Network Working Group   D. Purkayastha
Internet-Draft   A. Rahman
Intended status: Informational   D. Trossen
Expires: December 31, 2018   InterDigital Communications, LLC
                                                           T. Eckert
                                                             Huawei
                                                        June 29, 2018

BIER Multicast Overlay for HTTP Respone
draft-purkayastha-bier-multicast-http-response-00

Abstract

   HTTP Level multicast, using BIER, is described as a use case in BIER
   Use cases document. HTTP Level Multicast is used in today's video
   streaming and delivery services such as HLS, AR/VR etc., generally
   realized over IP Multicast. A realization of "HTTP Multicast" over
   "IP Multicast" is described and few problems are identified.
   Realization over BIER, through a BIER Multicast Overlay Layer, is
   described. How BIER Multicast Overlay operation improves over IP
   Multicast is also discussed.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

BIER Use Cases document [I-D.ietf-bier-use-cases] describes an "HTTP
Level Multicast" scenario, where HTTP Responses are carried over a
BIER multicast infrastructure to multiple clients. Especially rate-
adaptive HTTP solutions can benefit from the dynamic multicast group
membership changes enabled by BIER. For this, the "server side NAP
(Network Attachment Point), creates a list of outstanding client side
NAP (Network Attachment Point) requests for the same HTTP resource.
When the response is available, the list of NAPs with outstanding
client requests are converted into the BIER or BIER-TE bitstring and
used to send the HTTP response.

In this draft, we describe how this class of use cases can be
realized over IP Multicast and how the operation of the use case can
be improved if realized over BIER. The realization over BIER is
achieved through what is called in BIER "BIER Multicast overlay"
layer: The methods by which the sending BIER router knows what . The

requirements for BIER Multicast overlay layer is described. It also
describes the necessary functions that form the BIER multicast
overlay and the operations that enable the desired "HTTP Level
Multicast" behavior. One such operation is generating PATH ID
(represents the BFIR and BFER) based on named service relationship
and translating it to appropriate BIER header. We describe a list of
protocols needed for the realization of the individual operations.

We conclude with future steps and seek input from the WG.

