

BANANA
Internet Draft
Intended Category: Proposed Standard

N. Leymann
C. Heidemann
Deutsche Telekom AG
M. Zhang
B. Sarikaya
Huawei
M. Cullen
Painless Security
December 26, 2017

Expires: June 29, 2018

BANdwidth Aggregation for interNet Access (BANANA)
The Control Protocol of Bonding Tunnels
draft-leymann-banana-signaling-02.txt

Abstract

There is an emerging demand for solutions to bond multiple access links to provide subscribers with redundancy and load-sharing across these access links. BANdwidth Aggregation for interNet Access (BANANA) will specify such solutions.

In this document, a control protocol is specified to deliver configuration and control information between two peering BANANA boxes.

Status of this Memo

This Internet-Draft is submitted to IETF 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/lid-abstracts.html>

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

INTERNET-DRAFT

BANANA Signaling

December 26, 2017

Copyright and License Notice

Copyright (c) 2017 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	2
2.	Acronyms and Terminology	3
3.	The Single Operator Scenario	5
4.	Addressing	6
5.	Control Protocol Specification	7
5.1.	Message Formats	7
5.2.	Establishment of Bonding Tunnels	10
6.	The Edge to Edge Scenario	12
7.	Security Considerations	13
8.	IANA Considerations	13
9.	References	14
9.1.	Normative References	14
9.2.	Informative References	14
	Contributors	14
	Authors' Addresses	15

[1.](#) Introduction

Service Providers used to provide subscribers with separate access to their multiple networks. It has become desirable to bond access links to these networks together to offer access service to subscribers. When the traffic volume exceeds the bandwidth of the first connection, the excess amount can be offloaded to a secondary connection. So bonding service is able to provide subscribers with increased access capacity and improved reliability.

This memo will mainly work two scenarios out: the single operator scenario and the edge to edge scenario. There would be mainly implementation issues to make BANANA be applicable to more scenarios. For the single-operator scenario, the local BANANA box is a Customer

Premises Equipment (CPE) device initiating the two connections. The remote BANANA box resides in the provider's networks to terminate these bonded connections. For the edge to edge scenario, the two peering BANANA boxes are two CPE devices which might be operated by different providers.

This document specifies the control protocol between the two BANANA boxes. This control protocol adopts GRE (Generic Routing Encapsulation [[RFC2890](#)]) since GRE is widely supported in various networks. Approaches specified in this document might also be used by other tunneling technologies to achieve tunnel bonding. However, such variants are out of scope for this document.

For each heterogeneous connection between the two BANANA boxes, one GRE tunnel is set up. The local and remote BANANA box, respectively, serve as the common termination point of the tunnels at both ends. Those GRE tunnels are bonded together to form a logical IP link for the subscriber. This provides an overlay: the users' IP packets (inner IP) are encapsulated in GRE, which is in turn carried over IP (outer IP).

A tunnel bonding solution of BANANA may support more than two tunnels between the two BANANA boxes though the reference topologies in this memo choose to use two tunnels between the two BANANA boxes to depict such a solution.

[2.](#) Acronyms and Terminology

GRE: Generic Routing Encapsulation [[RFC2890](#)].

BRAS: Broadband Remote Access Server. Routes traffic to and from broadband remote access devices such as Digital Subscriber Line Access Multiplexers (DSLAMs) on an Internet Service Provider's (ISP's) network.

PGW: Packet Data Network Gateway. In the Long Term Evolution (LTE)

architecture for the Evolved Packet Core (EPC), acts as an anchor for user-plane mobility.

PDP: Packet Data Protocol. A packet transfer protocol used in wireless GPRS (General Packet Radio Service) / HSDPA (High-Speed Downlink Packet Access) networks.

PPPoE: Point-to-Point over Ethernet. A network protocol for encapsulating PPP frames inside Ethernet frames.

DNS: Domain Name System. A hierarchical distributed naming system for computers, services, or any resource connected to the Internet

Leymann, et al.

Expires June 29, 2018

[Page 3]

INTERNET-DRAFT

BANANA Signaling

December 26, 2017

or a private network.

DHCP: Dynamic Host Configuration Protocol. A standardized network protocol used on Internet Protocol (IP) networks for dynamically distributing network configuration parameters, such as IP addresses for interfaces and services.

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](#)].

[illegible]

A6: The endpoint for the IPv6 address that is used for the BANANA service by the Host.

