Interdomain Routing Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 24, 2014

N. Leymann, Ed.
C. Heidemann
Deutsche Telekom AG
X. Li
Huawei

October 21, 2013

GRE Notifications
draft-heileyli-gre-notifications-00

Abstract

GRE (Generic Routing Encapsulation) specifies a protocol for the encapsulation of an arbitrary protocol over another arbitrary network layer protocol.

This document describes extensions to manage multiple GRE tunnels over multiple access lines to one home network with the purpose to present a novel architecture using Hybrid Access (HA) networks. HA is designed to bundle multiple access technologies, e.g. fixed access and wireless access to one Internet connection. This enables higher bandwidth for end customers. The document describes the Hybrid Access network architecture and th extensions for GRE which are necessary to implement the HA architecture.

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 24, 2014.

Copyright Notice

Copyright (c) 2013 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

This document specifies a new GRE extension which allows the operators to have home networks that consist of at least two IP access lines to the same location from one network gateway. GRE tunnels are used to combine the multiple access lines together to one Internet connection. The combination from e.g. a wireless access (LTE) and a fixed line (DSL) access enables new use cases:.

1. Bandwidth on demand, if fixed line is full, bandwidth of mobile access is added on demand

2. Seamless handover, if one access lines fails the service can still be provided without interruption

3. Application dependent transport of different combined or non-combined access lines to one home network location

This extension contains signaling information to manage those multiple GRE connections over the multiple access lines. The lines can have different weight and qualities and have to be merged to one authorized connection. Therefore a signaling is needed for the following cases:

1. Signaling of the IP addresses of the access lines and IP addressed of the combination GRE tunnels

2. Signaling of the weight of one access line, a cheaper access line should be used

3. Signaling of keep alive to decide which GRE tunnels are active and can be used

4. Signaling of bypass traffic amount which should be bypassed from the tunnels

5. Signalling of deny and allowed messages

2. Use Cases

Recently, the existing network status can bringing changes to the operator's network: Higher bandwidth requirement and current limited usage of existing networks(e.g.wireless network: LTE, etc). There is a strong interest to integrate the existing network resource as a single Internet connection for end customers. The resulting network is called a Hybrid Access (HA) network. The HA network can be controlled by the user's Customer Premise Equipment (CPE). The connectivity in HA is being implemented by using a tunnel mechanism on top of the physical infrastructure.

This document described the HA architecture by illustrating DSL and LTE bundled. Nevertheless the solution is not limited to those technologies and can be easily applied to other scenarios. This document describes the base HA network architecture, while solution approach with extensions of GRE protocol [RFC2890] will realize HA network architecture.

With HA many deficiencies in the current operator's network, following use cases are enabled.

1. Bandwidth On Demand

Typically end customers are only connected using a single link (e.g. fixed or wireless) to the network of a service provider. If the required bandwidth exceeds the bandwidth provided by the link usually the link needs to be upgraded. However, using the current network deployment model such an upgrade is not always economically feasible. A different approach is needed.

In this model end customer are connected over one standard access line (e.g.fixed line: DSL, Cable, ...) primarily to the operator's service. In addition there are one or more another connections over the different access technologies, (e.g. wireless line: LTE). If the traffic exceeds the bandwidth of the primary access line the bundled connections can provide a higher bandwidth to the end customer.

2. Seamless Handover

In HA network, the customer has a CPE connection through the access network (e.g.fixed line: DSL, Cable, ...). The CPE also provides a back-up WAN connection through another network (e.g. wireless line: LTE), when the fixed line is unavailable.

The customer is using Internet service via fixed WAN connection. When the fixed connection gets disconnected unexpectedly, the ongoing service is automatically switched to the backup connection. The Internet service is not interrupted by this event and continues seamlessly by end-user.

3. Conventions and Terminology

3.1. Terminology

Bonding Tunnel: The bundling tunnel of both fixed access tunnel and wireless access tunnel. The bonding tunnel in HA is the connection on top of the physical infrastructure, terminated between CPE and HAAP.

