INTERNET-DRAFT
Intended Status: Informational

M. Cullen
Painless Security
N. Leymann
C. Heidemann
Deutsche Telekom AG
M. Boucadair
France Telecom
H. Deng
China Mobile
M. Zhang
B. Sarikaya
Huawei

Expires: January 5, 2017
July 4, 2016

Problem Statement: Bandwidth Aggregation for Internet Access
draft-zhang-banana-problem-statement-02.txt

Abstract

Bandwidth aggregation capabilities for Internet access services can significantly improve end user's Quality of Experience. Such capabilities are especially attractive in environments where multi-interfaced boxes become commonplace and can technically connect to various access networks, both wired and wireless.

This document describes the issues with bandwidth aggregation for Internet access. A set of requirements are derived from the said issues.

Copyright and License Notice

Table of Contents

1. Introduction

Bandwidth aggregation across multiple Internet access links (a.k.a., a bonding service) provides several useful features. The working text [WT-348] that is being edited by the Broadband Forum describes several use cases of a bonding service: Service Providers may use the bonding service to provide customers with increased access bandwidth and higher access reliability; Service delivery may be fostered to access the Internet by means of providing a LTE (Long Term Evolution) connection while the wired line is being constructed.

Although host-based bonding service is possible, the scope of this document is restricted to network-based bonding service. Also, a bonding service has to be operated by a single provider. Host-based or multi-provider cases will be standardized in other places, such as the MIF Working Group.

Design techniques that are being investigated, developed and sometimes deployed to facilitate bandwidth aggregation and improve the resiliency of access conditions raise several issues from various standpoints: traffic forwarding, routing and traffic engineering policies, security, etc. This document aims at presenting these issues regardless of the nature of the design technique. It also intends to derive requirements accordingly, and which should be addressed by any bandwidth aggregation technique. Typically, this is one of the inputs that are expected from the IETF community.

A common framework will be sketched, including required functional modules and interactions. The various solution proposals (e.g., GRE, LISP, MIP, MPTCP) can be viewed as applicability assessments of the framework. To support the bonding service, the problems to be addressed includes: addresses, traffic classification, distribution and recombination, bypassing, measurement, and policy control. To address these problems, we may work as a group to

- specify a sole encapsulation format;
- develop a common control plane;
- and define or suggest the algorithms to be used in packet reordering, flow control and congestion control.

2. Acronyms and Terminology

Bonding Service: A service that relies upon Bandwidth Aggregation capabilities that are meant to improve end user's quality of experience for Internet access services.

CPE: Customer Premises Equipment. An equipment which is the property of the network operator and located on the customer premises.

HG: Home Gateway. A CPE device that is enhanced to support the

simultaneous use of both fixed broadband and 3GPP access connections.

HAAP: Hybrid Access Aggregation Point. A logical function in Operator's network, terminating bonded connections while offering high speed Internet.

DHCP: Dynamic Host Configuration Protocol [RFC2131].

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].

3. Generic Reference Model

```
            +--+                             +-----+
            |  +----- bonding connection ----+     |
            |  |                             |     |
            |  |                             |     |
            |HG|i/f --- sub-connection ---i/f|HAAP|
            |  |            . . .            |     |
            |  |i/f --- sub-connection ---i/f|     |
            |  |                             |     |
            +--+                             +-----+
```

Figure 3.1: Reference model of the bonding service

Customers access the Internet using the bonding service which comprises of several key component functions as shown in Figure 3.1: the Home Gateway (HG) as one peer, the Hybrid Access Aggregation Point (HAAP) as the other peer, the bonding connection between the two peers and the sub-connections that logically make up the bonding connection. The bonding connection is usually established at the IP or the transport levels. Sub-connections are usually established at the IP or transport levels, but could be established at the link level.

4. Problem Areas

The following subsections list the problems that need to be solved by bonding service solutions.

