Stream: Internet Engineering Task Force (IETF)
RFC: 9541
Category: Standards Track
Published: February 2024
ISSN: 2070-1721
Authors: J. Rabadan, Ed. *Nokia*; S. Sathappan *Nokia*; K. Nagaraj *Nokia*; M. Miyake *Softbank*; T. Matsuda *Softbank*

RFC 9541
Flush Mechanism for Customer MAC Addresses Based on Service Instance Identifier (I-SID) in Provider Backbone Bridging EVPN (PBB-EVPN)

Abstract

Provider Backbone Bridging (PBB) can be combined with Ethernet Virtual Private Networks (EVPNs) to deploy Ethernet Local Area Network (E-LAN) services in large Multiprotocol Label Switching (MPLS) networks. That combination is what we refer to as "PBB-EVPN." Single-Active multihoming and per Service Instance Identifier (I-SID) load-balancing can be provided to access devices and aggregation networks. In order to speed up the network convergence in case of failures on Single-Active multihomed Ethernet Segments (ESs), PBB-EVPN defines a flush mechanism for Customer MACs (C-MACs) called "C-MAC flush" that works for different Ethernet Segment Backbone MAC (B-MAC) address allocation models. This document complements those C-MAC flush procedures for cases in which no PBB-EVPN ESs are defined (i.e., the attachment circuit is associated with a zero Ethernet Segment Identifier (ESI)) and the C-MAC flush requires I-SID-level granularity.

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/rfc9541.

Copyright Notice

Table of Contents

1. Introduction

[RFC7623] defines how Provider Backbone Bridging (PBB) can be combined with Ethernet Virtual Private Networks (EVPNs) to deploy E-LAN services in very large MPLS networks. [RFC7623] also describes how Single-Active multihoming and per-I-SID load-balancing can be provided to access devices and aggregation networks. When Access Ethernet and/or MPLS networks exist, [EVPN-VIRTUAL-ES] describes how virtual Ethernet Segments (ESs) can be associated with a group of Ethernet Virtual Circuits (EVCs) or even pseudowires (PWs). In order to speed up the network convergence in case of failures on Single-Active multihomed ESs, [RFC7623] defines a Customer MAC flush mechanism that works for different ES B-MAC address allocation models.

In some cases, the administrative entities that manage the access devices or aggregation networks do not demand multihomed ESs from the PBB-EVPN provider, but simply demand multiple single-homed ESs. Single-homed ESs use null ESIs, whereas multihomed ESs use non-null ESIs. If using single-homed ESs, the PBB-EVPN network is no longer aware of the redundancy offered by the access administrative entity. Figure 1 shows an example where the PBB-EVPN network provides four different Attachment Circuits (ACs) for I-SID1, with those ACs not being part of any ES or virtual ES. (Therefore, they are referred to as null virtual Ethernet Segments.)