Tunnel Transit IP: The outer IP of a GRE encapsulation.

DSL Tunnel: The GRE tunnel between CPE DSL WAN and HAAP. The tunnel transit IP is IP address of CPE DSL WAN interface and HAAP address. It is one subset tunnel of bonding tunnel.

Dynamic GRE: The dynamic stateful GRE tunnel.

Customer Premise Equipment (CPE): A device that connects multiple terminals to provide connectivity to the service providers network.

Hybrid Access (HA): Hybrid Access (HA) is the bundling of two or more access line over different technologies (e.g. DSL and LTE) to one Internet connection for end customers.

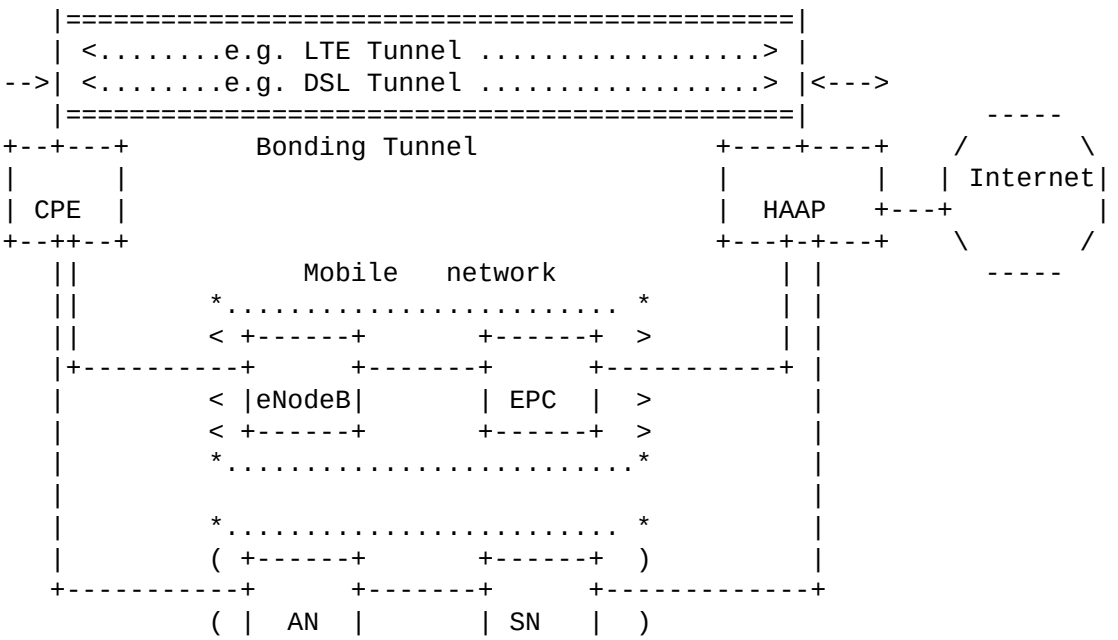
Hybrid Access Aggregation Point (HAAP): The HAAP which acts as a service termination and a service creation implements bonding mechanism and sets up a high speed Internet dual stack IP connection with CPE on top of two or more different access technologies.

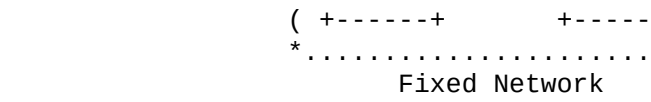
HA IP: The inner IP of a GRE encapsulation. This IP is assigned by HAAP to CPE; this is the IP for the Internet communication. Sometimes it is called tunnel IP in this document.

LTE Tunnel: The GRE tunnel between CPE LTE WAN and HAAP. The tunnel transit IP is IP address of CPE LTE WAN interface and HAAP address. It is one subset tunnel of bonding tunnel.

4. Hybrid Access Network Architecture

The basic idea of Hybrid Access is the bundling of the DSL and LTE access technologies. Figure 1 illustrates one example of Hybrid Access network.





