Using Media Encoding Networks to address MPEG-DASH video
draft-deen-naik-ggie-men-mpeg-dash-00

Abstract

This document describes an approach to using a Media Encoding Network of IPv6 Prefixes and Addresses as identifiers for MPEG-DASH encoded video. This is part of the GGIE Glass to Glass Internet Ecosystem effort for Internet Video.

This document is being discussed on the ggie@ietf.org mailing list.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on January 8, 2017.
1. Terminology

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].

2. Introduction

GGIE, the Glass to Glass Internet Ecosystem, described in [I-D.deen-daigle-ggie], is an effort to improve video’s use of the Internet though evolving and applying modern Internet networking technology to Interet video.
This document is a proposed Media Encoding Network organizational definition for MPEG-DASH encoded video. In the following sections, we describe a Media Encoding Network structure for MPEG-DASH content using IPv6 addresses as the address for MPEG-DASH video chunks, and organizing these addresses into a IPv6 subnet under a prefix.

A MPEG-DASH encoded video organized following this Media Encoding Network scheme is in turn referrable to using the assigned prefix, with each distinct encoding of the video being assigned a distinct prefix. Hence two copies of the same video encode would share the same prefix, while a different encode would have a different prefix.

Other Media Encoding Networks organizational definitions are possible for MPEG-DASH video. The simple organizational structure defined in this document is designed to work, in a backwards compatible manner, with existing MPEG-DASH video players.

2.1. Media Encoding Networks

One of the concepts being discussed in GGIE is that of a Media Encoding Network. As introduced in the GGIE Introduction [I-D.deen-daigle-ggie] document, a Media Encoding Network consists of the data elements of a audio-video encoding of a work organized following a distinct logical structure appropriate for efficiently transporting and accessing the data elements for the video asset. Network level identifiers are assigned to each of these elements under a shared prefix and following an address assignment plan appropriate for the type of encoding used for the AV data.

Media Encoding Networks is a generalized abstraction intended to be used with many different encoding and transport schemes.

GGIE recognizes that there is currently a great diversity of encoding and transports such as MPEG-DASH [DASH] and HTTP Live Streaming (HLS) [I-D.pantos-http-live-streaming] to name but two, with more continuing to be developed and introduced. Recognizing this diversity and innovative environment, GGIE proposes the Media Encoding Network as a reusable abstraction that can be tailored and defined with different logical organizations to support different environments, applications, and media encodings.

A Media Encoding Network is a logical entity that can be assigned a network level identifier enabling it to be referred to at a network device level and permitting devices and the network to worked cooperatively to optimize data transport and access choices.
3. MPEG-DASH Internet Video Concepts

A common technique used in the delivery of a media or video on the Internet via streaming services and CDNs is to break up an encoding of a video into chunks or media segments containing a fixed duration of video. MPEG-DASH [DASH] is an example of such an approach. The segments typically represent small portions of the video with 6-10 seconds of video playback being common. In most implementations, the segments of videos are identified by file names and served to clients using conventional web servers using HTTP GET requests.

Systems such as MPEG-DASH enable client players to switch between encodings of different quality levels of the video with higher quality encodings requiring large amounts of data, and conversely lower quality encodings requiring smaller amounts of data. The system coordinates each encoding to produce points of alignment called intra-coded frames or iFrames where a player can switch between different encodings without missing frames of the video playback. Thus, a player can adapt to changing network conditions without re-buffering or freezing of the playback.

When the encodings are broken into segments, the segments are organized such that the playback system can switch to a different encoding level from the version it has been playing by requesting the next segment of data holding the iFrame matching the next iFrame of the current encoding. In practice each segment of an encoding is an individual file stored on video or CDN server and playback consists of the player repeatedly requesting the next file in sequence from the server, with the file names following a consistent incremental naming scheme indicating an encoding identifier and a segment sequence identifier.

Typically, a video file is processed by an encoder to produce two or more different quality encodings with each encoded version being passed through a process to break into segment files with aligned iFrames and each file named with a name identifying the encoding and sequence number. This process requires coordination to create iFrame alignments and a consistent naming convention to allow players to transition between encodings and to iteratively access the next correct segment.

3.1. Internet Video playback as a network

Transitioning between segments is an example of a simple directed graph (or digraph). Each segment is a vertex or node and the naming convention defines an ordered directed traversal of the graph, and the iFrame aligned segments forming the edges of the graph. It is also possible to recognize that the directed graph behavior of a
player switching between segments can more generally be viewed as a network such as it is used on the Internet.

The network of segments can be identified using the IP addressing scheme from the Internet, in particular IPv6 is well suited for this due to the large number of addresses available in it’s 128-bit address space. IPv4 could also be used, but with only 32 bits of address space the available addresses would be quickly exhausted in practical use.

