

Guidelines for Implementation: DASH-IF Interoperability Points

November 15, 2018

DASH Industry Forum

Version 4.3



1 Scope

- 2 The scope of the DASH-IF InterOperability Points (IOPs) defined in this document is to provide support for
3 high-quality video distribution for over the top services using H.264/AVC and H.265/HEVC. Both live and
4 on-demand services are supported. The specified features enable relevant use cases including on-demand,
5 live services, ad insertion, trick modes, seek preview, content protection and subtitling. Extensions for multi-
6 channel audio and next generation audio with different codecs as well as extensions to video different
7 codecs and Ultra High Definition are defined.
- 8 Any identified bugs or missing features may be submitted through the DASH-IF issue tracker at [https://gi-
9 treports.com/issue/Dash-Industry-Forum/DASH-IF-IOP](https://gitreports.com/issue/Dash-Industry-Forum/DASH-IF-IOP).
- 10 Note that version 4.3 is published as an add on to v4.2, but with the next version it is expected that either
11 a multipart version or a restructured version will be generated, with major editorial updates. The new
12 version is expected to be available by spring of 2019.

1 Disclaimer

2 This is a document made available by DASH-IF. The technology embodied in this document may
3 involve the use of intellectual property rights, including patents and patent applications owned or
4 controlled by any of the authors or developers of this document. No patent license, either implied
5 or express, is granted to you by this document. DASH-IF has made no search or investigation for
6 such rights and DASH-IF disclaims any duty to do so. The rights and obligations which apply to
7 DASH-IF documents, as such rights and obligations are set forth and defined in the DASH-IF
8 Bylaws and IPR Policy including, but not limited to, patent and other intellectual property license
9 rights and obligations. A copy of the DASH-IF Bylaws and IPR Policy can be obtained at
10 <http://dashif.org/>.

11 The material contained herein is provided on an "AS IS" basis and to the maximum extent per-
12 mitted by applicable law, this material is provided AS IS, and the authors and developers of this
13 material and DASH-IF hereby disclaim all other warranties and conditions, either express, implied
14 or statutory, including, but not limited to, any (if any) implied warranties, duties or conditions of
15 merchantability, of fitness for a particular purpose, of accuracy or completeness of responses, of
16 workmanlike effort, and of lack of negligence.

17 In addition, this document may include references to documents and/or technologies controlled
18 by third parties. Those third party documents and technologies may be subject to third party rules
19 and licensing terms. No intellectual property license, either implied or express, to any third party
20 material is granted to you by this document or DASH-IF. DASH-IF makes no any warranty what-
21 soever for such third party material.

22 Note that technologies included in this document and for which no test and conformance materi-
23 al is provided, are only published as a candidate technologies, and may be removed if no test
24 material is provided before releasing a new version of this guidelines document. For the availa-
25 bility of test material, please check <http://www.dashif.org>.

1 Contents

2	GUIDELINES FOR IMPLEMENTATION: DASH-IF INTEROPERABILITY POINTS.....	1
3	SCOPE.....	1
4	DISCLAIMER.....	2
5	CONTENTS	3
6	LIST OF FIGURES.....	9
7	LIST OF TABLES	10
8	ACRONYMS, ABBREVIATIONS AND DEFINITIONS.....	11
9	REFERENCES	14
10	1. INTRODUCTION	1
11	2. CONTEXT AND CONVENTIONS	5
12	2.1. RELATION TO MPEG-DASH AND OTHER DASH SPECIFICATIONS	5
13	2.2. COMPATIBILITY AND EXTENSIONS TO EARLIER VERSIONS.....	5
14	2.2.1. <i>Summary of Version 3 Modifications</i>	5
15	2.2.2. <i>Backward-Compatibility Considerations</i>	8
16	2.3. USE OF KEY WORDS	9
17	2.3.1. <i>Background</i>	9
18	2.3.2. <i>Key Words</i>	9
19	2.3.3. <i>Mapping to DASH-IF Assets.....</i>	10
20	2.4. DEFINITION AND USAGE OF INTEROPERABILITY POINTS.....	10
21	2.4.1. <i>Profile Definition in ISO/IEC 23009-1.....</i>	10
22	2.4.2. <i>Usage of Profiles.....</i>	11
23	2.4.3. <i>Interoperability Points and Extensions.....</i>	11
24	3. USAGE OF MPEG-DASH	12
25	3.1. SCOPE	12
26	3.2. DASH FORMATS.....	12
27	3.2.1. <i>Introduction.....</i>	12
28	3.2.2. <i>Media Presentation Description constraints for v1 & v2 Clients.....</i>	14
29	3.2.3. <i>Segment format constraints.....</i>	15
30	3.2.4. <i>Presence of Attributes and Elements</i>	16
31	3.2.5. <i>MPD Dimension Constraints.....</i>	17
32	3.2.6. <i>Generic Metadata</i>	17
33	3.2.7. <i>DASH Timing Model.....</i>	17
34	3.2.8. <i>Bandwidth and Minimum Buffer Time</i>	19
35	3.2.9. <i>Trick Mode Support</i>	20
36	3.2.10. <i>Adaptation Set Constraints</i>	21
37	3.2.11. <i>Media Time Information of Segment.....</i>	23
38	3.2.12. <i>Content Offering with Periods</i>	25
39	3.2.13. <i>Adaptation Set Media Type</i>	26
40	3.2.14. <i>Seek Preview and Thumbnail Navigation</i>	27
41	3.2.15. <i>MPD Location and Reference Resolution.....</i>	27
42	3.3. CLIENT IMPLEMENTATION REQUIREMENTS AND GUIDELINES	28
43	3.3.1. <i>Overview.....</i>	28
44	3.3.2. <i>DASH Client Guidelines.....</i>	29
45	3.3.3. <i>Seamless switching.....</i>	29

1	3.3.4. DASH Client Requirements	30
2	3.4. TRANSPORT AND PROTOCOL-RELATED ISSUES.....	30
3	3.4.1. General.....	30
4	3.4.2. Server Requirements and Guidelines.....	30
5	3.4.3. Client Requirements and Guidelines.....	31
6	3.4.4. Transforming Proxies and Other Adaptation Middleboxes.....	31
7	3.5. SYNCHRONIZATION CONSIDERATIONS	31
8	3.6. CONSIDERATIONS FOR DIFFERENT SERVICES TYPES.....	32
9	3.7. CONSIDERATIONS ON AD INSERTION	32
10	3.8. SWITCHING ACROSS ADAPTATION SETS.....	32
11	3.9. ANNOTATION AND CLIENT MODEL FOR CONTENT SELECTION	33
12	3.9.1. Introduction.....	33
13	3.9.2. Adaptation Set Labeling Options for Selection.....	33
14	3.9.3. Content Model.....	41
15	3.9.4. Signalling Requirements and Recommendations.....	43
16	3.9.5. Client Processing Reference Model	47
17	3.10. ON-DEMAND SERVICES.....	52
18	3.10.1. Introduction	52
19	3.10.2. On-Demand Services with Live profiles.....	52
20	3.10.3. On-Demand Services with On-Demand profile	56
21	3.10.4. Mixed On-Demand Content	61
22	4. LIVE SERVICES	61
23	4.1. INTRODUCTION	61
24	4.2. OVERVIEW DYNAMIC AND LIVE MEDIA PRESENTATIONS	62
25	4.3. DYNAMIC SEGMENT DOWNLOAD.....	65
26	4.3.1. Background and Assumptions.....	65
27	4.3.2. Preliminaries.....	65
28	4.3.3. Service Offering Requirements and Guidelines	71
29	4.3.4. Client Operation, Requirements and Guidelines	80
30	4.3.5. Additional DVB-DASH alignment aspects.....	85
31	4.3.6. Considerations on live edge.....	85
32	4.4. SIMPLE LIVE SERVICE OFFERING INCLUDING MPD UPDATES.....	86
33	4.4.1. Background and Assumptions.....	86
34	4.4.2. Preliminaries.....	87
35	4.4.3. Service Offering Requirements and Guidelines	88
36	4.4.4. MPD-based Live Client Operation based on MPD	92
37	4.5. MPD AND SEGMENT-BASED LIVE SERVICE OFFERING.....	93
38	4.5.1. Preliminaries.....	93
39	4.5.2. Service Offering Requirements and Guidelines	95
40	4.5.3. Client Requirements and Guidelines.....	99
41	4.6. PROVISIONING OF LIVE CONTENT IN ON-DEMAND MODE.....	101
42	4.6.1. Scenario.....	101
43	4.6.2. Content Offering Requirements and Recommendations	101
44	4.6.3. Client Behavior	102
45	4.6.4. Transition Phase between Live and On-Demand	103
46	4.7. AVAILABILITY TIME SYNCHRONIZATION BETWEEN CLIENT AND SERVER.....	103
47	4.7.1. Background	103

