Stream: Internet Engineering Task Force (IETF)
RFC: 8611
Updates: 8029
Category: Standards Track
Published: June 2019
ISSN: 2070-1721
Authors:
N. Akiya *Big Switch Networks*
G. Swallow *SETC*
S. Litkowski *Orange*
B. Decraene *Orange*
J. Drake *Juniper Networks*
M. Chen *Huawei*

# Label Switched Path (LSP) Ping and Traceroute Multipath Support for Link Aggregation Group (LAG) Interfaces

## Abstract

This document defines extensions to the MPLS Label Switched Path (LSP) Ping and Traceroute mechanisms as specified in RFC 8029. The extensions allow the MPLS LSP Ping and Traceroute mechanisms to discover and exercise specific paths of Layer 2 (L2) Equal-Cost Multipath (ECMP) over Link Aggregation Group (LAG) interfaces. Additionally, a mechanism is defined to enable the determination of the capabilities supported by a Label Switching Router (LSR).

This document updates RFC 8029.

## Status of This Memo

This is an Internet Standards Track document.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Further information on Internet Standards is available in Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at https://www.rfc-editor.org/info/rfc8611.

## Table of Contents

# 1. Introduction

## 1.1. Background

The MPLS Label Switched Path (LSP) Ping and Traceroute mechanisms [RFC8029] are powerful tools designed to diagnose all available Layer 3 (L3) paths of LSPs, including diagnostic coverage of L3 Equal-Cost Multipath (ECMP). In many MPLS networks, Link Aggregation Groups (LAGs), as defined in [IEEE802.1AX], provide Layer 2 (L2) ECMP and are often used for various reasons. MPLS LSP Ping and Traceroute tools were not designed to discover and exercise specific paths of L2 ECMP. This produces a limitation for the following scenario when an LSP traverses a LAG:

- Label switching over some member links of the LAG is successful, but fails over other member links of the LAG.
- MPLS echo request for the LSP over the LAG is load-balanced on one of the member links that is label switching successfully.

With the above scenario, MPLS LSP Ping and Traceroute will not be able to detect the label-switching failure of the problematic member link(s) of the LAG. In other words, lack of L2 ECMP diagnostic coverage can produce an outcome where MPLS LSP Ping and Traceroute can be blind to label-switching failures over a problematic LAG interface. It is, thus, desirable to extend the MPLS LSP Ping and Traceroute to have deterministic diagnostic coverage of LAG interfaces.

The work toward a solution to this problem was motivated by issues encountered in live networks.

## 1.2. Terminology

The following acronyms/terms are used in this document:

- MPLS - Multiprotocol Label Switching.
- LSP - Label Switched Path.
- LSR - Label Switching Router.
- ECMP - Equal-Cost Multipath.
- LAG - Link Aggregation Group.
- Initiator LSR - The LSR that sends the MPLS echo request message.

- Responder LSR - The LSR that receives the MPLS echo request message and sends the MPLS echo reply message.

### 1.3. Requirements Language

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. Overview of Solution

This document defines a new TLV to discover the capabilities of a responder LSR and extensions for use with the MPLS LSP Ping and Traceroute mechanisms to describe Multipath Information for individual LAG member links, thus allowing MPLS LSP Ping and Traceroute to discover and exercise specific paths of L2 ECMP over LAG interfaces. The reader is expected to be familiar with the Downstream Detailed Mapping TLV (DDMAP) described in Section 3.4 of [RFC8029].

The solution consists of the MPLS echo request containing a DDMAP TLV and the new LSR Capability TLV to indicate that separate load-balancing information for each L2 next hop over LAG is desired in the MPLS echo reply. The responder LSR places the same LSR Capability TLV in the MPLS echo reply to provide acknowledgement back to the initiator LSR. It also adds, for each downstream LAG member, load-balancing information (i.e., multipath information and interface index). This mechanism is applicable to all types of LSPs that can traverse LAG interfaces. Many LAGs are built from peer-to-peer links, with router X and router X+1 having direct connectivity and the same number of LAG members. It is possible to build LAGs asymmetrically by using Ethernet switches between two routers. Appendix A lists some use cases for which the mechanisms defined in this document may not be applicable. Note that the mechanisms described in this document do not impose any changes to scenarios where an LSP is pinned down to a particular LAG member (i.e., the LAG is not treated as one logical interface by the LSP).

The following figure and description provide an example of an LDP network.

```
<----- LDP Network ----->

        +-------+
        |       |
A-------B=======C-------E
        |               |
        +-------D-------+

---- Non-LAG
==== LAG comprising of two member links
```

*Figure 1: Example LDP Network*

When node A is initiating LSP Traceroute to node E, node B will return to node A load-balancing information for the following entries:

1. Downstream C over Non-LAG (upper path).
2. First Downstream C over LAG (middle path).
3. Second Downstream C over LAG (middle path).
4. Downstream D over Non-LAG (lower path).