Legend:

AN	Access Node
CPE	Customer Premise Equipment
SN	Service Node
EPC	Evolved Packet Core
HAAP	Hybrid Access Aggregation Point

Figure 1: Hybrid Access Network Architecture

In the fixed network, users are served fixed services by the Access Node (AN) and Service Node (e.g. Broadband Network Gateway (BNG)). In the wireless network, cellular sites are connected to the mobile core network using mobile backhaul network and EPC core network. The new approach of Hybrid should take into consideration the fact that operators introduces additional network bandwidth resource with limited usage to users.

In the HA architecture, on the client site, CPE is used to implement the bonding mechanism for customers. On the network side, a device named as Hybrid Access Aggregation Point (HAAP) MUST be deployed. The HAAP which acts as a service termination and a service creation implements bonding mechanism and sets up a higher speed Internet dual stack IP connection with the CPE on top of both access technologies a.k.a., DSL and LTE. The HA connection between the end user’s CPE and the HAAP could be done by using the tunnel mechanism on top of the physical infrastructure. This document describes the extensions for dynamic GRE which are necessary.

The bonding tunnel between CPE and HAAP carries best effort traffic going to and coming from the public Internet. Particularly, based on the operator’s requirement, it is possible that not all traffic from the home network is routed into the bonding tunnel in order to ensure that existing services are not influenced by using HA. Certain traffic such as VoIP, IPTV traffic depending on its destination or QoS markings, needs to be routed via the fixed line interface or via the wireless interface instead to be routed into the bonding tunnel between CPE and HAAP. The CPE should implement an mechanism which can be used to specify exceptions (traffic which should not be routed into the tunnel). This mechanisms is out of scope of this document.

5. Solution Approach Overview

The bonding tunnel behavior is accomplished by implementing dynamic subset tunnels setup and bonding them together during the procedure. This section identifies the HA solution approaches that operators can leverage for deploying HA technologies, which is dynamic Generic Routing Encapsulation (GRE).

5.1. Dynamic GRE Definition

The dynamic GRE protocol is specified for encapsulation of the user traffic over multiple arbitrary network layer via bundling mechanism on CPE and HAAP. This section describes dynamic GRE protocol.

The dynamic GRE protocol transport layer carries GRE encapsulated Control messages, and GRE encapsulated Data messages. GRE Data messages encapsulate forwarded user frames. GRE Control messages are management messages exchanged between a CPE and a HAAP in HA architecture. The format of GRE protocol Control are defined in section 7.

The dynamic GRE protocol begins with a base access phase. CPE gets DSL WAN interface IP address through PPPoE from service node (e.g., BNG) or DHCP and LTE WAN interface IP address through Packet Data Protocol (PDP)[TS23.401] from PGW. Additionally, CPE obtains HAAP address for tunnel establishment. From the base access phase, a CPE discoveries the HAAP with which to establish the tunnels for HA.

Once the base access have be completed, GRE Request is initiated by CPE in order to begin bonding tunnels setup phase between CPE and HAAP. CPE setups the authorized LTE GRE tunnel before DSL GRE tunnel by sending GRE Request control messages via LTE WAN interface to HAAP. After, CPE obtains HA IP address from HAAP through DHCP over LTE GRE tunnel. Subsequently, authorized DSL GRE tunnel is established. During these exchanges, the CPE may receive some information in order to enable both tunnel bundled. GRE Accept/Deny identifies that GRE tunnel setup request is accepted/rejected.

When the CPE and HAAP have completed the bonding tunnels setup exchange, the customers have the single service connection through both access technologies infrastructure. Particularly,the specific traffic will send through the bonding tunnels and thus encapsulated by GRE. As long as the primary connection (e.g., DSL) is sufficient, traffic goes through the DSL GRE tunnel only. If traffic exceeds the bandwidth, traffic overflows to the LTE GRE tunnel.