- A4: The endpoint for the IPv4 address that is used for the BANANA service by the Host.
- C: The service endpoint of the bonding service at the local BANANA box. IP address or IP prefix of C is assigned by the DHCP from a pool of IP addresses maintained on the remote BANANA box.
- A1: The endpoint of the Gateway on the same Local Area Network (LAN) as the Host. The pre-configured IP address of A1 will be translated to the IP address of C by the Network Address Translator (NAT).
- F: The local endpoint of the first tunnel (Tunnel 1) at the local BANANA box.
- S: The local endpoint of the secondary tunnel (Tunnel 2) at the local BANANA box.
- R: The common remote endpoint for Tunnel 1 and Tunnel 2 at the remote BANANA box. The DNS server will resolve an URL provisioned by the Service Provider to be the IP address of R.
- I: The endpoint of the destination in the Internet.

Figure 3.1: Tunnel bonding for a single operator

If a Service Provider runs multiple networks, subscribers are eager to use those networks simultaneously for increased access capacity rather than just using a single network. As shown by the reference topology in Figure 3.1, the subscriber expects a significantly higher

access bandwidth from the bonding connection than from just the first connection. In other words, when the traffic volume exceeds the bandwidth of the first connection, the excess amount may be offloaded to the secondary connection.

One tunnel is established per-connection between the two BANANA boxes (see Figure 3.1). These tunnels are bonded together as if there is a single IP link provided between the two boxes for the subscriber who buys the local BANANA box.

Compared to per-flow load balancing mechanisms which are widely used nowadays, BANANA MUST support per-packet offloading approach. For per-flow load-balancing, the maximum bandwidth that may be used by a traffic flow is the bandwidth of an individual connection. While for per-packet offloading, a single flow may use the added-up bandwidth

of all the connections.

Although this memo depicts the tunnel bonding solution using reference topologies (see also [Section 6](#)) with two GRE tunnels between the two BANANA boxes, a tunnel bonding solution can support more than two tunnels between the two BANANA boxes.

[4.](#) Addressing

When the Host boots up, IP addresses of A4 and/or A6 will be assigned by the DHCP from a pool of IP addresses maintained on the local BANANA box. The Gateway IP addresses of A1 is locally configured and will be translated into the IP address of C which is assigned by the DHCP from a pool of IP addresses maintained on the remote BANANA box.

IPv6 address of A6 has the same prefix as C so that NAT function is unnecessary. The DHCP message that carries the IP address of C will be encapsulated as a GRE data packet ([\[BANANA-encap\]](#)) after Tunnel 1 is established.

When the local BANANA box boots up, IP addresses of F and S will be automatically assigned by network devices connected to the local BANANA box. As examples, if this network device is a Broadband Remote Access Server (BRAS), the local BANANA box gets an IP address through the Point-to-Point Protocol over Ethernet (PPPoE). If this network device is a Packet Data Network Gateway (PGW), the local BANANA box gets an IP address through the Packet Data Protocol (PDP).

In order to support automatic establishment of GRE tunnels, the IP address of F or S needs to be carried by the control protocol from the local BANANA box to the remote BANANA box.

The domain name of a remote BANANA box may be configured or obtained via the Wide Area Network (WAN) interface of the first or secondary connection based on gateway configuration protocols such as [\[TR-069\]](#).

The resolution of the remote BANANA box's domain name is requested via the WAN interface of the first or secondary connection. The Domain Name System (DNS) server will reply with the IP address of R which is assigned by DHCP from a pool of IP addresses maintained on the remote BANANA box.

A Service Provider might deploy multiple remote BANANA boxes in one site and place a branch router in front of these remote BANANA boxes.

The DNS server will resolve the URL to a pre-configured IP address of this branch router instead of the IP address of R. In this way, the tunnel setup request from the local box will reach this branch router instead of R. This branch router will adopt anycast mechanism to achieve load balancing and direct the tunnel setup request to one of the remote BANANA boxes. For this case, the IP address of R needs to be carried by the control protocol from the remote BANANA box to the local BANANA box for the purpose of automatic establishment of GRE tunnels.