This document defines:

- in Section 3, a mechanism to discover capabilities of responder LSRs;
- in Section 4, a mechanism to discover L2 ECMP information;
- in Section 5, a mechanism to validate L2 ECMP traversal;
- in Section 6, the LSR Capability TLV;
- in Section 7, the LAG Description Indicator flag;
- in Section 8, the Local Interface Index Sub-TLV;
- in Section 9, the Remote Interface Index Sub-TLV; and
- in Section 10, the Detailed Interface and Label Stack TLV.

# 3. LSR Capability Discovery

The MPLS Ping operates by an initiator LSR sending an MPLS echo request message and receiving back a corresponding MPLS echo reply message from a responder LSR. The MPLS Traceroute operates in a similar way except the initiator LSR potentially sends multiple MPLS echo request messages with incrementing TTL values.

There have been many extensions to the MPLS Ping and Traceroute mechanisms over the years. Thus, it is often useful, and sometimes necessary, for the initiator LSR to deterministically disambiguate the differences between:

- The responder LSR sent the MPLS echo reply message with contents C because it has feature X, Y, and Z implemented.
- The responder LSR sent the MPLS echo reply message with contents C because it has a subset of features X, Y, and Z (i.e., not all of them) implemented.
- The responder LSR sent the MPLS echo reply message with contents C because it does not have features X, Y, or Z implemented.

To allow the initiator LSR to disambiguate the above differences, this document defines the LSR Capability TLV (described in Section 6). When the initiator LSR wishes to discover the capabilities of the responder LSR, the initiator LSR includes the LSR Capability TLV in the MPLS echo request message. When the responder LSR receives an MPLS echo request message with the LSR Capability TLV included, if it knows the LSR Capability TLV, then it MUST include the LSR Capability TLV in the MPLS echo reply message with the LSR Capability TLV describing the features and extensions supported by the local LSR. Otherwise, an MPLS echo reply must be sent back to the initiator LSR with the return code set to "One or more of the TLVs was not understood", according to the rules defined in Section 3 of [RFC8029]. Then, the initiator LSR can send another MPLS echo request without including the LSR Capability TLV.

It is RECOMMENDED that implementations supporting the LAG multipath extensions defined in this document include the LSR Capability TLV in MPLS echo request messages.

## 3.1. Initiator LSR Procedures

If an initiator LSR does not know what capabilities a responder LSR can support, it can send an MPLS echo request message and carry the LSR Capability TLV to the responder to discover the capabilities that the responder LSR can support.

## 3.2. Responder LSR Procedures

When a responder LSR receives an MPLS echo request message that carries the LSR Capability TLV, the following procedures are used:

If the responder knows how to process the LSR Capability TLV, the following procedures are used:

- The responder LSR MUST include the LSR Capability TLV in the MPLS echo reply message.

- If the responder LSR understands the LAG Description Indicator flag:
  - Set the Downstream LAG Info Accommodation flag if the responder LSR is capable of describing the outgoing LAG member links separately; otherwise, clear the Downstream LAG Info Accommodation flag.
  - Set the Upstream LAG Info Accommodation flag if the responder LSR is capable of describing the incoming LAG member links separately; otherwise, clear the Upstream LAG Info Accommodation flag.

# 4. Mechanism to Discover L2 ECMP

## 4.1. Initiator LSR Procedures

Through LSR Capability Discovery as defined in Section 3, the initiator LSR can understand whether the responder LSR can describe incoming/outgoing LAG member links separately in the DDMAP TLV.

Once the initiator LSR knows that a responder can support this mechanism, then it sends an MPLS echo request carrying a DDMAP TLV with the LAG Description Indicator flag (G) set to the responder LSR. The LAG Description Indicator flag (G) indicates that separate load-balancing information for each L2 next hop over a LAG is desired in the MPLS echo reply. The new LAG Description Indicator flag is described in Section 7.

## 4.2. Responder LSR Procedures

When a responder LSR receives an MPLS echo request message with the LAG Description Indicator flag set in the DDMAP TLV, if the responder LSR understands the LAG Description Indicator flag and is capable of describing outgoing LAG member links separately, the following procedures are used, regardless of whether or not the outgoing interfaces include LAG interfaces:

- For each downstream interface that is a LAG interface:
  - The responder LSR MUST include a DDMAP TLV when sending the MPLS echo reply. There is a single DDMAP TLV for the LAG interface, with member links described using sub-TLVs.
  - The responder LSR MUST set the LAG Description Indicator flag in the DS Flags field of the DDMAP TLV.
  - In the DDMAP TLV, the Local Interface Index Sub-TLV, Remote Interface Index Sub-TLV, and Multipath Data Sub-TLV are used to describe each LAG member link. All other fields of the DDMAP TLV are used to describe the LAG interface.

- For each LAG member link of the LAG interface:
  - The responder LSR MUST add a Local Interface Index Sub-TLV (described in Section 8) with the LAG Member Link Indicator flag set in the Interface Index Flags field. It describes the interface index of this outgoing LAG member link (the local interface index is assigned by the local LSR).
  - The responder LSR MAY add a Remote Interface Index Sub-TLV (described in Section 9) with the LAG Member Link Indicator flag set in the Interface Index Flags field. It describes the interface index of the incoming LAG member link on the downstream LSR (this interface index is assigned by the downstream LSR). How the local LSR obtains the interface index of the LAG member link on the downstream LSR is outside the scope of this document.
  - The responder LSR MUST add a Multipath Data Sub-TLV for this LAG member link, if the received DDMAP TLV requested multipath information.