The dynamic GRE also provides commands exchange between CPE and HAAP for HA management. These may be included in GRE Notify message for

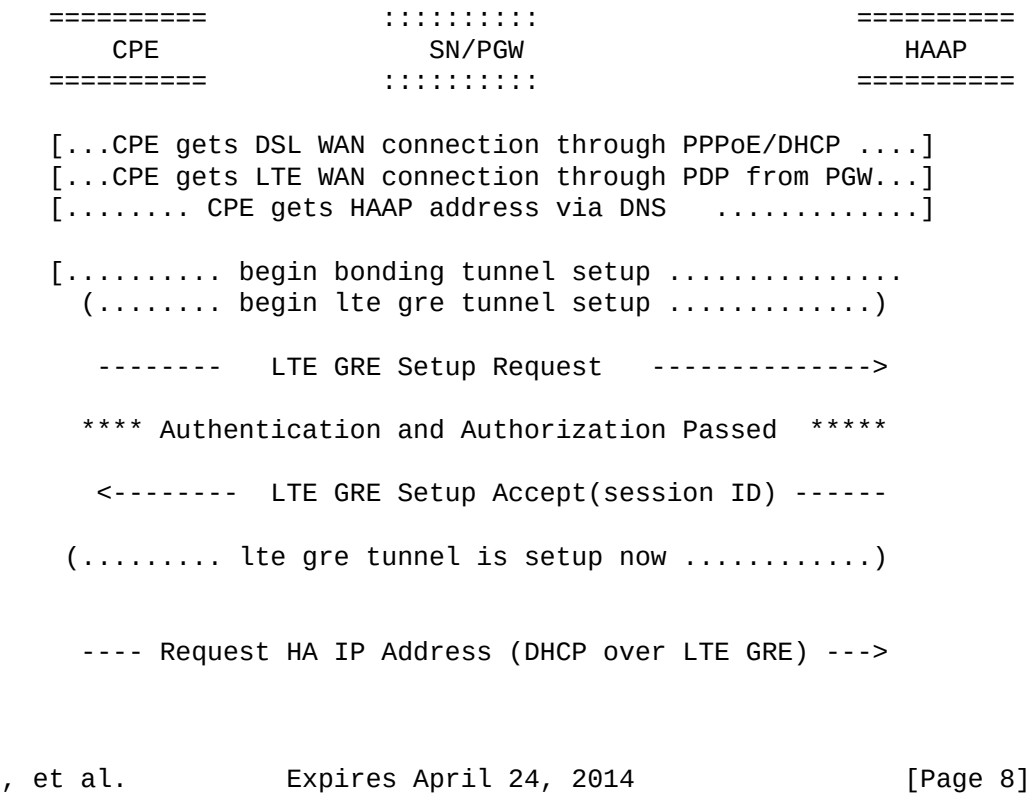
tunnel status/information changing between CPE and HAAP. These may include the bypass traffic amount which should be bypassed from the bonding tunnels.

The dynamic GRE protocol provides for a keep-alive feature that preserves the communication channel between the CPE and HAAP. If the tunnels fail to appear alive, the CPE will try to re-establish it. For example, if the DSL tunnel cannot be re-established, HA traffic will still run through the LTE tunnel only. If the LTE tunnel cannot be re-established, new Internet sessions will be established over native DSL. The DSL tunnel will be finally torn down after a period without service interruption.

For maintenance reasons, the GRE Tear Down message also can be used by CPE and HAAP to complete the HA architecture in out of service scenario.