1	4.7.2. <i>Service Provider Requirements and Guidelines</i>	103
2	4.7.3. <i>Client Requirements and Guidelines</i>	104
3	4.8. ROBUST OPERATION	105
4	4.8.1. <i>Background</i>	105
5	4.8.2. <i>Tools for Robust Operations</i>	105
6	4.8.3. <i>Synchronization Loss of Segmenter</i>	106
7	4.8.4. <i>Encoder Clock Drift</i>	106
8	4.8.5. <i>Segment Unavailability</i>	106
9	4.8.6. <i>Swapping across Redundant Tools</i>	107
10	4.8.7. <i>Service Provider Requirements and Guidelines</i>	107
11	4.8.8. <i>Client Requirements and Guidelines</i>	107
12	4.9. INTEROPERABILITY ASPECTS	107
13	4.9.1. <i>Introduction</i>	107
14	4.9.2. <i>Simple Live Operation</i>	107
15	4.9.3. <i>Main Live Operation</i>	108
16	4.10. TRICK MODE FOR LIVE SERVICES	109
17	4.10.1. <i>Introduction</i>	109
18	4.10.2. <i>Service Offering Requirements and Recommendations</i>	109
19	4.10.3. <i>Client Implementation Guidelines</i>	109
20	4.10.4. <i>Conversion for Live-to-VoD for Trick Mode Adaptation Sets</i>	110
21	4.11. DEPLOYMENT SCENARIOS	110
22	4.11.1. <i>Introduction</i>	110
23	4.11.2. <i>Reliable and Consistent-Delay Live Service</i>	110
24	4.11.3. <i>Relevant DASH-IF IOP Technologies</i>	112
25	4.11.4. <i>Proposed Service Configuration and Generation of the MPD and Segments based on a "Segment Stream"</i>	117
26	4.11.5. <i>Client Support Considerations</i>	121
27	5. AD INSERTION IN DASH	122
28	5.1. INTRODUCTION	122
29	5.1.1. <i>General</i>	122
30	5.1.2. <i>Definitions</i>	122
31	5.1.3. <i>DASH Concepts</i>	123
32	5.2. ARCHITECTURES	126
33	5.3. SERVER-BASED ARCHITECTURE	127
34	5.3.1. <i>Introduction</i>	127
35	5.3.2. <i>Mapping to DASH</i>	127
36	5.3.3. <i>Workflows</i>	130
37	5.3.4. <i>Examples</i>	135
38	5.3.5. <i>Use of query parameters</i>	136
39	5.4. APP-BASED ARCHITECTURE	137
40	5.4.1. <i>Introduction</i>	137
41	5.4.2. <i>Mapping to DASH</i>	138
42	5.4.3. <i>Workflows</i>	140
43	5.5. EXTENSIONS FOR AD INSERTION	141
44	5.5.1. <i>Asset Identifiers</i>	141
45	5.5.2. <i>Remote Periods</i>	141
46	5.5.3. <i>User-defined events</i>	142

1	5.6.	INTEROPERABILITY ASPECTS	142
2	5.6.1.	<i>Server-based Ad insertion</i>	142
3	5.6.2.	<i>App-based Ad Insertion</i>	143
4	6.	MEDIA CODING TECHNOLOGIES	143
5	6.1.	INTRODUCTION	143
6	6.2.	VIDEO	143
7	6.2.1.	<i>General</i>	143
8	6.2.2.	<i>DASH-specific aspects for H.264/AVC video</i>	144
9	6.2.3.	<i>DASH-specific aspects for H.265/HEVC video</i>	146
10	6.2.4.	<i>Video Metadata</i>	147
11	6.2.5.	<i>Adaptation Sets Constraints</i>	148
12	6.2.6.	<i>Tiles of thumbnail images</i>	149
13	6.3.	AUDIO	152
14	6.3.1.	<i>General</i>	152
15	6.3.2.	<i>DASH-specific aspects for HE-AACv2 audio</i>	152
16	6.3.3.	<i>Audio Metadata</i>	153
17	6.4.	AUXILIARY COMPONENTS	153
18	6.4.1.	<i>Introduction</i>	153
19	6.4.2.	<i>Subtitles and Closed Captioning</i>	153
20	6.4.3.	<i>CEA-608/708 in SEI messages</i>	154
21	6.4.4.	<i>Timed Text (IMSC1)</i>	156
22	6.4.5.	<i>Guidelines for side-loaded TTML and WebVTT files</i>	157
23	6.4.6.	<i>Annotation of Subtitles</i>	157
24	7.	CONTENT PROTECTION AND SECURITY	158
25	7.1.	INTRODUCTION	158
26	7.2.	HTTPS AND DASH	158
27	7.3.	BASE TECHNOLOGIES SUMMARY	159
28	7.4.	ISO BMFF SUPPORT FOR COMMON ENCRYPTION AND DRM	160
29	7.4.1.	<i>Box Hierarchy</i>	160
30	7.4.2.	<i>ISO BMFF Structure Overview</i>	162
31	7.5.	PERIODIC RE-AUTHORIZATION	163
32	7.5.1.	<i>Introduction</i>	163
33	7.5.2.	<i>Use Cases and Requirements</i>	163
34	7.5.3.	<i>Implementation Options</i>	164
35	7.6.	MPD SUPPORT FOR ENCRYPTION AND DRM SIGNALING	165
36	7.6.1.	<i>Introduction</i>	165
37	7.6.2.	<i>Use of the Content Protection Descriptor</i>	165
38	7.7.	ADDITIONAL CONTENT PROTECTION CONSTRAINTS	168
39	7.7.1.	<i>ISO BMFF Content Protection Constraints</i>	168
40	7.7.2.	<i>MPD Content Protections Constraints</i>	169
41	7.7.3.	<i>Other Content Protections Constraints</i>	170
42	7.7.4.	<i>Additional Constraints for Periodic Re-Authorization</i>	170
43	7.7.5.	<i>Encryption of Different Representations</i>	170
44	7.7.6.	<i>Encryption of Multiple Periods</i>	171
45	7.7.7.	<i>DRM System Identification</i>	171
46	7.7.8.	<i>Protection of Media Presentations that Include SD, HD and UHD Adaptation Sets</i>	171
47	7.7.9.	<i>Client Interactions with DRM Systems</i>	172

1	7.8. WORKFLOW OVERVIEW	173
2	8. DASH-IF INTEROPERABILITY POINTS	176
3	8.1. INTRODUCTION	176
4	8.2. DASH-AVC/264 MAIN	176
5	8.2.1. <i>Introduction</i>	176
6	8.2.2. <i>Definition</i>	176
7	8.3. DASH-AVC/264 HIGH.....	177
8	8.3.1. <i>Introduction</i>	177
9	8.3.2. <i>Definition</i>	177
10	8.4. DASH-IF IOP SIMPLE	177
11	8.4.1. <i>Introduction</i>	177
12	8.4.2. <i>Definition</i>	178
13	8.5. DASH-IF IOP MAIN.....	178
14	8.5.1. <i>Introduction</i>	178
15	8.5.2. <i>Definition</i>	178
16	9. MULTI-CHANNEL AUDIO EXTENSIONS.....	179
17	9.1. SCOPE	179
18	9.2. TECHNOLOGIES	179
19	9.2.1. <i>Dolby Multichannel Technologies</i>	179
20	9.2.2. <i>DTS-HD</i>	180
21	9.2.3. <i>MPEG Surround</i>	180
22	9.2.4. <i>MPEG-4 High Efficiency AAC Profile v2, level 6</i>	181
23	9.2.5. <i>MPEG-H 3D Audio</i>	182
24	9.2.6. <i>MPEG-D Unified Speech and Audio Coding</i>	183
25	9.3. CLIENT IMPLEMENTATION GUIDELINES	184
26	9.4. EXTENSIONS	184
27	9.4.1. <i>General</i>	184
28	9.4.2. <i>Dolby Extensions</i>	184
29	9.4.3. <i>DTS-HD Interoperability Points</i>	185
30	9.4.4. <i>MPEG Surround Interoperability Points</i>	187
31	9.4.5. <i>MPEG HE-AAC Multichannel Interoperability Points</i>	187
32	9.4.6. <i>MPEG-H 3D Audio Interoperability Points</i>	188
33	9.4.7. <i>MPEG-D USAC Interoperability Points</i>	188
34	10. DASH-IF UHD EXTENSIONS	188
35	10.1. INTRODUCTION	188
36	10.2. DASH-IF UHD HEVC 4k.....	189
37	10.2.1. <i>Introduction</i>	189
38	10.2.2. <i>Elementary Stream Requirements</i>	189
39	10.2.3. <i>Mapping to DASH</i>	191
40	10.2.4. <i>Compatibility Aspects</i>	194
41	10.3. DASH-IF IOP HEVC HDR PQ10.....	195
42	10.3.1. <i>Introduction</i>	195
43	10.3.2. <i>Elementary Stream Requirements</i>	195
44	10.3.3. <i>Mapping to DASH</i>	198
45	10.3.4. <i>Compatibility Aspects</i>	199
46	10.4. DASH-IF IOP UHD DUAL-STREAM (DOLBY VISION)	199
47	10.4.1. <i>Introduction</i>	199