1.1. Reference Deployment

Let us formulate the architecture of the BIER multicast overlay for
the scenario outlined in [I-D.ietf-bier-use-cases]. This overlay is
shown in Figure 1 below.

```
    +---------+   +------------+
    |         | |            |/
    +IP only  +---+  SR + BFER +-----|
    |receiver |   |   (CNAP)   |\    |
    |UE       |   +----/\------+     |
    +---------+        ||            |
                       ||     +----------+   +---------+
                       ||     |          |   |         |
             |-------- |  BFR      |---|   BFR    |------|
             |         |          |   |         |   |      |
             |         +----------+   +---------+      |
       +---------+                              +-------+
       |         |-------------------------------------->| BFIR  |
       + BIER TE +                              |  +    |
       |   PCE   |           +---------+   +-------+   |  SR   |
       |         |--||       |         |----| BFR   |----|(SNAP) |
       +---------+  ||       |  BFR    |    +-------+   |       |
                    ||       |         |              +-------+
                    ||     +---------+                  /|\
    +---------+   +------\/----+      |                   |
    |         | |            |/      |                   |
    +IP only  +---+  SR + BFER +------|              +----------+
    |receiver |   |   (CNAP)   |\                    | IP only  |
    +---------+   +------------+                     | Sender   |
                                                     |(Server)  |
                                                     +----------+
```
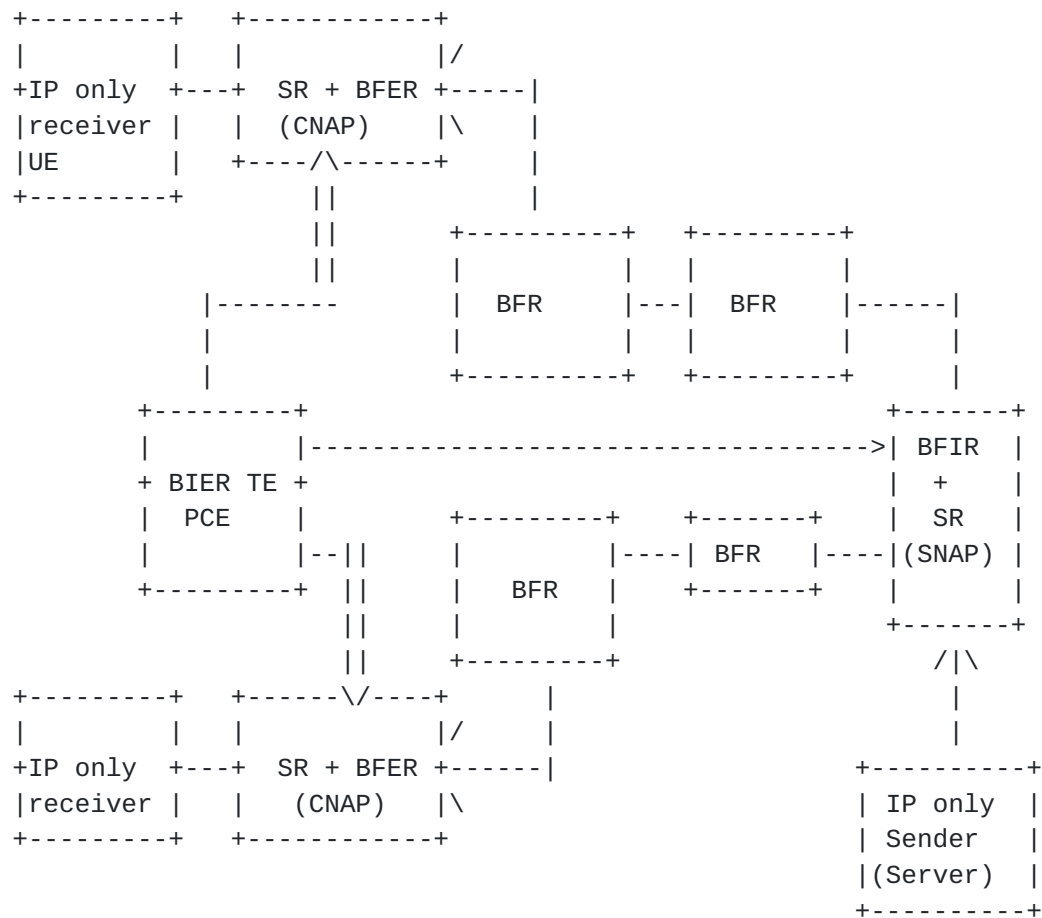
        [SR : Service Handler, CNAP : Client Network Attachment Point]
        [SNAP : Server Network Attachment Point]
        [PCE : Path Computation Element]

                        Figure 1: Deployment over BIER

    The multicast overlay is formed by the BFIR and BFER of the BIER
    layer and the additional SR (Service Handler) and PCE (Path
    Computation Element) elements shown in the figure. When connecting
    to a standard IP routed peering network, a special SR, such as Border
    Gateway may be used

    The Service Handler and BFER can be assumed to be collocated and can
    be viewed as Client Network Attachment Point (CNAP). Clients sends
    and receives HTTP transactions through CNAP.

    On the server side, the Service handling function can be part of the
    Server Network Attachment Point (SNAP). It also includes the BFIR
    function. SNAP is responsible for aggregating all HTTP Request and

sending BIER Multicasted HTTP response to multiple clients who
requested the same content.

As part of POINT/RIFE EU Horizon 2020 project, HTTP Level Multicast
use case has been executed on SDN based and ICN based underlay
network, as described in the [I-D.irtf-icnrg-deployment-guidelines].

"HTTP multicast" demonstrated benefits in HTTP-level streaming video
delivery, when deployed on POINT test bed with 80+ nodes. This draft
[I-D.irtf-icnrg-deployment-guidelines] also describes protocol
requirements to enable HTTP multicast to work on ICN underlay.

