

PCE Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 18, 2011

F. Zhang
Q. Zhao
Huawei Technologies
O. Gonzalez de Dios
Telefonica I+D
R. Casellas
CTTC
D. King
Old Dog Consulting
October 18, 2010

Extensions to
Path Computation Element Communication Protocol (PCEP)
for Hierarchical Path Computation Elements (PCE)
[draft-zhang-pcep-hierarchy-extensions-00](#)

Abstract

The hierarchical Path Computation Element (PCE) architecture defined in [[PCE-HIERARCHY-FWK](#)] allows the optimum sequence of domains to be selected, and the optimum end-to-end path to be derived through the use of a hierarchical relationship between domains.

This document defines the Path Computation Element Protocol (PCEP) extensions for the purpose of implementing hierarchical PCE procedures which are described in [[PCE-HIERARCHY-FWK](#)].

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

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

This Internet-Draft will expire on April 18, 2011.

Internet-Draft

October 2010

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction.....	3
2.	PCEP Extension Requirements.....	4
2.1.	New Objective Functions.....	4
2.2.	PCEP Request Qualifiers.....	4
2.3.	Discovery Between Parent and Child PCEs.....	4
2.3.1.	Parent PCE Capability Discovery.....	5
2.3.2.	PCE Domain and PCE ID Discovery.....	5
2.4.	Domain Connectivity Information Collection.....	5
2.5.	Error Case Handling.....	6
3.	PCEP Extensions.....	6
3.1.	Extensions to OPEN Object.....	7
3.1.1.	OF Codes.....	7
3.1.2.	OPEN Object Flags.....	7
3.1.3.	Domain-ID TLV.....	7
3.1.4.	PCE-ID TLV.....	8
3.1.5.	Procedures.....	8
3.2.	Extensions to RP Object.....	9
3.2.1.	RP Object Flags.....	9
3.2.2.	Domain-ID TLV.....	9
3.2.3.	Procedures.....	9
3.3.	Extensions to NOTIFICATION Object.....	9
3.3.1.	Notification Types.....	10
3.3.2.	Inter-domain Link TLV.....	10
3.3.3.	Inter-domain Node TLV.....	11
3.3.4.	Domain-ID TLV.....	11
3.3.5.	PCE-ID TLV.....	11

3.3.6. Procedures.....	12
3.4. Extensions to PCEP-ERROR Object.....	12
3.4.1. Hierarchy PCE Error-Type.....	12
3.4.2. Procedures.....	12
4. Manageability Considerations.....	13
5. IANA Considerations.....	13
6. Security Considerations.....	13
7. References.....	13
7.1. Normative References.....	13
7.2. Informative References.....	13

1. Introduction

[PCE-HIERARCHY-FWK] describes a hierarchy PCE architecture which can be used for computing the end-to-end paths of inter-domain MPLS Traffic Engineering (TE) and GMPLS Label Switched Paths (LSPs). In the hierarchy PCE architecture, the parent PCE can compute a domain path based on the domain connectivity information and the child PCE can compute the intra-domain path based on the domain topology information. The end-to-end domain path computing procedures can be abstracted as follows:

- o The PCC requests a child PCE to return an inter-domain path.
- o The child PCE forwards the request to the parent PCE.
- o The parent PCE computes one or multiple domain paths from the ingress domain to the egress domain.
- o The parent PCE sends the intra-domain path computation requests (between the domain border nodes) to the child PCEs which are responsible for the domains along the domain path(s).
- o The child PCEs return the intra-domain paths to the parent PCE.
- o The parent PCE constructs the end-to-end inter-domain path based on the intra-domain paths and returns the inter-domain path to the child PCE.
- o The child PCE forwards the inter-domain path to the PCC.

This document defines the PCEP extensions for the purpose of implementing hierarchy PCE procedures which are described in [PCE-HIERARCHY-FWK].

The document also uses a number of editor notes to describe options and alternative solutions. These options and notes will be removed before publication.