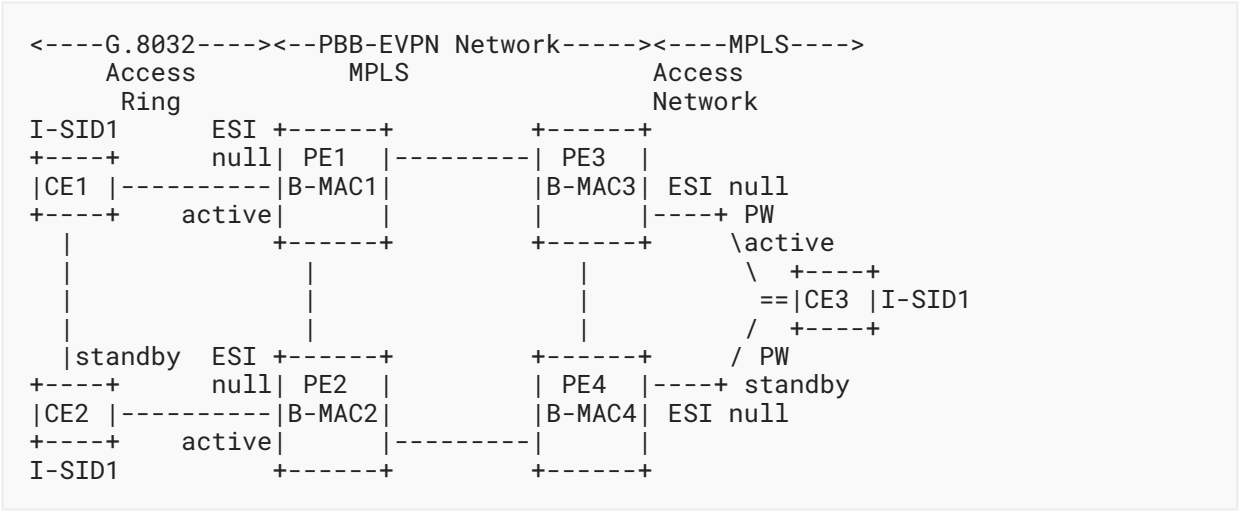

Figure 1: PBB-EVPN and Non-ES-Based Redundancy

In the example in Figure 1, CE1, CE2, and CE3 are attached to the same service, identified by I-SID1, in the PBB-EVPN PEs. CE1 and CE2 are connected to the PEs via "Ethernet ring protection switching" as specified in [G.8032], and their ACs to PE1 and PE2 are represented by a port and VLAN identifier. CE3 is dual-homed to PE3 and PE4 through an active/standby PW, and its AC to the PEs is represented by a PW. Each of the four PEs uses a dedicated Backbone MAC address as a source MAC address (B-MAC1, B-MAC2, B-MAC3, and B-MAC4, respectively) when encapsulating customer frames in PBB packets and forwarding those PBB packets to the remote PEs as per

[RFC7623]. There are no multihomed ESs defined in the PBB-EVPN network of the example; that is why the four ACs in Figure 1 show the text "ESI null", which means the Ethernet Segment Identifier on those ACs is zero. Since there are no multihomed ESs defined, the PEs keep their ACs active as long as the physical connectivity is established and the CEs are responsible for managing the redundancy, avoiding loops, and providing per-I-SID load-balancing to the PBB-EVPN network. Examples of CEs managing their own redundancy are described in [G.8032], or [RFC4762] for active/standby PWs.

For instance, in normal conditions, CE2 will block its link to CE1 and CE3 will block its forwarding path to PE4. In this situation, a failure in one of the redundant ACs will trigger the CEs to start using their redundant paths; however, those failures will not trigger any C-MAC flush procedures in the PEs that implement [RFC7623], since the PEs are not using the PBB-EVPN multihoming procedures. For example:

- If the active PW from CE3 (to PE3) fails and the failure is due to the entire PE3 node failing, then the procedures in [RFC7623] guarantee that all the remote PEs flush all the C-MACs associated with B-MAC3 (the B-MAC of PE3). In this case, CE3 detects the fault due to the PW going operationally down.
- However, if the active PW from CE3 (to PE3) fails (but PE3 is still operationally up), following the procedures in [RFC7623], neither PE3 nor PE4 issue a C-MAC flush message. Therefore, the remote PEs will continue pointing at PE3's B-MAC to reach CE3's C-MACs, until the C-MACs age out in the I-SID1 forwarding tables. In this case, PE3 may use any of the existing PW notifications so that CE3 switches the active PW to PE4.
- The same issue is exposed when the active PW from CE3 switches over from PE3 to PE4 due to a manual operation on CE3; that is, neither PE3 nor PE4 trigger any C-MAC flush notification and the remote PEs continue sending the traffic to PE3 until the C-MACs age out.

[RFC7623] provides a C-MAC flush solution based on a shared B-MAC update along with the MAC Mobility extended community where the sequence number is incremented. However, the procedure is only used along with multihomed ESs. Even if that procedure could be used for null ESs, as in the example of Figure 1, the Customer MAC flush procedure in [RFC7623] would result in unnecessary flushing of unaffected I-SIDs on the remote PEs, and subsequent flooding of unknown unicast traffic in the network. For instance, in the case that CE3 switches its active PW over to PE4, a potential solution reusing the existing C-MAC flush notifications in [RFC7623] is for PE3 to issue an update for the MAC/IP Advertisement route of B-MAC3 with a higher sequence number. However, this update would cause unnecessary Customer MAC flushing for all the I-SIDs attached to PE3 (supposing multiple I-SIDs in PE3) rather than for only I-SID1.