2. Conventions used in this document

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 [RFC2119].

3. Use cases

With the extensive use of "web technology", "distributed services"
and availability of heterogeneous network, HTTP has effectively
transitioned into the common transport or session layer for E2E and
multi-hop communication across the web that is also called Service
signaling. Multi-hop when using a sequence of HTTP instance such as
HTTP caches. The draft "On the use of HTTP as a Substrate"
[I-D.ietf-httpbis-bcp56bis], describes how HTTP is commonly used
among service instances to communicate with each other, thus
abstracting the lower layer details to application developers.

Referring to the BIER Use Cases [I-D.ietf-bier-use-cases], multicast
is used to scale out HLS (HTTP live streaming) to a large number of
receivers that use HTTP. This is used today in solutions like DOCSIS
hybrid streaming [TR_IPMC_ABR]. Multicast can speed up both live and
high-demand VoD streaming. Adaptive Bit Rate IPMC [TR_IPMC_ABR]
describes use of IP multicast towards the CMTS or a box beside it,
where the content is converted to HTTP/TCP to stream to the receivers
(e.g., homes). A server hosting the HLS content is shown as "NAP
Server". The gateways acting as receivers for the multicast from the
server are shown as "Client-NAP" (CNAP). Each CNAP can serve
multiple clients.

HTTP request and response used in media streaming services like HLS,
use HTTP response for delivery of content. In such scenarios, where
semi-synchronous access to the same resource occurs (such as watching
prominent videos over Netflix or similar platforms or live TV over
HTTP), traffic grows linearly with the number of viewers since the
HTTP-based server will provide an HTTP response to each individual

   viewer. This poses a significant burden on operators in terms of
   costs and on users in terms of likely degradation of quality.

   This solution is not limited to traditional TV broadcasting.
   Consider a virtual reality use case where several users are joining a
   VR session at the same time, e.g., centered around a joint event.
   Hence, due to the temporal correlation of the VR sessions, we can
   assume that multiple requests are sent for the same content at any
   point, particularly when viewing angles of VR clients are similar or
   the same. Due to availability of virtual functions and cloud
   technology, the actual end point from where content is delivered may
   change.

[4]. Realization over IP Multicast

   IPTV or Internet video distribution in CDNs, uses HTTP Level
   Multicast and realized over IP Multicast (IPMC). Many features of
   the IPTV service uses IPMC Group dependent state. Besides popular
   features like PIM, Mldp, in a variable bit rate encoded content
   source, content consumption also depends on group state.

   Assume clients that are consuming the same content (such as a TV
   program) and that this content has for each block (typically segments
   worth 2 seconds of content) a set of outstanding requests from its
   clients. When IP Multicast is used in the domain, such as in
   aforementioned pre-existing solutions like in Cablelabs/DOCSIS
   [TR_IPMC_ABR], all possible blocks of the content have to be mapped
   to some IP multicast group, and the CNAP will need to know the
   mapping of block to groups. For example, a live stream may have 11
   different bitrates available. In the most simple Block to IP
   multicast group mapping scheme, there could be 11 multicast groups,
   one for all the blocks of one bitrate (note that this is not
   necessarily done in deployments of this solution, but we consider it
   here for the purpose of explanation).

   If the multicast domain and especially the links into the CNAP has
   enough bandwidth, this solution work well with IP multicast. As soon
   as there is at least one Client connected to a CNAP for one
   particular content, the CNAP would join all 11 multicast groups for
   this content.

[4.1]. Mapping to Requirements

   To realize "HTTP Level Multicast" over "IP Multicast", some
   additional functions needs to be supported in an intermediate
   (overlay) layer.

Support of mapping between FQDN based end points, Multicast Address.
Creating multicast group from FQDN based end points.

Control mechanism related to time when to start sending response as
the multicast group is created. It is required that the source
should not send response immediately to the Multicast address. Wait
for some time to build the group sufficiently and then send response.