1	10.4.2. <i>Definition</i>	199
2	10.4.3. <i>Mapping to DASH</i>	203
3	11. DASH-IF VP9 EXTENSIONS	204
4	11.1. INTRODUCTION	204
5	11.2. DASH-SPECIFIC ASPECTS FOR VP9 VIDEO.....	205
6	11.2.1. <i>General</i>	205
7	11.2.2. <i>Bitstream Switching</i>	205
8	11.3. DASH-IF VP9 EXTENSION IOPs.....	206
9	11.3.1. <i>DASH-IF VP9-HD</i>	206
10	11.3.2. <i>DASH-IF VP9-UHD</i>	206
11	11.3.3. <i>DASH-IF VP9-HDR</i>	206
12	ANNEX A EXAMPLES FOR PROFILE SIGNALLING.....	1
13	ANNEX B LIVE SERVICES - USE CASES AND ARCHITECTURE	2
14	B.1 BASELINE USE CASES.....	2
15	B.1.1 <i>Use Case 1: Live Content Offered as On-Demand</i>	2
16	B.1.2 <i>Use Case 2: Scheduled Service with known duration and Operating at live edge</i>	2
17	B.1.3 <i>Use Case 3: Scheduled Service with known duration and Operating at live edge and time shift buffer</i>	2
18	B.1.4 <i>Use Case 4: Scheduled Live Service known duration, but unknown Segment URLs</i>	2
19	B.1.5 <i>Use Case 5: 24/7 Live Service</i>	2
20	B.1.6 <i>Use Case 6: Approximate Media Presentation Duration Known</i>	2
21	B.2 BASELINE ARCHITECTURE FOR DASH-BASED LIVE SERVICE	3
22	B.3 DISTRIBUTION OVER MULTICAST	3
23	B.4 TYPICAL PROBLEMS IN LIVE DISTRIBUTION.....	4
24	B.4.1 <i>Introduction</i>	4
25	B.4.2 <i>Client Server Synchronization Issues</i>	4
26	B.4.3 <i>Synchronization Loss of Segmenter</i>	5
27	B.4.4 <i>Encoder Clock Drift</i>	5
28	B.4.5 <i>Segment Unavailability</i>	5
29	B.4.6 <i>Swapping across Redundant Tools</i>	6
30	B.4.7 <i>CDN Issues</i>	6
31	B.4.8 <i>High End-to-end Latency</i>	6
32	B.4.9 <i>Buffer Management & Bandwidth Estimation</i>	7
33	B.4.10 <i>Start-up Delay and Synchronization Audio/Video</i>	7
34	B.5 ADVANCED USE CASES.....	7
35	B.5.1 <i>Introduction</i>	7
36	B.5.2 <i>Use Case 7: Live Service with undetermined end</i>	7
37	B.5.3 <i>Use Case 8: 24/7 Live Service with canned advertisement</i>	7
38	B.5.4 <i>Use case 9: 24x7 live broadcast with media time discontinuities</i>	7
39	B.5.5 <i>Use case 10: 24x7 live broadcast with Segment discontinuities</i>	7
40	ANNEX C DOLBY VISION STREAMS WITHIN THE ISO BASE MEDIA FILE FORMAT.....	8
41	C.1 INTRODUCTION	8
42	C.2 DOLBY VISION CONFIGURATION BOX AND DECODER CONFIGURATION RECORD	8
43	C.2.1 <i>Definition</i>	8
44	C.2.2 <i>Syntax</i>	8
45	C.2.3 <i>Semantics</i>	9
46	C.3 DOLBY VISION SAMPLE ENTRIES	9

1	C.3.1 Definition.....	9
2	C.3.2 Syntax.....	10
3	C.3.3 Semantics	10
4	C.6 DOLBY VISION FILES.....	10
5	C.7 DOLBY VISION TRACK IN A SINGLE FILE	10
6	C.7.1 Constraints on EL Track	10
7	C.7.2 Constraints on the ISO base media file format boxes	11
8	C.7.2.1 Constraints on Movie Fragments	11
9	C.7.2.2 Constraints on Track Fragment Random Access Box	11
10	ANNEX D SIGNALING DOLBY VISION PROFILES AND LEVELS.....	13
11	D.1 DOLBY VISION PROFILES AND LEVELS.....	13
12	D.1.1 Dolby Vision Profiles.....	13
13	D.1.1.1 Dolby Vision Profile String format.....	14
14	D.1.2 DOLBY VISION LEVELS.....	14
15	B.1.2.1 Dolby Vision Level String Format.....	15
16	B.1.3 Dolby Vision Codec Profile and Level String	15
17	B.1.3.1 Device Capabilities	16
18	ANNEX E DISPLAY MANAGEMENT MESSAGE.....	17
19	E.1 INTRODUCTION	17
20	E.2 SYNTAX AND SEMANTICS	18
21	ANNEX F COMPOSING METADATA MESSAGE.....	23
22	F.1 INTRODUCTION	23
23	F.2 SYNTAX AND SEMANTICS	23
24	ANNEX G SAMPLE DUAL-LAYER MPD.....	25
25		

26 List of Figures

27	Figure 1 Overview Timing Model.....	18
28	Figure 2 DASH aspects of a DASH-AVC/264 client compared to a client supporting the union of 29 DASH ISO BMFF live and on-demand profile.....	29
30	Figure 3 Content Model for DASH Multitrack.....	41
31	Figure 4 Different Client Models	64
32	Figure 5 Segment Availability on the Server for different time NOW (blue = valid but not yet 33 available segment, green = available Segment, red = unavailable Segment)	76
34	Figure 6 Simple Client Model	80
35	Figure 7 Advanced Client Model	99
36	Figure 8 Example Deployment Architecture	111
37	Figure 9 Loss scenarios	115
38	Figure 10: Different properties of a segment stream	118
39	Figure 11: XLink resolution	123
40	Figure 12: Server-based architecture	127
41	Figure 13 Using an Asset Identifier	129
42	Figure 14: Live Workflow.....	131

1	Figure 15: Ad Decision.....	133
2	Figure 16: Example of MPD for "Top Gun" movie	136
3	Figure 17: App-based architecture	137
4	Figure 18 Inband carriage of SCTE 35 cue message.....	139
5	Figure 19: In-MPD carriage of SCTE 35 cue message.....	139
6	Figure 20: Linear workflow for app-driven architecture.....	140
7	Figure 21: Visualization of box structure for single key content	161
8	Figure 22: Visualization of box structure with key rotation	162
9	Figure 23: PSSH with version numbers and KIDs.	165
10	Figure 24 Logical Roles that Exchange DRM Information and Media.....	173
11	Figure 25 Example of Information flow for DRM license retrieval	175
12	Figure 26 Overview of Dual-stream System	200
13	Figure 27 Typical Deployment Scenario for DASH-based live services.....	3
14	Figure 28 Typical Deployment Scenario for DASH-based live services partially offered through 15 MBMS (unidirectional FLUTE distribution).....	4