This document describes an extension of the Customer MAC flush procedures in [RFC7623], so that in the failure example above, PE3 can trigger a Customer MAC flush notification that makes PE1, PE2, and PE4 flush all the Customer MACs associated with PE3's B-MAC3 and (only) I-SID1. In order to do so, this specification introduces the encoding of the I-SID in the MAC/IP Advertisement routes advertised for the B-MACs. The C-MAC flush procedure explained in this document is referred to as "PBB-EVPN I-SID-based C-MAC flush" and can be used in PBB-EVPN networks with null or non-null (virtual) ESs.

This specification assumes that the reader is familiar with the procedures in [RFC7623].

1.1. Abbreviations

AC: Attachment Circuit

B-MAC: Backbone MAC

CE: Customer Edge

C-MAC: Customer MAC

ES: Ethernet Segment

ESI: Ethernet Segment Identifier

EVI: EVPN Instance

EVPN: Ethernet Virtual Private Network (as in [RFC7432])

I-SID: Service Instance Identifier

MAC: Media Access Control

MAC-VRF: MAC Virtual Routing and Forwarding

PBB-EVPN: Provider Backbone Bridging and EVPN (as in [RFC7623])

PE: Provider Edge

1.2. Terminology and Conventions

Familiarity with the terminology in [RFC7623] is expected.

B-MAC/0 route: This is an EVPN MAC/IP Advertisement route that uses a B-MAC in the MAC address field and a zero Ethernet Tag ID.

B-MAC/I-SID route: This is an EVPN MAC/IP Advertisement route that uses a B-MAC in the MAC address field and an I-SID in the Ethernet Tag field. It is used to notify remote PEs about the required C-MAC flush procedure for the C-MACs associated with the advertised B-MAC and I-SID.

G.8032: Refers to Ethernet ring protection switching as described in [G.8032].

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. Solution Requirements

The following requirements are followed by the C-MAC flush solution described in this document:

a. The solution **MUST** prevent packet-loss scenarios in case of failures on null ES ACs (Attachment Circuits not associated with an ES; that is, their corresponding ESI is zero) when the access device or access network is responsible for the redundancy.
b. This extension described in this document **MUST** work with Single-Active non-null ESs and virtual ESs, irrespective of the PE B-MAC address assignment (dedicated per-ES B-MAC or shared B-MAC, as in [RFC7623]).
c. In case of failure on the egress PE, the solution **MUST** provide a C-MAC flush notification at the B-MAC and I-SID granularity level.
d. The solution **MUST** provide a reliable C-MAC flush notification in PBB-EVPN networks that use Route Reflectors (RRs). MAC flushing needs to be provided to all affected I-SIDs in spite of the BGP flush notification messages being aggregated at the RR.
e. The solution **MUST** coexist in [RFC7623] networks where there are PEs that do not support this specification.
f. The solution **SHOULD** be enabled or disabled by an administrative option on a per-PE and per-I-SID basis (as opposed to always being enabled for all the I-SIDs).

3. EVPN BGP Encoding for I-SID-Based C-MAC Flush

The solution does not use any new BGP attributes but reuses the MAC Mobility extended community as an indication of C-MAC flush (as in [RFC7623]) and encodes the I-SID in the Ethernet Tag field of the EVPN MAC/IP advertisement route. As a reference, Figure 2 shows the MAC Mobility extended community and the EVPN MAC/IP advertisement route that are used as specified in [RFC7432] and used in this document as a C-MAC flush notification message.