4.1. Addressing

HG and HAAP need to allocate addresses to the interfaces attached to each sub-connection as well as the whole bonding connection. IPv4, IPv6 or dual-stack operation ought to be supported. Sub-connections can be de-multiplexed by their interface addresses. Upon bootstrap,

these connection addresses are acquired by the other end by means of a specific protocol, such as DHCP or MPTCP. The connection addresses of HG may thus be dynamically assigned by the HAAP.

4.2. Traffic Classification

Traffic classification occurs before the flows or packets are distributed among the connections. HG and HAAP should support the classification function that marks a flow or packet which is to be further processed by the traffic distribution function. Classification criteria include IP addresses, port numbers, etc. Traffic classification policies can be defined by end users and service providers and must be enforced by the HG and HAAP equipments.

4.3. Traffic Distribution

For traffic that is to be distributed across multiple sub-connections, equal load balancing generally applies, possibly inferred by the bandwidth that is available in each link that supports sub-connection. Unequal load balancing should be supported as well. Traffic may be distributed across sub-connections as a function of their available bandwidth. Traffic may also be split in such a way that whenever one sub-connection is saturated, then traffic is forwarded over another sub-connection.

There are two kinds of traffic distribution methods for the bonding service: per-flow load balancing and per-packet load sharing. The per-flow load balancing method is used to be widely adopted in core IP networks. It is suitable for the scenario where there are a large number of flows to be distributed. For end users, usually there are few number of applications to be transmitted over the bonded connections. Per-flow load balance techniques might not be able to achieve a fine grained load distribution, so that the per-packet load sharing is necessary.

For the per-flow use case, the load can be distributed using hashing methods. For the per-packet use case, the coloring mechanism specified in [RFC2698] can be used to classify customer's IP packets, both upstream and downstream, and packets will then be forwarded over the appropriate sub-connections. For example, packets colored as green are forwarded over one sub-connection and packets colored as yellow are forwarded over another sub-connection. For scenarios that rely upon more than two sub-connections, multiple levels of coloring mechanism could be implemented.

4.4. Traffic Recombination

For the per-packet use case, the recombination function at the

receiver sides reorders packets to preserve the integrity of the communication. The sender needs to mark each packet with a sequence number. The sequence number MUST be set per the whole bandwidth aggregation rather than per sub-connection so that all packets forwarded over several sub-connections actually share the same reordering buffer.

For the per-flow use case, an acknowledge number field appears in the packets in order to support congestion and flow control. In order to support such control, the sender needs also to maintain a sending buffer. See Section 3.3.2 of [RFC6824] for an example.

4.5. Bypass

Service Providers may require some traffic to bypass the bonding service. For example, some delay sensitive applications such as live TV broadcasting carried over a lossy sub-connection would impair customers' Quality of Experience. Service providers could then make sure that such traffic is forwarded over a set of wired sub-connections only, thereby disregarding low-rate mobile connections, for example.

There are two types of bypass: the bypassing traffic are transmitted on a sub-connection out of all the sub-connections between HG and HAAP; the bypassing traffic is still transmitted on a sub-connection between HG and HAAP, just that the occupied bandwidth of the bypassing traffic on this sub-connection can not used for bandwidth aggregation anymore. In either case, the bypassing traffic would not be under control of the bonding service scheme.

HG and HAAP needs to exchange information about bypassing, such as the application types that need to bypass the bonding service and the bandwidth occupied by the bypassing traffic (See also Section 4.6).

4.6. Measurement

HG and HAAP need to monitor and exchange performance parameters of the bonding service, including performance parameters that pertain to each sub-connection that belongs to the same connection. Such parameters include (but are not necessarily limited to):

- Operating state: The operating state has to be measured by control messages. When a sub-connection fails, this sub-connection has to be removed from the bonding connection.

- End-to-end delay and packet delay variation: The measurement of this parameter is used by flow and congestion control algorithms for per-flow and per-packet distribution purposes. The probing

packet could be piggy-backed by data packets or could be carried by out-of-band control messages.