[1.1.](#) Terminology

This document uses the terminology defined in [\[RFC4655\]](#) and [\[RFC5440\]](#) and [\[PCE-HIERARCHY-FWK\]](#).

[1.2.](#) Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [\[RFC2119\]](#).

Zhang, et al.

[Page 3]

Internet-Draft

October 2010

[2.](#) PCEP Extension Requirements

[2.1.](#) New Objective Functions

For inter-domain path computation, there are three new objective functions which are defined in section 1.3.1 of [\[PCE-HIERARCHY-FWK\]](#).

- o Minimize the number of boundary nodes used.
- o Limit the number of domains crossed.
- o Disallow domain re-entry.

During the PCEP session establishment procedure, the PCE needs to be capable of indicating the objective functions (OF) capability in the Open message. This information can be, in turn, announced by child PCEs and used for selecting the PCE when a PCC want a path that satisfies a certain inter-domain objective function.

When a PCC requests a PCE to compute an inter-domain path, the PCC needs also to be capable of indicating the new objective functions for inter-domain path.

For the reasons described above, new OF codes need to be defined for the new inter-domain objective functions. Then the PCE can notify its new inter-domain objective functions to the PCC by carrying them in the OF-list TLV which is carried in the OPEN object. The PCC can

specify which objective function code to use, which is carried in the OF object when requesting a PCE to compute an inter-domain path.

[2.2.](#) PCEP Request Qualifiers

As described in section 5.8.1 of [[PCE-HIERARCHY-FWK](#)], support of the H-PCE architecture will introduce two new qualifications as follows:

- o It must be possible for a child PCE to indicate that the request it sends to a parent PCE should be satisfied by a domain sequence only, that is, not by a full end-to-end path. This allows the child PCE to initiate per-domain or backward recursive path computation.
- o A parent PCE needs to be able to ask a child PCE whether a particular node address (the destination of an end-to-end path) is present in the domain that the child PCE serves.

To meet the above requirements, the PCEP PCReq message should be extended.

[2.3.](#) Discovery Between Parent and Child PCEs

In the H-PCE architecture, the parent PCE does not need to be aware

Zhang, et al.

[Page 4]

Internet-Draft

October 2010

of each child domain topology. Therefore, it is possible that the parent PCE does not join the IGP instance of the child PCE domain, i.e. there is no IGP discovery mechanism between the parent PCE and child PCE.

Therefore there must be a discovery mechanism for basic PCE information between the parent and child PCEs. In this case, PCEP needs to provide discovery mechanisms that do not rely on IGP announcement/discovery procedures.

Editors note. A child PCE could forward the topology within PCNtf messages or any other mechanisms, without an IGP adjacency. Further discussion of the discovery mechanism and scope will be discussed in later versions of this document.

[2.3.1.](#) Parent PCE Capability Discovery

As described in [[PCE-HIERARCHY-FWK](#)], during the PCEP session establishment procedure, the child PCE needs to be capable of indicating to the parent PCE whether it requests the parent PCE

capability or not. The parent PCE needs also to be capable of indicating whether its parent capability can be provided to the child PCE or not.

[2.3.2.](#) PCE Domain and PCE ID Discovery

A PCE domain is a single domain with an associated PCE. it is possible for a PCE to manage multiple domains. The PCE domain may be an IGP area or AS.

The PCE ID is an IPv4 and/or IPv6 address that is used to reach the parent/child PCE. It is RECOMMENDED to use an address that is always reachable if there is any connectivity to the PCE.

The PCE ID information and PCE domain identifiers may be provided during the PCEP session establishment procedure or the domain connectivity information collection procedure.

[2.4.](#) Domain Connectivity Information Collection

As described in [[PCE-HIERARCHY-FWK](#)], the parent PCE builds the domain topology map either from configuration or from information received from each child PCE. A child PCE may report its neighbor domain connectivity to its parent PCE. It is reasonable to use PCEP PCNtf message to do this procedure. If an IGP adjacency is established between parent and children, it could be used for this purpose.