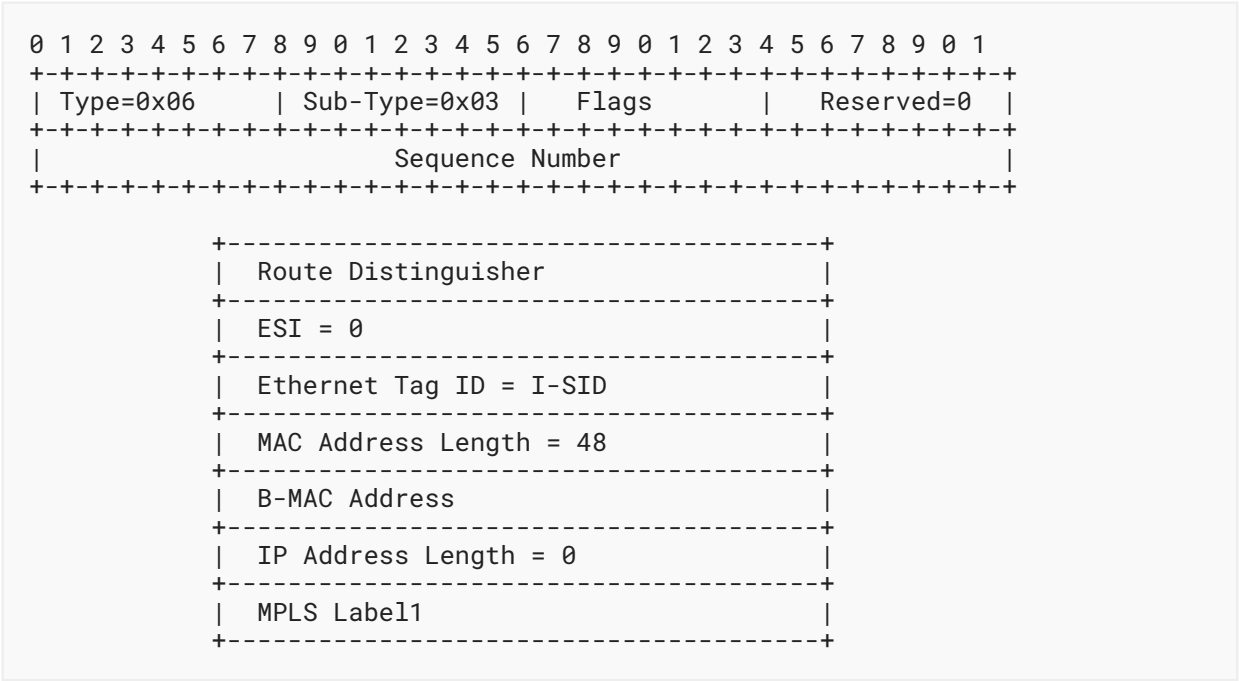

Figure 2: C-MAC Flush Notification Encoding: B-MAC/I-SID Route

Where:

- The route's route distinguisher and route targets are the ones corresponding to its EVI. Alternatively to the EVI's Route Target (RT), the route **MAY** be tagged with an RT auto-derived from the Ethernet Tag (I-SID) instead. [RFC7623] describes how the EVPN MAC/IP Advertisement routes can be advertised along with the EVI RT or an RT that is derived from the I-SID.
- The Ethernet Tag encodes the I-SID. This indicates to the PE that it must flush the C-MACs for that encoded I-SID, upon reception of the route.
- The MAC address field encodes the B-MAC address. This indicates to the PE that it must flush the C-MACs associated with that encoded B-MAC, upon reception of the route.
- The MAC Mobility extended community is used as in [RFC7623], where an increment in the sequence number between two updates for the same B-MAC/I-SID will be interpreted as a C-MAC flush notification for the corresponding B-MAC and I-SID.

All the other fields are set and used as defined in [RFC7623]. This document will refer to this route as the "B-MAC/I-SID route", as opposed to the EVPN MAC/IP Advertisement route used in [RFC7623] that contains a specific B-MAC and the Ethernet Tag ID set to zero. This document uses the term "B-MAC/0 route" to represent a B-MAC route advertised with an Ethernet Tag ID = 0.

Note that this B-MAC/I-SID route will be accepted and reflected by any RR as specified in [RFC7432], since no new attributes or values are used. A PE receiving the route will process the received B-MAC/I-SID update only in the case of supporting the procedures described in this document.

