Considerations and IETF Protocol and Documentation Usage for IEEE 802 Parameters", BCP 141, RFC 7042, DOI 10.17487/RFC7042, October 2013, <<https://www.rfc-editor.org/info/rfc7042>>.
- [RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <<https://www.rfc-editor.org/info/rfc7348>>.

- [RFC7942] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", BCP 205, RFC 7942, DOI 10.17487/RFC7942, July 2016, <<https://www.rfc-editor.org/info/rfc7942>>.
- [RFC8446] Rescorla, E., "The Transport Layer Security (TLS) Protocol Version 1.3", RFC 8446, DOI 10.17487/RFC8446, August 2018, <<https://www.rfc-editor.org/info/rfc8446>>.
- [TR-384] Broadband Forum, "Cloud Central Office Reference Architectural Framework", BBF TR-384, 2018.

Authors' Addresses

Shujun Hu
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: hushujun@chinamobile.com

Donald Eastlake, 3rd
Futurewei Technologies
1424 Pro Shop Court
Davenport, FL 33896
USA

Phone: +1-508-333-2270
Email: d3e3e3@gmail.com

Mach (Guoyi) Chen
Huawei Technologies
Huawei Bldg., No. 156 Beiqing Road
Beijing 100095 China

Email: mach.chen@huawei.com

Fengwei Qin
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: qinfengwei@chinamobile.com

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: lizhenqiang@chinamobile.com

Tee Mong Chua
Singapore Telecommunications Limited
31 Exeter Road, #05-04 Comcentre Podium Block
Singapore City 239732
Singapore

Email: teemong@singtel.com

Daniel Huang
ZTE

Email: huang.guangping@zte.com.cn

Copyright, Disclaimer, and Additional IPR Provisions

Copyright (c) 2019 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.

rtgwg
Internet-Draft
Intended status: Informational
Expires: April 25, 2019

S. Hu
China Mobile
V. Lopez
Telefonica
F. Qin
Z. Li
China Mobile
T. Chua
Singapore Telecommunications Limited
Donald. Eastlake
M. Wang
J. Song
Huawei
October 22, 2018

Requirements for Control Plane and User Plane Separated BNG Protocol
draft-cuspd-rtgwg-cusp-requirements-03

Abstract

This document introduces the Control Plane and User Plane separated BNG (Broadband Network Gateway) architecture and defines a set of associated terminology. It also specifies a set of protocol requirements for communication between the BNG-CP and the BNG-UPs in the Control Plane and User Plane Separated BNG.

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 <https://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 April 25, 2019.

Copyright Notice

Copyright (c) 2018 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 (<https://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	2
2. Concept and Terminology	3
2.1. Terminology	3
3. CU Separated BNG Model	3
3.1. Internal interfaces between the CP and UP	5
4. The usage of CU separation BNG protocol	6
5. Control Plane and User Plane Separation Protocol Requirements	7
5.1. Transmit information tables	7
5.2. Message Priority	7
5.3. Reliability	7
5.4. Support for Secure Communication	8
5.5. Version negotiation	8
5.6. Capability Exchange	9
5.7. CP primary/backup capability	9
5.8. Event Notification	9
5.9. Query Statistics	10
6. Security Considerations	10
7. IANA Considerations	10
8. References	10
8.1. Normative References	10
8.2. Informative References	11
Authors' Addresses	11

1. Introduction

A Broadband Network Gateway (BNG) is an Ethernet-centric IP edge router and the aggregation point for user traffic. To provide centralized session management, flexible address allocation, high scalability for subscriber management capacity, and cost-efficient redundancy, the CU separated BNG is introduced [TR-384]. The CU separated Service Control Plane could be virtualized and centralized;

it is responsible for user access authentication and sending forwarding entries to user planes. The routing control and forwarding plane, i.e. BNG user plane (local), could be distributed across the infrastructure.

This document introduces the Control Plane and User Plane separated BNG architecture and modeling. This document also defines the protocol requirements for Control Plane and User Plane Separated BNG (CUSP).

2. Concept and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2.1. Terminology

BNG: Broadband Network Gateway. A broadband remote access server (BRAS, B-RAS or BBRAS) that routes traffic to and from broadband remote access devices such as digital subscriber line access multiplexers (DSLAM) on an Internet service provider's (ISP) network. BRAS can also be referred to as a Broadband Network Gateway (BNG).

CP: Control Plane. The CP is a user control management component which manages UP's resources such as the user entry and user's QoS policy.

CUSP: Control Plane and User Plane Separated BNG Protocol.

UP: User Plane. UP is a network edge and user policy implementation component. The traditional router's Control Plane and forwarding plane are both preserved on BNG devices in the form of a user plane.

3. CU Separated BNG Model

Figure 1 shows the architecture of CU separated BNG

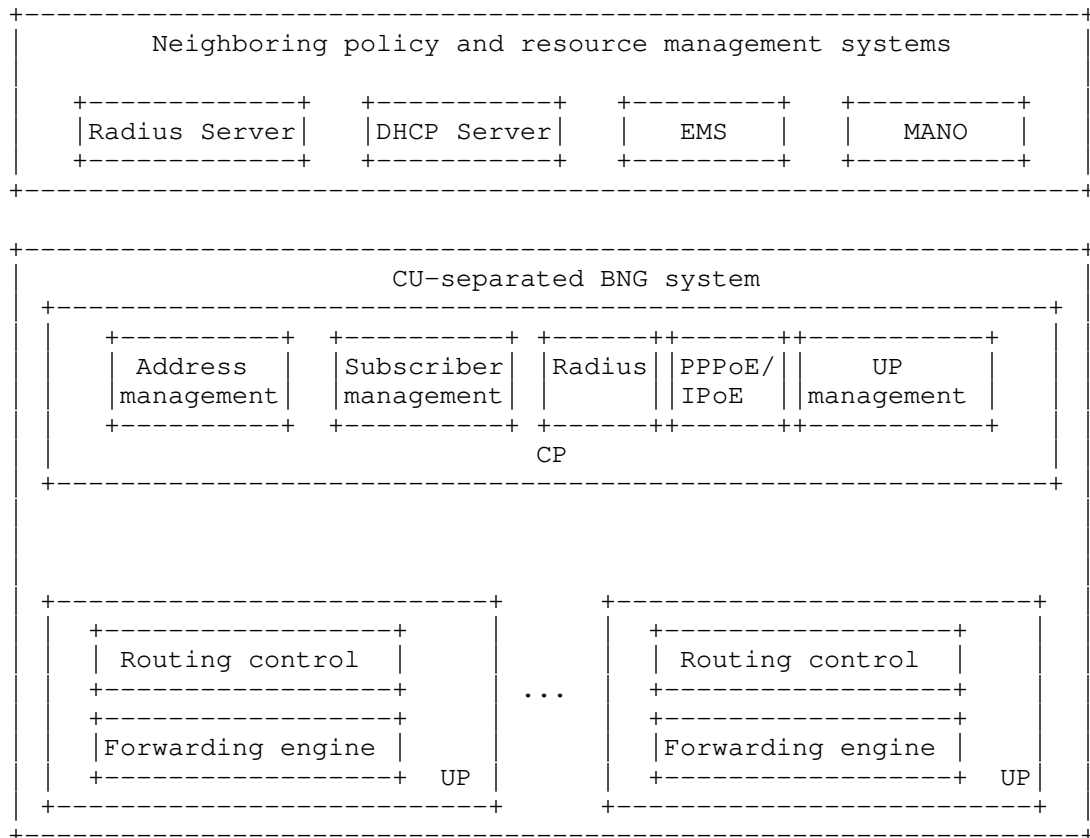


Figure 1. Architecture of CU Separated BNG

Briefly, a CU separated BNG is made up of a Control Plane (CP) and a set of User Planes (UPs) [TR-384], [I-D.cuspd-t-rtgwg-cu-separation-bng-deployment]. The Control Plane is a user control management component which manages UP's resources such as the user entry and user's Quality of Service (QoS) policy, for example, the access bandwidth and priority management. This Control Plane could be virtualized and centralized. The functional modules inside the BNG Service Control Plane can be implemented as Virtual Network Functions (VNFs) and hosted in a Network Function Virtualization Infrastructure (NFVI). The User Plane Management module in the BNG control plane centrally manages the distributed BNG user planes (e.g. load balancing), as well as the setup, deletion, update, and maintenance of channels between control planes and user planes [TR-384], [I-D.cuspd-t-rtgwg-cu-separation-bng-deployment]. The User Plane (UP) is a network edge and user policy implementation component. It can support the forwarding plane functions on traditional BNG devices,

such as traffic forwarding, QoS, and traffic statistics collection, and it can also support the control plane functions on traditional BNG devices, such as routing, multicast, etc [TR-384], [I-D.cuspdrt-gwg-cu-separation-bng-deployment].

3.1. Internal interfaces between the CP and UP

To support communication between the Control Plane and User Plane, several interfaces are involved. Figure 2 illustrates the three internal interfaces of CU Separated BNG.

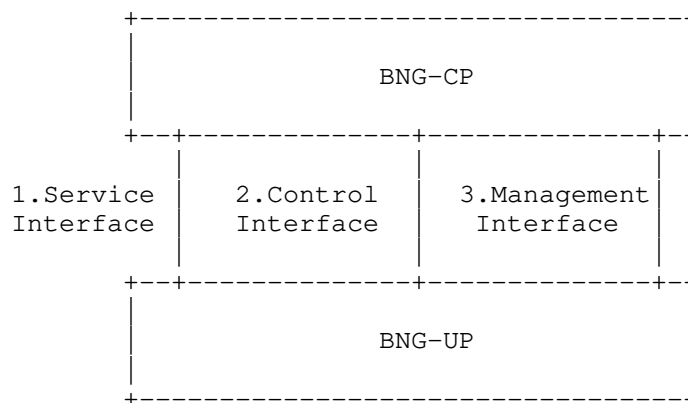


Figure 2. Interfaces between the BNG-CP and the BNG-UP

Service interface: The CP and UP use this interface to establish VXLAN tunnels with each other and transmit PPPoE and IPoE packets over the VXLAN tunnels.

Control interface: The CP uses this interface to deliver service entries, and the UP uses this interface to report service events to the CP.

Management interface: The CP uses this interface to deliver configurations to the UP. This interface uses NETCONF.

The CUSP (Control plane and User plane Separated BNG protocol) defines the control interface, and specifies the communication between the centralized control plane and user planes. This protocol should be designed to support establishing and maintaining a conversation between CP and UPs, and transporting the tables that are specified in [draft-cuspdrt-gwg-cu-separation-infor-model].

4. The usage of CU separation BNG protocol

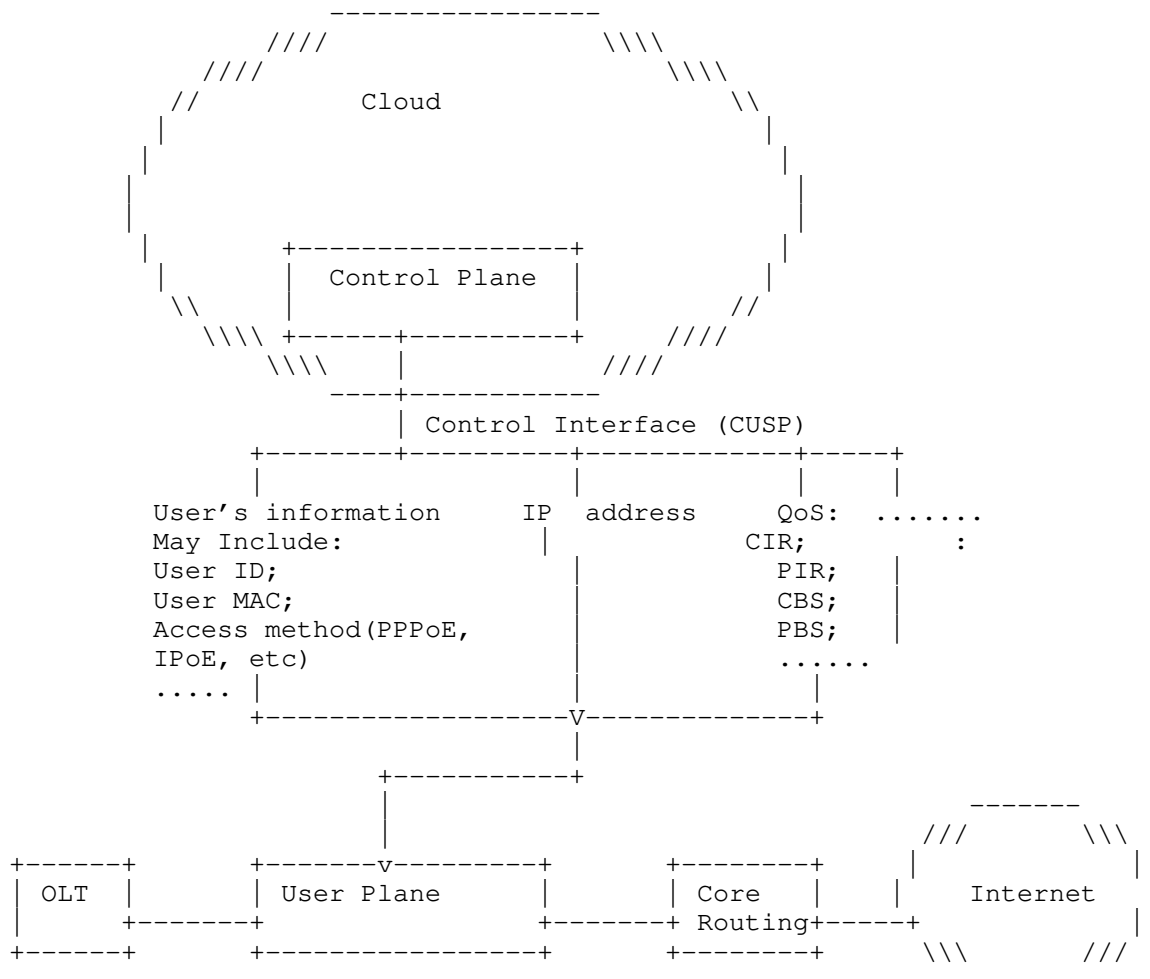


Figure 3. CU Separation BNG protocol usage

As shown in Figure 3, when users access the BNG network, the control plane solicits user information (such as user's ID, user's MAC, user's access methods, for example via PPPoE/IPoE), associates users with available bandwidth which is reported by User planes, and, based on the service's requirement, generates a set of tables, which may include user's information, UP's IP segment, and QoS, etc. Then the control plane can transmit these tables to the User planes. User

planes receive these tables, parse them, and then perform corresponding actions.