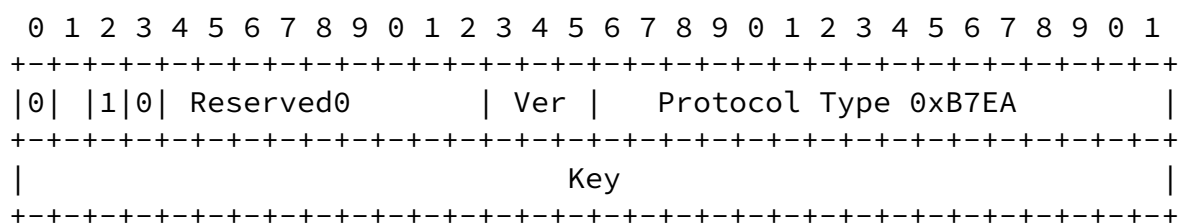
[5.](#) Control Protocol Specification

The control protocol of BANANA is designed to exchange configuration and control information between the two BANANA boxes, such as IP prefixes of local links (see [Section 4](#)), the link properties and status and the information needed by the encapsulations.

[5.1.](#) Message Formats

Messages of the control protocol are delivered as GRE encapsulated packets and routed via the same GRE tunnels as GRE data packets. All control messages are sent in network byte order (high-order bytes first). The GRE Protocol Type field is used to distinguish control packets from GRE data packets. The new GRE Protocol Type (0xB7EA) is allocated for this purpose. GRE packets with a Protocol Type that equals to this number will be consumed by the receiving BANANA box rather than forward further.

The standard GRE header as per [[RFC2890](#)] with Checksum Present bit and Sequence Number Present bit set to zero and Key Present bit set to one is used in this memo. This means the Checksum, the Sequence Number and the Reserved1 fields are not present. So, the format of the GRE header for control messages of is as follows:



The remote BANANA box generates a random number to be carried as the Key field of each control message by the local BANANA box. Except the first GRE Tunnel Setup Request message, the Key field of all control messages originated by the local BANANA box MUST be set to this random number. The remote BANANA box uses the value of the Key to authenticate the local BANANA box.

The general format of the entire control message is as follows:

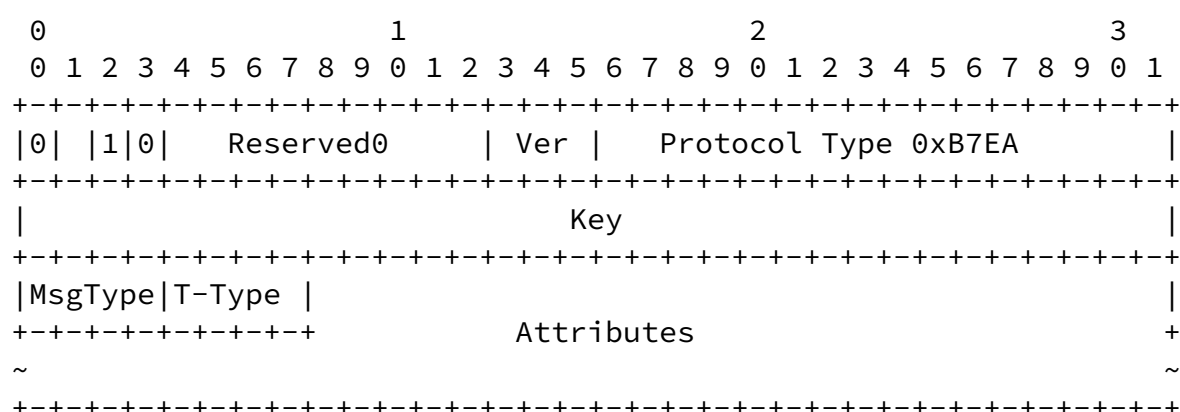


Figure 5.1: Format of Control Messages of GRE Tunnel Bonding

- o MsgType (4 bits)
Message Type. The control message family contains the following six types of control messages (not including "Reserved"):

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

- GRE Tunnel Setup Request
The local BANANA box uses the GRE Tunnel Setup Request message to request that the remote BANANA box establishes the GRE tunnels.

It is sent out from the local BANANA box's F and S interfaces (see Figure 3.1).

- GRE Tunnel Setup Accept

The remote BANANA box uses the GRE Tunnel Setup Accept message as the response to the GRE Tunnel Setup Request message. This message indicates the acceptance of the tunnel establishment and carries parameters of the GRE tunnels.