Support of IGMP signaling between User device, NAPs and Multicast
Router.

4.2. Problems

If the number of clients on a CNAP for a particular program is large,
the approach will work fairly well, because the likelihood that each
of the 11 bitrates of a content is necessary for at least one Client
is then fairly high.

When the number of receivers is not very large, IP multicast runs
into two issues. If all the bitrates for the content are sent across
the same group, then many of the bitrates may not be required and
would have to be received unnecessarily and dropped by the CNAP. If
each bitrate was sent on a different IP multicast group, the CNAP
could dynamically join/leave each multicast group based on the known
receivers, but that would create an extremely high and undesirable
amount of IP multicast signaling protocol activity (PIM/IGMP) that is
easily overloading the network

For efficiency reasons, the CNAP would need to dynamically join to
only those bitrate steams where it does have outstanding requests,
therefore achieving the best efficiency. This would mean in the
worst case that a CNAP would need to send for each new block, aka.:
every two second for every client one IGMP/PIM leave and one IGMP/PIM
join towards the upstream router to get a block for an appropriate
bitrate (or changed content) whenever bitrate or content on a client
have changed. This high rate of control-plane signaling between CNAP
and routers, and even between routers inside the multicast Domain is
a major pain point and may easily prohibit deployment of these
solutions because in many network devices, the performance of PIM/
IGMP is not scaled for continuous change in forwarding. Even worse,
the limit may not simply be the CPU performance of the routers
control plane, but a limitation in the number of changes in
forwarding that the forwarding plane units (NPU/ASICs) can support.

5. **Realization over BIER**

5.1. **Description of a "BIER Multicast Overlay" to support HTTP Multicast**

   The Service Handler (as in Figure 1) in BIER Multicast Overlay,
   process the FQDN in the service request. At the service level, e.g.
   HTTP service, the fixed relationship among consumer and providers may
   be abstracted using "Service Names", and the changing relationship at
   the Service execution endpoints can be managed at the "multicast
   overlay" level, handing out the exact locations where service request
   or response needs to be sent to BIER layer.


```
        +-------------+        +-----------+       +-----------+
        |             |        |           |       | PATH ID   |
        | Service name|        | Multicast |       | translates|
        | [producer,  |------->| Overlay   |------>| to BIER   |
        |  consumer]  |        | Layer     |       | header    |
        |             |        |           |       |           |
        +-------------+        +-----------+       +-----------+
```
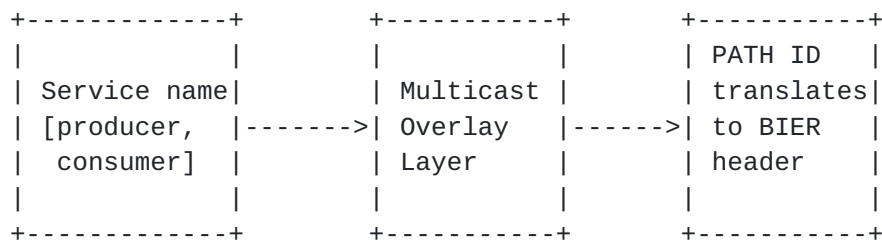

                Figure 2: Service name to Path ID translation

   We illustrate this using HTTP URI as service names. It should be
   noted, other identifiers can also be used as service name, such as IP
   address. In the example illustration, other layers such as TCP, IP
   has been abstracted. Outside BIER domain we terminate TCP/IP session
   to extract the URI. URI is processed by the "multicast overlay"
   layer to generate PATH IDENTIFIER, which is used as BIER header.
   Once the BIER header is determined and added at the BFIR, the rest of
   the transport layers is assumed to be any underlay technology as
   supported by BIER.