5. Control Plane and User Plane Separation Protocol Requirements

This section specifies the requirements for the CU separation protocol.

5.1. Transmit information tables

The Control Plane and User Plane Separation Protocol **MUST** allow the CP to send tables to each User Plane device.

a) The current BNG service requires that the UP should support at least 2000 users being accessed every second. And every user requires at least 2000 bytes. To achieve high performance, the CU Separation protocol **SHOULD** be lightweight.

b) CU separation protocol should support data encoded as either XML or binary. It allows user information data to be read, saved, and manipulated with tools specific to XML or binary.

c) In order to provide centralized session management, high scalability for subscriber management capacity, and cost-efficient redundancy, batching ability should be provided. The CU Separation protocol should be able to group an ordered set of commands to a UP device. Each such group of commands **SHOULD** be sent to the UP in as few messages as practical. Furthermore, the protocol **MUST** support the ability to specify if a command group **MUST** have all-or-nothing semantics.

d) The CU Separation protocol **SHOULD** be able to support at least hundreds of UP devices and tens of thousands of ports. For example, the protocol field sizes corresponding to UP or port numbers **SHALL** be large enough to support the minimum required numbers. This requirement does not relate to the performance of the system as the number of UPs or ports in the system grows.

5.2. Message Priority

The CU Separation protocol **MUST** provide a means to express the protocol message priorities.

5.3. Reliability

Heartbeat is a periodic signal generated by hardware or software to test for some aspects of normal operation or to synchronize other parts of network system.

In the CU separation BNG, a heartbeat is sent between CP and UPs at a regular interval on the order of seconds. If the CP/UP does not receive a heartbeat for a time--usually a few heartbeat intervals--the CP/UP that should have sent the heartbeat is assumed to have failed.

The CU separation protocol should support some kind of heartbeat monitoring mechanism. And this mechanism should have ability to distinguish whether the interruption is an actual failure. For example, in some scenarios (i.e. CP/UP update, etc), the connection between the UP and CP need to be interrupted. In this case, the interruption should not be reported.

5.4. Support for Secure Communication

As mentioned above, CP may send some information tables to the UP which may be critical to the network function (e.g, User Information, IPv4/IPv6 information) and may reflect the business information (e.g, QoS, service level agreements, etc). Therefore, supporting the integrity of all CU Separation protocol messages and protecting against man-in-the-middle attacks MUST be supported.

The CP Separation protocol should support security in a variety of scenarios. For example, the connections between the CP and UPs could be dedicated lines, VPNs within one domain, or could cross several domains, that is, cross third party networks. Thus it is likely that more than one security mechanism SHOULD be supported. TLS and IPsec are good candidates for such mechanisms.

5.5. Version negotiation

The CU separated BNG may consist of different vendors' devices implementing different versions of protocol. Therefore, the CU separation protocol MUST provide some mechanisms to perform the version negotiation.

Version negotiation is the process that the CU separated BNG's Control-Plane uses to evaluate the protocol versions supported by both the control-plane and the user-plane devices. Then a suitable protocol version is selected for communication in CUSP. The process is a "negotiation" because it requires identifying the most recent protocol version that is supported by both the control-plane and the user-plane devices or determining that they have no version in common.

5.6. Capability Exchange

The UP Capability Report displays the device's profile, service capability, and other assigned capabilities within the CU separated BNG. The CU separation protocol should MUST provide some mechanism to exchange the UP device's capabilities.

5.7. CP primary/backup capability

A backup CP for failure recovery is required for the CU separated BNG network. And the CUSP should provide some mechanism to implement the backup CP:

- a) In some scenarios, there may be two CP devices both declaring the primary CP. Thus the CUSP should support or associate with some mechanisms to determine which CP is the primary device.
- b) In the scenario of the primary CP down, the CUSP should support switching between primary and backup CP.

5.8. Event Notification

The CUSP protocol SHOULD be able to asynchronously notify the CP of events on the UP such as failures and changes in available resources and capabilities. Some scenarios that may initiate event notifications are listed below.

- a) Sending response message: As mentioned above, the control plane solicits users' information, associates them with available bandwidth, and generates a set of tables based on the service's requirement. Then the control plane transmits these tables to the corresponding User plane. The UP should respond with an event notification to inform the CP that the tables are received.
- b) User trace: The user trace mechanism can support the Control Plane tracing and monitoring the network status for users (for example the real-time bandwidth, etc), to help debug the user's application. Therefore, the UPs SHOULD be able to notify the CP with the User trace message.
- c) Sending statistics parameters: In CU separation BNG, the User-plane will report the traffic statistics parameters to the Control-plane, such as the ingress packets, ingress bytes, egress packets, egress bytes, etc. These parameters can help measure the BNG network performance. Available network resources can be allocated basing on the statistics parameters by the BNG-CP. Therefore, the UPs SHOULD be able to notify the CP with statistics parameters.

d) Report the result of User Detect: "User Detect" message will be send periodically to detect user dial-up and disconnect. The UPs SHOULD be able to notify the CP with the result of User Detect.

5.9. Query Statistics

The CUSP protocol MUST provide a means for the CP to be able to query statistics (performance monitoring) from the UP.

6. Security Considerations

As this is an Informational requirements document, detailed technical Security Considerations are not included. However, Section 5.4 covers general security requirements and Section 5.7 covers backup requirements relevant to some denial of service scenarios.

7. IANA Considerations

This document requires no IANA actions.

8. References

8.1. Normative References

- [I-D.cuspdtdt-rtgwg-cu-separation-bng-deployment]
Gu, R., Hu, S., and Z. Wang, "Deployment Model of Control Plane and User Plane Separation BNG", draft-cuspdtdt-rtgwg-cu-separation-bng-deployment-00 (work in progress), October 2017.
- [I-D.cuspdtdt-rtgwg-cu-separation-infor-model]
Wang, Z., Gu, R., Lopezalvarez, V., and S. Hu, "Information Model of Control-Plane and User-Plane separation BNG", draft-cuspdtdt-rtgwg-cu-separation-infor-model-00 (work in progress), February 2018.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

8.2. Informative References

- [TR-384] Broadband Forum, "Cloud Central Office Reference Architectural Framework", BBF TR-384, January. 2018.

Authors' Addresses

Shujun Hu
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: shujun_hu@outlook.com

Victor Lopez
Telefonica
Sur 3 building, 3rd floor, Ronda de la Comunicacion s/n
Madrid 28050
Spain

Email: victor.lopezalvarez@telefonica.com

Fengwei Qin
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: qinfengwei@chinamobile.com

Zhenqiang Li
China Mobile
32 Xuanwumen West Ave, Xicheng District
Beijing, Beijing 100053
China

Email: lizhenqiang@chinamobile.com

Tee Mong Chua
Singapore Telecommunications Limited
31 Exeter Road, #05-04 Comcentre Podium Block
Singapore City 239732
Singapore

Email: teemong@singtel.com

Donald Eastlake, 3rd
Huawei
1424 Pro Shop Court
Davenport, FL 33896
USA

Email: d3e3e3@gmail.com

Michael Wang
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: wangzitao@huawei.com

Jun Song
Huawei
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Email: song.jun@huawei.com

Internet Area WG
Internet Draft
Intended status: Informational
Expires: April 30, 2019

Praveen Muley
Wim Henderickx
Nokia
Geng Liang
China Mobile
Hans Liu
D-Link Corp
Loris Cardullo
Jonathan Newton
Vodafone
SungHoon Seo
Korea Telecom
Sagiv Draznin
Verizon Wireless
Basavaraj Patil
AT&T
October 22, 2018

Network based Bonding solution for Hybrid Access
draft-muley-network-based-bonding-hybrid-access-03

Abstract

In order to address increasing bandwidth demands, operators are considering bundling of multiple heterogeneous access networks (Hybrid access) for residential and enterprise customers. This document describes a solution for Hybrid access and covers the use case scenarios.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

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 April 30, 2019.

Copyright Notice

Copyright (c) 2018 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.....	3
2. Terminology.....	3
3. Reference Architecture.....	4
4. Network Based Bonding Solution Overview.....	5
4.1. Separate BNG and PGW.....	5
4.2. Integrated BNG and SGW/PGW.....	6
5. HAG Function.....	7
5.1. Address Assignment.....	7
5.1.1. Separate BNG and PGW.....	7
5.1.2. Integrated BNG and SGW/PGW.....	8
5.2. Setup and Tunnel Management.....	9
5.3. Traffic distribution policies.....	10
5.4. Path Management.....	11
5.5. Backward compatibility.....	12
6. Applicability in Mobile networks.....	12
7. Inter-working with MP-TCP.....	14
8. Security Considerations.....	14
9. IANA considerations.....	15
10. References.....	15
10.1. Normative References.....	15
10.2. Informative References.....	15
11. Acknowledgments.....	16

1. Introduction

To address the increasing demand of bandwidth by residential and enterprise customers, operators are looking for alternatives that can avoid rebuilding of the existing fixed access networks.

In Hybrid Access network, a Customer Premise Equipment (CPE) is connected to heterogeneous access networks (e.g. DSL, LTE etc) simultaneously. Traffic is distributed in flexible manner over these heterogeneous links thus increasing the bandwidth capacity of a residential or an enterprise customer.

This document describes a solution to implement the network based bonding Hybrid Access architecture. The solution is generic enough that it is applicable to fixed as well mobile nodes with multiple Access technologies.

2. Terminology

All mobility related terms are to be interpreted as defined in [RFC5213] and [RFC5844]. Additionally, this document uses the following terms

IFOM IP flow mobility

NB-IFOM Network based IFOM

ePDG Evolved Packet Data Gateway (defined in 3GPP [24.302])

RR Routing Rule

HAG Hybrid Access Gateway

PcRF Policy and Charging Rules Function

NBF Network based Bonding Function

MCP Multi-path conversion point (defined in [NAMPTCP])

3. Reference Architecture

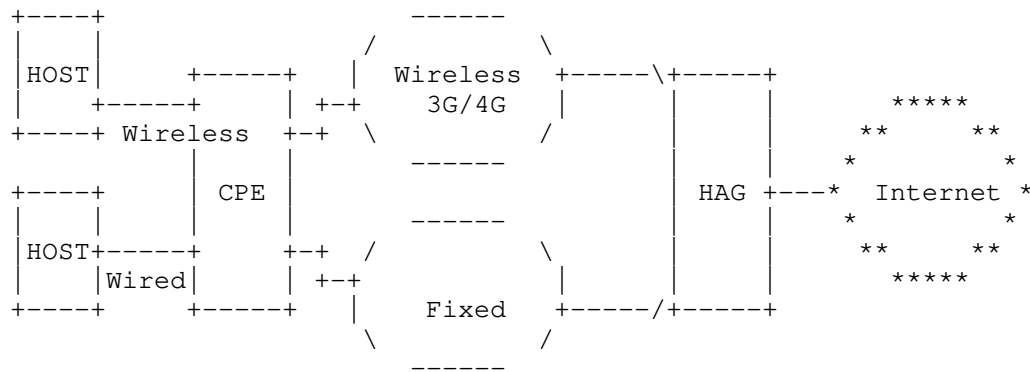


Figure 1 Network based bonding Hybrid Access Architecture

A CPE with HAG Figure 1 shows the network based bonding hybrid access architecture. In this architecture, HAG with network bonding function is deployed at the remote side of the CPE. The HAG receives the downstream traffic from internet and can apply the policies to distribute downstream traffic towards the CPE over available paths.

An in-band control protocol between the CPE and the HAG MAY be used to negotiate the traffic distribution policies for uplink traffic.

However, there SHOULD be flexibility to download the traffic distribution policies OUT-of-band.

Traffic distribution policies on CPE and HAG can have independent packet-based behavior.

Operators can have flexibility to distribute flows over multiple paths or associate affinity of flow to a particular access type. Traffic policies can also be applied taking into account the access networks link status such as latency, state etc.

Operator can also apply policy to fill one access link first before utilizing other (MAX-FILL). Affinity to one access MAY be due to cost or application characteristics. In this case the distribution of traffic is adjusted dynamically based on the load.

Behavior to adjust on moving around flows or packet is a matter of local policy.