Based on the procedures described above, every LAG member link will have a Local Interface Index Sub-TLV and a Multipath Data Sub-TLV entry in the DDMAP TLV. The order of the sub-TLVs in the DDMAP TLV for a LAG member link MUST be Local Interface Index Sub-TLV immediately followed by Multipath Data Sub-TLV, except as follows. A LAG member link MAY also have a corresponding Remote Interface Index Sub-TLV. When a Local Interface Index Sub-TLV, a Remote Interface Index Sub-TLV, and a Multipath Data Sub-TLV are placed in the DDMAP TLV to describe a LAG member link, they MUST be placed in the order of Local Interface Index Sub-TLV, Remote Interface Index Sub-TLV, and Multipath Data Sub-TLV. The blocks of Local Interface Index, Remote Interface Index (optional), and Multipath Data Sub-TLVs for each member link MUST appear adjacent to each other and be in order of increasing local interface index.

A responder LSR possessing a LAG interface with two member links would send the following DDMAP for this LAG interface:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
~  DDMAP fields describing LAG interface (DS Flags with G set)  ~
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Local Interface Index Sub-TLV of LAG member link #1           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Remote Interface Index Sub-TLV of LAG member link #1          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Multipath Data Sub-TLV LAG member link #1                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Local Interface Index Sub-TLV of LAG member link #2           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Remote Interface Index Sub-TLV of LAG member link #2          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Multipath Data Sub-TLV LAG member link #2                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Label Stack Sub-TLV                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 2: Example of DDMAP in MPLS Echo Reply*

When none of the received multipath information maps to a particular LAG member link, then the responder LSR MUST still place the Local Interface Index Sub-TLV and the Multipath Data Sub-TLV for that LAG member link in the DDMAP TLV. The value of the Multipath Length field of the Multipath Data Sub-TLV is set to zero.

## 4.3. Additional Initiator LSR Procedures

The procedures in Section 4.2 allow an initiator LSR to:

- Identify whether or not the responder LSR can describe outgoing LAG member links separately, by looking at the LSR Capability TLV.
- Utilize the value of the LAG Description Indicator flag in DS Flags to identify whether each received DDMAP TLV describes a LAG interface or a non-LAG interface.
- Obtain multipath information that is expected to traverse the specific LAG member link described by the corresponding interface index.

When an initiator LSR receives a DDMAP containing LAG member information from a downstream LSR with TTL=n, then the subsequent DDMAP sent by the initiator LSR to the downstream LSR with TTL=n+1 through a particular LAG member link MUST be updated according to the following procedures:

- The Local Interface Index Sub-TLVs MUST be removed in the sending DDMAP.

- If the Remote Interface Index Sub-TLVs were present and the initiator LSR is traversing over a specific LAG member link, then the Remote Interface Index Sub-TLV corresponding to the LAG member link being traversed SHOULD be included in the sending DDMAP. All other Remote Interface Index Sub-TLVs MUST be removed from the sending DDMAP.
- The Multipath Data Sub-TLVs MUST be updated to include just one Multipath Data Sub-TLV. The initiator LSR MAY just keep the Multipath Data Sub-TLV corresponding to the LAG member link being traversed or combine the Multipath Data Sub-TLVs for all LAG member links into a single Multipath Data Sub-TLV when diagnosing further downstream LSRs.
- All other fields of the DDMAP are to comply with procedures described in [RFC8029].