5.1.1. **BIER Multicast Overlay Componentst**

   With reference to Figure 1, the following components are part of BIER
   Multicast Overlay Layer.

   o SR : The Service handler terminates application level protocols,
      extracts the URI. It processes the URI in order to determine the
      PATH ID, which is used to send the HTTP Request.

   o Optional PCE : Path Computation Element keeps track of all service
      execution end points and how to reach them. SR may interact with
      PCE to obtain PATH information

o Interface functions to BFIR where the PATH ID is mapped to BIER
   header. An Interface to the BFER is likely not required because
   the BFER will only receive the traffic that they need and should
   be able to derive from the BIER payload which subset of its
   receivers need to get an HTTP encapsulated version of a particular
   reply.

5.1.2. BIER Multicast Overlay Operations

As shown in Figure 3, the "Multicast overlay function" includes a
function called PCE (Path Computation Element function), which is
responsible for selecting the correct multicast end point and
possibly realizing path policy enforcement. The result of the
selection is a BIER path identifier, which is delivered to the SR
upon initial path computation request (or provided to the ingress
router BFIR to be added as BIER header) (i.e., when sending a
request to or response for a specific URL for the first time). The
path identifier is utilized for any future request for a given URL-
based request.

All service end points indicate availability to the PCE through a
registration procedure, the PCE will instruct all SRs to invalidate
previous path identifiers to the specific URL. This may result in an
initial path computation request at the next service request
forwarding. Through this, the newly registered service endpoint
might be utilized if the policy-governed path computation selects
said service instance.

```
+-------+    +------+----+   +--------+                   +----+-----+
|Apps   |    |Apps---->  |   | PCE    |                   |    | APP |
|layer  |--->|layer | SR |   +---/\---+                   | SR-->    |
|prot   |    |prot  |    |   |   ||                       |    | LYR |
+-------+    +------+----+   +--------+   +---------+      +----+-----+
|   L2  |    |       L2   |-->|Forwarder|-->|Forwarder|-->|    L2    |
+-------+    +------+----+   +---------+   +---------+     +----------+
                    |--------BIER DOMAIN -------|
```
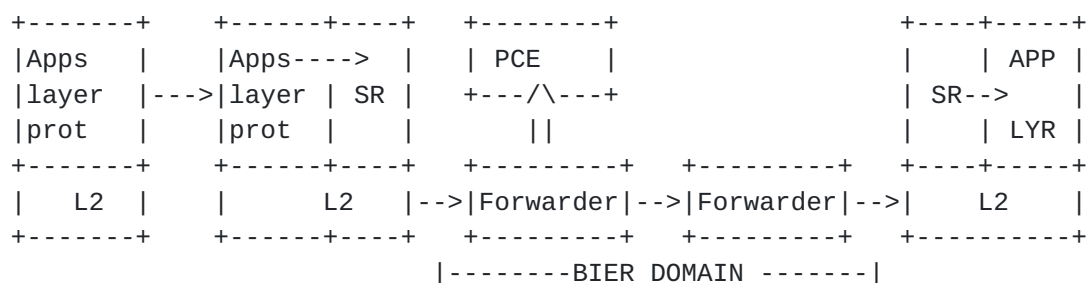
                 Figure 3: Protocol for Multicast Overlay Layer

In the diagram shown above, an HTTP request is sent by an IP-based
device towards the FQDN of the server defined in the HTTP request.

At the client facing SR, the HTTP request is terminated at the HTTP
level at a local HTTP proxy. We assume termination on the client
side at Layer 3 and above protocols, such as TCP. Server side SR at

the egress, terminates any transport protocol on the outgoing
(server) side. These terminating functions are assumed to be part of
the client/server SR. The SR obtains the destination "Service Name"
from the received HTTP request.