4. Network Based Bonding Solution Overview

4.1. Separate BNG and PGW

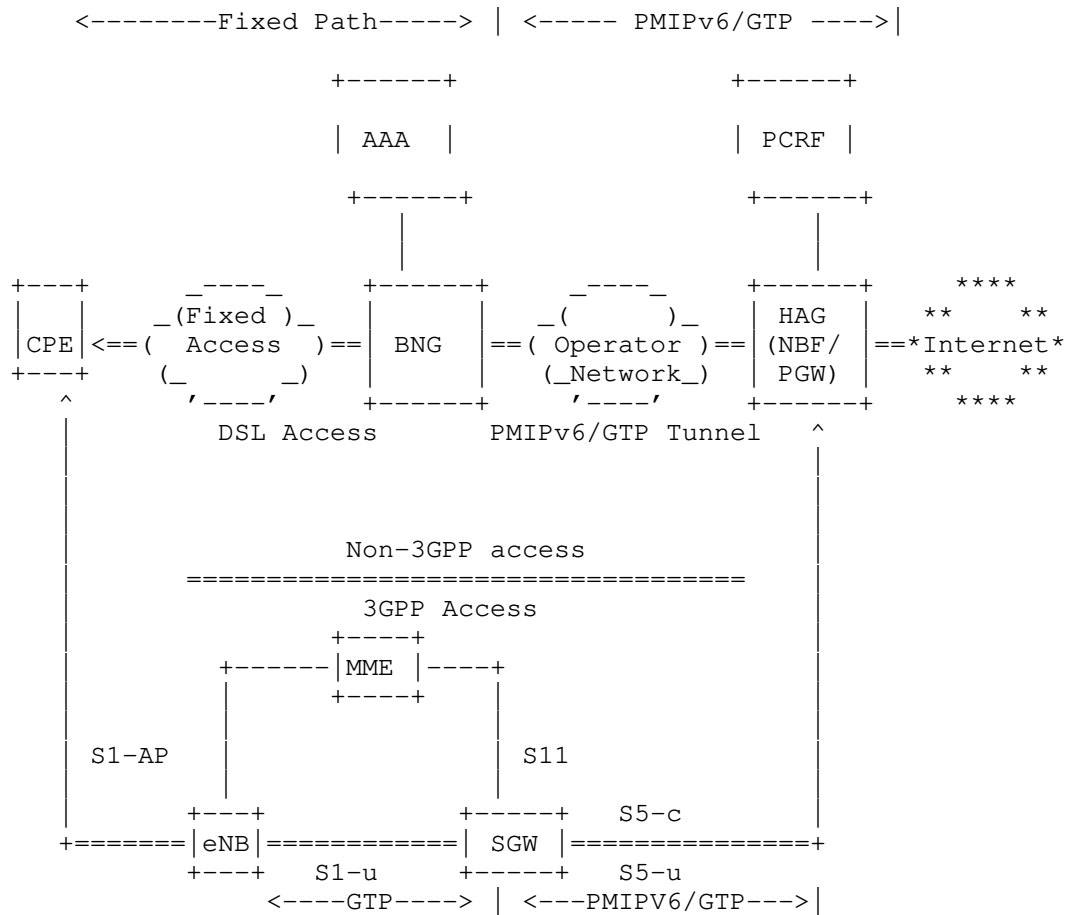


Figure 2 Hybrid access service in Fixed mobile convergence

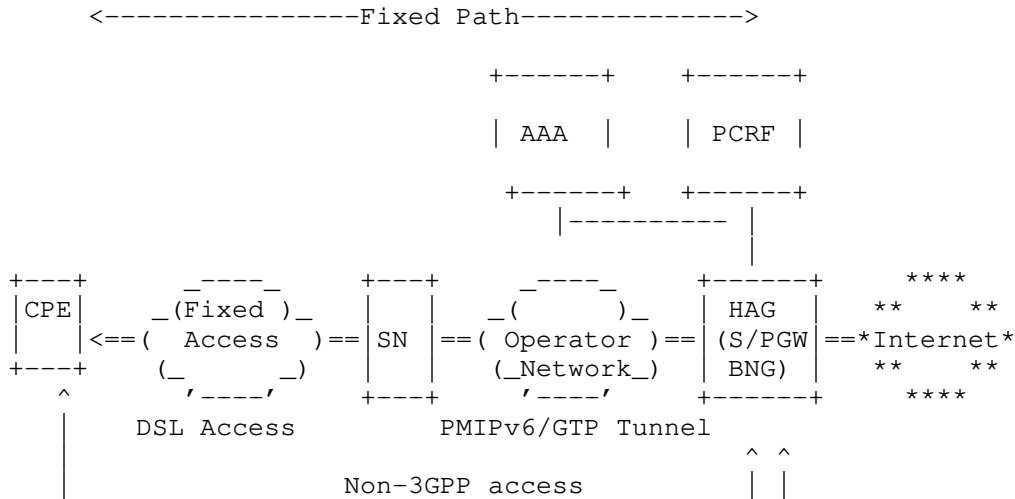
In Figure 2, CPE (either home or enterprise) is connected to internet via fixed access network using DSL as well as wireless access network using 4G cellular network.

The fixed access network BNG is connected to the PGW using 3GPP s2b reference point [TS23401]. The 4G cellular network is connected to the same PGW using S5 reference point (GPRS Tunneling Protocol (GTP) or Proxy Mobile IPV6 (PMIPV6) [RFC5213]) as specified by the 3GPP system architecture [TS23401].

The 3GPP as well non-3GPP access is bonded in CPE and the HAG which consist of NBF and PGW function. The bonding at HAG is achieved by allocating the same "IP address" when LTE access is setup on s5 and fixed (DSL) access over s2b.

The packet distribution policies applied to the bonded session on the HAG and CPE. Policies applied on HAG helps steering downlink traffic on specific access type or distribute percentage of traffic across both access types on per flow basis or per packet basis. Similarly policies applied CPE helps steering uplink traffic on specific access type or distribute percentage of traffic on per flow basis or per packet basis.

4.2. Integrated BNG and SGW/PGW



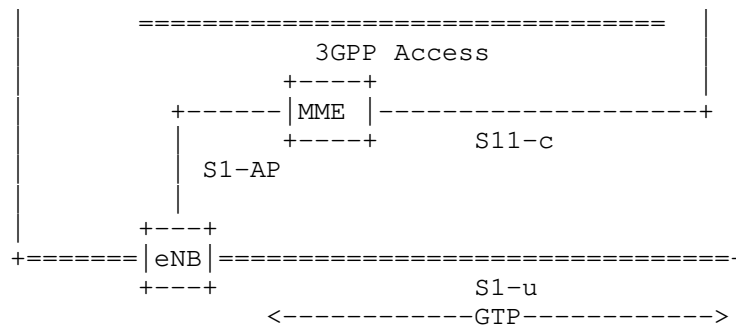


Figure 3 Integrated BNG,SGW,PGW with HAG

In Figure 3 , CPE is connected to internet through HAG by fixed and wireless access. HAG consist of BNG,SGW/PGW and NBF function.

HAG performs address assignment for all access types and acts as IP anchor point for IP services.

5. HAG Function

5.1. Address Assignment



5.1.1. Separate BNG and PGW

Following are steps for address allocation when BNG and PGW are separate. HAG in this case performs the NBF and PGW function.

[...CPE obtains LTE WAN IF address "A" during Pdp from HAG...]

(...CPE performs LTE attach for IMSI "X" APN "Y"...)

(...HAG allocates address "A" from APN.....)

[...CPE obtains DSL WAN IF address "A" during PPPoE from HAG...]

(...CPE begins the PPPoE setup with BNG.....)
(...BNG authenticates the CPE)
(...BNG receives all the 3GPP attributes from AAA server...)
(...BNG signals on s2b to HAG for address allocation.....)
(...HAG receives the s2b attach for APN "Y" with same IMSI.)
(...HAG finds session for IMSI "X" in APN "Y" RAT=LTE.....)
(...HAG bonds the LTE session with s2b session.....)
(...HAG returns address "A" in S2b response to BNG.....)
(...BNG stitches the PPPoE session with s2b session)
(...BNG returns the address "A" in PPPoE/DHCP to CPE.....)

HAG performs Address assignment for all access type which acts as anchor point for IP services.

APN "Y" on HAG is configured with property of "bonding" so that it can accept request from another access type for the same IMSI within same APN for same Pdp type. This helps in bonding the session with another access type session instead of treating it as handover.

BNG performs authentication of CPE. As part of authentication, it also receives the 3GPP attributes like IMSI, APN and HAG information from AAA server. It uses (3GPP) S2b reference point in [TS23402], specified by the 3GPP system architecture to get IP address from HAG and stitches the fixed access (PPPoE/IPoE) session with the s2b session both in control plane and data-plane.

The CPE remains unchanged as it uses standard method of IP address management for IPv4 and IPv6, on LTE link as well as DSL link.

5.1.2. Integrated BNG and SGW/PGW

Following are the steps for address allocation when BNG, SGW and PGW function is integrated along with the HAG function

[...CPE obtains LTE WAN IF address "A" during Pdp from PGW/HAG...]

(...CPE performs LTE attach for IMSI "X" APN "Y"...)

(...HAG allocates address "A" from APN.....)

[...CPE obtains DSL WAN IF address "A" during PPPoE from BNG/HAG..]

(...CPE begins the PPPoE setup with on BNG.....)

(...BNG authenticates the CPE)

(...BNG receives all the 3GPP attributes from AAA server...)

(...BNG/HAG finds session for IMSI "X" in APN "Y" RAT=LTE..)

(...BNG bonds the PPPoE session with LTE session.....)

(...BNG returns the address "A" in PPPoE/DHCP to CPE.....)

Address assignment is done in the HAG for all access type which acts as anchor point for IP services.

APN "Y" on HAG is configured with property of "bonding" so that it can accept request from another access type for the same IMSI within same APN for same Pdp type. This helps in bonding the session with another access type.

BNG performs authentication of CPE. As part of authentication, it also receives the 3GPP attributes like IMSI, APN and PGW information from AAA server. BNG detects that the PGW is local and hence internally bonds the fixed (PPPoE/IPoE) session with the LTE session with the same IMSI and APN. As part of bonding it uses the same IP allocated for the LTE session and sends back in PPPoE response or waits for DHCP to request for the address in the DHCP response. Traffic distribution policies are applied to the bonded LTE and fixed (PPPoE/IPoE) session to distribute the traffic.

The CPE remains unchanged as it uses standard method of IP address management for IPv4 and IPv6, on LTE link as well as DSL link.

5.2. Setup and Tunnel Management

There is no extra tunnel apart from the link tunnels representing each access used in this solution.

Any link can be setup first. The link that sets up access tunnel first gets the IP address from HAG. The link which comes later is bonded in HAG with the control plane to the existing access tunnel and the same IP address is returned to the later tunnel setup.

BNG stitches the fixed (PPPoE/IPoE) tunnel to the s2b tunnel setup towards the HAG. As part of it, it maps the setup and tear down event of the fixed (PPPoE/IPoE) tunnel to the s2b tunnel and vice versa.

5.3. Traffic distribution policies

As mentioned in earlier section, traffic distribution policies for upstream traffic is applied at CPE where as the downstream policies are applied at HAG. Given that single IP is allocated to all access type in this solution, it greatly helps to do flow mobility within the accesses.

Traffic distribution can be done on per flow basis, per MP-TCP sub-flow basis or on per packet. Flow based traffic distribution avoids out-of-order packets resulting out of differential latencies on each access tunnel and doesn't require buffering resources at the CPE or HAG to re-order the packets.

Policies applied in CPE can be downloaded out-of-band using ANDSF mechanism. Some CPEs are capable of sending initial uplink traffic on access type using random hashing but are able to move the flow to the access type chosen by network for the downlink traffic of that flow. Such CPEs need zero to minimal traffic policy configuration.

Traffic distribution policies applied at HAG for downlink traffic distribution can help in distributing flows or packets using hashing.

Traffic policies MUST have the flexibility to configure the amount of percentage of traffic to be steered over a given access type. This allows addressing the use case where operator MAY want to send a particular type of traffic over a specific access type (Video over DSL). In this case a video rule with affinity of DSL access can be set to steer 100 percent of traffic over DSL link whereas traffic matching any-any rule can be set to steer 50% over DSL and 50 % over LTE.

Traffic policies MUST allow asymmetric affinity association of access type for upstream and downstream traffic which allows splitting of a flow in upstream and downstream direction. Applying

such policies operators can use LTE for uplink where as fixed (DSL) for downlink. Studies of such configuration have shown application performance improvement over use same access for an application.

For the use case where a desired access link bandwidth is filled first (MAX-FILL) and use of second link is for the bandwidth overflow, one can use flow based or packet based approach for traffic distribution. The desirability of preferred access can be due to cheap access path or link characteristics for the given application.

To fulfill this requirement, two rate Three color marker (trTCM) can be used. Each access link uses token buckets to meter the packets as per configuration both at CPE and the HAG. Colored based policy is applied at CPE and HAG to steer packets to an access based on color. For ex. Green packets are steered to DSL if that is the preferred access, whereas yellow packets are steered over LTE access.

If flow based distribution is used, then on reaching certain thresholds there MUST be flexibility to move the flows from preferred access (say DSL) to another (LTE) by changing the percentage distribution. However, moving of FAT flows MAY result in under utilization of preferred access link. Similarly once the threshold drops, the traffic can be move back to preferred access by reverting the percentage distribution.

To avoid FAT flow distribution issues, packet based traffic distribution can be used to fully utilize the preferred access. Packets sent over different access for the same flow can reach out-of-order at the receiving end, due to differential transport latencies. Hence receiving end needs buffering and re-ordering capabilities to deliver flow packets correctly to an application.