5.2. Bonding Tunnel Establishement Overview

This section describes the bonding tunnel establishment process message exchanges between CPE and HAAP. The annotated ladder diagram shows the CPE on the left, the HAAP on the right. The dynamic GRE state mechanism is defined in detail in Section 6. Note that in the Authentication step, the authentication required certain messages are aggregated into a single step, which is denoted via an asterisk line in Figure 2



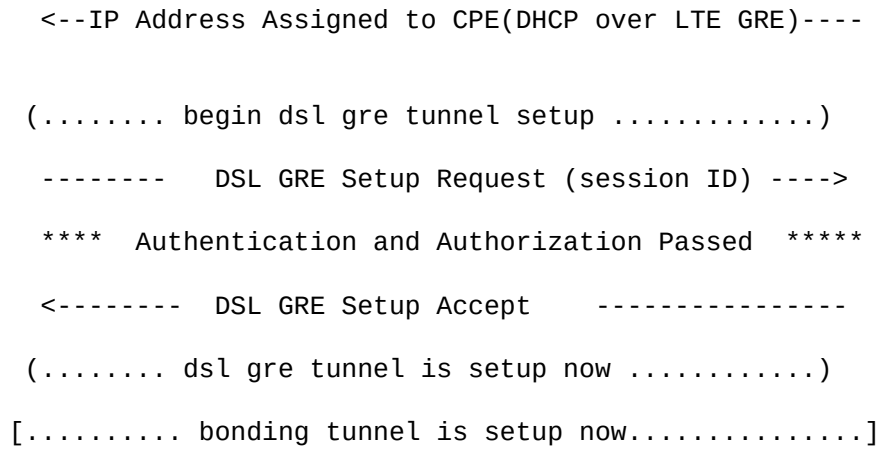


Figure 2: GRE Tunnel Establishment

At the end of the illustrated GRE control messages exchange, the bonding tunnel between CPE and HAAP is setup completely by binding LTE Tunnel and DSL Tunnel with same session ID, defined in Section 7.2.

After, the CPE and HAAP are securely exchanging GRE Control messages for tunnel keepalive (GRE Hello) and management (GRE Notify).

The GRE Notify message is used to inform status/information changing information between CPE and HAAP. A notify acknowledgement (ACK) and retransmission mechanism can be used to provide certain level reliable transport capability. When receiving end receives a notify packet, it will send back a GRE notify packet without any attributions appended to the sending end immediately as acknowledgement for the notify packet. When sending end doesn't receive the notify ACK message in after a specific seconds, sending end treats it as lost of notify message and will retransmit the notify message.

When sending end not receiving the notify ACK for a certain notify message for continually specific times, sending end treats it as sending failure and tunnel failure also, sending end will tear down the GRE tunnel which sent the notify message. If the CPE is the sending end, the CPE will tear down the tunnel over which the notify packet was send, and try to re-establish the tunnel. If HAAP is the sending end, the HAAP will tear down the corresponding GRE tunnel and wait for CPE to reestablish it.

This illustration is provided to clarify the protocol operation, and does not include any possible error conditions and all the control packets. Section 6 provides a detailed description of the corresponding state machine.

6. Dynamic GRE State Machine Definition

The following state diagram (Figure 3) represents the life cycle of HA bonding tunnel.