This is really a simple evolution of the way MPEG-DASH chunks are organized today as files with names such as MOVIE-SEGMENT-00, MOVIE-SEGMENT-01,... and so on. In practical terms, this scheme simply replaces the ASCII filename, with a 128-bit number represented as HEX digits. In this way, this scheme remains compatible with existing CDN serving of MPEG-DASH video.

4. MPEG-DASH Video Chunk Addressing

Staying consistent with Media Encoding Networks being a generic abstraction, the more generic term Shard is used in place of the MPEG-DASH specific Chunk for individual units of encoded video data.

IPv6 addresses [RFC4291] are specified in and are broken into two parts that split the available 128 bits of address space as follows:

```
+--------------+----------------+
|      Prefix  | Interface id   |
+--------------+----------------+
```

Figure 1: IPv6 Address

One addressing approach to naming segments can be as follows:

```
n bits          128-n bits
+----------------+-----------------+-------------+
| Encoding Prefix | Sub-Encoding id | Shard id    |
+----------------+-----------------+-------------+
```

Figure 2: Proposed

Which consists of an Encoding Prefix that is uniquely assigned to a set of aligned MPEG-DASH encodings of the video, a sub-encoding id which identifies a particular encoding, and the id of the individual shard of encoded video data.
The encoding prefix permits a set of encodings to be associated with one another. Grouping a set of encodings of a video under a shared Encoding Prefix permits referencing all the segments of a group of encodings as a single entity under the Encoding Prefix.

The sub-encoding id groups the shards of a single sub-encoding together under an identifier to permit managing the collection of segments as a single entity.

Shards that share MPEG iFrame alignment share the same Shard id. This then defines a network layout with shards for each different bit-rate organized sequentially and contiguously under a shared sub-encoding subnet and shards with aligned iFrames being organized with the same shard id across sub-encoding subnets.

5. Video Playback

This approach permits the Prefix to identify a particular group of encodings of a video. Each encoding has an assigned series of addresses consisting of the prefix, followed by the series of address bits that uniquely identify the shard. All the playback pathways are preserved in this addressing scheme of the edges of the graph.

The above approach works well for a video that is encoded by one party that can coordinate the encoding process, to produce aligned iFrames, and assign the common encoding prefix and segment assignments for the network.

A playback device can be provided the Prefix for the network, and can iterate through the segments to play the video. It can jump between sub-encode subnets to select different quality or vary the bit rate of the playback.

6. Implementation

For the evaluation of this scheme, a prototype video streaming service implementing this approach was developed. In particular, it provides an Electronic Program Guide (EPG) and uses an open-source HTML5 video player with MPEG-DASH. Instead of providing the player with HTTP URIs for each segment of video, our this prototype uses global IPv6 addresses. This change is transparent to the host operating system, the HTML5 video player, and the network. The service backend is implemented in Python and utilizes other open source components. A demonstration at IETF96 is planned to be shown during Bits-n-Bytes.
7. Conclusion and Next Steps

This draft proposes a Media Encoding Network addressing scheme for MPEG-DASH Internet video using IPv6 addresses. It is an example that can built upon to define other more complex Media Encoding Network schemes for MPEG-DASH and other encoding/transports.

8. Acknowledgements

9. IANA Considerations

   None (yet).

10. Security Considerations

    None (yet).

11. References

11.1. Normative References

[I-D.deen-daigle-ggie]
Deen, G. and L. Daigle, "Glass to Glass Internet Ecosysten
Introduction", draft-deen-daigle-ggie-01 (work in
progress), June 2016.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119,
DOI 10.17487/RFC2119, March 1997,

Architecture", RFC 4291, DOI 10.17487/RFC4291, February

11.2. Informative References

[DASH] ISO, "Dynamic adaptive streaming over HTTP (DASH) -- Part
1: Media presentation description and segment formats",
<http://www.iso.org/iso/home/store/catalogue_ics/
catalogue_detail_ics.htm?csnumber=65274>.

[I-D.pantos-http-live-streaming]
Pantos, R. and W. May, "HTTP Live Streaming", draft-
pantos-http-live-streaming-19 (work in progress), April
2016.
Authors’ Addresses

Glenn Deen
Comcast-NBCUniversal

Email: rgd.ietf@gmail.com

Gaurav Naik
Drexel University

Email: gn@drexel.edu

John Jason Brzozowski
Comcast

Email: John_Brzozowski@cable.comcast.com

Leslie Daigle
Thinking Cat Enterprises LLC

Email: ldaigle@thinkingcat.com

Bill Rose
WJR Consulting

Email: brose@wjconsulting.com

Mark Townsley
Cisco
Paris

Email: townsley@cisco.com