5.4. Path Management

This solution relies of existing mechanism of Path management for wireless (LTE) and fixed (PPPoE/IPoE) tunnels.

In case of failure of any access tunnel the traffic MUST be switched to the alternate available access tunnel based on the traffic distribution policies.

5.5. Backward compatibility

This solution does not introduce any new protocol extensions. In this solution the CPE uses ANDSF routing rules to do the traffic distribution downloaded off-band in the CPE.

The policies at the HAG are either local configured or downloaded from PCRF. The existing service (ex. IPTV traffic MUST remain on DSL access) remains untouched by configuring appropriate traffic distribution policies. The exact configuration of those policies is outside the scope of this document.

6. Applicability in Mobile networks

A mobile node (MN) (also called User Equipment UE) connected to a 3GPP access network specified by the 3GPP system architecture [TS23401] is connected over the S5 reference point (Proxy Mobile IPV6 (PMIPV6) [RFC5213] or GPRS Tunneling Protocol (GTP)) to the PGW where the mobile node's session is anchored.

The (3GPP) S2b reference point in [TS23402], specified by the 3GPP system architecture defines a mechanism for allowing the mobile node (MN) attached to an "untrusted" non-3GPP IP access network to securely connect to a 3GPP network and access IP services. In this scenario, the mobile node establishes an IPsec ESP tunnel [RFC4303] to the security gateway called evolved packet data gateway (ePDG) and which in turn establishes a GPRS Tunneling Protocol (GTP) [TS23402] or Proxy Mobile IPV6 (PMIPV6) [RFC5213] tunnel to the packet data gateway (PGW) [TS23402] where the mobile node's session is anchored.

The figure below shows the hybrid access figure where the mobile node is connected to 3GPP and non-3GPP access simultaneously getting access to IP services via a PGW.

Figure 4 Hybrid access service in Mobile network

In the hybrid access architecture, an User equipment (UE) is connected to multiple access technology at the same time. It MAY connect to same network or different IP network based on the operator service. A mobile node with Third Generation Partnership Project (3GPP) access technology such as LTE, UMTS and non-3GPP access such as WIFI having simultaneous network connections is a use case of hybrid access in mobile networks.

As shown in Figure 4, the LTE access is bonded with the WIFI access and the same IP address is allocated on s2b as well as

s5 3gpp reference point. As discussed above, traffic distribution policies can be applied to steer traffic over specific access type or distribute over both access type to increase the bandwidth for the mobile node.

In some mobile networks, WIFI is preferred access since it's cheap, in that case policies described in MAX-FILL can be applied.

In some mobile networks, mobile nodes are configured to prefer WIFI access as local break out policy. However it's been observed that if mobile node has LTE access as well WIFI access available and if the mobile node connects to WIFI access over the s2b reference point to the same PGW, the PGW treats it as 3GPP to non-3GPP access handover and disconnecting the LTE access. But since mobile node is configured to be always connected over LTE access, mobile node reconnects over LTE and the PGW treats it as non-3GPP to 3GPP access handover disconnecting the WIFI access. This results in ping-pong effect. Since both accesses are simultaneously connected, in this solution, it helps in addressing the ping-pong issue as well.

7. Inter-working with MP-TCP

When used flow based hashing, it is possible that a FAT flow may cause to over congest the access link. To address FAT flow issues operator can deploy a MCP with the NBF. Operator in that case can apply policy to ensure the FAT flow traffic is split among small multi-path flows which can be seamlessly moved between the access types based on traffic distribution policies.

Inter-working helps operators in using MP-TCP for selective traffic thus ensuring effective utilization of buffering resources both at CPE as well as at MCP.

8. Security Considerations

The security considerations applicable while deploying the access types independently remains same while deploying network based bonding hybrid access architecture. This specification does not introduce any new security vulnerabilities.

9. IANA considerations

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.
- [24.008] 3GPP, "Technical specification Group Core Network and Terminals: Mobile radio interface Layer 3 specification; Core network protocols; Stage 3"
- [24.301] 3GPP, "Technical specification Group Core Network and Terminals: Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3"
- [NAMPTCP] M.Bouchadair et al. "draft-nam-mptcp-deployment-considerations-00", (work in progress), October 2016

10.2. Informative References

- [RFC5213] Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, August 2008.
- [RFC5844] Wakikawa, R. and S. Gundavelli, "IPv4 Support for Proxy Mobile IPv6", RFC 5844, May 2010.
- [TS23402] 3GPP, "Architecture enhancements for non-3GPP accesses", .
- [TS23401] 3GPP, General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access.

11. Acknowledgments

The authors are thankful for the detailed review and valuable feedback provided by Guiu Fabregas and Laurent Thiebaud.

Authors' Addresses

Praveen Muley
Nokia
805. E. Middle Field Rd.
Mountain View, CA, 94043

Email: praveen.muley@nokia.com

Wim Henderickx
Nokia
Copernicuslaan 50
Antwerp 2018
Belgium

Email: wim.henderickx@nokia.com

Geng Liang
China Mobile
32 Xuanwumen West Street,
Xicheng District, Beijing, 100053,
China

Email: gengliang@chinamobile.com

Hans Liu
D-Link Corporation
289, Sinhu 3rd Road,
Neihu District, Taipei City, 11494,
Taiwan, R.O.C.

Email: hans_liu@dlink.com.tw

Loris Cardullo
Vodafone
Italy

Email: Loris.Cardullo@vodafone.com

Jonathan Newton
Vodafone
United Kingdom

Email: Jonathan.Newton@vodafone.com

SungHoon Seo
Korea Telecom
South Korea

Email: sh.seo@kt.com

Sagiv Draznin
Verizon Wireless
USA

Email: Sagiv.Draznin@VerizonWireless.com

Basavaraj Patil
AT&T
2900 W. Plano Pkwy
Plano, Texas 75075
USA

Email: bp801n@att.com

gwg
Internet Draft
Intended status: Informational
Expires: September 11, 2019

S. Wadhwa
K. DeSmedt
Nokia
R. Shinde
Reliance Jio
J. Newton
Vodafone
R. Hoffman
TELUS
P. Muley
Nokia
Subrat Pani
Juniper Networks
Mar 11, 2019

Architecture for Control and User Plane Separation on BNG
draft-wadhwa-rtgwg-bng-cups-03.txt

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

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

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

This Internet-Draft will expire on September 11, 2019.

Copyright Notice

Copyright (c) 2019 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.

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.

Abstract

This document discusses separation of subscriber-management control plane and data-plane for BNG. Traditionally, the BNG provides aggregation of fixed access nodes (such as DSLAM and OLTs) over Ethernet and provides subscriber management and traffic management functions for residential subscribers. The BNG has however evolved to become a multi-access edge device that also provides termination of subscribers over fixed-wireless and hybrid access. Therefore, this document proposes interfaces between control and user-plane of a BNG that can support multi-access BNG.

Table of Contents

1. Introduction.....	3
1.1. Requirements Language.....	3
2. CUPS for BNG.....	3
2.1. Convergence.....	5
3. Interfaces for CUPS.....	6
3.1. In-band Signaling Channel.....	7
3.2. State Control Interface.....	8
3.2.1. Session level state management.....	8
3.2.2. Session level event notifications.....	14
3.2.3. Node level management.....	15

3.2.4. Node level event notifications.....	16
3.3. Management Interface.....	17
4. Protocol Selection for CUPS Interfaces.....	17
5. Address Pool Management.....	19
6. Security Considerations.....	19
7. IANA Considerations.....	19
8. References.....	20
8.1. Normative References.....	20
8.2. Informative References.....	20

1. Introduction

This document describes requirements and architecture for separation of subscriber management control plane and user plane for the BNG. In rest of the document the control plane is referred to as CP, user plane as UP, and the separation is referred to as CUPS (control and user plane separation). The draft describes the functional decomposition between CP and UP, and applicability of CUPS to a BNG that can support multiple access technologies such as fixed (DSL or Fiber), fixed-wireless (LTE,5G) and hybrid access i.e. simultaneous fixed and wireless access described in BBF [WT378]. The subsequent sections of the draft also define the interfaces required between CP and UP and briefly discusses a candidate base protocol for these interfaces.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. CUPS for BNG

In a CUPS architecture, signaling to setup subscriber sessions CP terminates signaling to setup subscriber sessions, and interfaces with the UP to create forwarding state for these sessions on the UP.

For fixed access subscribers, the CP terminates the signaling protocols (e.g. DHCP, PPPoE, SLAAC) from the customer premise, performs authorization/authentication with AAA Server, participates in address assignment, and then interfaces with the UP to create state related to forwarding and SLA management for the subscriber sessions on the UP. A subscriber session is a single IP connection, such as an IPoE or PPPoE session. The session can be single-stack (IPv4 or IPv6 only), or dual-stack (both IPv4 and IPv6). A CPE can

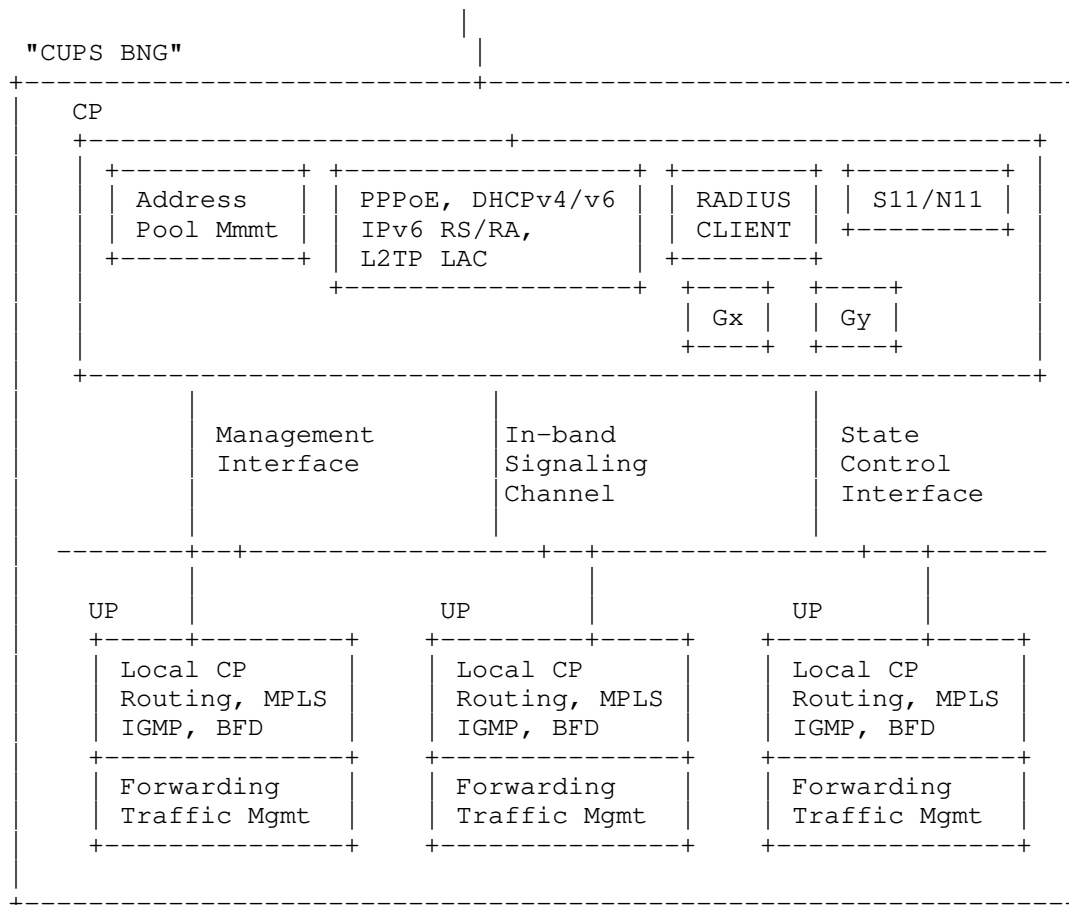
have multiple sessions, if multiple IP connections are required (e.g. on per service, or one per device behind the CPE).

The CP also processes solicited or unsolicited event notifications from the UP e.g. periodic accounting updates, usage reports, or session inactivity notifications. The interface between CP and UP that is used by the CP to manage session related forwarding state on the UP is being referred to as "state control interface". Asynchronous event notifications from UP to CP are also part of this interface.

In typical fixed access deployments, signaling (e.g. DHCPv4/v6, PPPoE, ICMPv6 RS/RA) to setup the subscriber sessions is in-band, and hence the UP receives the signaling messages from the customer premise. The UP should transparently forward (unmodified) in-band control messages as received from the customer premise to the CP and return messages from CP to the customer premise. Therefore, an in-band signaling channel is required between UP and CP. With a typical "CUPS BNG" deployment, the CP and UP are connected over a network, and the in-band signaling channel must be over a tunnel.

The UP performs forwarding and traffic management for the subscriber sessions. The infrastructure routing and signaling is done on the local control plane of the UP for fast convergence on network topology changes. In rest of the document the term "UP" is used generically for both functions performed by the local control plane on the UP and the data-plane.