If no local BIER forwarding information exists to the client side SR,
a path computation entity (PCE) is consulted, which calculates a
unicast path from the BFIR to which the client SR is connected to the
BFER to which the server SR is connected. The PCE provides the
forwarding information (Path ID) to the client SR, which in turn
caches the result. The Client SR may forward the Path ID to BFIR,
which creates the BIER header.

```
        +------------+--------------+
        |            |              |
        | BIER HEADER | HTTP REQUEST |
        |            | [ENCODED IN  |
        |            | TEXT]        |
        |            |              |
        +------------+--------------+
```
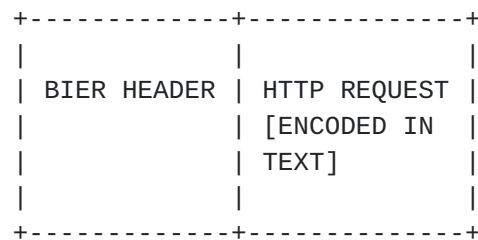
        Figure 4: Encapsulation of Service Request

Ultimately, the "HTTP Request" encapsulated by BIER header, as shown
in above diagram, is forwarded by the client SR towards the server-
facing SR via the local BFIR. We assume a (TCP-friendly) transport
protocol being used for the transmission between client and server SR
while not mandating the use of TCP for this transmission. A suitable
transport or Layer 2 encapsulation, as supported by BIER layer, is
added to the above payload as shown in the following diagram.

```
        +-------------+------------+--------------+
        |             |            |              |
        | Transport L2| BIER HEADER | HTTP REQUEST |
        |   HEADER    |            | [ENCODED IN  |
        |             |            | TEXT]        |
        |             |            |              |
        +------------+------------+--------------+
```

        Figure 5: Transport Encapsulation of BIER payload

Upon arrival of an HTTP request at the server SR, it forwards the
HTTP request as a well-formed HTTP request locally to the server.

If no BIER forwarding information exists for the reverse direction
towards the requesting client SR, this information is requested from
the PCE, similar to the operation in forward direction.

Upon arrival of any further client SR request at the server SR to an
HTTP request whose response is still outstanding, the client SR is
added to an internal request table. Optionally, the request is
suppressed from being sent to the server.

Upon arrival of an HTTP response at the server SR, the server SR
consults its internal request table for any outstanding HTTP requests
to the same request. The server SR retrieves the stored BIER
forwarding information for the reverse direction for all outstanding
HTTP requests and determines the path information to all client SRs
through a binary OR over all BIER forwarding identifiers with the
same SI field. This newly formed joint BIER multicast response
identifier is used to send the HTTP response across the network.

5.2. Achieving Multicast Responses

BIER makes the solution scalable. Instead of IP multicast with IGMP/
PIM, BIER is being used between Server NAP (SNAP) and CNAP, the SNAP
simply coalesces the forwarded HTTP requests from the CNAP, and
determines for every requested block the set of CNAPs requesting it.
A set of CNAPs corresponds to a set of bits in the BIER-bitstring,
one bit per CNAP. The SNAP then sends the block into BIER with the
appropriate bitstring set.

This completely eliminates any dynamic multicast signaling between
CNAP and SNAP. It also avoid sending of any unnecessary data block,
which in the IP multicast solution is pretty much unavoidable.

Furthermore, using the approach with BIER, the SNAP can also easily
control how long to delay sending of blocks. For example, it may
wait for some percentage of the time of a block (e.g, 50% = 1
second), therefore ensuring that it is coalescing as many requests
into one BIER multicast answer as possible.

5.3. BIER Traffic Enginnering

BIER-TE (BIER Traffic Engineering [I-D.ietf-bier-te-arch]) forwards
and replicates packets like BIER based on a BitString in the packet
header. Where BIER forwards and replicates its packets on shortest
paths towards BFER, BIER-TE allows (and requires) to also use bits in
the bitstring to indicate the paths in the BIER domain across which
the BIER-TE packets are to be sent. This is done to support Traffic
Engineeringfor BIER packets via explicit hop-by-hop and/or loose hop

forwarding of BIER-TE packets. A BIER-TE controller calculates
explicit paths for this packet forwarding.

The Multicast Flow Overlay operates as in BIER. Instead of
interacting with the BIER layer, it interacts with the BIER-TE
Controller.