Figure 3 is an example that shows how to use the DDMAP TLV to send a notification about which member link (link #1 in the example) will be chosen to send the MPLS echo request message to the next downstream LSR:

```
  0                   1                   2                   3
  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 ~  DDMAP fields describing LAG interface (DS Flags with G set)  ~
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |[OPTIONAL] Remote Interface Index Sub-TLV of LAG member link #1|
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |             Multipath Data Sub-TLV LAG member link #1         |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                       Label Stack Sub-TLV                     |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 3: Example of DDMAP in MPLS Echo Request*

# 5. Mechanism to Validate L2 ECMP Traversal

Section 4 defines the responder LSR procedures to construct a DDMAP for a downstream LAG. The Remote Interface Index Sub-TLV that describes the incoming LAG member links of the downstream LSR is optional, because this information from the downstream LSR is often not available on the responder LSR. In such case, the traversal of LAG member links can be validated with procedures described in Section 5.1. If LSRs can provide the Remote Interface Index Sub-TLVs, then the validation procedures described in Section 5.2 can be used.

### 5.1. Incoming LAG Member Links Verification

Without downstream LSRs returning Remote Interface Index Sub-TLVs in the DDMAP, validation of the LAG member link traversal requires that the initiator LSR traverses all available LAG member links and takes the results through additional logic. This section provides the mechanism for the initiator LSR to obtain additional information from the downstream LSRs and describes the additional logic in the initiator LSR to validate the L2 ECMP traversal.

#### 5.1.1. Initiator LSR Procedures

An MPLS echo request carrying a DDMAP TLV with the Interface and Label Stack Object Request flag and LAG Description Indicator flag set is sent to indicate the request for Detailed Interface and Label Stack TLV with additional LAG member link information (i.e., interface index) in the MPLS echo reply.

#### 5.1.2. Responder LSR Procedures

When it receives an echo request with the LAG Description Indicator flag set, a responder LSR that understands that flag and is capable of describing the incoming LAG member link SHOULD use the following procedures, regardless of whether or not the incoming interface was a LAG interface:

- When the I flag (Interface and Label Stack Object Request flag) of the DDMAP TLV in the received MPLS echo request is set:
  - The responder LSR MUST add the Detailed Interface and Label Stack TLV (described in Section 10) in the MPLS echo reply.
  - If the incoming interface is a LAG, the responder LSR MUST add the Incoming Interface Index Sub-TLV (described in Section 10.1.2) in the Detailed Interface and Label Stack TLV. The LAG Member Link Indicator flag MUST be set in the Interface Index Flags field, and the Interface Index field set to the LAG member link that received the MPLS echo request.

These procedures allow the initiator LSR to utilize the Incoming Interface Index Sub-TLV in the Detailed Interface and the Label Stack TLV to derive, if the incoming interface is a LAG, the identity of the incoming LAG member.

#### 5.1.3. Additional Initiator LSR Procedures

Along with procedures described in Section 4, the procedures described in this section will allow an initiator LSR to know:

- The expected load-balance information of every LAG member link, at LSR with TTL=n.

- With specific entropy, the expected interface index of the outgoing LAG member link at TTL=n.
- With specific entropy, the interface index of the incoming LAG member link at TTL=n+1.

Depending on the LAG traffic division algorithm, the messages may or may not traverse different member links. The expectation is that there's a relationship between the interface index of the outgoing LAG member link at TTL=n and the interface index of the incoming LAG member link at TTL=n+1 for all entropies examined. In other words, the messages with a set of entropies that load-balances to outgoing LAG member link X at TTL=n should all reach the next hop on the same incoming LAG member link Y at TTL=n+1.

With additional logic, the initiator LSR can perform the following checks in a scenario where it (a) knows that there is a LAG that has two LAG members, between TTL=n and TTL=n+1, and (b) has the multipath information to traverse the two LAG member links.

The initiator LSR sends two MPLS echo request messages to traverse the two LAG member links at TTL=n+1:

- Success case:
  - One MPLS echo request message reaches TTL=n+1 on LAG member link 1.
  - The other MPLS echo request message reaches TTL=n+1 on LAG member link 2.

  The two MPLS echo request messages sent by the initiator LSR reach the immediate downstream LSR from two different LAG member links.

- Error case:
  - One MPLS echo request message reaches TTL=n+1 on LAG member link 1.
  - The other MPLS echo request message also reaches TTL=n+1 on LAG member link 1.
  - One or both MPLS echo request messages cannot reach the immediate downstream LSR on whichever link.

  One or two MPLS echo request messages sent by the initiator LSR cannot reach the immediate downstream LSR, or the two MPLS echo request messages reach at the immediate downstream LSR from the same LAG member link.

Note that the procedures defined above will provide a deterministic result for LAG interfaces that are back-to-back connected between LSRs (i.e., no L2 switch in between). If there is an L2 switch between the LSR at TTL=n and the LSR at TTL=n+1, there is no guarantee that every incoming interface at TTL=n+1 can be traversed, even when traversing every outgoing LAG member link at TTL=n. Issues resulting from LAG with an L2 switch in between are further described in Appendix A. LAG provisioning models in operator networks should be considered when analyzing the output of LSP Traceroute that is exercising L2 ECMPs.

### 5.2. Individual End-to-End Path Verification

When the Remote Interface Index Sub-TLVs are available from an LSR with TTL=n, then the validation of LAG member link traversal can be performed by the downstream LSR of TTL=n+1. The initiator LSR follows the procedures described in Section 4.3.

The DDMAP validation procedures for the downstream responder LSR are then updated to include the comparison of the incoming LAG member link to the interface index described in the Remote Interface Index Sub-TLV in the DDMAP TLV. Failure of this comparison results in the return code being set to "Downstream Mapping Mismatch (5)".

## 6. LSR Capability TLV

This document defines a new TLV that is referred to as the LSR Capability TLV. It MAY be included in the MPLS echo request message and the MPLS echo reply message. An MPLS echo request message and an MPLS echo reply message MUST NOT include more than one LSR Capability TLV. The presence of an LSR Capability TLV in an MPLS echo request message is a request that a responder LSR includes an LSR Capability TLV in the MPLS echo reply message, with the LSR Capability TLV describing features and extensions that the responder LSR supports.

The format of the LSR Capability TLV is as below:

LSR Capability TLV Type is 4. Length is 4. The LSR Capability TLV has the following format:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      LSR Capability Flags                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 4: LSR Capability TLV*

Where:

The Type field is 2 octets in length, and the value is 4.

The Length field is 2 octets in length, and the value is 4.

The LSR Capability Flags field is 4 octets in length; this document defines the following flags:

```
   0                   1                   2                   3
   0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
1
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+
  |                 Reserved (Must Be Zero)                   |U|
D|
  +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-
+-+
```

This document defines two flags. The unallocated flags MUST be set to zero when sending and ignored on receipt. Both the U and the D flag MUST be cleared in the MPLS echo request message when sending and ignored on receipt. Zero, one, or both of the flags (U and D) MAY be set in the MPLS echo reply message.