A typical deployment architecture for CUPS includes a centralized CP running as a VNF interacting with multiple BNG UP instances that may be more distributed than the CP and could run as VNF or PNF. In this model, the CP and UP association is 1:N. This composite system containing CP VNF and one or more UP instances is referred to as a "CUPS BNG" in rest of the document. For operational ease, the CP MUST provide a single point for control and management for the entire "CUPS BNG". It MUST expose a single interface on behalf of the "CUPS BNG" to external systems such as AAA servers, OSS/BSS, Policy and charging servers. The CP VNF MUST support scale-out in order to cope with growth in number of subscriber sessions and/or increase in number of UP instances in the "CUPS BNG". Figure 1 below shows the functional components and interfaces for a "CUPS BNG".

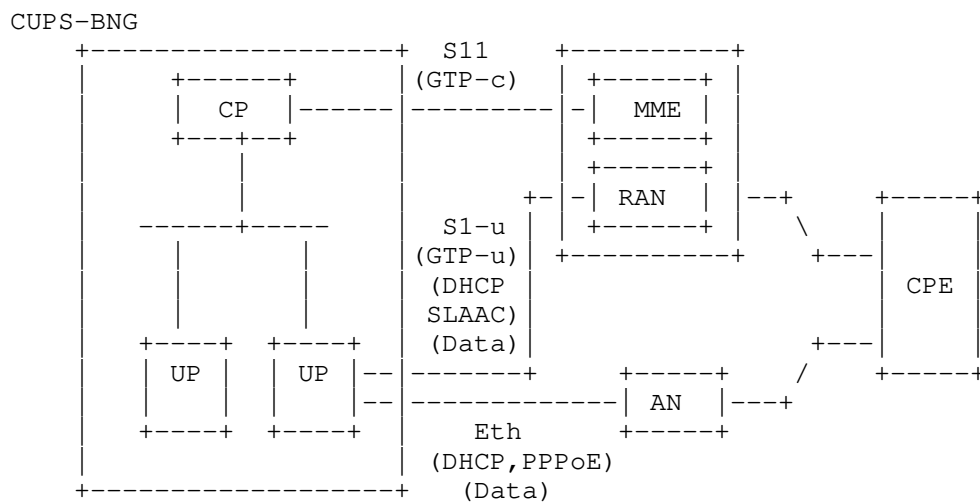


CUPS BNG System

2.1. Convergence

A single BNG can support subscribers over fixed, "fixed-wireless" or hybrid access. When a residential gateway has fixed-wireless access (LTE or 5G), then the BNG participates in 3GPP signaling with an MME

or AMF (i.e. support 3GPP S11 and N11 interfaces) to setup connections from (NG)RAN. With Hybrid access the customer premise initiates both fixed and wireless connections. The BNG in this case aggregates subscribers over Ethernet from fixed access nodes (DSLAMs and OLTs), but simultaneously terminates connections from (NG)RAN by participating in signaling with MME or AMF (S11/N11 interface). These deployment models are drivers for fixed-mobile convergence. It is important to ensure that the interfaces between CP and UP for CUPS can support not only fixed L2 access, but also the converged access scenario shown in Figure 2. One key requirement on the CP in these cases is the need to participate in 3GPP signaling (which is out-of-band) to setup the data-path. The data-path is a GTP-u (GPRS Tunneling protocol - User Plane) tunnel from the RAN (i.e. S1-u interface for LTE) as described in 3GPP [TS29281], and it terminates on the UP. It carries data traffic but also subscriber signaling messages (e.g. DHCPv4, DHCPv6, SLAAC) from the customer premise. The UP therefore still requires an in-band signaling channel to transport these protocol messages to the CP.



"CUPS BNG" with Converged Access

3. Interfaces for CUPS

A "CUPS BNG" MUST support the following interfaces between CP and UP, as shown in the figure in section 2.

3.1. In-band Signaling Channel

Section 2 describes the need for a signaling channel between CP and UP to transport in-band control messages between CP and the customer premise. Following are some key requirements for this interface.

- . The UP MUST pass the access circuit identifier over which the signaling messages are received as meta-data to the CP. This includes port, VLAN tags, tunnel endpoint IPs, any tunnel identifiers such as GTP TEID, MPLS labels, L2TP tunnel-id etc. The UP MUST also pass the L2 or L3 transport service that the access circuit is associated with. In case the control message PDU is carried in an Ethernet frame, then the UP SHOULD pass the received Ethernet frame to the CP. Both access circuit identifier and information in the Ethernet header are required by the CP to construct successful response packet (control message) back towards the customer premise. The access circuit identifier MUST be reflected from CP to UP, so UP can identify the access circuit over which it needs to send the CP's response packet. In the control message sent from UP to CP, the UP MUST also include the local MAC address associated with access circuit. This is because certain control messages from the customer premise are destined to a broadcast MAC (e.g. DHCP DISCOVER) or multicast (e.g. ICMPv6 RS), so CP cannot infer the local MAC from these messages. Certain messages also require the local MAC address to be inserted in the message (e.g. Link-Layer address in ICMPv6 RA messages)
- . The CP MUST be able to control the UP to forward only specific control messages to the CP.
- . The CP MUST be able to control the UP to block certain control messages received on a particular access circuit.
- . The CP MUST be able to control the UP to limit the rate of control messages (of specified type) to be sent by the UP.
- . The CP MUST be able to prioritize reception of certain control messages over others in a granular manner (e.g. prioritize DHCP RENEWS over DISCOVERS or prioritize PPP Keepalive over other messages).
- . The in-band signaling channel MUST support both fixed and converged access as described in section 2.1. The tunnel used for transporting these messages should therefore support both Ethernet and IP payloads.

3.2. State Control Interface

The CP and UP can exchange state at two levels using the "state control interface". One is at the node level and includes node-level information such as supported features, software releases, available resources, and operational state (e.g. active, failed, or overloaded). The other is at the subscriber session level. Subscriber session is described in section 2. The session level state includes basic forwarding and traffic management rules per session, that need to be provided by the CP to the UP in order to control per session forwarding and traffic management on the UP. It also includes state that triggers routing related actions on the UP. The session level state can include asynchronous event notifications from UP to CP, such as notifications to report per session usage (periodically or based on thresholds), notification to report session inactivity, and session liveness.

The interactions between CP and UP over "state control interface" can be categorized as:

- o Session level state management
- o Session level event notifications
- o Node level management
- o Node level event notifications

Following sub-sections provide more details on these interactions. The interactions between CP and UP over "state control interface" are modeled via abstract request/response messages between CP and UP. These messages will need to be defined as part of the protocol specification for this interface.

The protocol selected to implement this interface MUST support both fixed access and converged access (described in section 2.1) on BNG

3.2.1. Session level state management

Once the CP has successfully authorized and/or authenticated the subscriber session, and completed address assignment, it uses the "state control interface" to install forwarding and related state for the session on the forwarding path of the UP. This is abstracted as a "session create request" call from CP to UP, as shown in the figure below. The UP MUST ack or NACK via a response back to CP.

Since BNG can support different access types (e.g. fixed L2 access, or tunneled L3 in case of fixed-wireless, or a combination in case of hybrid access), it is important that the forwarding state information for the subscriber sessions, sent from CP to UP, can be specified as flexible packet matching rules and set of actions related to forwarding and traffic management. The UP should be able to use these match rules and actions to derive various lookup tables and processing in the forwarding path to forward traffic to and from the CPE.

The basic forwarding state in upstream direction (i.e. access to network) and downstream direction (i.e. network to access) fundamentally consists of session identification and one or more actions. Following shows a logical representation of a directive from CP to UP to install basic forwarding state on the UP for fixed L2 access (i.e. access from DSLAM or OLTs over Ethernet).

Direction Upstream - Access to Network:

Subscriber-identification: Port/VLAN-tag(s) + subscriber-MAC

Action: remove encapsulation, IP FIB lookup, forward to network.

Direction Downstream - Network to Access:

Subscriber-identification: IP address

Action: lookup IP DA, build encapsulation using Port/VLAN-tag(s)+ subscriber-MAC, forward to access.

Optionally, the IP address assigned to the CPE can also be provided for subscriber-identification (e.g. for anti-spoofing) in the upstream direction.

In case of PPPoE sessions, the subscriber-identification for upstream direction and encapsulation for downstream direction also includes the PPPoE session-id.

Based on the directive from CP to UP (as shown in the example above), the UP can then populate appropriate tables in the forwarding path, e.g. subscriber lookup tables, IP-FIB, and ARP or IPv6 Neighbor discovery table. It can also program the packet processing in both upstream and downstream direction based on the specified actions.

In case of "fixed-wireless" access, the access circuit is a GTP-u tunnel. In this case there is no physical interface (or port), and hence the CP MUST provide a tunnel definition to the UP to use as access circuit in upstream direction, and encapsulation in downstream direction. The tunnel definition will include the tunnel endpoint IP, and TEID that is established via out-of-band signaling

between the CP and the customer premise. It can also include the routing context for transporting the tunnel.

In addition to setting up the forwarding state as directed by the CP, the UP also needs to announce in routing the aggregate prefixes from which the CP assigns IPv4 and IPv6 addresses (or prefixes) to the CPEs. The CP SHOULD provide these aggregate prefixes to the UP as part session state. In case the aggregate prefixes are not provided, the UP MUST announce individual CPE addresses in routing, or it MAY try to aggregate in case addresses for multiple CPEs are from a contiguous address space.

The CPE can have a routed subnet behind it (aka framed-route). CP can learn the framed-routes during authentication/authorization. The CP should provide the framed-route to the UP as part of session state. The UP MUST install this route in the forwarding path and associate it with the forwarding state of the corresponding subscriber session. It should also announce this in routing towards the Network.

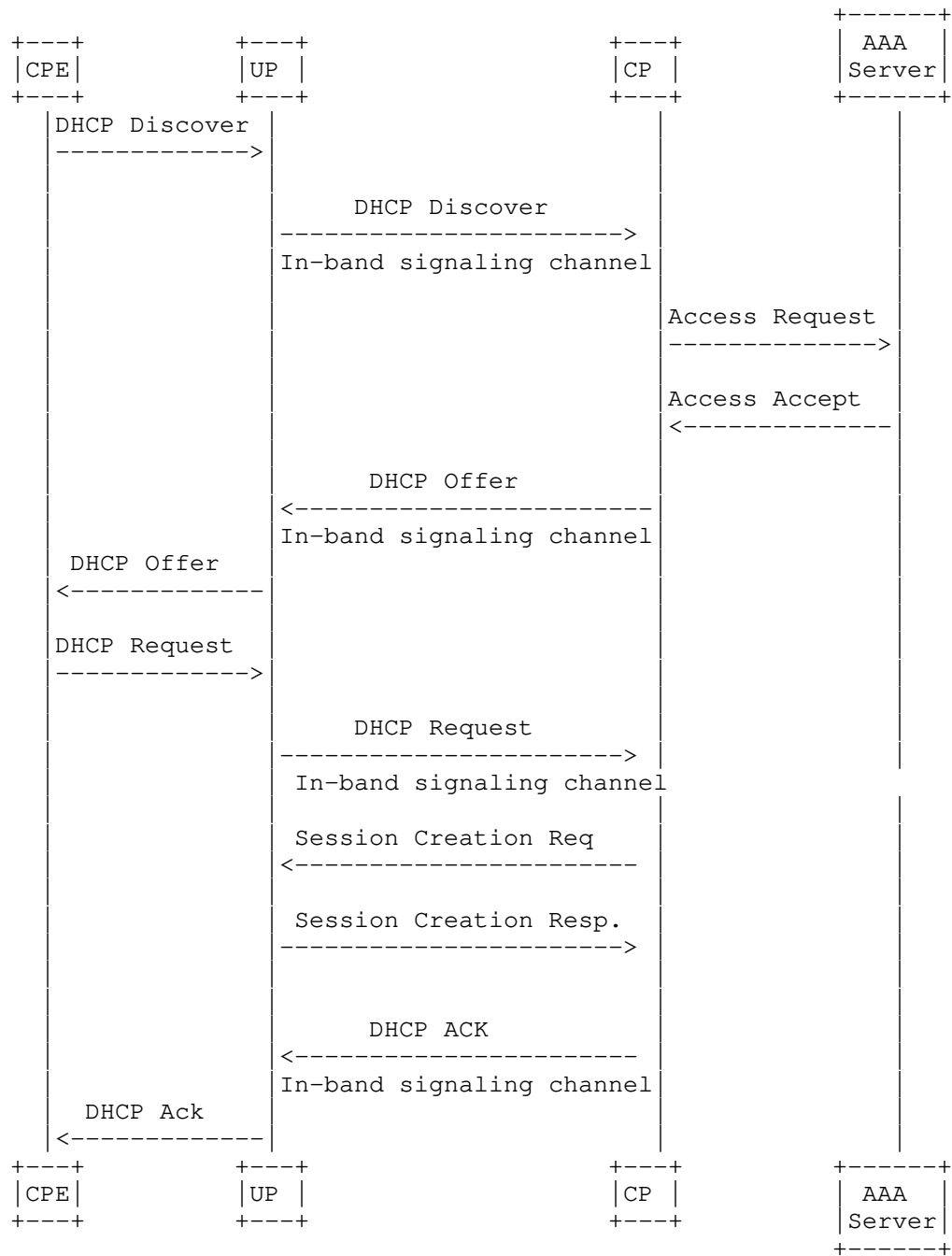
The CP MUST also provide to the UP the address assigned as IP gateway address to the CPEs in DHCP. The UP MUST locally configure this address appropriately, such that it can respond to ARP requests for this address from the CPEs.