16 List of Tables

17	Table 1 DASH-IF Interoperability Points.....	1
18	Table 2 DASH-IF Interoperability Point Extensions	2
19	Table 3 Identifiers and other interoperability values defined in this Document.....	4
20	Table 4 Adaptation Set Attributes and Elements and Usage in DASH-IF IOPs (see ISO/IEC 21 23009-1 [4]).....	34
22	Table 5 -- Information related to Segment Information for Using Live Profiles for On-Demand 23 Services	52
24	Table 6 -- Information related to Segment Information for On-Demand Profile	58
25	Table 7 Main features and differences of simple and main live services	61
26	Table 8 -- Information related to Segment Information and Availability Times for a dynamic service 27	66
28	Table 9 – Basic Service Offering	72
29	Table 10 – Basic Service Offering	75
30	Table 11 Multi-Period Service Offering.....	76
31	Table 12 – Service Offering with Segment Timeline	78
32	Table 13 – Information related to Live Service Offering with MPD-controlled MPD Updates	87
33	Table 14 – Basic Service Offering with MPD Updates.....	88
34	Table 15 – Service Offering with Segment Timeline and MUP greater than 0	91
35	Table 16 – Service Offering with MPD and Segment-based Live Services.....	94
36	Table 17 InbandEventStream@value attribute for scheme with a value 37 "urn:mpeg:dash:event:2012".....	96
38	Table 18 – Basic Service Offering with Inband Events	98
39	Table 19 H.264 (AVC) Codecs parameter according to RFC6381 [10].....	144
40	Table 20 Signaling of HEVC IRAP Pictures in the ISOBMFF and in DASH	146
41	Table 21 Codecs parameter according to ISO/IEC 14496-15 [9].....	147

1	Table 22 HE-AACv2 Codecs parameter according to RFC6381 [10].....	152
2	Table 23 Subtitle MIME type and codecs parameter according to IANA and W3C registries...	157
3	Table 24 Boxes relevant for DRM systems	162
4	Table 25 Dolby Technologies: Codec Parameters and ISO BMFF encapsulation	179
5	Table 26: DTS Codec Parameters and ISO BMFF encapsulation	180
6	Table 27 Codecs parameter according to RFC6381 [10] and ISO BMFF encapsulation for MPEG Surround codec.....	181
8	Table 28 Codecs parameter according to RFC6381 [10] and ISO BMFF encapsulation	181
9	Table 29 Codecs parameter and ISO BMFF encapsulation.....	183
10	Table 30 Codecs parameters according to RFC 6381 [10] and ISO BMFF encapsulation for MPEG-D USAC	184
12	Table 31: Compound Content Management SEI message: HEVC (prefix SEI NAL unit with nal_unit_type = 39, payloadType=4).....	202
14	Table 32: UserID: user identifier.....	202
15	Table 33 Sample table box hierarchy for the EL track of a dual-track Dolby Vision file.....	11
16		

17 Acronyms, abbreviations and definitions

18 For acronyms, abbreviations and definitions refer to ISO/IEC 23009-1 [4]. Additional defini-
19 tions may be provided in the context of individual sections.

20 In addition, the following abbreviations and acronyms are used in this document:

21	AAC	Advanced Audio Coding
22	AFD	Active Format Description
23	AST	Availability Start Time
24	AVC	Advanced Video Coding
25	BL	Base Layer
26	BMFF	Base Media File Format
27	CDN	Content Delivery Network
28	CEA	Consumer Electronics Association
29	CT	Composition Time
30	DECE	Digital Entertainment Content Ecosystem
31	DRM	Digital Rights Management
32	DSI	Decoder Specific Information
33	DT	Decode Time
34	DTV	Digital Television
35	DVB	Digital Video Broadcasting

1	DVS	Digital Video Subcommittee
2	ECL	Entitlement Control License
3	EDL	Encoding Decision List
4	EL	Enhancement Layer
5	EME	Encrypted Media Extension
6	EML	Entitlement Management License
7	EPT	Earliest Presentation Time
8	FCC	Federal Communications Commission
9	GOP	Group-of-Pictures
10	HD	High-Definition
11	HDR	High Dynamic Range
12	HDR10	DASH-IF HDR 10 bit
13	HDMI	High-Definition Multimedia Interface
14	HE-AAC	High Efficiency AAC
15	HEVC	High-Efficiency Video Coding
16	KID	common Key Identifier
17	IAB	International Advertising Bureau
18	IDR	Instantaneous Decoder Refresh
19	IOP	InterOperability Point
20	ISO	International Standards Organization
21	HDR	High Dynamic Range
22	HEVC	High Efficiency Video Coding
23	HFR	High Frame Rate
24	HTTP	HyperText Transport Protocol
25	MBT	Minimum Buffer Time
26	MHA	MPEG-H 3D Audio
27	MPEG	Moving Pictures Experts Group
28	MUP	Minimum Update Period
29	NAL	Network Abstraction Layer
30	OTT	Over-The-Top
31	PCM	Pulse Code Modulation
32	PIFF	Protected Interoperable File Format
33	PPS	Picture Parameter Set
34	PQ	Perceptual Quantization

1	PS	Parametric Stereo
2	PT	Presentation Time
3	PTO	Presentation Time Offset
4	PVR	Personal Video Recorder
5	RFC	Request for Comments
6	SAP	Stream Access Point
7	SAET	Segment Availability End Time
8	SAST	Segment Availability Start Time
9	SBR	Spectral Band Replication
10	SCTE	Society of Cable Telecommunications Engineers
11	SD	Standard Definition
12	SDR	Standard Dynamic Range
13	SEI	Supplemental Enhancement Information
14	SMPTE	Society of Motion Picture and Television Engineers
15	SPD	Suggested Presentation Delay
16	SPS	Sequence Parameter Set
17	TSB	Time Shift Buffer depth
18	TT	Timed Text
19	TTML	Timed Text Markup Language
20	UHD	Ultra-High Definitions
21	URL	Universal Resource Location
22	USAC	Unified Speech and Audio Coding
23	UTC	Universal Time Clock
24	UUID	Universally Unique Identifier
25	VAST	Video Ad Serving Template
26	VES	Video Elementary Stream
27	VP9	Video Project 9
28	VPS	Video Parameter Set
29	VUI	Video Usability Information
30	WCG	Wide Colour Gamut
31	xHE-AAC	Extended High Efficiency AAC: a specific USAC profile

1 References

2 Notes:

- 3 1) If appropriate, the references refer to specific versions of the specifications. However, implementers are
4 encouraged to check later versions of the same specification, if available. Such versions may provide
5 further clarifications and corrections. However, new features added in new versions of specifications are
6 not added automatically.
7 2) Specifications not yet officially available are marked in *italics*.
8 3) Specifications considered informative only are marked in *Arial*

9 [1] DASH-IF DASH-264/AVC Interoperability Points, version 1.0, available at
10 <http://dashif.org/w/2013/06/DASH-AVC-264-base-v1.03.pdf>

11 [2] DASH-IF DASH-264/AVC Interoperability Points, version 2.0, available at
12 <http://dashif.org/w/2013/08/DASH-AVC-264-v2.00-hd-mca.pdf>

13 [3] ISO/IEC 23009-1:2012/Cor.1:2013 Information technology -- Dynamic adaptive streaming
14 over HTTP (DASH) -- Part 1: Media presentation description and segment formats.

15 Note: this document is superseded by reference [4], but maintained as the initial version of
16 this document is provided in the above reference.

17 [4] ISO/IEC 23009-1:2014 Information technology -- Dynamic adaptive streaming over
18 HTTP (DASH) -- Part 1: Media presentation description and segment formats. Including:

- 19 • ISO/IEC 23009-1:2014/Cor 1:2015
- 20 • ISO/IEC 23009-1:2014/Cor 2:2017
- 21 • ISO/IEC 23009-1:2014/Amd 1:2015 High Profile and Availability Time
22 Synchronization
- 23 • ISO/IEC 23009-1:2014/Amd 2:2015 Spatial relationship description, gen-
24 eralized URL parameters and other extensions
- 25 • ISO/IEC 23009-1:2014/Amd 3:2016 Authentication, MPD linking,
26 Callback Event, Period Continuity and other Extensions.
- 27 • ISO/IEC 23009-1:2014/DAm 4:2016 Segment Independent SAP Signal-
28 ling (SISSI), MPD chaining, MPD reset and other extensions.

29 All the above is expected to be rolled into a third edition of ISO/IEC 23009-1 as:

- 30 • ISO/IEC 23009-1:2018 Information technology -- Dynamic adaptive
31 streaming over HTTP (DASH) -- Part 1: Media presentation description and
32 segment formats. [Note: Expected to be published by end of 2018. The draft third edi-
33 tion is available in the MPEG document [m44441](#).]