- Maximum sub-connection bandwidth: This parameter may be used to determine the amount of traffic that can be distributed across all or a subset of sub-connections.

- Bypassing bandwidth: This parameter has to be periodically monitored. Usually, this is measured for the opposite end to figure out the available sending bandwidth. For example, the HG reports the downward bypassing bandwidth used in a sub-connection so that the HAAP can calculate the remaining downward bandwidth of this sub-connection.

4.7. Policy Control

Operators and customers may control the bonding service with policies. These policies will be instantiated into parameters or actions that impact traffic classification, distribution, combination, measurement and bypassing. Such policies may consist in:

- Defining traffic filter lists used by the traffic classification function.

- Per-flow or per-packet forwarding policies.

- Operators may determine the maximum allowed size (See MAX_PERFLOW_BUFFER in [RFC2890]) of the buffer for reordering. They may also define the maximum time (See OUTOFORDER_TIMER in [RFC2890]) that a packet can stay in the buffer for reordering. These parameters may pact traffic recombination.

- Operators may specify the frequency for detecting a sub-connection and the detection retry times before a sub-connection can be declared as "failed". Operators may specify maximum value of the difference between two measured one-way transit delays.

- Operators and customers may specify the service types to bypass the bonding service.

5. Requirements

Requirements for the bonding service are described in this section. Also, some additional requirements are listed for discussion in the Appendix.

The solution MUST apply for both IPv4 and IPv6 traffic.

The solution MUST NOT require any new capability to support bonding service from the host.

In the bonding service, forwarding traffic flows over various interfaces may have a negative impact on customers' experience (e.g., hazardous log outs, broken HTTPS associations, etc.). The solution should be carefully designed, so that

- activating the solution MUST NOT impact the stability, availability of the services delivered to the customer compared to customers who access the same service whose traffic is forwarded along a single path.

"Roles" (primary or backup) should be assigned to each supported network interface type (e.g., fixed or mobile access). This role is assigned by the network operator (e.g., fixed access can be set as the "primary"). Note that there may be more than two sub-connections to support primary and backup roles. A default setting can be considered.

- Setting of the role of the sub-connections SHOULD NOT be changed by the user.

The solution should provide Service Providers with means to enforce policy control of the bonding service. For example,

- the solution MUST allow to rate limit the traffic on a per-interface basis.

- the solution MUST support means to enable/disable the activation of multiple interfaces at the HG.

- the solution MUST support means to instruct the HG about the logic for mounting interfaces.

- the solution MUST support means to bind a given traffic to a subset of interfaces.

For the sake of policy enforcement or analytics at the network side,

- the solution MAY ease correlating flows, conveyed over multiple access networks, and which belong to the same customer.

Some services such as VoIP may be available over all active interfaces, but distinct logins and credentials may be used.

- The HG SHOULD be provided with clear instructions about which account to use to place outgoing sessions. For the sake of

simplicity, it is RECOMMENDED to use the login/credentials that are independent of the underlying access network used to access the service.

6. Related IETF Work

Bonding service designs can rely upon several solutions. The following subsections recap the work that has been or is being conducted by the IETF in this area. Note that solutions are listed in an alphabetic order. No preference order should be assumed by the reader.

6.1. GRE Tunnel Bonding

GRE Tunnel Bonding [GRE-HA] uses per-packet traffic distribution to achieve a fine-grained load sharing among the sub-connections. Out-of-sequence packets may be received so that reordering function is provided. IP packets are encapsulated in the GRE header which is in turn encapsulated in an outer IP header and forwarded over the sub-connections. The Sequence Number field of the GRE header is used to number the packets at the sender side, while the receiver uses of this sequence number to reorder the packets.

A new control plane is defined. Control messages are transported in the same GRE tunnels that are used to transport data packets. The control messages and data packets are distinguished by the GRE Protocol Type 0xB7EA.