The session state on the UP is always controlled by the CP i.e. the UP just follows the directive from the CP to install, modify and delete the session state. In addition to the basic forwarding state, the CP can also associate, update and disassociate other related state with the session e.g. state related to:

- . Filtering
- . SLA management
- . Statistics collection
- . Credit control
- . Traffic mirroring
- . Traffic Steering
- . NAT
- . Application aware policies

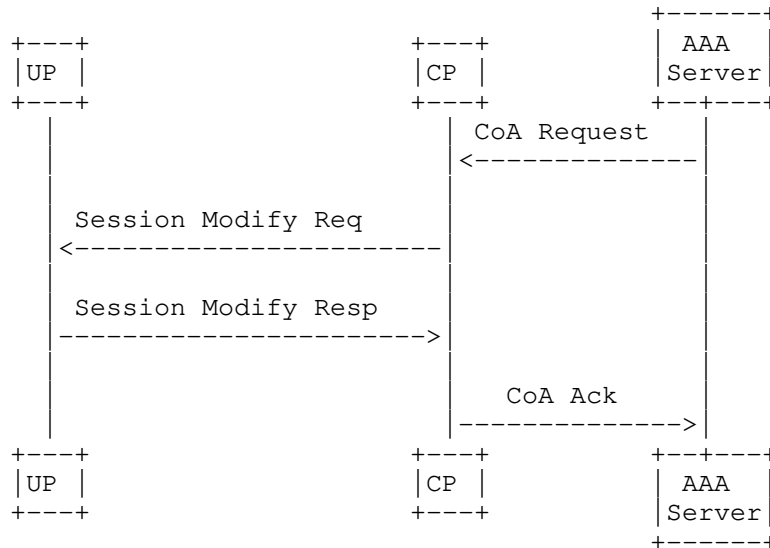
BNG deployments use hierarchical QoS (H-QoS) models which follows from a combination of link-layer over-subscription, multi-service networks and multiple layers of aggregation. For example, a common hierarchy exists of at least a QoS layer per access-node, and per

CPE. The CP MUST provide SLA management information to the UP per CPE. This includes applicable QoS parameters (e.g. rates, queues, markings) and the QoS hierarchy to which the CPE belongs. The CP may choose to signal this via a QoS policy that is locally pre-configured on the UP.



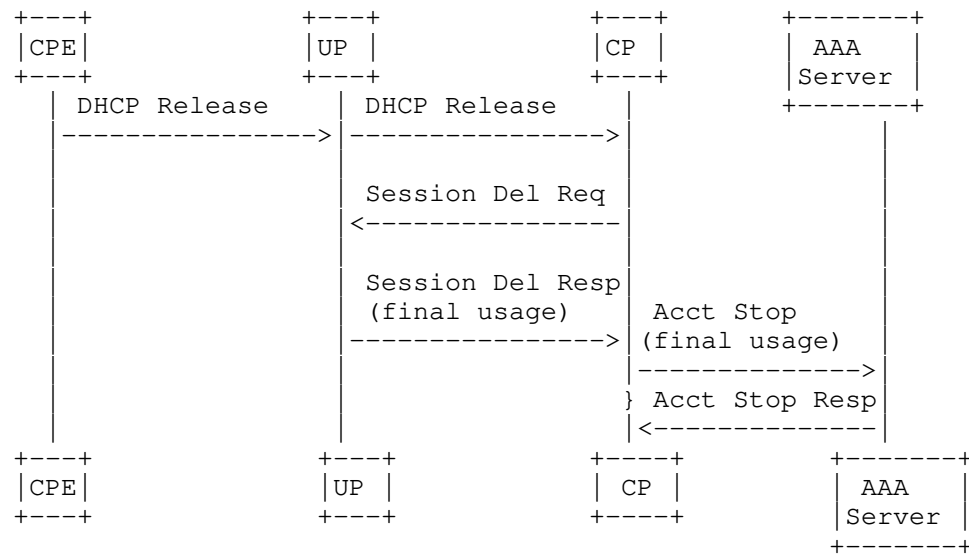
Session Creation Sequence

CP can trigger update of session state on the UP, triggered by re-authentication or CoA from AAA or policy-server, as show in the figure below.



Session Modification

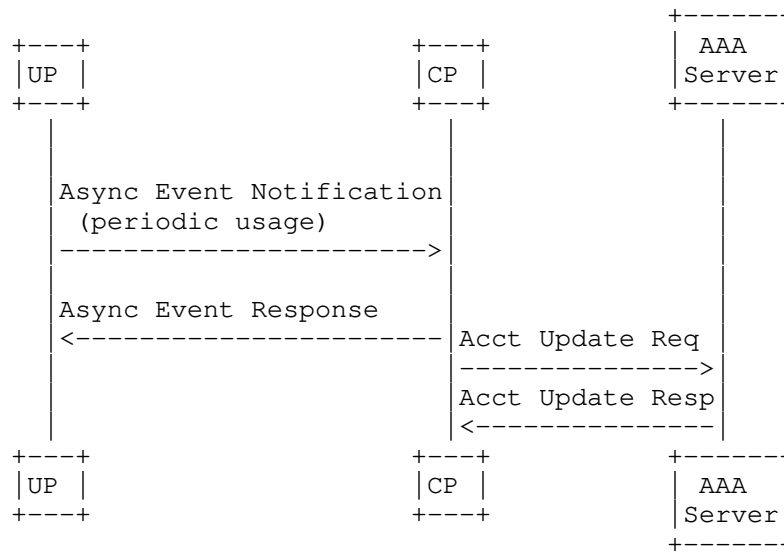
CP can trigger the deletion of session state based on signaling messages (as shown in the figure below), administrative action or disconnect-message initiated from the AAA server.



Session Deletion

3.2.2. Session level event notifications

UP can asynchronously generate Session level event notifications to the CP. An example of asynchronous notification is periodic usage reporting from UP to the CP, so that the CP can report the usage to a AAA server via interim accounting-updates. The CP can set the periodicity of this notification on the UP based on interim accounting interval configured by the operator on the CP.



Async Event Notification for periodic usage

Following are some other examples requiring asynchronous notifications from UP to CP.

- o Threshold based usage reporting
- o Inactivity timeout
- o Subscriber unreachability detection

The protocol for "state control interface" MUST support asynchronous notifications from UP to CP.

3.2.3. Node level management

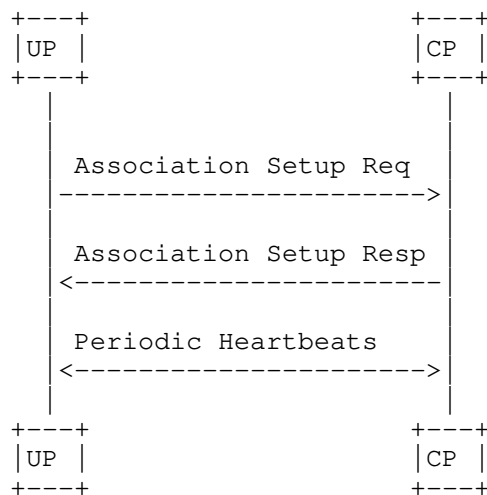
There needs to exist a concept of association between CP and UP. When the CP or UP comes online it should setup an association with configured or discovered peers via a message exchange. In association setup, the nodes should be able to exchange supported capabilities, version of software, load/overload information, and resource information. Also, any node-wide parameters can be exchanged during association setup.

No session state related messages should be accepted from the peer by either CP or UP unless an association exists.

Either node should be able to update the association to report changed feature capabilities, overload condition, resource exhaustion or any other node-wide parameters.

The UP should be able to request a graceful association release from the CP. In this case the CP should delete all sessions from that UP and process the final stats report for each session and send it in accounting-stop to the AAA server. During this process the CP MUST not create new sessions on the UP. Once all sessions are successfully deleted, the CP should release the association.

There needs to be a periodic node-level heartbeat exchange between CP and UP to detect if the peer is reachable and active. If peer is determined to be down based on heartbeat messages, then all the data-plane session state associated with the peer should be deleted.



Node Association Setup and Maintenance

3.2.4. Node level event notifications

There needs to be support for asynchronous node level event notifications from UP to CP. Example includes switchover

notification in case ports or UP failures when UP node level warm-standby redundancy is enabled. Based on this notification, the CP can create session state for all the sessions associated with the failure domain on the new primary UP.

3.3. Management Interface

The CP MUST provide a single point for local management of "CUPS BNG" system to the operator. This requires a management interface between CP and each of its associated UPs for pushing configuration to the UP and retrieving operational state from the UP. The interface MUST minimally include BNG specific configuration and state.

The Management interface SHOULD support transactional configuration from CP to UPs and SHOULD support state retrieval, both based on a well-defined data schema. The management interface SHOULD support unsolicited signaling of state changes (events) from UP to CP i.e. MUST provide telemetry for events. Either gNMI or NETCONF can be considered as acceptable candidates for model driven management interface.

4. Resiliency

"CUPS BNG" system MUST be protected against failure of CP VNF and MUST be able to recover the session state without operator intervention and reliance on CPEs. This can be achieved by providing redundancy for processing resources within CP VNF and maintaining redundant instance of session state.

Protection against UP failures based on 1:1 UP (hot-redundancy) and N:M (warm-redundancy) SHOULD be supported. For 1:1 hot-redundancy the CP needs to create data-plane state for sessions on both UPs that form a redundant pair, using the "state control interface". The CP needs to ensure the data-plane state for a session stays synchronized between the two nodes. A given session's data-plane should only be active on one UP in the pair, which serves as active UP for the session. However, sessions that share the redundant UP pair can be distributed between the two UPs for active forwarding.

N:M warm-redundancy ($N > M$) can be supported via creation of data-plane state on the designated backup chassis after the failure has been detected. This would result in longer failover times than 1:1 hot-redundancy.

Redundant network connectivity between CP and UPs MUST be supported. In the "CUPS BNG" architecture, it is important to configure redundant connectivity that doesn't share fate.

5. Protocol Selection for CUPS Interfaces

It is important that the selected protocol for "state control interface" between CP and UP works not just for fixed access but also works for converged access on BNG. 3GPP has defined PFCP (Packet Forwarding Control Protocol) in [TS29244] as the interface between CP and UP for LTE gateways. This protocol is suited for large scale state management between CP and UP. Following are some of the key attributes of this protocol:

- o It supports management of forwarding and QOS enforcement state on the UP from CP. It also supports usage reporting from UP to CP.
- o It is over UDP transport and doesn't suffer from any HOL blocking.
- o It provides reliable operation based on request/response with message sequencing and retransmissions.
- o It provides an overload control procedure where overload on UP can be handled gracefully.
- o The protocol is extensible and allows addition of new IEs.

For fixed access BNG, the protocol requires simple extensions in form of additional IEs. The required extensions are mainly due to fact that typically a fixed access BNG requires tighter control over L2 behavior and manages access and subscriber using L2 identifiers (such as VLANs and MAC addresses), whereas mobile access works in terms of L3, either routed or tunneled.

The details of the protocol as applicable to the BNG and the required extensions will be defined in a separate draft.

[TS29244] also describes an in-band signaling channel based on GTP-u tunnel between CP and UP. GTP-u (GPRS Tunneling protocol - User Plane) is defined in 3GPP [TS29281] and defines a tunneling protocol which carries IP payloads. The protocol runs over a UDP/IP stack and uses UDP port number 2152. Data within a tunnel can be multiplexed based on Tunnel Endpoint Identifiers (TEIDs). The protocol supports optional sequence numbers. The protocol supports extension headers to allow development of new features. GTP-u tunnels are signaled between CP and UP, and it is possible

to associate filters to block certain control packets from being forwarded from UP to CP. The payload type carried by GTP-u can be extended to Ethernet (via payload type in extension header). The tunnel encapsulation can also be extended similarly to carry any required meta-data.

6. Address Pool Management

The CP MUST support management of IPv4 and IPv6 address pools, where each pool can contain one or more subnets. The pool management MUST support pool selection based on one or more of the following criteria:

- o UP
- o Access port on the UP.
- o Redundancy domain on the UP (e.g. set of access ports that share fate with respect to switchovers due to failures, when UP node level redundancy is enabled).
- o Service (e.g. HSI, VoIP, IPTV etc.).
- o Location (e.g. based on circuit-id/remote-id or part of circuit-id/remote-id in DHCP and PPPoE).

Pool management on CP SHOULD NOT statically link subnets to UPs but SHOULD dynamically allocate subnets to UP based on load i.e. on-demand, and signal allocated subnets using the "state control interface" as described in section 3.2.1. This allows for better IP resource utilization and less subnet fragmentation.

7. Security Considerations

For security between CP and UP, Network Domain Security (NDS) as defined in [TS33210] can be considered. As per NDS, the network can be split into security domains. Communication within a single security domain is considered secure, and protocols can operate without any additional security. When communication has to cross security domains, then IPSEC can be used.

8. IANA Considerations

None.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [TS29244] 3GPP, "Interface between the Control Plane and the User Plane Nodes", TS 29.244 15.2.0, June 2018, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3111>.
- [TS29281] 3GPP, "General Packet Radio System (GPRS) Tunneling Protocol User Plane (GTPv1-U)", TS 29.281 15.3.0, June 2018, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1699>.
- [TS33210] 3GPP, "Network Domain Security (NDS); IP network layer security", TS 33.210 15.0.0, June 2018, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2279>.
- [WT378] BBF, "Nodal Requirements for Hybrid Access Broadband Networks", WT-378, 2018.