- GRE Tunnel Setup Deny

The remote BANANA MUST send the GRE Tunnel Setup Deny message to the local BANANA box if the GRE Tunnel Setup Request from this local BANANA box is denied. The local BANANA box MUST terminate the GRE tunnel setup process as soon as it receives the GRE Tunnel Setup Deny message.

- GRE Tunnel Hello

After the first or the second GRE tunnel is established (see Figure 3.1), the local BANANA box begins to periodically send out GRE Tunnel Hello messages via the tunnel; the remote BANANA box acknowledges the local BANANA box's messages by returning GRE Tunnel Hello messages received from the local BANANA box. This continues until the tunnel is terminated.

- GRE Tunnel Tear Down

The remote BANANA box can terminate a GRE tunnel by sending the GRE Tunnel Tear Down message to the local BANANA box via the tunnel. After receiving the GRE Tunnel Tear Down message, the local BANANA removes the IP address of R (see Figure 3.1).

- GRE Tunnel Notify

The two BANANA boxes use the GRE Tunnel Notify message, which is transmitted through either the first or the second GRE tunnel, to notify each other about their status regarding the two GRE tunnels, the information for the bonding tunnels, the actions that need to be taken, etc.

Normally, the receiver just sends the received attributes back as the acknowledgement for each GRE Tunnel Notify message. If the size of the attribute is too large, an acknowledgement attribute for it need to be defined.

- o T-Type (4 bits)

Tunnel Type. Set to 0001 if the control message is sent via the first GRE tunnel. Set to 0010 if the control message is sent via the secondary GRE tunnel. Values 0000 and values from 0011 through

1111 are reserved for future use and MUST be ignored on receipt.

o Attributes

The Attributes field includes the attributes that need to be carried in the control message. Each Attribute has the following format:

```
+---+---+---+---+
|Attribute Type |           (1 byte)
+---+---+---+---+---+---+---+---+---+---+---+---+
| Attribute Length |       (2 bytes)
+---+---+---+---+---+---+---+---+---+---+---+---+
| Attribute Value  ~       (variable)
+---+---+---+---+---+---+---+---+---+---+---+---+
```

- Attribute Type

The Attribute Type specifies the type of the attribute.

- Attribute Length

Attribute Length indicates the length of the Attribute Value in bytes.

- Attribute Value

The Attribute Value includes the value of the attribute.

Specific attributes will be defined in [[BANANA-attributes](#)].

[5.2.](#) Establishment of Bonding Tunnels

One major goal of BANANA signaling is to enable the automatic set up of GRE tunnels which used to be set up manually. After the IP addresses of tunnel endpoints have been acquired, the local BANANA box starts the following procedure to set up the bonding tunnels.

The local BANANA box will firstly set up the secondary GRE tunnel (Tunnel 2 in Figure 3.1) and then the first GRE tunnel (Tunnel 1 in Figure 3.1). If the secondary GRE tunnel cannot be established successfully, the local BANANA box will not set up the first GRE tunnel since it's more economical to transmit traffic over a raw link than over a GRE tunnel.

The local BANANA box sends the GRE Tunnel Setup Request message to the remote BANANA box via the endpoint S. The outer source IP address for this message is the IP address of S, while the outer destination IP address is the IP address of the branch router (if anycast is not used, the outer destination IP address would be IP address of R). The remote BANANA box with the highest priority (e.g., the one that the local BANANA box has the least-cost path to reach) in the group of remote BANANA boxes, which receives the GRE Tunnel Setup Request message, will initiate the procedure for

authentication and authorization to check whether the local BANANA box is trusted by the network device attached to S.

If the authentication and authorization succeed, the remote BANANA box sets the IP address of S, which is obtained from the GRE Tunnel Setup Request message (i.e., its outer source IP address), as the destination endpoint IP address of the secondary GRE tunnel and replies to the endpoint of the local BANANA box's secondary GRE tunnel with the GRE Tunnel Setup Accept message in which an IP address of R (e.g., an IP address of a Line Card in the remote BANANA box) and a Session ID randomly generated by the remote BANANA box are carried as attributes. The outer source IP address for this message is the IP address of R or the IP address of the branch router, while the outer destination IP address is the IP address of S. Otherwise, the remote BANANA box MUST send to the Local BANANA box's endpoint of the secondary GRE tunnel the GRE Tunnel Setup Deny message, and the local BANANA box MUST terminate the tunnel setup process once it receives the GRE Tunnel Setup Deny message.