4. Solution Description

Figure 1 will be used in the description of the solution. CE1, CE2, and CE3 are connected to ACs associated with I-SID1, where no (multihomed) ESs have been enabled, and the ACs and PWs are in active or standby state as per Figure 1.

Enabling or disabling I-SID-based C-MAC flush **SHOULD** be an administrative choice on the system that **MAY** be configured per I-SID (I-Component, Service Instance Component), as opposed to being configured for all I-SIDs. When enabled on a PE:

a. The PE will be able to generate B-MAC/I-SID routes as C-MAC Flush notifications for the remote PEs.
b. The PE will be able to process B-MAC/I-SID routes received from remote PEs.

The PE **MUST** follow the procedures in [RFC7623] for C-MAC flush. This specification provides some additional procedures when I-SID-based C-MAC flush is enabled.

This C-MAC flush specification is described in three sets of procedures:

- I-SID-based C-MAC flush activation
- C-MAC flush notification generation upon AC failures
- C-MAC flush process upon receiving a C-MAC flush notification

4.1. I-SID-Based C-MAC Flush Activation Procedures

The following behavior **MUST** be followed by the PBB-EVPN PEs following this specification. Figure 1 is used as a reference.

- As in [RFC7623], each PE advertises a shared B-MAC in a B-MAC/0 route (with B-MAC1, B-MAC2, B-MAC3, and B-MAC4 in the MAC address field, respectively). This is the B-MAC that each PE will use as B-MAC SA (Source Address) when encapsulating the frames received on any local single-homed AC. Each PE will import the received B-MAC/0 routes from the remote PEs and will install the B-MACs in its B-component (Backbone Component) MAC-VRF. For instance, PE1 will advertise B-MAC1/0 and will install B-MAC2, B-MAC3, and B-MAC4 in its MAC-VRF.
- Assuming I-SID-based C-MAC flush is activated for I-SID1, the PEs will advertise the shared B-MAC with I-SID1 encoded in the Ethernet Tag. That is, PE1 will advertise B-MAC1/1 and will receive B-MAC2/1, B-MAC3/1, and B-MAC4/1. The receiving PEs **MUST** use these B-MAC/I-SID routes only for C-MAC flush procedures and they **MUST NOT** be used to add/withdraw any B-MAC entry in the MAC-VRFs. As per [RFC7623], only B-MAC/0 routes can be used to add/withdraw B-MACs in the MAC-VRFs.
- The above procedure **MAY** also be used for dedicated B-MACs (B-MACs allocated per ES).

4.2. C-MAC Flush Generation

If, for instance, there is a failure on PE1's AC, PE1 will generate an update including B-MAC1/1 along with the MAC Mobility extended community where the Sequence Number has been incremented. The reception of the B-MAC1/1 with an increment in the sequence number will trigger the C-MAC flush procedures on the receiving PEs.

- An AC going operationally down **MUST** generate a B-MAC/I-SID with a higher Sequence Number. If the AC going down makes the entire local I-SID go operationally down, the PE will withdraw the B-MAC/I-SID route for the I-SID.
- An AC going operationally up **SHOULD NOT** generate any B-MAC/I-SID update, unless it activates its corresponding I-SID, in which case the PE will advertise the B-MAC/I-SID route.
- An AC receiving a G.8032 flush notification or a flush message in any other protocol from the access network **MAY** propagate it to the remote PEs by generating a B-MAC/I-SID route update with a higher Sequence Number.

4.3. C-MAC Flush Process upon Receiving a C-MAC Flush Notification

A PE receiving a C-MAC flush notification will follow these procedures:

- A received B-MAC/I-SID route (with a non-zero I-SID) **MUST NOT** add/remove any B-MAC to/from the MAC-VRF.
- An update of a previously received B-MAC/I-SID route with an increment Sequence Number **MUST** flush all the C-MACs associated with that I-SID and B-MAC. C-MACs associated with the same I-SID but different B-MAC **MUST NOT** be flushed.
- A received B-MAC/I-SID withdraw (with a non-zero I-SID) **MUST** flush all the C-MACs associated with that B-MAC and I-SID.