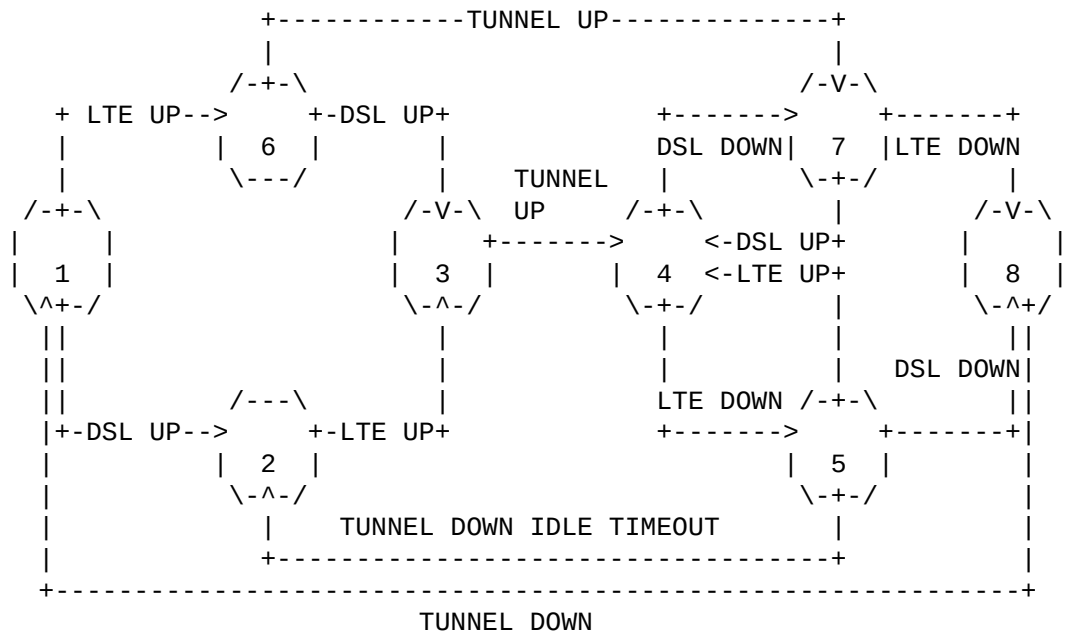


Figure 3: GRE State Machine

The various states are described as below:

State No.	DSL Tunnel	LTE Tunnel	Bonding Tunnel
1	Down	Down	Down
2	Up	Down	Down
3	Up	Up	Down
4	Up	Up	Up
5	Up	Down	Up
6	Down	Up	Down
7	Down	Up	Up
8	Down	Down	Up

Tunnel / GRE States

7. Dynamic Packet Format

This section describes GRE encapsulated control messages definitions and attributes which can be optionally carried in the GRE control messages.

7.1. Dynamic GRE Control Messages

The GRE control messages are defined according to [RFC2890]. The proposed GRE header of the control messages has the following format (see Figure 4):

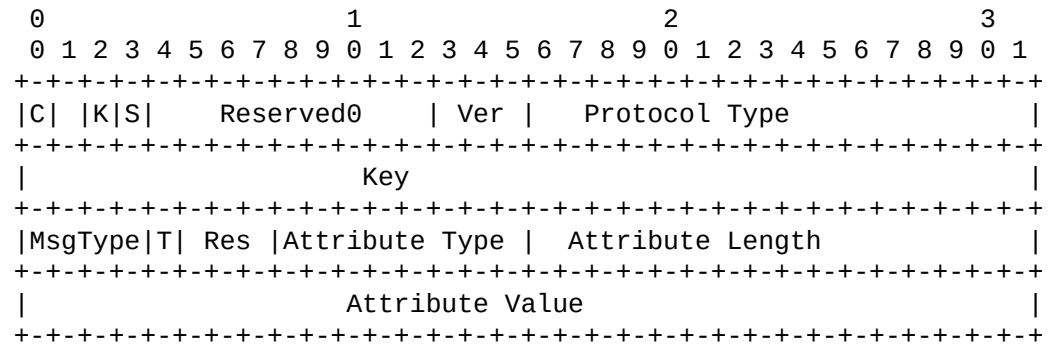


Figure 4: GRE Header Format

Protocol Type (2 octets)

The Protocol Type field identifies the dynamic GRE protocol. The value is TBD.

Message Type (MesType) (4 bits)

The Message Type field identifies the dynamic GRE protocol control messages in the HA network. The value is TBD. The existing control message types are listed below. Additional values may be defined in the future.

Control Message Family	Type
GRE Setup Request	1
GRE Setup Accept	2
GRE Setup Deny	3
GRE Hello	4
GRE Tear Down	5
GRE Notify	6
Reserved	0,7-15

Figure 5: GRE Control Messages

Tunnel Type (T) (1 bit)

It indicates this control message for the type of the subset tunnel in HA. For example, if the Tunnel Type (T) bit is set to 1, then this control message is for the DSL tunnel shown in Figure 5. Otherwise it indicates that this is for the LTE tunnel shown in Figure 5.