After the secondary GRE tunnel is successfully set up, the local BANANA box will obtain the C address (see Figure 3.1) over the tunnel from the remote BANANA box through the Dynamic Host Configuration Protocol (DHCP). After that, the local BANANA box starts to set up the first GRE tunnel. It sends a GRE Tunnel Setup Request message via F, carrying the aforementioned Session ID received from the remote BANANA box. The outer source IP address for this message is the IP address of F, while the outer destination IP address is the IP address of R. The remote BANANA box, which receives the GRE Tunnel Setup Request message, will initiate the procedure for authentication and authorization in order to check whether the local BANANA box is trusted by the network device attached to F.

If the authentication and authorization succeed, the remote BANANA sets the IP address of F, which is obtained from the GRE Tunnel Setup Request message (i.e., its outer source IP address), as the destination endpoint IP address of the GRE tunnel and replies to the endpoint F with the GRE Tunnel Setup Accept message. The outer source IP address for this message is the IP address of R, while the outer destination IP address is the IP address of F. In this way, the two tunnels with the same Session ID can be used to carry traffic from the same user. That is to say, the two tunnels are "bonded" together. Otherwise, if the authentication and authorization fail, the remote BANANA box MUST send the GRE Tunnel Setup Deny message to the tunnel endpoint F. Meanwhile, it MUST send the GRE Tunnel Tear Down message to the tunnel endpoint S. The local BANANA box MUST terminate the tunnel setup process once it receives the GRE Tunnel Setup Deny message and MUST tear down the secondary GRE tunnel that has already been set up once it receives the GRE Tunnel Tear Down

message.

6. The Edge to Edge Scenario

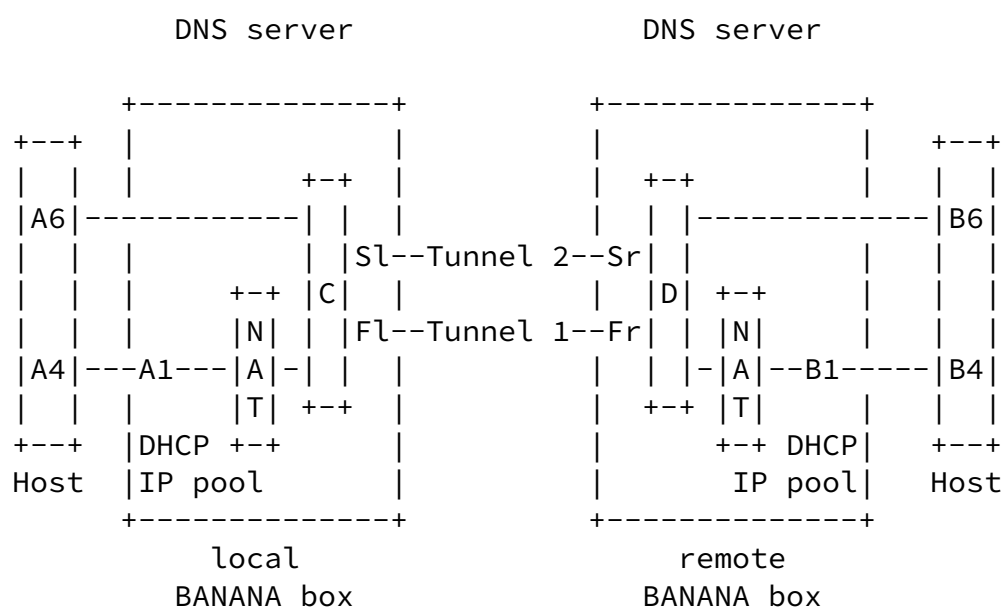


Figure 6.1: Tunnel bonding for the edge to edge scenario

The applicability of bonding tunnels is not limited to the single operator scenario. This section explains how bonding tunnels are

adapted to the edge to edge scenario. By and large, the adaptations are implementation issues.

o Addressing

IP addresses of B4, B6, B1, Fr and Sr are obtained in the same way as A4, A6, A1, Fl and Sl respectively, as in the single operator scenario.

C and D are the service endpoints of the bonding service at the two BANANA box respectively. IP addresses of C and D will be assigned from the locally configured IP pool via DHCP rather than be assigned remotely from the peering BANANA box.