GRE tunnel bonding has been implemented and deployed. Flow and congestion control could be supported within the tunnel through extending the GRE header, though it is currently out of the scope.

6.2. LISP

LISP has the basic capability to support the bonding service [LISP-HA] [ILNP]. LISP can be used to enforce traffic engineering based upon static load balancing that is not agnostic to link characteristics.

Packet-based traffic distribution has been considered in [LISP-HA] as well. The detail specification of the reordering mechanism based on a "Reorder" flag is left for future work.

6.3. Mobile IP

[MIP-HA] investigates how to support the bonding service by using IP mobility protocols. By treating the bonding service as a special scenario of IP mobility, some mechanisms (such as redundancy and load

balancing) that are supported by IP mobility protocols could be reused. However, IP mobility protocols have to be tailored in order to mitigate the complexity of their operation, let alone their scalability.

A new multipath binding option is proposed as an extension of PMIPv6 in [MIP-HA]. This option can be used to exchange the binding information, such as the Access Technology Type [RFC5213], the Interface Label and Binding ID, between peers.

Currently, per-flow traffic management is well supported by IP mobility protocols (such as [RFC6088] and [RFC6089]) while packet-based traffic distribution is left for future work.

6.4. Multipath TCP

MPTCP provides the ability to establish a communication over multiple paths, by means of sub-flow establishment and operation [RFC6824]. However, the traditional MPTCP is a host-based technology therefore it's out the scope of this document. What is considered as a candidate technology to support the bonding service is the MPTCP proxy mechanism. There are some implementations and deployments.

The MPTCP proxy operates at the transport layer and locates in the operator's network. A MPTCP proxy terminates a user's TCP flow and reinitiates MPTCP sub-flows towards the other MPTCP proxy. The other MPTCP proxy will terminate the MPTCP sub-flows and restore the user's TCP flow. The MPTCP protocol suite provides features to manage the state of sub-flows. [MPTCP-HA] discuss MPTCP proxy deployment concerns and also specifies an extension to transport UDP datagrams in MPTCP packets. UDP traffic can therefore be forwarded over a MPTCP connection.

6.5. Network Based Layer-2 Tunneling

Network Based Layer-2 Tunneling assigns a single IP address at each peer for the bonding connection. Layer-2 tunnels are set up per sub-connections. Layer 2 load balancing techniques, such as Link Aggregation [802.1AX] can be used to achieve the traffic distribution function. Packet-based distribution might be supported as well. However, per-packet distribution may cause the packets to be received as out-of-sequence, which is an issue that is yet-to-be-addressed by the Network Based Layer-2 Tunneling.

7. Security Considerations

The bonding service adds new capabilities. It also introduces new threats to the network. For example, traffic sent on unsecured sub-

connections would be easily intercepted by an attacker who performs man-in-the-middle attack. The multi-path communication may be leveraged to perform Denial of Service (DoS) attack or Distributed Denial of Service (DDoS) attack (e.g., based upon flooding traffic) that may jeopardize the aggregation gateway as well as the access equipment and end station operation.

These kind of new security issues should be carefully considered in designing solutions that aim to address the problems of Bandwidth Aggregation for Internet Access.

8. IANA Considerations

No IANA action is required in this document. RFC Editor: please remove this section before publication.

9. Acknowledgements

Authors would like to thank the comments and suggestions from Christian Jacquenet and Li Xue.

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.

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, March 1997.

[RFC2689] Bormann, C., "Providing Integrated Services over Low-bitrate Links", RFC 2689, September 1999.

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

10.2. Informative References

[WT-348] Broadband Forum Work on "Hybrid Access for Broadband Networks" (WT-348), October 21, 2014, <http://datatracker.ietf.org/liaison/1355/>.

[GRE-HA] N. Leymann, C. Heidemann, M. Zhang, et al, "GRE Tunnel Bonding", draft-zhang-gre-tunnel-bonding, work in progress.