```
   Flag  Name and Meaning
   ----  ----------------

      U  Upstream LAG Info Accommodation

         An LSR sets this flag when the LSR is capable of
describing
         a LAG member link in the Incoming Interface Index Sub-
TLV
         in the Detailed Interface and Label Stack TLV.

      D  Downstream LAG Info Accommodation

         An LSR sets this flag when the LSR is capable of
describing
         LAG member links in the Local Interface Index Sub-TLV
and
         the Multipath Data Sub-TLV in the Downstream Detailed
         Mapping TLV.
```

# 7. LAG Description Indicator Flag: G

This document defines a new flag, the G flag (LAG Description Indicator), in the DS Flags field of the DDMAP TLV.

The G flag in the MPLS echo request message indicates the request for detailed LAG information from the responder LSR. In the MPLS echo reply message, the G flag MUST be set if the DDMAP TLV describes a LAG interface. It MUST be cleared otherwise.

The G flag is defined as below:

The Bit Number is 3.

```
     0 1 2 3 4 5 6 7
    +-+-+-+-+-+-+-+-+
    | MBZ |G|E|L|I|N|
    +-+-+-+-+-+-+-+-+
```

```
Flag  Name and Meaning
----  ----------------

   G  LAG Description Indicator

      When this flag is set in the MPLS echo request, the
responder
      LSR is requested to respond with detailed LAG
information.
      When this flag is set in the MPLS echo reply, the
corresponding
      DDMAP TLV describes a LAG interface.
```

# 8. Local Interface Index Sub-TLV

The Local Interface Index Sub-TLV describes the interface index assigned by the local LSR to an egress interface. One or more Local Interface Index sub-TLVs MAY appear in a DDMAP TLV.

The format of the Local Interface Index Sub-TLV is below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                     Local Interface Index                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 5: Local Interface Index Sub-TLV*

Where:

- The Type field is 2 octets in length, and the value is 4.
- The Length field is 2 octets in length, and the value is 4.
- The Local Interface Index field is 4 octets in length; it is an interface index assigned by a local LSR to an egress interface. It's normally an unsigned integer and in network byte order.

# 9. Remote Interface Index Sub-TLV

The Remote Interface Index Sub-TLV is an optional TLV; it describes the interface index assigned by a downstream LSR to an ingress interface. One or more Remote Interface Index sub-TLVs MAY appear in a DDMAP TLV.

The format of the Remote Interface Index Sub-TLV is below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |            Length             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                    Remote Interface Index                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 6: Remote Interface Index Sub-TLV*

Where:

- The Type field is 2 octets in length, and the value is 5.
- The Length field is 2 octets in length, and the value is 4.

- The Remote Interface Index field is 4 octets in length; it is an interface index assigned by a downstream LSR to an ingress interface. It's normally an unsigned integer and in network byte order.

## 10. Detailed Interface and Label Stack TLV

The Detailed Interface and Label Stack TLV MAY be included in an MPLS echo reply message to report the interface on which the MPLS echo request message was received and the label stack that was on the packet when it was received. A responder LSR MUST NOT insert more than one instance of this TLV into the MPLS echo reply message. This TLV allows the initiator LSR to obtain the exact interface and label stack information as it appears at the responder LSR.

Detailed Interface and Label Stack TLV Type is 6. Length is K + Sub-TLV Length (sum of Sub-TLVs). K is the sum of all fields of this TLV prior to the list of Sub-TLVs, but the length of K depends on the Address Type. Details of this information is described below. The Detailed Interface and Label Stack TLV has the following format:

```
    0                   1                   2                   3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |             Type              |            Length             |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | Address Type  |             Reserved (Must Be Zero)           |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                   IP Address (4 or 16 octets)                 |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |                   Interface (4 or 16 octets)                  |
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   .                                                               .
   .                      List of Sub-
 TLVs                        .
   .                                                               .
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 7: Detailed Interface and Label Stack TLV*

The Detailed Interface and Label Stack TLV format is derived from the Interface and Label Stack TLV format (from [RFC8029]). Two changes are introduced. The first is that the label stack is converted into a sub-TLV. The second is that a new sub-TLV is added to describe an interface index. The other fields of the Detailed Interface and Label Stack TLV have the same use and meaning as in [RFC8029]. A summary of these fields is as below:

Address Type

The Address Type indicates if the interface is numbered or unnumbered. It also determines the length of the IP Address and Interface fields. The resulting total length of the initial part of the TLV is listed as "K Octets". The Address Type is set to one of the following values:

```
        Type #          Address Type           K Octets
        ------          ------------           --------
             1          IPv4 Numbered                16
             2          IPv4 Unnumbered              16
             3          IPv6 Numbered                40
             4          IPv6 Unnumbered              28
```

IP Address and Interface

IPv4 addresses and interface indices are encoded in 4 octets; IPv6 addresses are encoded in 16 octets.