34 In addition, the following documents are under preparation in MPEG:

- 35 • ISO/IEC 23009-1:2014/DCor 3:2018 [Note: Expected to be published by mid of
36 2019. The study of the COR is available as an output document [w17951](#).]
- 37 • ISO/IEC 23009-1:2014/DAm 5:2018 Device Information and other exten-
38 sions. 2018 [Note: Expected to be published by mid of 2019. The DAM is available as
39 an output document [w18057](#).]

-
- 1 [5] ISO/IEC 23009-2:2014: Information technology -- Dynamic adaptive streaming
2 over HTTP (DASH) -- Part 2: Conformance and Reference.
3 [6] ISO/IEC 23009-3:2014: Information technology -- Dynamic adaptive streaming
4 over HTTP (DASH) -- Part 3: Implementation Guidelines.
5 [7] ISO/IEC 14496-12:2015 Information technology -- Coding of audio-visual objects -- Part
6 12: ISO base media file format. This also includes amendments and corrigendas, for de-
7 tails see here: <https://www.iso.org/standard/68960.html>
8 [8] ITU-T Recommendation H.264 (04/2017): "Advanced video coding for generic audiovis-
9 ual services" | ISO/IEC 14496-10:2010: "Information technology – Coding of audio-vis-
10 ual objects – Part 10: Advanced Video Coding".
11 [9] ISO/IEC 14496-15:2017:2015: Information technology -- Coding of audio-visual objects
12 -- Part 15: Carriage of network abstraction layer (NAL) unit structured video in ISO base
13 media file format.
14 [10] IETF RFC 6381, The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types, Au-
15 gust 2011.
16 [11] ISO/IEC 14496-3:2009 - Information technology -- Coding of audio-visual objects -- Part
17 3: Audio with Corrigendum 1:2009, Corrigendum 2:2011, Corrigendum 3:2012, Amend-
18 ment 1:2009, Amendment 2:2010, Amendment 3:2012, and Amendment 4:2014.
19 [12] ISO/IEC 14496-14:2003/Amd 1:2010 Information technology -- Coding of audio-visual
20 objects -- Part 14: The MP4 File Format
21 [13] 3GPP (2005-01-04). "ETSI TS 126 401 V6.1.0 (2004-12) - Universal Mobile Tel-
22 ecommunications System (UMTS); General audio codec audio processing func-
23 tions; Enhanced aacPlus general audio codec; General description (3GPP TS
24 26.401 version 6.1.0 Release 6)"
25 [14] ANSI/CEA-708-E: Digital Television (DTV) Closed Captioning, August 2013
26 [15] 3GPP TS 26.245: "Transparent end-to-end Packet switched Streaming Service
27 (PSS); Timed text format"
28 [16] W3C Timed Text Markup Language 1 (TTML1) (Second Edition) 24 September 2013.
29 [17] SMPTE ST 2052-1:2013 "Timed Text Format (SMPTE-TT)",
30 <https://www.smpte.org/standards>
31 [18] W3C WebVTT - The Web Video Text Tracks,— <http://dev.w3.org/html5/webvtt/>
32 [19] ITU-T Recommendation H.265 (02/2018): "Advanced video coding for generic audiovis-
33 ual services" | ISO/IEC 23008-2:2015/Amd 1:2015: " High Efficiency Coding and Media
34 Delivery in Heterogeneous Environments – Part 2: High Efficiency Video Coding",
35 downloadable here: <http://www.itu.int/rec/T-REC-H.265>
36 [20] EBU Tech 3350, "EBU-TT, Part 1, Subtitling format definition", July 2012,
37 <http://tech.ebu.ch/docs/tech/tech3350.pdf?vers=1.0>
38 [21] IETF RFC 7230, Hypertext Transfer Protocol (HTTP/1.1): Message Syntax and Routing,
39 June 2014.
40 [22] IETF RFC 7231, Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content, June
41 2014.

-
- 1 [23] IETF RFC 7232, Hypertext Transfer Protocol (HTTP/1.1): Conditional Requests, June
2 2014.
- 3 [24] IETF RFC 7233, Hypertext Transfer Protocol (HTTP/1.1): Range Requests, June 2014.
- 4 [25] IETF RFC 7234, Hypertext Transfer Protocol (HTTP/1.1): Caching, June 2014.
- 5 [26] IETF RFC 7235, Hypertext Transfer Protocol (HTTP/1.1): Authentication, June 2014.
- 6 [27] SMPTE RP 2052-10-2013: Conversion from CEA-608 Data to SMPTE-TT
7 <https://www.smpte.org/standards>
- 8 [28] SMPTE RP 2052-11-2013: Conversion from CEA 708 to SMPTE-TT
9 <https://www.smpte.org/standards>
- 10 [29] ISO/IEC 14496-30:2014, "Timed Text and Other Visual Overlays in ISO Base Media
11 File Format". Including:
12 ISO/IEC 14496-30:2014, Cor 1:2015
13 ISO/IEC 14496-30:2014, Cor 2:2016
- 14 [30] ISO/IEC 23001-7:2016: "Information technology -- MPEG systems technologies -- Part
15 7: Common encryption in ISO base media file format files".
- 16 [31] DASH Industry Forum, Test Cases and Test Vectors: <http://testassets.dashif.org/>.
- 17 [32] DASH Industry Forum, "Guidelines for Implementation: DASH-AVC/264 Conformance
18 Software", <http://dashif.org/conformance.html>.
- 19 [33] DASH Identifiers Repository, available here: <http://dashif.org/identifiers>
- 20 [34] DTS 9302J81100, "Implementation of DTS Audio in Media Files Based on ISO/IEC
21 14496", <http://www.dts.com/professionals/resources/resource-center.aspx>
<http://www.dts.com/professionals/resources/resource-center.aspx>
- 22 [35] ETSI TS 102 366 v1.2.1, Digital Audio Compression (AC-3, Enhanced AC-3) Standard
23 (2008-08)
- 24 [36] MLP (Dolby TrueHD) streams within the ISO Base Media File Format, version 1.0, Sep-
25 tember 2009.
- 26 [37] ETSI TS 102 114 v1.3.1 (2011-08), "DTS Coherent Acoustics; Core and Extensions with
27 Additional Profiles"
- 28 [38] ISO/IEC 23003-1:2007 - Information technology -- MPEG audio technologies -- Part 1:
29 MPEG Surround
- 30 [39] DTS 9302K62400, "Implementation of DTS Audio in Dynamic Adaptive Streaming over
31 HTTP (DASH)", <http://www.dts.com/professionals/resources/resource-center.aspx>
- 32 [40] IETF RFC5905, "Network Time Protocol Version 4: Protocol and Algorithms Specifica-
33 tion," June 2010.
- 34 [41] IETF RFC 6265: "HTTP State Management Mechanism", April 2011.
- 35 [42] ETSI TS 103 285 v.1.1.1: "MPEG-DASH Profile for Transport of ISO BMFF Based DVB
36 Services over IP Based Networks".
- 37 [43] ANSI/SCTE 128-1 2013: "AVC Video Constraints for Cable Television, Part 1 - Cod-
38 ing", available here: <http://www.scte.org/documents/pdf/Stand->
39 ards/ANSI_SCTE%20128-1%202013.pdf
- 40