9.2. Informative References

- [RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <https://www.rfc-editor.org/info/rfc2131>.
- [RFC2516] Mamakos, L., Lidl, K., Evarts, J., Carrel, D., Simone, D., and R. Wheeler, "A Method for Transmitting PPP Over Ethernet (PPPoE)", RFC 2516, DOI 10.17487/RFC2516, February 1999, <https://www.rfc-editor.org/info/rfc2516>.
- [RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <https://www.rfc-editor.org/info/rfc3315>.

Authors' Addresses

Sanjay Wadhwa
Nokia
777 East Middlefield Road
Mountain View
USA

Email: Sanjay.wadhwa@nokia.com

Killian De Smedt
Nokia
Copernicuslaan 50
Antwerp
Belgium

Email: Killian.de_smedt@nokia.com

Rajesh Shinde
Reliance Jio Infocomm Ltd.
Reliance Corporate Park
Thane Belapur Road, Ghansoli
Navi Mumbai 400710
India

Email: Rajesh.A.Shinde@ril.com

Jonathan Newton
Vodafone
Waterside House
Bracknell
United Kingdom

Email: jonathan.newton@vodafone.com

Ryan Hoffman
TELUS
1525 10th Ave SW
Calgary, Alberta
Canada

Email: ryan.hoffman@telus.com

Praveen Muley
Nokia
805. E. Middle Field Rd.
Mountain View, CA, 94043
USA

Email: praveen.muley@nokia.com

Subrat Pani
Juniper Networks

10 Technology Park Dr.
Westford, MA
USA

Email: spani@juniper.net

gwg
Internet Draft
Intended status: Informational
Expires: September 11, 2019

S. Wadhwa
K. DeSmedt
P. Muley
Nokia
R. Shinde
Reliance Jio
J. Newton
Vodafone
R. Hoffman
TELUS
S. Pani
Juniper Networks
March 11, 2019

Requirements for Protocol between Control and User Plane on BNG
draft-wadhwa-rtgwg-bng-cups-protocol-requirements-02.txt

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), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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

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

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

This Internet-Draft will expire on September 11, 2019.

Copyright Notice

Copyright (c) 2019 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.

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.

Abstract

Traditionally, the BNG provides aggregation of fixed access nodes (such as DSLAM and OLTs) over Ethernet and provides subscriber management and traffic management functions for residential subscribers. The BNG has however evolved to become a multi-access edge device that also provides termination of subscribers over fixed-wireless and hybrid access. An overall architecture and interfaces required between separated control and user-plane for a multi-access BNG are described in draft-wadhwa-rtgwg-bng-cups-01.txt. This document discusses requirements for protocol between subscriber-management control-plane and user-plane for BNG to achieve separation.

Contents

1. Introduction.....	3
1.1. Requirements Language.....	3
2. Requirements for "CUPS protocol".....	3
2.1. State Control Interface Requirements.....	5
2.2. Extensibility.....	8
2.3. Scalability and Performance.....	9
2.4. Transport Protocol.....	10
2.5. In-band Control Channel Requirements.....	10
2.6. Resiliency.....	12
2.7. Security.....	13
3. "CUPS protocol" candidate.....	13
4. Security Considerations.....	14
5. Management Interface Requirements.....	14

6. IANA Considerations.....	15
7. References.....	15
7.1. Normative References.....	15
7.2. Informative References.....	16

1. Introduction

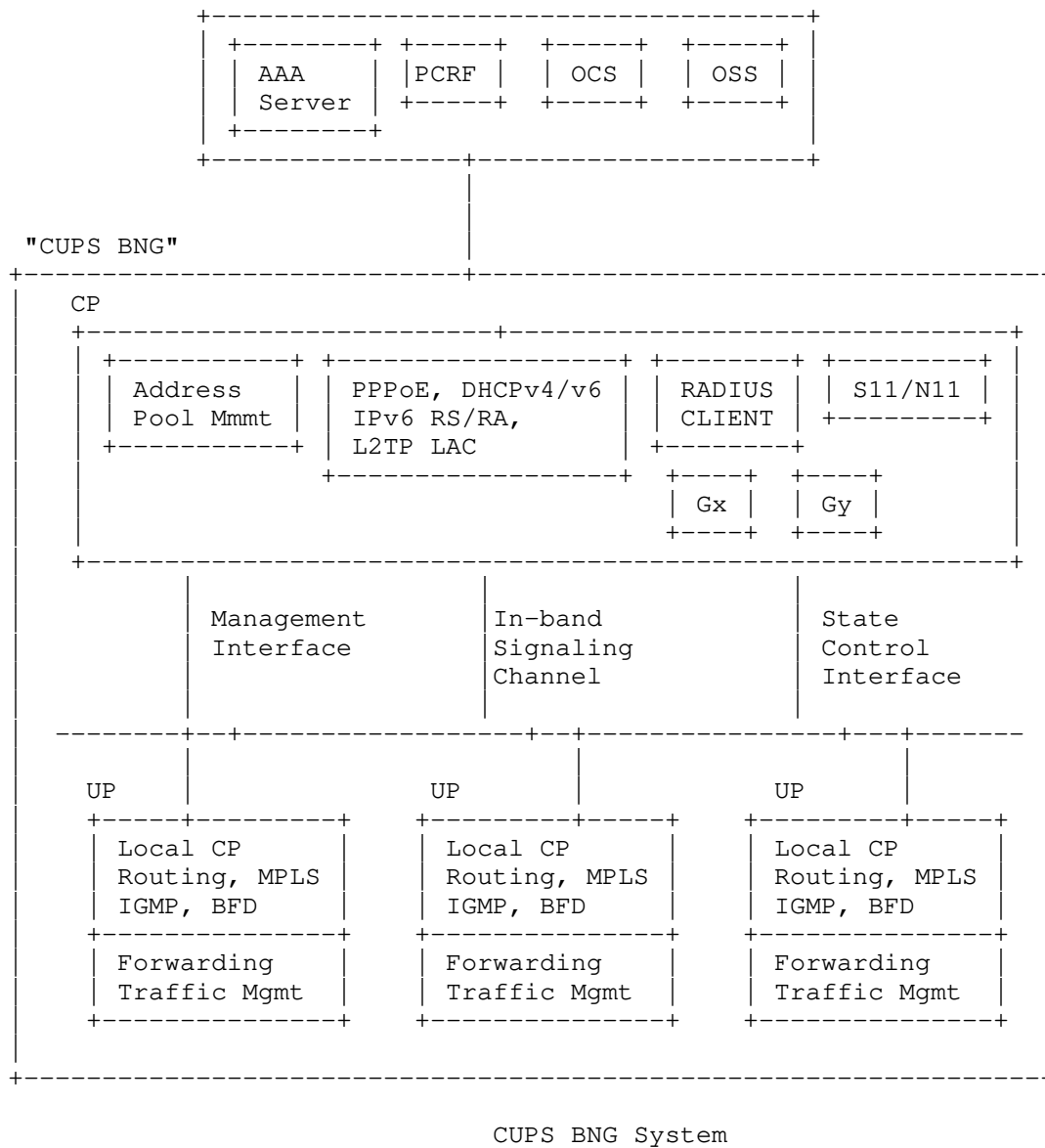
This document describes a set of requirements for protocol between subscriber-management control and user plane for BNG, that need to be met, in order to achieve separation. In rest of the document the control plane is referred to as CP, user plane as UP, and the separation is referred to as CUPS (control and user plane separation). The protocol between control and user-plane to achieve separation is referred to as "CUPS protocol". These requirements should form the basis for "CUPS protocol" selection. The functional decomposition between CP and UP, and applicability of CUPS to a BNG that can support multiple access technologies such as fixed (DSL or Fiber), fixed-wireless (LTE, 5G) and hybrid access are described in [CUPS].

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Requirements for "CUPS protocol"

[CUPS] defines overall operation and architecture for control and user-plane separation on BNG. It also defines key functional interfaces between CP and UP, as shown in Fig 1, to realize the separation. "CUPS protocol" MUST provide support for information exchange to realize the "state control interface" and "in-band signaling channel" as defined in [CUPS].



2.1. State Control Interface Requirements

- . "CUPS protocol" MUST support convergence on BNG, where the CPEs terminating connections on the BNG can have fixed-access (e.g. xDSL/PON/Ethernet), fixed-wireless access (LTE/5G) or hybrid-access (i.e. combined fixed and wireless access).
- . "CUPS protocol" MUST support messages and information exchange for node level management. There needs to exist a concept of association between CP and UP. When the CP or UP comes online it should setup an association with the configured or discovered peers via a message exchange. In association setup, the nodes should be able to exchange supported capabilities, version of software, load/overload information, and resource information. Also, any node-wide parameters can be exchanged during association setup.
- . "CUPS protocol" MUST allow either node to update the association to report changed feature capabilities, overload condition, resource exhaustion or any other node-wide parameters.
- . "CUPS protocol" MUST provide support for UP to request a graceful association release from the CP.
- . "CUPS protocol" MUST support periodic node-level heartbeat exchange between CP and UP to detect if the peer is reachable and active.
- . "CUPS protocol" MUST support exchange of messages and information elements (IEs) between CP and UP for session level state management on the UP.

A subscriber session is a single IP connection, such as an IPoE or PPPoE session. A CPE can have multiple sessions, if multiple IP connections are required (e.g. one per service, or one per device behind the CPE). The session level state on the UP, managed from the CP includes:

 - o Data-plane state for forwarding data traffic from subscriber sessions in upstream direction (access to network), and downstream direction (network to access).
 - o Forwarding state related to in-band control plane messages (such as messages for DHCP, PPPoE, SLAAC) that are forwarded

- from CPE to CP via the UP (in upstream direction), and from CP to CPE via the UP (in downstream direction).
- . In addition to the basic forwarding state, the "CUPS protocol" MUST support messages and information elements (IEs) for CP to associate, update and disassociate other data-plane related state with the session e.g. state related to:
 - o Filtering
 - o SLA management
 - o Statistics collection
 - o Credit control (usage monitoring and reporting)
 - o Traffic mirroring for legal intercept
 - o NAT
 - o Application (L4-L7) aware policies
 - . Depending on the type of access and the network between access-nodes and the BNG, the subscriber traffic from the CPEs can be encapsulated and transported over an L2 connection or over an L3 tunnel. Common scenarios for fixed access include Ethernet (q-in-q,.1q), L2oGRE, L2TPv3, VxLAN, and MPLS PW. For fixed-wireless the access is over a GTP tunnel (as defined in [CUPS]). The tunnel transport for L3 tunneled subscriber traffic can IPv4 or IPv6. The subscriber traffic itself can be IPv4, IPv6 or PPPoE. In case of PPPoE, the BNG can terminate PPPoE or tunnel it over L2TP to another gateway. The data-plane on the BNG decapsulates the upstream (access->network) traffic and routes it towards the network in appropriate routing-context, and optionally perform NAT before routing. It determines the subscriber for downstream (network->access) IP traffic, encapsulates it appropriately before forwarding towards the access. In addition, it does traffic-management and SLA management, maintains traffic statistics and optionally monitors and reports usage. The "CUPS protocol" MUST be able to carry state from CP to UP for IPv4, IPv6 and PPPoE sessions, for various flavors of transport connections mentioned above.
 - . Given the variety of access types on the CPE and type of transport networks between access-nodes and BNG (as outlined above) , the "CUPS protocol" MUST specify forwarding state information for the subscriber sessions, for both data and in-band control, as flexible packet matching rules and set of actions related to forwarding and traffic management, rather than just fixed-format lookup tables understood by particular UP implementation. Using the flexible match rules and actions conveyed in the "CUPS protocol" IEs, the UP should unambiguously be able to derive

various lookup tables and processing in the forwarding path to forward traffic to and from the CPE. The basic forwarding state in upstream direction (i.e. access to network) and downstream direction (i.e. network to access) fundamentally consists of session identification and one or more actions. Following shows a logical representation of a directive from CP to UP to install basic forwarding state on the UP for fixed L2 access (i.e. access from DSLAM or OLTs over Ethernet).

- o Direction Upstream - Access to Network:
 - . Subscriber-session identification: Port/VLAN-tag(s) + subscriber-MAC + Session IP address + PPPoE Session-ID
 - . Action: remove encapsulation (i.e. Ethernet and PPPoE/PPP headers), apply policer, do IP FIB lookup, forward to network.
- o Direction Downstream - Network to Access:
 - . Subscriber-session identification: IP address
 - . Action: apply subscriber-shaper, build encapsulation using (PPPoE session-id and Port/VLAN-tag(s)+ subscriber-MAC), forward to access.

Examples of actions and processing related to forwarding and traffic management include encapsulation/decapsulation, table lookups, drop, forward, mirror, count, redirect, police, classify, queue, shape etc.

- . In addition to packet-matching rules and actions to setup data-path on the UP, the "CUPS protocol" MUST allow CP to specify subscriber routing and IP interface related information. This includes the following:
 - o Aggregate IPv4 subnets and IPv6 prefixes that are used for assigning addresses or prefixes (e.g. IPv6 delegated-prefix) to subscribers on a UP. These are announced in routing by the UP to draw downstream traffic.
 - o UE's IP address and subnet mask.
 - o Default gateway IP address within the subscriber subnets. This is used to draw upstream traffic from the CPEs and the UP is required to respond to ICMP requests for this address from the CPEs.
 - o Subnets for network behind a CPE (also known as framed-routes).