If the interface upon which the echo request message was received is numbered, then the Address Type MUST be set to IPv4 Numbered or IPv6 Numbered, the IP Address MUST be set to either the LSR's Router ID or the interface address, and the Interface MUST be set to the interface address.

If the interface is unnumbered, the Address Type MUST be either IPv4 Unnumbered or IPv6 Unnumbered, the IP Address MUST be the LSR's Router ID, and the Interface MUST be set to the index assigned to the interface.

Note: Usage of IPv6 Unnumbered has the same issue as [RFC8029], which is described in Section 3.4.2 of [RFC7439]. A solution should be considered and applied to both [RFC8029] and this document.

### 10.1. Sub-TLVs

This section defines the sub-TLVs that MAY be included as part of the Detailed Interface and Label Stack TLV. Two sub-TLVs are defined:

```
          Sub-Type    Sub-TLV Name
          ---------   ------------
            1         Incoming Label Stack
            2         Incoming Interface Index
```

#### 10.1.1. Incoming Label Stack Sub-TLV

The Incoming Label Stack Sub-TLV contains the label stack as received by an LSR. If any TTL values have been changed by this LSR, they SHOULD be restored.

Incoming Label Stack Sub-TLV Type is 1. Length is variable, and its format is as below:

```
     0                   1                   2                   3
     0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |             Type              |            Length             |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                 Label                 | TC  |S|      TTL      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    .                                                               .
    .                                                               .
    .                                                               .
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |                 Label                 | TC  |S|      TTL      |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 8: Incoming Label Stack Sub-TLV*

#### 10.1.2. Incoming Interface Index Sub-TLV

The Incoming Interface Index Sub-TLV MAY be included in a Detailed Interface and Label Stack TLV. The Incoming Interface Index Sub-TLV describes the index assigned by a local LSR to the interface that received the MPLS echo request message.

Incoming Interface Index Sub-TLV Type is 2. Length is 8, and its format is as below:

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|             Type              |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Interface Index Flags      |       Reserved (Must Be Zero) |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                   Incoming Interface Index                    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

*Figure 9: Incoming Interface Index Sub-TLV*

Interface Index Flags

The Interface Index Flags field is a bit vector with following format.

```
 0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Reserved (Must Be Zero)   |M|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

One flag is defined: M. The remaining flags MUST be set to zero when sending and ignored on receipt.

```
Flag  Name and Meaning
----  ----------------

   M  LAG Member Link Indicator

      When this flag is set, the interface index described in
this
      sub-TLV is a member of a LAG.
```

Incoming Interface Index

An Index assigned by the LSR to this interface. It's normally an unsigned integer and in network byte order.

# 11. Rate-Limiting on Echo Request/Reply Messages

An LSP may be over several LAGs. Each LAG may have many member links. To exercise all the links, many echo request/reply messages will be sent in a short period. It's possible that those messages may traverse a common path as a burst. Under some circumstances, this might cause congestion at the common path. To avoid potential congestion, it is RECOMMENDED that implementations randomly delay the echo request and reply messages at the initiator LSRs and responder LSRs. Rate-limiting of ping traffic is further specified in Section 5 of [RFC8029] and Section 4.1 of [RFC6425], which apply to this document as well.

# 12. Security Considerations

This document extends the LSP Traceroute mechanism [RFC8029] to discover and exercise L2 ECMP paths to determine problematic member link(s) of a LAG. These on-demand diagnostic mechanisms are used by an operator within an MPLS control domain.

[RFC8029] reviews the possible attacks and approaches to mitigate possible threats when using these mechanisms.

To prevent leakage of vital information to untrusted users, a responder LSR MUST only accept MPLS echo request messages from designated trusted sources via filtering the source IP address field of received MPLS echo request messages. As noted in [RFC8029], spoofing attacks only have a small window of opportunity. If an intermediate node hijacks these messages (i.e., causes non-delivery), the use of these mechanisms will determine the data plane is not working as it should. Hijacking of a responder node such that it provides a legitimate reply would involve compromising the node itself and the MPLS control domain. [RFC5920] provides additional MPLS network-wide operation recommendations to avoid attacks. Please note that source IP address filtering provides only a weak form of access control and is not, in general, a reliable security mechanism. Nonetheless, it is required here in the absence of any more robust mechanisms that might be used.

# 13. IANA Considerations

## 13.1. LSR Capability TLV

IANA has assigned value 4 (from the range 0-16383) for the LSR Capability TLV from the "TLVs" registry under the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA-MPLS-LSP-PING].

| Type | TLV Name | Reference |
|---|---|---|
| 4 | LSR Capability | RFC 8611 |

### 13.1.1. LSR Capability Flags

IANA has created a new "LSR Capability Flags" registry. The initial contents are as follows:

| Value | Meaning | Reference |
|---|---|---|
| 31 | D: Downstream LAG Info Accommodation | RFC 8611 |
| 30 | U: Upstream LAG Info Accommodation | RFC 8611 |
| 0-29 | Unassigned | |

Assignments of LSR Capability Flags are via Standards Action [RFC8126].