-
- 1 [44] IETF RFC 2119, "Key words for use in RFCs to Indicate Requirement Levels",
2 April 1997.
- 3 [45] ISO: "ISO 639.2, Code for the Representation of Names of Languages — Part 2: alpha-3
4 code," as maintained by the ISO 639/Joint Advisory Committee (ISO 639/JAC),
5 http://www.loc.gov/standards/iso639-2/iso639jac.html; JAC home page:
6 http://www.loc.gov/standards/iso639-2/iso639jac.html; ISO 639.2 standard online:
7 <http://www.loc.gov/standards/iso639-2/langhome.html>.
- 8 [46] CEA-608-E, Line 21 Data Service, March 2008.
- 9 [47] IETF RFC 5234, "Augmented BNF for Syntax Specifications: ABNF", January 2008.
- 10 [48] SMPTE ST 2086:2014, "Mastering Display Color Volume Metadata Supporting High
11 Luminance And Wide Color Gamut Images"
- 12 [49] ISO/IEC 23001-8:2016, "Information technology -- MPEG systems technologies -- Part
13 8: Coding-independent code points", available here: http://standards.iso.org/ittf/PubliclyAvailableStandards/c069661_ISO_IEC_23001-8_2016.zip
- 15 [50] IETF RFC 7164, "RTP and Leap Seconds", March 2014.
- 16 [51] void
- 17 [52] IAB Video Multiple Ad Playlist (VMAP), available at <http://www.iab.net/media/file/VMAPv1.0.pdf>
- 19 [53] IAB Video Ad Serving Template (VAST), available at <http://www.iab.net/media/file/VASTv3.0.pdf>
- 21 [54] ANSI/SCTE 35 2015, Digital Program Insertion Cueing Message for Cable
- 22 [55] ANSI/SCTE 67 2014, Recommended Practice for SCTE 35 Digital Program Insertion
23 Cueing Message for Cable
- 24 [56] ANSI/SCTE 214-1, MPEG DASH for IP-Based Cable Services, Part 1: MPD Constraints
25 and Extensions
- 26 [57] ANSI/SCTE 214-3, MPEG DASH for IP-Based Cable Services, Part 3: DASH/FF Profile
- 27 [58] EIDR ID Format - EIDR: ID Format, v1.2, March 2014, available at http://eidr.org/documents/EIDR_ID_Format_v1.2.pdf
- 29 [59] Common Metadata, TR-META-CM, ver. 2.0, January 3, 2013, available at
30 http://www.movie labs.com/md/md/v2.0/Common_Metadata_v2.0.pdf
- 31 [60] IETF RFC 4648, "The Base16, Base32, and Base64 Data Encodings", October 2006.
- 32 [61] W3C TTML Profiles for Internet Media Subtitles and Captions 1.0 (IMSC1), Editor's
33 Draft 03 August 2015, available at: <https://dvcs.w3.org/hg/ttml/raw-file/tip/ttml-ww-profiles/ttml-ww-profiles.html>
- 35 [62] W3C TTML Profile Registry, available at: <https://www.w3.org/wiki/TTML/CodecsRegistry>
- 37 [63] ETSI TS 103 190-1 v1.2.1, "Digital Audio Compression (AC-4); Part 1: Channel based
38 coding".
- 39 [64] ISO/IEC 23008-3:2018, Information technology -- High efficiency coding and media de-
40 livery in heterogeneous environments -- Part 3: 3D audio.

-
- 1 [65] IETF RFC 5246, "The Transport Layer Security (TLS) Protocol, Version 1.2", August
2 2008.
- 3 [66] IETF RFC 4337, "MIME Type Registration for MPEG-4", March 2006.
- 4 [67] SMPTE: "Digital Object Identifier (DOI) Name and Entertainment ID Registry (EIDR)
5 Identifier Representations," RP 2079-2013, Society of Motion Picture and Television En-
6 gineers, 2013.
- 7 [68] SMPTE: "Advertising Digital Identifier (Ad-ID®) Representations," RP 2092-1, Society
8 of Motion Picture and Television Engineers, 2015.
- 9 [69] W3C Encrypted Media Extensions - <https://www.w3.org/TR/encrypted-media/>.
- 10 [70] void
- 11 [71] SMPTE ST 2084:2014, "Mastering Display Color Volume Metadata Supporting High
12 Luminance and Wide Color Gamut Images"
- 13 [72] ISO/IEC 23001-8:2013, "Information technology -- MPEG systems technologies -- Part
14 8: Coding-independent code points", available here: <http://standards.iso.org/ittf/Public->
15 [AvailableStandards/c062088 ISO IEC 23001-8 2013.zip](http://standards.iso.org/ittf/PubliclyAvailableStandards/c062088_ISO_IEC_23001-8_2013.zip)
- 16 [73] Recommendation ITU-R BT.709-6 (06/2015): "Parameter values for the HDTV standards
17 for production and international programme exchange".
- 18 [74] Recommendation ITU-R BT.2020-1 (06/2014): "Parameter values for ultra-high defini-
19 tion television systems for production and international programme exchange".
- 20 [75] ETSI TS 101 154 v2.2.1 (06/2015): "Specification for the use of Video and Audio
21 Coding in Broadcasting Applications based on the MPEG-2 Transport Stream."
- 22 [76] ETSI TS 103 285 v1.1.1 (05/2015): "Digital Video Broadcasting (DVB); MPEG-DASH
23 Profile for Transport of ISO BMFF Based DVB Services over IP Based Networks."
- 24 [77] 3GPP TS 26.116 (03/2016): "Television (TV) over 3GPP services; Video Pro-
25 files."
- 26 [78] DECE (05/2015): "Common File Format & Media Formats Specification",
27 <http://uvcentral.com/sites/default/files/files/PublicSpecs/CFFMediaFormat->
28 [2_2.pdf](http://uvcentral.com/sites/default/files/files/PublicSpecs/CFFMediaFormat-2_2.pdf)
- 29 [79] Ultra HD Forum: Phase A Guidelines, version 1.1, July 2015
- 30 [80] Recommendation ITU-R BT.2100-1 (07/2016): "Image parameter values for high dy-
31 namic range television for use in production and international programme exchange".
- 32 [81] SMPTE ST 2086:2014, "Mastering Display Color Volume Metadata Supporting High
33 Luminance And Wide Color Gamut Images"
- 34 [82] SMPTE ST 2094-1:2016, "Dynamic Metadata for Color Volume Transform – Core
35 Components"
- 36 [83] SMPTE ST 2094-10:2016, "Dynamic Metadata for Color Volume Transform – Appli-
37 cation #1"
- 38 [84] Recommendation ITU-R BT.1886: "Reference electro-optical transfer function for flat
39 panel displays used in HDTV studio production"
- 40 [85] ETSI DGS/CCM-001 GS CCM 001 "Compound Content Management"
- 41 [86] VP9 Bitstream & Decoding Process Specification. <https://storage.googleapis.com/downloads.webmproject.org/docs/vp9/vp9-bitstream-specification-v0.6-20160331-draft.pdf>

-
- 1 [87] VP Codec ISO Media File Format Binding <https://www.webmproject.org/vp9/mp4/>
2
3 [88] ETSI TS 143 433-1, "High-Performance Single Layer High Dynamic Range (HDR) Sys-
4 tem for use in Consumer Electronics devices; Part 1: Directly Standard Dynamic Range
5 (SDR) Compatible HDR System (SL-HDR1)"
6 [89] ST 2094-40:2016 - SMPTE Standard - Dynamic Metadata for Color Volume Transform
7 — Application #4, Sept. 2016
8 [90] CTA 861-G - CTA Standard: A DTV Profile for Uncompressed High Speed Digital In-
9 terfaces, Nov. 2016
10 [91] ISO/IEC 23000-19:2018 - Information technology -- Coding of audio-visual objects --
11 Part 19: Common media application format (CMAF) for segmented media, Amend-
12 ment 2: xHE-AAC and other media profiles

1

2 1. Introduction

3 This document defines DASH-IF's InterOperability Points (IOPs). The document includes IOPs
 4 for only this version of the document. For earlier versions, please refer to version 1 [1] and version
 5 2 [2] of this document. DASH-IF recommends to deprecate the IOPs in previous versions and
 6 deploy using one of the IOPs and extensions in this document.

7 As a historical note, the scope of the initial DASH-AVC/264 IOP, issued with version 1 of this
 8 document [1] was the basic support high-quality video distribution over the top. Both live and on-
 9 demand services are supported.

10 In the second version of this document [2], HD video (up to 1080p) extensions and several multi-
 11 channel audio extensions are defined.

12 In this third version of the DASH-IF IOP document, two new DASH-264/AVC IOPs are defined.
 13 Detailed refinements and improvements for DASH-IF live services and for ad insertion were added
 14 in these IOPs. One of these IOP is the superset of the simpler one. Additionally, two corresponding
 15 IOPs are defined to also support HEVC [19]. In both cases, AVC and HEVC, the more advanced
 16 IOP adds additional requirements on the DASH client to support segment parsing to achieve en-
 17 hancement of live services. This structuring separates the Media Profiles from DASH features.

18 In the fourth version, beyond minor improvements, corrections and alignment with MPEG-DASH
 19 third edition, the key additions are extensions for next generation audio and UHD/HDR video.

20 This document defines the IOPs in Table 1 and Extensions in Table 2. The Implementation Guide-
 21 line's version in which each IOP or Extension was added is also provided in the tables.