[RFC6824] Ford, A., Raiciu, C., Handley, M., and O. Bonaventure, "TCP Extensions for Multipath Operation with Multiple

Addresses", RFC 6824, January 2013.

[MPTCP-HA] M. Boucadair and C. Jacquenet, "An MPTCP Option for Network-Assisted MPTCP Deployments: Plain Transport Mode", draft-boucadair-mptcp-plain-mode, work in progress.

[MIP-HA] P. Seite, A. Yegin and S. Gundavelli, "Multihoming support for Residential Gateways", draft-seite-dmm-rg-multihoming, work in progress.

[RFC5213] Gundavelli, S., Ed., Leung, K., Devarapalli, V., Chowdhury, K., and B. Patil, "Proxy Mobile IPv6", RFC 5213, August 2008.

[RFC6088] Tsirtsis, G., Giarreta, G., Soliman, H., and N. Montavont, "Traffic Selectors for Flow Bindings", RFC 6088, January 2011.

[RFC6089] Tsirtsis, G., Soliman, H., Montavont, N., Giaretta, G., and K. Kuladinithi, "Flow Bindings in Mobile IPv6 and Network Mobility (NEMO) Basic Support", RFC 6089, January 2011.

[802.1AX] IEEE, "IEEE Standard for Local and Metropolitan Area Networks - Link Aggregation", 802.1AX-2014, 24 December 2014.

[LISP-HA] M. Menth, A. Stockmayer and M. Schmidt, "LISP Hybrid Access", draft-menth-lisp-ha, work in progress.

[ILNP] "ILNP - Identifier-Locator Network Protocol", online available: http://ilnp.cs.st-andrews.ac.uk/

Appendix A. Additional Requirements

The following requirements are listed as record and may subject to change.

- The solution MUST be valid for any kinds of interfaces that need to be aggregated. No dependency to the underlying media should be assumed.

- The solution MUST comply with servers policy regarding IP addresses that are bound to (HTTP session) cookies.

- The solution MUST NOT break TLS associations.

- Activating the solution MUST NOT have negative impacts on the service usability (including the HG management).

- Service degradation MUST NOT be observed when enabling the solution.

- Enabling the solution MUST increase the serviceability of the HG. In particular, the solution MUST allow for the HG to always establish a network attachment when the primary connectivity is out of service.

- The solution SHOULD NOT alter any mechanism, to aggregate available resources or to ensure a service continuity among multiple access points, that is supported by end-devices connected to the HG.

- The HG MUST bind the DNS server(s) discovered during the network attachment phase to the interface from which the information was received.

- The HG MUST bind the service information (e.g., SIP Proxy Server) discovered during the network attachment phase to the interface from which the information was received.

- When sending the traffic via a given interface, the HG MUST use as source address an address (or an address from a prefix) that was assigned for that interface.

- For protocols such as RTP/RTCP, the same IP address MUST be used for both RTP and RTCP sessions.

- Dedicated tools SHOULD be provided to the customer to assess the aggregated capacity (instead of link-specific). This can be included as part of the HG UI, a dedicated portal, etc.

Author's Addresses

Margaret Cullen
Painless Security
14 Summer St. Suite 202
Malden, MA 02148 USA

EMail: margaret@painless-security.com

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: +4961515812721
EMail: heidemannc@telekom.de

Mohamed Boucadair
France Telecom
Rennes 35000
France

EMail: mohamed.boucadair@orange.com

Hui Deng
China Mobile
53A,Xibianmennei Ave.,
Xuanwu District,
Beijing 100053
China

EMail: denghui@chinamobile.com

Mingui Zhang
Huawei Technologies
No.156 Beiqing Rd. Haidian District,
Beijing 100095 P.R. China

EMail: zhangmingui@huawei.com

Behcet Sarikaya
Huawei USA
5340 Legacy Dr. Building 3
Plano, TX 75024

EMail: sarikaya@ieee.org