There are two types of domain border for providing the domain connectivity information:

- o Domain border is a TE link, e.g. the inter-AS TE link which connects two ASs.
- o Domain border is a node, e.g. the IGP ABR which connects two IGP areas.

For the inter-AS TE links, the following information needs to be notified to the parent PCE:

- o Identifier of advertising child PCE.
- o Identifier of PCE's domain.

- o Identifier of the link.
- o TE properties of the link (metrics, bandwidth)
- o Other properties of the link (technology-specific).
- o Identifier of link end-points.
- o Identifier of adjacent domain.

For the ABR, the following information needs to be notified to the parent PCE:

- o Identifier of the ABR.
- o Identifier of the IGP Area IDs.

[2.5.](#) Error Case Handling

A PCE that is capable of acting as a parent PCE might not be configured or willing to act as the parent for a specific child PCE. This fact could be determined when the child sends a PCReq that requires parental activity (such as querying other child PCEs), and could result in a negative response in a PCEP Error (PCErr) message and indicate the hierarchy PCE error types.

[3.](#) PCEP Extensions

3.1. Extensions to OPEN Object

3.1.1. OF Codes

There are three new OF codes defined here for H-PCE:

- o Name: Minimize the number of Boundary Nodes used (MBN)
Objective Function Code: (to be assigned by IANA, recommended 9)

Zhang, et al.

[Page 6]

Description: Find a path P such that passes through the least boundary nodes.

- o Name: Minimize the number of Transit Domains (MTD) 10)
Objective Function Code: (to be assigned by IANA, recommended
Description: Find a path P such that passes through the least

transit domains.

- o Name: Disallow Domain Re-entry (DDR)
Objective Function Code: (to be assigned by IANA, recommended 11)
Description: Find a path P such that does not entry a domain more than once.

3.1.2. OPEN Object Flags

There are two OPEN object flags defined here for H-PCE:

- o Parent PCE request bit (to be assigned by IANA, recommended bit 0):
if set it means the child PCE wishes to use the peer PCE as a parent PCE.
- o Parent PCE indication bit (to be assigned by IANA, recommended bit 1):
if set it means the PCE can be used as a parent PCE by the peer PCE.

Editors Note. It is possible that a parent PCE will also act as a child PCE.

3.1.3. Domain-ID TLV

The type of Domain-ID TLV is to be assigned by IANA (recommended 7). The length is 8 octets. The format of this TLV is defined below:

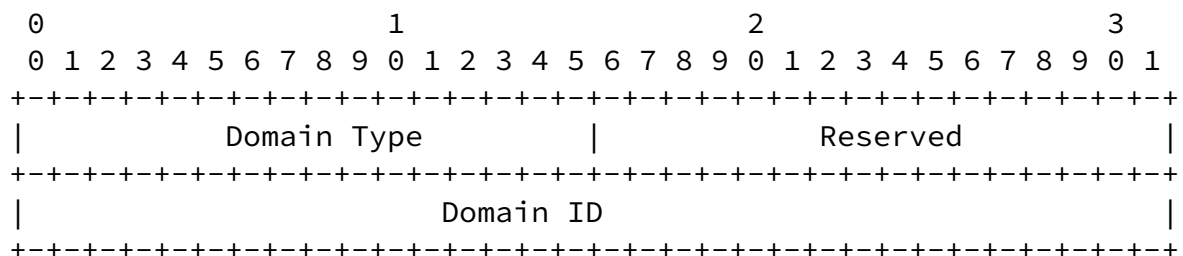


Figure 1: Domain-ID TLV

Domain Type (8 bits): Indicates the domain type. There are two types of domain defined currently:

- o Type=1: the Domain ID field carries an IGP Area ID.
- o Type=2: the Domain ID field carries an AS number.

Domain ID (32 bits): Indicates an IGP Area ID or AS number.

Editors note. It maybe necessary to support 64 bit domain IDs.