Note that the C-MAC flush procedures described in [RFC7623] for B-MAC/0 routes are still valid and a PE receiving [RFC7623] C-MAC flush notification messages **MUST** observe the behavior specified in [RFC7623].

5. Conclusions

The I-SID-based C-MAC flush solution described in this document has the following benefits:

a. The solution solves packet-loss scenarios in case of failures on null ES ACs, since the C-MAC flush procedures are independent of the ES definition.
b. This extension can also be used with Single-Active non-null ESs and virtual ESs, irrespective of the PE B-MAC address assignment (dedicated per-ES B-MAC or shared B-MAC).
c. It provides a C-MAC flush notification at B-MAC and I-SID granularity level, therefore flushing a minimum number of C-MACs and reducing the amount of unknown unicast flooding in the network.

d. It provides a reliable C-MAC flush notification in PBB-EVPN networks that use RRs. RRs will propagate the C-MAC flush notifications for all the affected I-SIDs, irrespective of the order in which the notifications make it to the RR.

e. The solution can coexist in a network with systems supporting or not supporting this specification. Non-supporting systems ignore the B-MAC/I-SID routes; however, they still follow the C-MAC flush procedures in [RFC7623].

6. Security Considerations

Security considerations described in [RFC7623] apply to this document.

In addition, this document suggests additional procedures that can be activated on a per I-SID basis and generate additional EVPN MAC/IP Advertisement routes in the network. The format of these additional EVPN MAC/IP Advertisement routes is backwards compatible with the procedures in [RFC7623] and should not create any issues for receiving PEs that do not follow this specification. However, the additional routes may consume extra memory and processing resources on the receiving PEs. Because of that, it is **RECOMMENDED** to activate this feature only when necessary (when multihomed networks or devices are attached to the PBB-EVPN PEs), and not by default in any PBB-EVPN PE.

7. IANA Considerations

This document has no IANA actions.

8. References

8.1. 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>.

[RFC7432] Sajassi, A., Ed., Aggarwal, R., Bitar, N., Isaac, A., Uttaro, J., Drake, J., and W. Henderickx, "BGP MPLS-Based Ethernet VPN", RFC 7432, DOI 10.17487/RFC7432, February 2015, <https://www.rfc-editor.org/info/rfc7432>.

[RFC7623] Sajassi, A., Ed., Salam, S., Bitar, N., Isaac, A., and W. Henderickx, "Provider Backbone Bridging Combined with Ethernet VPN (PBB-EVPN)", RFC 7623, DOI 10.17487/RFC7623, September 2015, <https://www.rfc-editor.org/info/rfc7623>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

8.2. Informative References

[EVPN-VIRTUAL-ES] Sajassi, A., Brissette, P., Schell, R., Drake, J., and J. Rabadan, "EVPN Virtual Ethernet Segment", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-virtual-eth-segment-14, 23 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-virtual-eth-segment-14>.

[G.8032] ITU-T, "Ethernet ring protection switching", ITU-T Recommendation G.8032/Y.1344, March 2020, <https://www.itu.int/rec/T-REC-G.8032-202003-I/en>.

[RFC4762] Lasserre, M., Ed. and V. Kompella, Ed., "Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling", RFC 4762, DOI 10.17487/RFC4762, January 2007, <https://www.rfc-editor.org/info/rfc4762>.

Acknowledgments

The authors want to thank Vinod Prabhu, Sriram Venkateswaran, Laxmi Padakanti, and Ranganathan Boovaraghavan for their review and contributions.

Authors' Addresses

Jorge Rabadan (EDITOR)
Nokia
520 Almanor Avenue
Sunnyvale, CA 94085
United States of America
Email: jorge.rabadan@nokia.com

Senthil Sathappan
Nokia
520 Almanor Avenue
Sunnyvale, CA 94085
United States of America
Email: senthil.sathappan@nokia.com

Kiran Nagaraj
Nokia
520 Almanor Avenue
Sunnyvale, CA 94085
United States of America
Email: kiran.nagaraj@nokia.com

M. Miyake
Softbank
Email: masahiro.miyake@g.softbank.co.jp

T. Matsuda
Softbank
Email: taku.matsuda@g.softbank.co.jp