Attribute Type (1 octet)

The Attribute Type indicates the type of the appended attribution in the control message. The attribute value pair is defined in Section 7.2.

Attribute Length (2 octets)

The Attribute Length field indicates the length of the attribute by byte.

Attribute Value (variable)

The Attribute Value field includes the value of the attribute.

7.2. Dynamic GRE Protocol Messages Attributes

This section defines the attributions that are included in dynamic GRE protocol control messages. Attributions are used to carry information needed during bonding tunnel setup and management procedure. Every attribution is identified by the Type, Length, Value field. All of the message attributions in this document use the same format, shown as below in Figure 6.

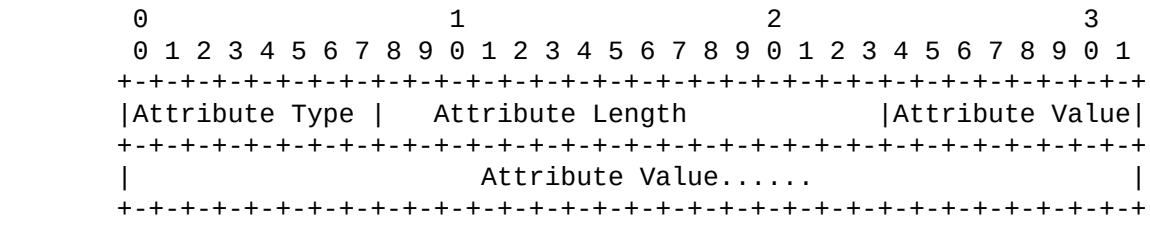


Figure 6: GRE Message Attributes

The 8-bits Type field identifies the type of the appended attribution carried in Attribute Value field in the control messages header. The type field values are allocated right now listed in this section as follows:

o Session ID

This is 32 bits value is generated by HAAP, and unified within a HAAP, to identify a certain subscriber. This value is using to binding the DSL tunnel and LTE tunnel together for individual HA user. HAAP generates a session ID for a HA user's and sends this value to CPE via LTE GRE setup accept, then CPE carries this value in the DSL GRE setup request, defined in section 4.2. With this information, HAAP binds the DSL tunnel and LTE tunnel together, then bonding GRE tunnel is achieved accordingly. When LTE recovery from failure with DSL tunnel exists, the re-establish LTE tunnel request need carry the Session ID attribute.

Type: 4 for Session ID

Length: 4 Octets

Value: 32 bits value generated by HAAP.

o Bypass Traffic Rate

This attribute is used to notify HAAP of the downstream bypass traffic on CPE via DSL Notify control message from CPE. HAAP will calculate the available DSL bandwidth for DSL GRE tunnel based on this information. CPE and HAAP can decide the bypass traffic amount which should be bypassed from the combination tunnels. The unit of this value is kbps.

Type: 6 for Bypass traffic rate

Length: 4 Octets

Value: Downstream bypass traffic on CPE.

o Hello Interval

This is the configuration which is assigned to the CPE by the HAAP. Configure the Hello signaling checking period on CPE from HAAP. The unit of this value is second. This attribution is carried in GRE Setup Accept control message for LTE tunnel.

Type: 14 for Hello Interval

Length: 4 Octets

Value: Hello Interval

8. Overflow Bonding Operations

In HA network, the user traffic can be transferred to wireless access (Overflow tunnel) when the fixed line (Primary tunnel) bandwidth is not any more sufficient.

There are two types of overflow bonding mechanisms, packet-based balancing and stream-based balancing. Balancing per stream can require the DSL bandwidth threshold configuration. The steam only runs over the DSL or LTE link. Balancing of per packet will use DSL bandwidth as soon as DSL bandwidth is not full. It is possible that the same stream can run the different DSL and LTE link at the same time. In HA network, packet-based balancing is proposed for efficiency.