[3.1.4.](#) PCE-ID TLV

The type of PCE-ID TLV is to be assigned by IANA (recommended 8). The length is 4. The format of this TLV is defined below:

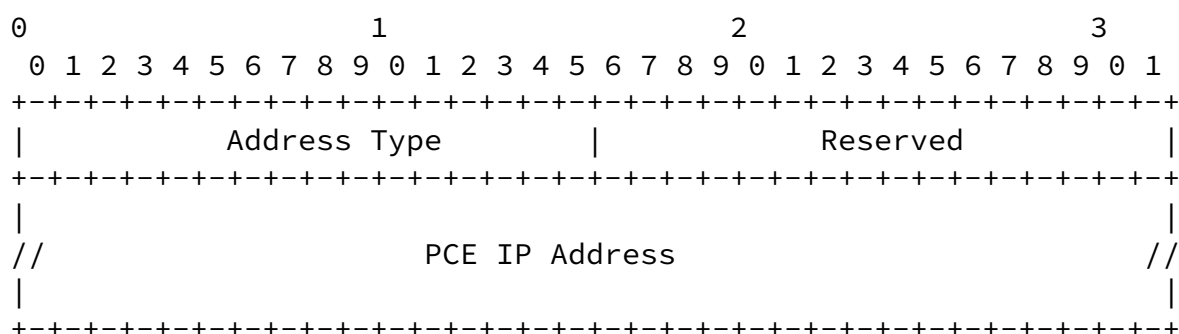


Figure 2: PCE ID TLV

Address Type (16 bits): Indicates the address type of PCE IP Address. 1 means IPv4 address type, 2 means IPv6 address type.

PCE IP Address: Indicates the reachable address of a PCE.

[3.1.5.](#) Procedures

The OF codes defined in this document can be carried in the OF-list TLV of the OPEN object. If the OF-list TLV carries the OF codes, it means that the PCE is capable of implementing the corresponding objective functions. This information can be used for selecting a proper parent PCE when a child PCE wants to get a path that satisfies a certain objective function.

If a child PCE wants to use the peer PCE as a parent, it can set the parent PCE request bit in the OPEN object carried in the Open message during the PCEP session creation procedure. If the peer PCE does not want to provide the parent function to the child PCE, it must send a PCErr message to the child PCE and clear the parent PCE indication bit in the OPEN object.

If the parent PCE can provide the parent function to the peer PCE, it may set the parent PCE indication bit in the OPEN object carried in the Open message during the PCEP session creation procedure.

The PCE may also report its PCE ID and list of domain ID to the peer PCE by specifying them in the PCE-ID TLV and List of Domain-ID TLVs in the OPEN object carried in the Open message during the PCEP session creation procedure.

Internet-Draft

October 2010

[3.2.](#) Extensions to RP Object

[3.2.1.](#) RP Object Flags

- o Domain Path Request bit (to be assigned by IANA, recommended bit 17): if set it means the child PCE wishes to get the domain sequence.
- o Destination Domain Query bit (to be assigned by IANA, recommended bit 16): if set it means the parent PCE wishes to get the destination domain ID.

[3.2.2.](#) Domain-ID TLV

The format of this TLV is defined in [section 2.1.3](#). This TLV can be carried in an OPEN object to indicate a (list of) managed domains, or carried in a RP object to indicate the destination domain ID when a child PCE responds to the parent PCE's destination domain query by a PCRep message.

Editors note. In some cases, the Parent PCE may need to allocate a node which is not necessarily the destination node.

[3.2.3.](#) Procedures

If a child PCE only wants to get the domain sequence for a multi-domain path computation from a parent PCE, it can set the Domain Path Request bit in the RP object carried in a PCReq message. The parent PCE which receives the PCReq message tries to compute a domain sequence for it. If the domain path computation succeeds the parent PCE sends a PCRep message which carries the domain sequence in the ERO to the child PCE. The domain sequence is specified as AS or AREA ERO sub-objects (type 32 for AS [[RFC3209](#)] or type. Otherwise it sends a PCReq message which carries the NO-PATH object to the child PCE.