22 Note that all version 1 IOPs are also defined in version 2 and therefore referencing version [2] is
 23 sufficient.

24

Table 1 DASH-IF Interoperability Points

Interoperability Point	Identifier	Ver- sion	R e f e r e n c e
DASH-AVC/264	http://dashif.org/guidelines/dash264	1.0	[2] , .3
DASH-AVC/264 SD	http://dashif.org/guidelines/dash264#sd	1.0	[2] , 7

DASH-AVC/264 HD	http://dashif.org/guidelines/dash264#hd	2.0	[2] ,	3 8 .3
DASH-AVC/264 main	http://dashif.org/guidelines/dash264main	3.0	8 .2	
DASH-AVC/264 high	http://dashif.org/guidelines/dash264high	3.0	8 .3	
DASH-IF IOP simple	http://dashif.org/guidelines/dash-if-simple	3.0	8 .4	
DASH-IF IOP main	http://dashif.org/guidelines/dash-if-main	3.0	8 .5	
DASH-IF IOP on-demand	http://dashif.org/guidelines/dash-if-on-demand	4.3	3 .1 0 .3	
DASH-IF IOP mixed on-demand	http://dashif.org/guidelines/dash-if-mixed	4.3	3 .1 0 .4	

1

2 Note that all extensions defined in version 2 of this document are carried over into version 3 without any modifications. In order to maintain a single document, referencing in Table 2 is restricted
 3 to this document.

4

5 **Table 2 DASH-IF Interoperability Point Extensions**

Extension	Identifier	Version	Section
DASH-IF multichannel audio extension with Enhanced AC-3	http://dashif.org/guidelines/dashif#ec-3	2.0	9.4.2.3
DASH-IF multichannel extension with Dolby TrueHD	http://dashif.org/guidelines/dashif#mlpa	2.0	9.4.2.3
DASH-IF multichannel extension with AC-4	http://dashif.org/guidelines/dashif#ac-4	3.1	9.4.2.3

DASH-IF multichannel audio extension with DTS Digital Surround	http://dashif.org/guidelines/dashif#dtsc	2.0	9.4.3.3
DASH-IF multichannel audio extension with DTS-HD High Resolution and DTS-HD Master Audio	http://dashif.org/guidelines/dashif#dtsh	2.0	9.4.3.3
DASH-IF multichannel audio extension with DTS Express	http://dashif.org/guidelines/dashif#dtse	2.0	9.4.3.3
DASH-IF multichannel extension with DTS-HD Lossless (no core)	http://dashif.org/guidelines/dashif#dtsl	2.0	9.4.3.3
DASH-IF multichannel audio extension with MPEG Surround	http://dashif.org/guidelines/dashif#mps	2.0	9.4.4.3
DASH-IF multichannel audio extension with HE-AACv2 level 4	http://dashif.org/guidelines/dashif#heaa_c-mc51	2.0	9.4.5.3
DASH-IF multichannel audio extension with HE-AACv2 level 6	http://dashif.org/guidelines/dashif#heaa_c-mc71	2.0	9.4.5.3
DASH-IF multichannel audio extension with MPEG-H 3D Audio	http://dashif.org/guidelines/dashif#mpeg_h-3da	4.2	9.4.6.3
DASH-IF audio extension with USAC	http://dashif.org/guidelines/dashif#cxha	4.3	9.4.7.3
DASH-IF UHD HEVC 4k	http://dashif.org/guidelines/dash-if-uhd#4k	4.0	10.2
DASH-IF HEVC HDR PQ10	http://dashif.org/guidelines/dash-if-uhd#hdr-pq10	4.0	10.3
DASH-IF UHD Dual-Stream (Dolby Vision)	http://dashif.org/guidelines/dash-if-uhd#hdr-pq10	4.1	10.4
DASH-IF VP9 HD	http://dashif.org/guidelines/dashif#vp9	4.1	11.3.1
DASH-IF VP9 UHD	http://dashif.org/guidelines/dash-if-uhd#vp9	4.1	11.3.2

DASH-IF VP9 HDR	<code>http://dashif.org/guide-lines/dashif#vp9-hdr</code>	4.1	11.3.3
	<code>http://dashif.org/guide-lines/dash-if-uhd#vp9-hdr</code>		

1
2 In addition to the Interoperability points in Table 1 and extensions in Table 2, this document also
3 defines several other identifiers and other interoperability values for functional purposes as docu-
4 mented in Table 3.

5
6 **Table 3 Identifiers and other interoperability values defined in this Document**

Identifier	Semantics	Type	Section
<code>http://dashif.org/identifiers/vast30</code>	Defines an event for signalling events of VAST3.0	Event	5.6
<code>http://dashif.org/guidelines/trickmode</code>	Defines a trick mode Adaptation Set.	Functionality	3.2.9
<code>http://dashif.org/guidelines/clearKey</code>	Defines name space for the Laurl element in W3C	Namespace	7.6.2.4
<code>e2719d58-a985-b3c9-781a-b030af78d30e</code>	UUID for W3C Clear Key with DASH	Content Protection	7.6.2.4
<code>http://dashif.org/guidelines/last-segment-number</code>	Signaling last segment number	Functionality	4.4.3.6
<code>http://dashif.org/guidelines/thumb-nail_tile</code>	Signalling the availability of the thumbnail tile adaptation set	Functionality	6.2.6

7
8 DASH-IF supports these guidelines with test and conformance tools:

- 9 • DASH-IF conformance software is available for use online at <http://dashif.org/conformance.html> [32]. The software is based on an open-source code. The frontend source code
10 and documentation is available at: <https://github.com/Dash-Industry-Forum/Conformance-Software>. The backend source code is available at: <https://github.com/Dash-Industry-Forum/Conformance-and-reference-source>.
- 11 • DASH-IF test assets (features, test cases, test vectors) along with the documentation are
12 available at <http://testassets.dashif.org> [31].
- 13 • DASH Identifiers for different categories can be found at <http://dashif.org/identifiers/> [33].
14 DASH-IF supporters are encouraged that external identifiers are submitted for documen-
15 tation there as well. Note also that DASH-IF typically tries to avoid defining identifiers.
16 Identifiers in *italics* are subject to discussion with other organizations and may be depre-
17 cated in a later version.

18 Technologies included in this document and for which no test and conformance material is pro-
19 vided, are only published as a candidate technology and may be removed if no test material is
20 provided before releasing a new version of this guidelines document.

1 **2. Context and Conventions**

2 **2.1. Relation to MPEG-DASH and other DASH specifications**

3 Dynamic Adaptive Streaming over HTTP (DASH) is initially defined in the first edition of
4 ISO/IEC 23009-1 which was published in April 2012 and some corrections were done in 2013 [1].
5 In May 2014, ISO/IEC published the second version of ISO/IEC 23009-1 [4] that includes additional
6 features and provide additional clarifications. The initial two versions of this document
7 where based on the first edition of ISO/IEC 23009-1. This version is based on the second edition
8 of ISO/IEC 23009-1, i.e. ISO/IEC 23009-1:2014 including Cor.3 and Amd.3 [4]. This means that
9 also for all interoperability points that were initially defined in earlier versions of the document,
10 also now the second edition serves as the reference. Backward-compatibility across different edition
11 is handled by MPEG-DASH in ISO/IEC 23009-1 [4]. Note that this document also refers to
12 technologies in draft corrigenda and draft amendments of MPEG. For this version, in particular
13 Draft Amd.4 of ISO/IEC 23009-1:2014, and Draft Cor.3 of ISO/IEC 23009-1:2014 are of relevance.
14

15 This document was generated in close coordination with DVB-DASH [42]. The tools and features
16 are aligned to the extent considered reasonable. To support implementers, this document attempts
17 to highlight any differences and/or further restrictions or extensions when compared to DVB-
18 DASH. However, as a disclaimer, this coverage is not considered complete.