- . The "CUPS protocol" MUST provide support for CP to specify session level HQOS related information to the UP. A common QoS hierarchy on BNG consists of at least a QoS layer per access-node, and per CPE. "CUPS protocol" MUST provide support for CP to specify QoS parameters (e.g. rates, queues, markings) and the QoS hierarchy to which the CPE belongs, to the UP. The CP may choose to signal this via a QoS policy that is locally pre-configured on the UP. "CUPS protocol" MUST provide support for CP to specify HQOS-policy that the session is associated with.
- . "CUPS protocol" MUST support asynchronous session level event notifications from UP to CP. Session level asynchronous notifications include:
 - o Periodic usage-reports
 - o Threshold based usage-reports
 - o Inactivity timeout
 - o Subscriber unreachability detection
- . "CUPS protocol" MUST support asynchronous node level event notifications from UP to CP. Example includes switchover notification in case ports or UP failures when node level redundancy is enabled.

2.2. Extensibility

- . "CUPS protocol" MUST support exchange of software version and feature capabilities when a node level association is setup between a CP and UP.
- . "CUPS protocol" MUST encode information in messages as TLVs.
- . "CUPS protocol" MUST allow extension to defined Information Elements (IEs) i.e. it MUST allow adding new information to existing IEs while maintaining backwards compatibility.
- . "CUPS protocol" MUST allow addition of new IEs exchanged in protocol messages.
- . "CUPS protocol" MUST support vendor specific IEs (modelled as TLVs) by carving out TLV space for vendor specific extensions.

- . "CUPS protocol" processing on UP MUST support graceful handling when an unknown TLV is received. The UP MUST ignore unknown TLV and continue with normal message processing. This ensures the CP MAY send non-mandatory TLVs to the UP. However, CP MUST only send mandatory TLVs if it knows the UP will accept it (based on local configuration or based on capability exchange during association setup). A TLV is considered mandatory if session state cannot be installed or updated without it.

2.3. Scalability and Performance

- . A single CP VNF can control multiple UP nodes. Each UP can support its maximum scale of subscriber sessions as allowed by its data-plane. External control plane running as a VNF can horizontally scale-out as needed with the growth in CUPS system-wide subscriber scale. In typical deployments CP may be centralized whereas the UPs may be distributed, with multiple L2 or L3 hops between CP and UPs. There are scenarios where a large number of sessions may be getting created or deleted close in time via "CUPS protocol". It is important that latency to bring subscribers online is minimized. The transport protocol chosen for "CUPS protocol" MUST NOT suffer from head-of-line (HOL) blocking where transport of messages related to one subscriber can be adversely impacted by messages being exchanged for other subscribers.
- . "CUPS protocol" MUST limit chattiness by minimizing number of messages required to create fully functional subscriber on the UP with complete forwarding, traffic management, HQOS, and routing state. Ideally, a single request/response message exchange between CP and UP should be able to create subscriber with all the required state in the data-plane. The "CUPS protocol" message that creates the subscriber session MUST therefore be able to signal IEs for all the required subscriber state.
- . To further reduce latency the protocol MUST be binary encoded.
- . "CUPS protocol" MUST allow dynamic scale-out for control plane VNF with the growth in subscriber scale of the CUPS system, as more UPs are added to the CUPS system or more ports are enabled on a UP in a CUPS system.

- . The "CUPS Protocol" MUST allow mechanism to provide balancing of processing load amongst compute resources of control-plane VNF that supports dynamic scale-out.
- . "CUPS protocol" SHOULD support signaling of overload state and optionally overload mitigation parameters from UP to CP, when UP determines the incoming signaling from CP is exceeding (or about to exceed) its nominal processing capacity. Overload mitigation can include a temporary message throttling on CP towards UP. Mitigation parameters can include message rate and validity time for the specified rate.

2.4. Transport Protocol

- . As mentioned in section 2.3, the transport protocol used for "CUPS protocol" MUST NOT suffer from HOL blocking. Therefore, TCP is not an option for the transport protocol.
- . Ideally, the transport protocol SHOULD preserve message boundary with datagram semantics and should be available or easily implementable on any simple forwarding devices. Therefore, UDP is the preferred option.
- . "CUPS protocol" MUST therefore support reliability and ordering for exchanged messages. The reliability and ordering can be based on request/response with message sequencing and re-transmissions.

2.5. In-band Control Channel Requirements

- . "CUPS protocol" MUST support setting up of control channel between UP and CP for transporting in-band control messages (e.g. DHCPv4/v6 and PPPoE) received on the UP (from CPEs) to the CP, and for return messages sent from CP to the UP (destined to CPEs).
- . There can be a L3 network between CP and UPs. Therefore, L3 tunneling is required between CP and UP to carry messages for in-band control plane protocols. "CUPS protocol" MUST support exchange of tunnel identifiers between CP and UP.

- . Because L2 access setup is in-band, control plane messages will arrive on the UP before any per-session state is learned. Therefore, "CUPS protocol" MUST support messages and information exchange to install forwarding state related to in-band control plane messages that do not match any existing subscriber session. These messages should be forwarded to the CP over a common default control channel.
- . The in-band control channel setup by "CUPS protocol" MUST have support for UP to pass access-circuit identifier over which the signaling messages are received from the CPEs. Based on type of access, access-circuit identifier can include port/VLAN tags or tunnel identifiers which includes tunnel endpoint IPs and de-multiplexers such as GTP TEID, MPLS labels, L2TP tunnel-id etc. "CUPS protocol" MUST support setting up logically separate control channels for in-band control messages per access-circuit.
- . In case of fixed-access CPEs with Ethernet based network between access-nodes and BNG, the control messages are received in Ethernet frames. The Ethernet frame carrying the control messages received on UP MUST be carried over the control channel to the CP, as outlined in [CUPS]. In case of fixed-wireless access, control messages (e.g. DHCPv4 and DHCPv6) are received on the UP over GTP-u tunnel from the RAN. The GTP-u tunnel directly carries IP payload. Therefore, control channel setup via "CUPS protocol" MUST support transporting both Ethernet and IP payloads.
- . "CUPS protocol" MUST provide support for CP to specify the control protocols that should be forwarded by the UP over in-band control channel to the CP.
- . The "CUPS protocol" SHOULD have support for CP to specify rate-limits for specific control protocols and optionally specific messages within a control protocol, that the UP should enforce.
- . The "CUPS protocol" SHOULD provide support for CP to direct the UP to drop certain control messages received on a particular access-circuit.

- . The "CUPS protocol" SHOULD provide support for CP to prioritize reception of certain control messages over others.

2.6. Resiliency

- . "CUPS protocol" MUST allow support for both 1:1 (hot standby) and N:M (warm standby) UP node level redundancy.
- . "CUPS protocol" MUST provide support for CP to specify the "redundancy domain" that a subscriber session is associated with during session level state creation on the UP. The "redundancy domain" is set of resources that share fate with respect to switchover on failure, e.g. a set of VLANs on a port, or a set of ports on a UP, or entire UP. "CUPS protocol" MUST also provide support for CP to provide relevant parameters to UP about the "redundancy domains". The UPs can then locally perform failure detection and switchover for the redundancy domains.
- . The "CUPS protocol" MUST provide support for UP to notify the CP about switchover event. This notification must be on the granularity of "redundancy domain" on a UP.
- . For warm standby redundancy, "CUPS protocol" MUST provide support for CP to create session level state on the backup UP node(s) for all subscribers associated with the impacted "redundancy domain".
- . "CUPS protocol" MUST support in-service software upgrade (ISSU) on UPs. The protocol MUST provide support for UP to notify CP when it is completed ISSU to the new software release.

2.7. Security

"CUPS protocol" MUST be compatible with proven security mechanisms such as IPSEC or DTLS to satisfy following security requirements:

- . Data-integrity and confidentiality MUST be ensured for the information exchanged via "CUPS protocol".
- . Protection against man-in-the-middle attacks MUST be provided.
- . Anti-replay protection MUST be provided.

3. "CUPS protocol" candidate

3GPP has defined PFCP (Packet Forwarding Control Protocol) in [TS29244] as the interface between CP and UP for LTE gateways. This protocol is suited for large scale state management between CP and UP and can be extended for BNG providing converged access. The protocol provides a good base for satisfying the requirements outlined in this draft for BNG "CUPS protocol". Following are some of the key attributes of this protocol/

- . It supports management of forwarding and QOS enforcement state on the UP from CP.
- . It also supports usage reporting from UP to CP.
- . It is over UDP transport and doesn't suffer from any HOL blocking.
- . It provides reliable operation based on request/response with message sequencing and retransmissions.
- . It provides support for graceful handling of overload on UP.
- . The protocol is extensible and allows addition of new IEs.
- . For fixed access BNG, the protocol requires simple extensions in the form of additional IEs. The required extensions are mainly due to fact that typically a fixed access BNG requires tighter control over L2 behavior and manages access and subscriber using L2 identifiers (such as VLANs and MAC

addresses), whereas mobile access works in terms of L3, either routed or tunneled.

- . [TS29244] also describes an in-band signaling channel based on GTP-u tunnel between CP and UP. GTP-u (GPRS Tunneling protocol User Plane) is defined in 3GPP [TS29281] and defines a tunneling protocol which carries IP payloads. The protocol runs over a UDP/IP stack and uses UDP port number 2152. Data within a tunnel can be multiplexed based on Tunnel Endpoint Identifiers (TEIDs). The protocol supports optional sequence numbers. The protocol supports extension headers to allow development of new features. GTP-u tunnels are signaled between CP and UP, and it is possible to associate filters to forward or block certain control packets from UP to CP. The payload type carried by GTP-u can be extended to Ethernet (via payload type in extension header). The tunnel encapsulation can also be extended by introducing an additional NSH (network services header) to carry any required meta-data.

4. Security Considerations

For security between CP and UP, Network Domain Security (NDS) as defined in [TS33210] can be considered. As per NDS, the network can be split into security domains. Communication within a single security domain is considered secure, and protocols can operate without any additional security. When communication has to cross security domains, then IPSEC can be used.

5. Management Interface Requirements

- . The CP MUST provide a single point for management of "CUPS BNG" system to the operator.
- . Management interface for the CUPS system MUST provide support for both configuration of UPs, and state retrieval. The interface MUST minimally support BNG specific configuration and state.
- . Management interface SHOULD support transactional configuration from CP to UPs, based on a well-defined data schema. Transactional configuration may be achieved by editing a candidate configuration on the UP which is subsequently activated (commit) or by providing the whole transaction in a

single command. In case UP data-stores are used, it MUST be possible for the CP to lock a data-store for exclusive access.

- . The management interface SHOULD support transaction confirmation, where an unconfirmed transaction gets reverted automatically after a timeout even if the transaction succeeded. This is to avoid configuration errors where a valid configuration breaks communication between UP and CP, requiring on-site intervention.
- . The management interface SHOULD support state retrieval based on a well-defined data schema. This includes retrieval for any state that is not signaled via the state control interface.
- . The management interface SHOULD support unsolicited signaling of state changes (events) from UP to CP i.e. SHOULD provide telemetry for events. Even while state changes are sent unsolicited, the CP SHOULD be able to subscribe to a specific subset of state it is interested in.
- . The management interface MUST provide security through an existing mechanism such as (D)TLS or IPSEC to guarantee confidentiality and authenticity and protect against replay and man in the middle attacks.

6. IANA Considerations

None.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [CUPS] Wadhwa, S. et al., "Architecture for control and user plane separation on BNG, July 2019.
<https://datatracker.ietf.org/doc/draft-wadhwa-rtgwg-bng-cups/>

- [TS29244] 3GPP, "Interface between the Control Plane and the User Plane Nodes", TS 29.244 15.2.0, June 2018, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3111>.
- [TS29281] 3GPP, "General Packet Radio System (GPRS) Tunneling Protocol User Plane (GTPv1-U)", TS 29.281 15.3.0, June 2018, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=1699>.
- [TS33210] 3GPP, "Network Domain Security (NDS); IP network layer security", TS 33.210 15.0.0, June 2018, <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2279>.

7.2. Informative References

- [RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <https://www.rfc-editor.org/info/rfc2131>.
- [RFC2516] Mamakos, L., Lidl, K., Evarts, J., Carrel, D., Simone, D., and R. Wheeler, "A Method for Transmitting PPP Over Ethernet (PPPoE)", RFC 2516, DOI 10.17487/RFC2516, February 1999, <https://www.rfc-editor.org/info/rfc2516>.
- [RFC3315] Droms, R., Ed., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, DOI 10.17487/RFC3315, July 2003, <https://www.rfc-editor.org/info/rfc3315>.

Authors' Addresses

Sanjay Wadhwa
Nokia
777 East Middlefield Road
Mountain View
USA

Email: Sanjay.wadhwa@nokia.com

Killian De Smedt
Nokia
Copernicuslaan 50
Antwerp
Belgium

Email: Killian.de_smedt@nokia.com

Praveen Muley
Nokia
805. E. Middle Field Rd.
Mountain View, CA, 94043
USA

Email: praveen.muley@nokia.com

Rajesh Shinde
Reliance Jio Infocomm Ltd.
Reliance Corporate Park
Thane Belapur Road, Ghansoli
Navi Mumbai 400710
India

Email: Rajesh.A.Shinde@ril.com

Jonathan Newton
Vodafone
Waterside House
Bracknell
United Kingdom

Email: jonathan.newton@vodafone.com

Ryan Hoffman
TELUS
1525 10th Ave SW
Calgary, Alberta
Canada

Email: ryan.hoffman@telus.com

Subrat Pani
Juniper Networks
10 Technology Park Dr.
Westford, MA
USA

Email: spani@juniper.net