The parent PCE can set the Destination Domain Query bit in a PCReq message to query the destination (which is specified in the END-POINTS objects) domain ID from a child PCE. If the child PCE knows the destination(s) domain ID, it sends a PCRep message to the parent PCE and specifies the domain ID in the Domain-ID TLV which is carried in the RP object. Otherwise it sends a PCRep message with a NO-PATH object to the parent PCE.

[3.3.](#) Extensions to NOTIFICATION Object

Because there will not be too many PCEP sessions between the child PCE(s) and parent PCE, it is recommended that the PCEP sessions between them keeping alive all the time . Then the child PCE can report all of the domain connectivity information to the parent PCE

Zhang, et al.

[Page 9]

Internet-Draft

October 2010

when the PCEP session is established successfully. It can also notify the parent PCE to update or delete the domain connectivity information when it detects the changes.

[3.3.1.](#) Notification Types

There is a new notification type defined in this document :

- o Domain Connectivity Information notification-type (to be assigned by IANA, recommended 3).

Notification-value=0: sent from the parent to the child to query all of the domain connectivity information maintained by the child PCE.

Notification-value=1: sent from the child to the parent to update the domain connectivity information maintained by the child PCE.

Notification-value=2: sent from the child to the parent to delete the domain connectivity information maintained by the child PCE.

[3.3.2.](#) Inter-domain Link TLV

IGP in each neighbor domain can advertise its inter-domain TE link capabilities [[RFC5316](#)], [[RFC5392](#)]. This information can be collected by the child PCEs and forwarded to the parent PCE. PCEP Inter-domain Link TLV is used for carrying the inter-domain TE link attributes for this purpose. Each Inter-domain Link TLV can carry the attributes of one inter-domain link at the most.

The type of Inter-domain Link TLV is to be assigned by IANA (recommended 9). The length is variable. The format of this TLV is defined 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

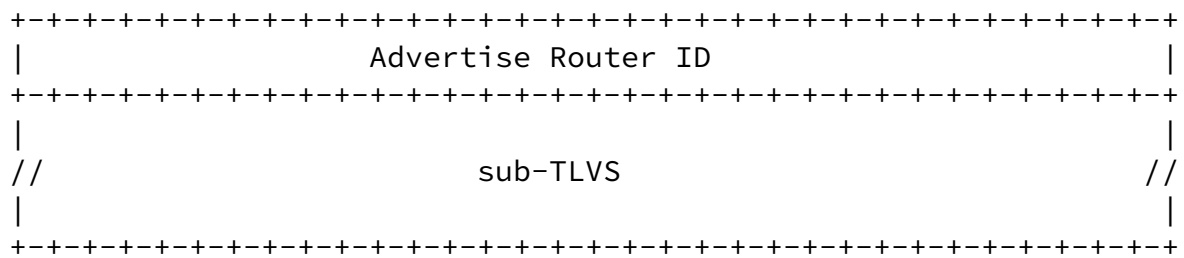


Figure 3: Inter-domain Link TLV

Advertise Router ID (32 bits): indicates the router ID which advertises the TE LSA or LSP.

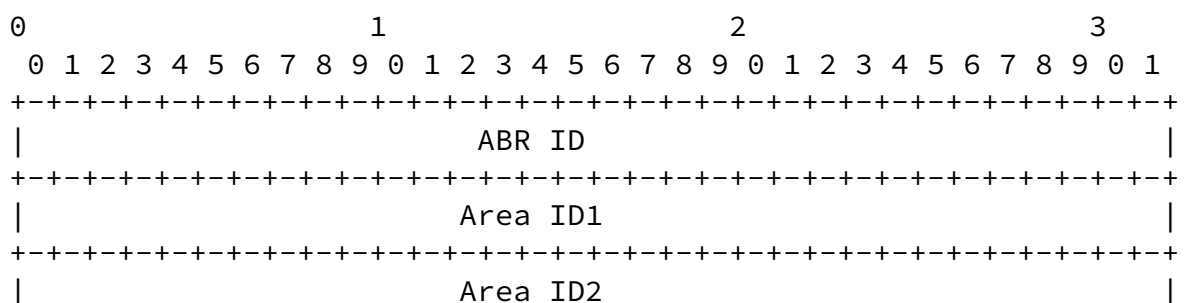
Sub-TLVs: the OSPF sub-TLVs for a TE link which defined in [[RFC5392](#)] and other associated OSPF RFCs. It is noted that if the IGP is IS-IS for the child domain the sub-TLVs must be converted to the OSPF sub-TLVs format when sending this information to the parent PCE through PCEP PCNtf message.