The packet-based overflow balancing is based on the Single Rate Three Color Marker (srTCM) and Two Rate Three Color Marker (trTCM) which is defined in[RFC2697] and [RFC2698].The packet is marked if the packet is overflowed or not. The CIR (Committed Information Rate) is equal to the DSL bandwidth minus several layers' overhead. The CBS (Committed Burst Size) provides the burst capability which can help TCP achieve the committed bandwidth. This mechanism is used on both HAAP for downstream overflow bonding and CPE for upstream overflow bonding.

Then the colored based policy routing is executed for packet-based balancing, user's packet will be routed into the corresponding tunnel based on color. For example,Yellow color packet will be routed to LTE GRE tunnel; green color packet will be routed into DSL GRE tunnel. At this stage, the GRE IP header will be added.

Additionally, during the packet-based balancing, reorder mechanism are need for both HA downstream and upstream. On the downstream, the packets encapsulated in GRE will come from DSL GRE tunnel and/or LTE GRE tunnel. The packets will be sent to a buffer for reordering. If the GRE sequence number is not continuous, the packets will be buffered until the missing sequence packet has arrived or the buffering time has expired. After reordering, the GRE header will be removed and the packet will be sent to the ordinary CPE processing. Upstream direction reordering is performed on HAAP using the same mechanism as downstream.

In order to ensure that existing services are not influenced by using HA, it is possible that certain traffic does not be routed through the tunnel, but directly over the corresponding interfaces. This is necessary in case the tunnel and the HAAP are not supporting QoS (e.g. for IPTV or VoIP services) and Multicast (for IPTV). Some customers delay-sensitive traffic (like Internet gaming) need to be

sent through the fixed line interface to ensure quality of experience depending on customer requirements.

This bypass behavior is accomplished by implementing a routing table which routes traffic which needs to bypass the tunnel through its relevant interfaces. For example, IPTV related traffic might be routed over the fixed line interface to ensure the use of QoS.

9. IANA Considerations

IANA is requested to allocate one code TBD for the dynamic GRE protocol.

10. Security Considerations

In the whole processing of HA, security of control messages MUST be guaranteed. The CPE discovers the HAAP by resolving the HAAP address over DNS. This protects the CPE against connections to foreign HAAP, if the DNS service and the domain name in the CPE isn't corrupted.

The CPE should be prevented against receiving GRE notifications without a valid session In the whole processing of end to end HAAP session establishing and GRE notification signaling, the source IP address for session establishment from CPE MUST be strictly verified, including IP address authentication and identification at the HAAP side. Any authentication mechanism with credential or checking the IP address is feasible.

GRE notification key poisoning Every new session at the HAAP generates a magic number, which is encapsulated in the key field of the GRE header and will be carried in the signalling messages and data traffic for verification by comparing the Magic Number in the message and the Magic Number in the local session table. Traffic without a valid Magic Number and outer IP address will be discarded on the HAAP. Magic number is used for both control message and data message security.

For data traffic security, it is also proposed to use IP address validation to protect against IP Spoofing attacks.

11. Acknowledgements

Many thanks to Dennis Kusidlo.

12. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2697] Heinanen, J. and R. Guerin, "A Single Rate Three Color Marker", RFC 2697, September 1999.

[RFC2698] Heinanen, J. and R. Guerin, "A Two Rate Three Color Marker", RFC 2698, September 1999.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, September 2000.

[TS23.401] , "3GPP TS23.401, General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access", September 2013.

Authors' Addresses

Nicolai Leymann (editor)
Deutsche Telekom AG
Winterfeldtstrasse 21-27
Berlin 10781
Germany

Phone: +49-170-2275345
Email: n.leymann@telekom.de

Cornelius Heidemann
Deutsche Telekom AG
Heinrich-Hertz-Strasse 3-7
Darmstadt 64295
Germany

Phone: +4961515812721
Email: heidemannc@telekom.de

Xue Li
Huawei
NO.156 Beiqing Rd. Z-park, Shi-Chuang-Ke-Ji-Shi-Fan-Yuan
Beijing, HaiDian District 100095
China

Email: xueli@huawei.com