In this draft, "Name-based" service forwarding over BIER, is
described to handle changes in service execution end points and
manage adhoc relationship in a multicast group. BIER-TE is another
way of doing this, while integrated with BIER architecture. The PCE
function described earlier in the BIER Multicast Overlay, may become
part of BIER-TE Controller. The SR function in the CNAP and SNAP
communicates with BIER TE controller. SR sends the service name to
the controller, which process the request using the PCE function and
returns the "bitstring" to be used as BIER header for delivery of the
HTTP response to multiple clients.

6. Requirements

A realization for the "HTTP multicast" use case may have the
following requirements:

o MUST support multiple FQDN-based service endpoints to exist in the
   overlay

o MUST send FQDN-based service requests at the network level to a
   suitable FQDN-based service endpoint via policy-based selection of
   appropriate path information

o MUST allow for multicast delivery of HTTP response to same HTTP
   request URI

o MUST provide direct path mobility, where the path between the
   egress and ingress Service Routers(SR) can be determined as being
   optimal (e.g., shortest path or direct path to a selected
   instance), is needed to avoid the use of anchor points and further
   reduce service-level latency

7. Next Steps

This Applicability Statement document describes how name based
service forwarding can be realized over BIER. Name based service
forwarding helps in handling the change of service execution end
points. This document describes the functionalities in the multicast
overlay layer to enable the name based forwarding. We would like to
get feedback from the WG if this is relevant and would like to get
support to continue this work. We will like to elaborate further on

the functions of the overlay layer and list the involved protocols by
the next IETF after Montreal. We also plan to implement in our H2020
projects.

8. IANA Considerations

This document requests no IANA actions.

9. Security Considerations

TBD.

10. Informative References

[I-D.ietf-bier-te-arch]
          Eckert, T., Cauchie, G., Braun, W., and M. Menth, "Traffic
          Engineering for Bit Index Explicit Replication (BIER-TE)",
          draft-ietf-bier-te-arch-00 (work in progress), January
          2018.

[I-D.ietf-bier-use-cases]
          Kumar, N., Asati, R., Chen, M., Xu, X., Dolganow, A.,
          Przygienda, T., Gulko, A., Robinson, D., Arya, V., and C.
          Bestler, "BIER Use Cases", draft-ietf-bier-use-cases-06
          (work in progress), January 2018.

[I-D.ietf-httpbis-bcp56bis]
          Nottingham, M., "On the use of HTTP as a Substrate",
          draft-ietf-httpbis-bcp56bis-05 (work in progress), May
          2018.

[I-D.irtf-icnrg-deployment-guidelines]
          Rahman, A., Trossen, D., Kutscher, D., and R. Ravindran,
          "Deployment Considerations for Information-Centric
          Networking (ICN)", draft-irtf-icnrg-deployment-
          guidelines-03 (work in progress), June 2018.

[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>.

[TR_IPMC_ABR]
          CableLabs, "IP Multicast Adaptive Bit Rate Architecture
          Technical Report", OC-TR-IP-MULTI-ARCH-V01-141112 C01,
          October 2016, <https://community.cablelabs.com/wiki/plugin
          s/servlet/cablelabs/alfresco/
          download?id=51b3c11a-3ba4-40ab-b234-42700e0d4669;1.0>.

Authors' Addresses

   Debashish Purkayastha
   InterDigital Communications, LLC
   Conshohocken
   USA

   Email: Debashish.Purkayastha@InterDigital.com

   Akbar Rahman
   InterDigital Communications, LLC
   Montreal
   Canada

   Email: Akbar.Rahman@InterDigital.com

   Dirk Trossen
   InterDigital Communications, LLC
   64 Great Eastern Street, 1st Floor
   London EC2A 3QR
   United Kingdom

   Email: Dirk.Trossen@InterDigital.com
   URI: http://www.InterDigital.com/

   Toerless Eckert
   Huawei USA - Futurewei Technologies Inc.
   2330 Central Expy
   Santa Clara 95050
   USA

   Email: tte+ietf@cs.fau.de