Each inter-domain link is identified by the combination of advertise router ID and the link local IP address or link local unnumbered identifier. The PCNtf message which is used for notifying the parent PCE to update or delete a inter-domain link must contain the information identifies a TE link exclusively.

3.3.3. Inter-domain Node TLV

The Inter-domain Node TLV carries only the two adjacent domain ID and the router (IGP ABR) ID.

The type of Inter-domain Node Information TLV is to be assigned by IANA (recommended 10). The length is variable . The format of this TLV is defined below:



re-sending the PCNtf message(s) with the newly information.

When the child PCE detects a deletion of domain connectivity (e.g., the inter-domain link TLV is aged out), it must notify the parent PCE to delete the inter-domain link by sending the PCNtf message. The notification-type is 3 and notification-value is 2 in the NOTIFICATION object.

[3.4.](#) Extensions to PCEP-ERROR Object

[3.4.1.](#) Hierarchy PCE Error-Type

A new PCEP Error-Type is allocated for hierarchy PCE (to be assigned by IANA, recommended 11):

Error-Type	Meaning
11	Hierarchy PCE error Error-value=1: parent PCE capability can not be provided

[3.4.2.](#) Procedures

When a specific child PCE sends a PCReq to a peer PCE that requires parental activity and the peer PCE does not want to act as the parent for it, the peer PCE should send a PCErr message to the child PCE and specify the error-type (11) and error-value (1) in the PCEP-ERROR object.

[4.](#) Manageability Considerations

TBD.

[5.](#) IANA Considerations

TBD.

[6.](#) Security Considerations

TBD.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.
- [RFC5440] Vasseur, JP., Ed., and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", [RFC 5440](#), March 2009.

7.2. Informative References

- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.
- [RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", [RFC 4655](#), August 2006.
- [RFC5316] M. Chen, R. Zhang, X. Duan, " ISIS Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering ", [RFC 5316](#), December 2008.
- [RFC5392] M. Chen, R. Zhang, X. Duan, " OSPF Extensions in Support of Inter-Autonomous System (AS) MPLS and GMPLS Traffic Engineering ", [RFC 5316](#), January 2009.
- [PCE-HIERARCHY-FWK] D. King, A. Farrel, " The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS ", [draft-king-pce-hierarchy-fwk-05](#), September 2010.

Zhang, et al.

[Page 13]

Internet-Draft

October 2010

Authors' Addresses

Fatai Zhang
Huawei Technologies
F3-5-B R&D Center, Huawei Base
Bantian, Longgang District
Shenzhen 518129 P.R.China
Phone: +86-755-28972912

Email: zhangfatai@huawei.com

Quintin Zhao
Huawei Technology
125 Nagog Technology Park
Acton, MA 01719
US
Email: qzhao@huawei.com

Oscar Gonzalez de Dios
Telefonica I+D
Emilio Vargas 6, Madrid
Spain
Email: ogondio@tid.es

Ramon Casellas
CTTC - Centre Tecnologic de Telecomunicacions de Catalunya
Av. Carl Friedrich Gauss n7
Castelldefels, Barcelona 08860
Spain
Email: ramon.casellas@cttc.es

Daniel King
Old Dog Consulting
Email: daniel@olddog.co.uk