## 13.2. Local Interface Index Sub-TLV

IANA has assigned value 4 (from the range 0-16383) for the Local Interface Index Sub-TLV from the "Sub-TLVs for TLV Type 20" subregistry of the "TLVs" registry in the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA-MPLS-LSP-PING].

| Sub-Type | Sub-TLV Name | Reference |
|---|---|---|
| 4 | Local Interface Index | RFC 8611 |

### 13.2.1. Interface Index Flags

IANA has created a new "Interface Index Flags" registry. The initial contents are as follows:

| Bit Number | Name | Reference |
| --- | --- | --- |
| 15 | M: LAG Member Link Indicator | RFC 8611 |
| 0-14 | Unassigned | |

Assignments of Interface Index Flags are via Standards Action [RFC8126].

Note that this registry is used by the Interface Index Flags field of the following sub-TLVs:

- The Local Interface Index Sub-TLV, which may be present in the Downstream Detailed Mapping TLV.
- The Remote Interface Index Sub-TLV, which may be present in the Downstream Detailed Mapping TLV.
- The Incoming Interface Index Sub-TLV, which may be present in the Detailed Interface and Label Stack TLV.

## 13.3. Remote Interface Index Sub-TLV

IANA has assigned value 5 (from the range 0-16383) for the Remote Interface Index Sub-TLV from the "Sub-TLVs for TLV Type 20" subregistry of the "TLVs" registry in the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA-MPLS-LSP-PING].

| Sub-Type | Sub-TLV Name | Reference |
| --- | --- | --- |
| 5 | Remote Interface Index | RFC 8611 |

## 13.4. Detailed Interface and Label Stack TLV

IANA has assigned value 6 (from the range 0-16383) for the Detailed Interface and Label Stack TLV from the "TLVs" registry in the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA-MPLS-LSP-PING].

| Type | TLV Name | Reference |
| --- | --- | --- |
| 6 | Detailed Interface and Label Stack | RFC 8611 |

### 13.4.1. Sub-TLVs for TLV Type 6

RFC 8029 changed the registration procedures for TLV and sub-TLV registries for LSP Ping.

IANA has created a new "Sub-TLVs for TLV Type 6" subregistry under the "TLVs" registry of the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA-MPLS-LSP-PING].

This registry conforms with RFC 8029.

The registration procedures for this sub-TLV registry are:

```
Range        Registration Procedure  Note
-----        ----------------------  -----
0-16383      Standards Action        This range is for mandatory
                                     TLVs or for optional TLVs that
                                     require an error message if
                                     not recognized.
16384-31743  RFC Required            This range is for mandatory
                                     TLVs or for optional TLVs that
                                     require an error message if
                                     not recognized.
31744-32767  Private Use             Not to be assigned
32768-49161  Standards Action        This range is for optional TLVs
                                     that can be silently dropped if
                                     not recognized.
49162-64511  RFC Required            This range is for optional TLVs
                                     that can be silently dropped if
                                     not recognized.
64512-65535  Private Use             Not to be assigned
```

The initial allocations for this registry are:

```
Sub-Type     Sub-TLV Name             Reference Comment
--------     ------------             --------- -------
0            Reserved                 RFC 8611
1            Incoming Label Stack     RFC 8611
2            Incoming Interface Index RFC 8611
3-31743      Unassigned
31744-32767                           RFC 8611  Reserved for
                                                Private Use
32768-64511  Unassigned
64512-65535                           RFC 8611  Reserved for
                                                Private Use
```

Note: IETF does not prescribe how the Private Use sub-TLVs are handled; however, if a packet containing a sub-TLV from a Private Use ranges is received by an LSR that does not recognize the sub-TLV, an error message MAY be returned if the sub-TLV is from the range 31744-32767, and the packet SHOULD be silently dropped if it is from the range 64511-65535.

### 13.4.2. Interface and Label Stack Address Types

The Detailed Interface and Label Stack TLV shares the Interface and Label Stack Address Types with the Interface and Label Stack TLV. To reflect this, IANA has updated the name of the registry from "Interface and Label Stack Address Types" to "Interface and Label Stack and Detailed Interface and Label Stack Address Types".

## 13.5. DS Flags

IANA has assigned a new bit number from the "DS Flags" subregistry of the "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters" registry [IANA-MPLS-LSP-PING].

Note: the "DS Flags" subregistry was created by [RFC8029].

```
Bit number Name                                        Reference
---------- ----------------------------------------   ---------
     3     G: LAG Description Indicator                RFC 8611
```

# 14. References

## 14.1. 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 , <https://www.rfc-editor.org/info/rfc2119>.

**[RFC8029]** Kompella, K., Swallow, G., Pignataro, C., Ed., Kumar, N., Aldrin, S., and M. Chen, "Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures", RFC 8029, DOI 10.17487/RFC8029, March 2017 , <https://www.rfc-editor.org/info/rfc8029>.