Domain names of the two BANANA boxes may be configured or obtained via [\[TR-069\]](#). A query of the domain names will be resolved to the IP address of C or D by the DNS server .

o Establishment of Bonding Tunnels

The local BANANA box will send the GRE Tunnel Setup message to the remote BANANA box using IP address of D as the outer destination IP address and using IP address of Sl as the outer source IP address.

When the remote BANANA box replies the local BANANA box with the GRE Tunnel Accept message, the outer source IP address for this message is set to the IP address of Sr or D, while the outer destination IP address is the IP address of Sl. In the GRE Tunnel Accept message, the IP address of Sr, the IP address of Fr and a Session ID randomly generated by the remote BANANA box will be carried as attributes. Tunnel 2 would be set up between Sl and Sr.

For Tunnel 1, the local BANANA box will use the IP address of Fr as the outer destination IP address and IP address of Fl as the outer source IP address to send the GRE Tunnel Setup message to the remote BANANA box. In this message, the Session ID received from the remote BANANA box will be carried as an attribute. The remote BANANA box will reply the local BANANA box with a GRE Tunnel Setup Accept message. The outer source IP address for this message is the IP address of Fr while the outer destination IP address for this message is the IP address of Fl. Tunnel 1 would be set up between Fl and Fr. Since Tunnel 1 and Tunnel 2 use the same Session ID, they would be

bonded together to carry traffic from the same user.

For the edge to edge scenario, a BANANA box can either be "local" or "remote". The IP addresses of the service endpoint is used to break the tie. The BANANA box with the smaller IP address will be regarded as "local" while the BANANA box with the larger IP address will be regarded as "remote".

7. Security Considerations

Malicious devices controlled by attackers may intercept the control messages sent on the GRE tunnels. Later on, the rogue devices may fake control messages to disrupt the GRE tunnels or attract traffic from the target local BANANA box.

As a security feature, the Key field of the GRE header of the control messages is generated as a 32-bit cleartext password, except for the first GRE Setup Request message per bonding connection sent from the local BANANA box to the remote BANANA box, whose Key field is filled with all zeros. The remote BANANA box and the local BANANA validate the Key value and the outer source IP address, and they discard any packets with invalid combinations.

8. IANA Considerations

IANA need not assign anything for this memo. The GRE Protocol Type, the Ethertype for the GRE Channel of the BANANA signaling, is set to 0xB7EA, which is under the control of the IEEE Registration Authority. However, IANA has updated the "IEEE 802 Numbers" IANA web page [[802Type](#)], which is of primarily historic interest.

9. References

9.1. 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, <<http://www.rfc-editor.org/info/rfc2119>>.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", [RFC 2890](#), DOI 10.17487/RFC2890, September 2000,

<<http://www.rfc-editor.org/info/rfc2890>>.

[TR-069] Broadband Forum, "CPE WAN Management Protocol", Issue: 1 Amendment 5, November 2013, <https://www.broadband-forum.org/technical/download/TR-069_Amendment-5.pdf>.

[BANANA-encap]

N. Leymann, C. Heidemann, et al, "BANdwidth Aggregation for interNet Access (BANANA) The Data Plane of Bonding Tunnels", [draft-leymann-banana-data-encap](#), work in progress.

[BANANA-attributes]

N. Leymann, C. Heidemann, et al, "BANdwidth Aggregation for interNet Access (BANANA) Attributes for the Control Protocol of Bonding Tunnels", [draft-leymann-banana-signaling-attributes](#), work in progress.

9.2. Informative References

[802Type] IANA, "IEEE 802 Numbers",
<<http://www.iana.org/assignments/ieee-802-numbers>>.

Contributors

Li Xue
Individual
Email: xueli_jas@163.com

Zhongwen Jiang
Huawei Technologies
Email: jiangzhongwen@huawei.com

Leymann, et al.

Expires June 29, 2018

[Page 14]

INTERNET-DRAFT

BANANA Signaling

December 26, 2017

Authors' Addresses

Nicolai Leymann
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: +49-6151-5812721
Email: heidemannc@telekom.de

Mingui Zhang
Huawei Technologies
No. 156 Beiqing Rd.
Haidian District
Beijing 100095
China
Email: zhangmingui@huawei.com

Behcet Sarikaya
Huawei USA
5340 Legacy Dr. Building 3
Plano, TX 75024
United States of America
Email: sarikaya@ieee.org

Margaret Cullen
Painless Security
14 Summer St. Suite 202
Malden, MA 02148
United States of America
Email: margaret@painless-security.com