**[RFC8126]** Cotton, M., Leiba, B., and T. Narten, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 8126, DOI 10.17487/RFC8126, June 2017 , <https://www.rfc-editor.org/info/rfc8126>.

**[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>.

### 14.2. Informative References

**[IANA-MPLS-LSP-PING]** IANA, "Multiprotocol Label Switching (MPLS) Label Switched Paths (LSPs) Ping Parameters", July 2019 , <https://www.iana.org/assignments/mpls-lsp-ping-parameters/>.

**[IEEE802.1AX]** IEEE, "IEEE Standard for Local and metropolitan area networks - Link Aggregation", IEEE Std. 802.1AX, July 2019.

**[RFC5920]** Fang, L., Ed., "Security Framework for MPLS and GMPLS Networks", RFC 5920, DOI 10.17487/RFC5920, July 2010 , <https://www.rfc-editor.org/info/rfc5920>.

**[RFC6425]** Saxena, S., Ed., Swallow, G., Ali, Z., Farrel, A., Yasukawa, S., and T. Nadeau, "Detecting Data-Plane Failures in Point-to-Multipoint MPLS - Extensions to LSP Ping", RFC 6425, DOI 10.17487/RFC6425, November 2011 , <https://www.rfc-editor.org/info/rfc6425>.

**[RFC7439]** George, W., Ed. and C. Pignataro, Ed., "Gap Analysis for Operating IPv6-Only MPLS Networks", RFC 7439, DOI 10.17487/RFC7439, January 2015 , <https://www.rfc-editor.org/info/rfc7439>.

## Appendix A. LAG with Intermediate L2 Switch Issues

Several flavors of provisioning models that use a "LAG with L2 switch" and the corresponding MPLS data-plane ECMP traversal validation issues are described in this appendix.

## A.1. Equal Numbers of LAG Members

```
R1 ==== S1 ==== R2
```

The issue with this LAG provisioning model is that packets traversing a LAG member from Router 1 (R1) to intermediate L2 switch (S1) can get load-balanced by S1 towards Router 2 (R2). Therefore, MPLS echo request messages traversing a specific LAG member from R1 to S1 can actually reach R2 via any of the LAG members, and the sender of the MPLS echo request messages has no knowledge of this nor any way to control this traversal. In the worst case, MPLS echo request messages with specific entropies will exercise every LAG member link from R1 to S1 and can all reach R2 via the same LAG member link. Thus, it is impossible for the MPLS echo request sender to verify that packets intended to traverse a specific LAG member link from R1 to S1 did actually traverse that LAG member link and to deterministically exercise "receive" processing of every LAG member link on R2. (Note: As far as we can tell, there's not a better option than "try a bunch of entropy labels and see what responses you can get back", and that's the same remedy in all the described topologies.)

## A.2. Deviating Numbers of LAG Members

```
                ____
R1 ==== S1 ==== R2
```

There are deviating numbers of LAG members on the two sides of the L2 switch. The issue with this LAG provisioning model is the same as with the previous model: the sender of MPLS echo request messages has no knowledge of the L2 load-balancing algorithm nor entropy values to control the traversal.

## A.3. LAG Only on Right

```
R1 ---- S1 ==== R2
```

The issue with this LAG provisioning model is that there is no way for an MPLS echo request sender to deterministically exercise both LAG member links from S1 to R2. And without such, "receive" processing of R2 on each LAG member cannot be verified.

## A.4. LAG Only on Left

```
R1 ==== S1 ---- R2
```

The MPLS echo request sender has knowledge of how to traverse both LAG members from R1 to S1. However, both types of packets will terminate on the non-LAG interface at R2. It becomes impossible for the MPLS echo request sender to know that MPLS echo request messages intended to traverse a specific LAG member from R1 to S1 did indeed traverse that LAG member.

## Acknowledgements

The authors would like to thank Nagendra Kumar and Sam Aldrin for providing useful comments and suggestions. The authors would like to thank Loa Andersson for performing a detailed review and providing a number of comments.

The authors also would like to extend sincere thanks to the MPLS RT review members who took the time to review and provide comments. The members are Eric Osborne, Mach Chen, and Yimin Shen. The suggestion by Mach Chen to generalize and create the LSR Capability TLV was tremendously helpful for this document and likely for future documents extending the MPLS LSP Ping and Traceroute mechanisms. The suggestion by Yimin Shen to create two separate validation procedures had a big impact on the contents of this document.

## Authors' Addresses

**Nobo Akiya**
Big Switch Networks
Email: nobo.akiya.dev@gmail.com

**George Swallow**
Southend Technical Center
Email: swallow.ietf@gmail.com

**Stephane Litkowski**
Orange
Email: stephane.litkowski@orange.com

**Bruno Decraene**
Orange
Email: bruno.decraene@orange.com

**John E. Drake**
Juniper Networks
Email: jdrake@juniper.net

**Mach(Guoyi) Chen**
Huawei
Email: mach.chen@huawei.com