19 **2.2. Compatibility and Extensions to Earlier Versions**

20 **2.2.1. Summary of Version 3 Modifications**

21 Version 3 of this document applies the following modifications compared to version 2 [2]:

- 22 • Reference to the second edition of ISO/IEC 23009-1 including amendment 1 and cor.1 [4],
23 as well as Amendment 3 [4].
- 24 • Add an explicit statement in DASH-264/AVC to forbid time code wrap around
- 25 • Definition on the usage of key words in clause 2.3.
- 26 • Add more constraints on the usage of Trick Modes for improved interoperability in clause
27 3.2.9.
- 28 • Add more constraints on the Representations in one Adaptation Set in clause 3.2.10, espe-
29 cially for the case when the bitstream switching is true.
- 30 • Add additional details on the usage of HTTP in clause 3.4.
- 31 • Add H.265/HEVC as a codec and create IOPs for inclusion of this codec.
- 32 • Add CEA-608/708 closed captioning in SEI messages in clause 6.4.3.
- 33 • Detailed description of simple and main live operation, with the latter including segment
34 parsing in clause 4.
- 35 • Detailed description of server-based and app-based ad insertion in clause 5
- 36 • General editorial updates and clarifications
- 37 • Updates and clarification to clause 7 on DRM and common encryption.
- 38 • Update to references

-
- 1 • Relaxation of the audio encoding requirements in clause 6.3.2.
2 • Add clarification on the usage of the minimum buffer time and bandwidth in clause 3.2.8.
3 • Add an informative clause on the timing model of DASH in clause 3.2.7.
4 • Relax the use of the 'lmsg' brand for signaling the last segment in clause 3.6.
5 • Simplification of the codecs table.

6 Version 3.1 of this document applies the following modifications compared to version 3

- 7 • Further updates to references
8 • Several editorial corrections and clarifications
9 • Obsolete reference to RFC2616 and refer to the new set of RFCs for HTTP/1.1
10 • A clause is added on how to operate with transforming proxies and other adaptation mid-
11 dle-boxes in clause 3.4.4.
12 • Considerations on how to operate at the live edge for a live service
13 • The addition of the availability time offset to the description of the live service.
14 • The explicit exclusion of on-request based xlink for dynamic services.
15 • Clarifications of HEVC signaling for DASH in Table 20 based on feedback from MPEG.
16 • Clarify relation between SMPTE-TT and IMSC1 in clause 6.4.2 and 6.4.4.
17 • Add extension for audio codec AC-4.

18 Version 3.2 of this document applies the following modifications compared to version 3.1

- 19 • Further updates to references.
20 • Several editorial corrections and clarifications.
21 • Small clarification updates on the timing model in clause 3.2.7.
22 • Added support for switching across Adaptation Sets, in particular for H.264/AVC and
23 H.265/HEVC switching in clause 3.8.
24 • Added a clarification on the value of @audioSamplingRate for AAC SBR in clause
25 6.3.2.
26 • Added a clause on the use of HTTPS with DASH in clause 7.2.
27 • Moved the test DRM to the test vector document 7.
28 • Added support for key rotation in clause 7.
29 • Add extension for MPEG-H audio in clause 9.

30 Version 3.3 of this document applies the following modifications compared to version 3.2

- 31 • Identifiers are summarized in the Introduction 1.
32 • References are updated
33 • Several editorial corrections and clarifications are applied
34 • Guidelines and permitted attributes when converting a live service to on-demand are doc-
35 umented, for details see clause 4.6.2.
36 • Added clarification on the format of remote elements in alignment with DCOR.3 of
37 ISO/IEC 23009-1:2014 [4].

-
- 1 • Addressed the differentiation of different content types through mime types, codecs and
2 Roles rather than the use @contentType. See clause 3.2.13 for more details.
3 • Added guidelines on how to use clear keys in the context of DASH, see clause 7.7.
4 • Provided guidelines on the timing model when using side-car subtitle files in DASH, see
5 clause 6.4.4.
6 • Update period continuity to align with ISO/IEC 23009-1:2014/Amd.3 [4].
7 • Update the callback event to align with ISO/IEC 23009-1:2014/Amd.3 [4].
8 • Update the MPD anchors to align with ISO/IEC 23009-1:2014/Amd.3 [4].
9 • Take into account the updated in ISO/IEC 23009-1:2014/DCOR.3 [4].

10 Version 4.0 of this document applies the following modifications compared to version 3.3

- 11 • References are updated
12 • Several editorial corrections and clarifications are applied
13 • Update Adaptation Set Switching to align with ISO/IEC 23009-1:2014/Amd.3 [4].
14 • Add a recommendation on the setting of **Period@start** in case of multi-period content
15 in clause 4.3.3.3.1.
16 • Add an external Representation Index for On-demand content in clause 3.2.1.
17 • Add Period connectivity to multi-period content to align with ISO/IEC 23009-
18 1:2014/Amd.3 [4] in clause 3.2.12.
19 • Add clarifications and further recommendations and restriction to the use of the Time Syn-
20 chronization in clause 4.7.
21 • Align Early Terminated Period with ISO/IEC 23009-1:2014/Amd.3 [4] in clause 4.8.3.
22 • Removal of interoperability requirements for CEA-708 in clause 6.4.3.4.
23 • References to EIDR and Ad-ID for Asset Identifiers in clause 5.5.1.
24 • Addition of W3C Clear Key with DASH in clause 7.6.2.4.
25 • Addition of guidelines for trick modes in live services in clause 4.10.
26 • Addition of UHD/HDR Extensions in clause 10.

27 Version 4.1 of this document applies the following modifications compared to version 4.0

- 28 • Several editorial fixes and updates.
29 • Guidelines for Robust Live Services in clause 4.11.
30 • Updates to W3C Clear Key usage in clause 7.6.2.4.
31 • Addition of last segment number signaling in clause 4.4.3.5 and 4.4.3.6.
32 • Addition of Dolby Vision dual layer in clause 10.4 and referenced Annexes
33 • Seek Preview and Thumbnail Navigation in clause 3.2.9 and clause 6.2.6
34 • Clarification on **AudioChannelConfiguration** element in clause 6.3.3.2
35 • Annotation and client model for content selection defined in clause 3.9.
36 • Remote referencing and use of query parameters for ad insertion in clause 5.3.2.1.3 and
37 5.3.5, respectively.

-
- 1 • Default KID Clarifications in clause 7.5.3.4.
2 • Addition of VP9 Video to DASH-IF IOP in clause 11.
3 • Updates to the references and separation of informative and normative references.
4 • Client and DRM System Interactions in clause 7 and in particular in clause 7.7.9

5 Version 4.2 of this document applies the following modifications compared to version 4.1

- 6 • Several editorial fixes and updates.
7 • Updates to the usage of HTTPS in clause 7.2 and replacing RFC 2818 with RFC 5246.
8 • Clarification on initialization range for on-demand content in clause 3.2.1.
9 • Clarifications on permitted updates in MPD updates in clause 4.4.3.3.
10 • Clarifications to MPD Content Protection Constraints in clause 7.7.2.
11 • Addition of Reference Resolution in clause 3.2.15.
12 • Removal of Representation Index as it is not defined for ISO BMFF until now.
13 • Addition of a clause that clarifies how a live service can be changed to an On-demand
14 service using the same MPD URL in clause 4.6.4.
15 • Clarification on the SegmentTemplate substitution parameters in clause 4.3.2.2.8.

16 Version 4.3 of this document applies the following modifications compared to version 4.2:

- 17 • Several editorial fixes and updates.
18 • Renaming of clause 3 to Usage of MPEG-DASH
19 • Adapting the default parameters for the ISO BMFF in clause 3.2.1
20 • Added restrictions on the usage of XML duration in clause 3.2.7.4
21 • Added details on MPD Location and Reference Resolution in clause 3.2.15.
22 • Addition of On-Demand Service description in clause 3.10 including mixed Media Presen-
23 tations
24 • Added a processing model for availability time offset when present on multiple levels.
25 • Provided an explanation on “change lead time” in clause 4.11.4.1
26 • Added clarification to Bitstream switching requirements in clause 6.2.5.2.
27 • Added a restriction on segmented subtitles in clause 6.4.5.
28 • Added correction to MPD Content Protection Constraints in clause 7.7.2.
29 • Updated the MPEG-H 3D Audio sample entry in clause 9.2.5.2.
30 • Addition of MPEG-D Unified Speech and Audio Coding in clause 9.2.6 and 9.4.7.
31 • Addition of HDR dynamic metadata for SDR in clause 10.2.2.4.
32 • Addition of HDR dynamic metadata for HDR PQ in clause 10.3.2.4.

33

34 **2.2.2. Backward-Compatibility Considerations**

35 Generally, content can be offered such that it can be consumed by version 2 and version 3 clients.
36 In such a case the restricted authoring should be used and it should be accepted that version 2

clients may ignore certain Representations and Adaptation Sets. Content Authors may also consider the publication of two MPDs, but use the same segment formats.

In terms of compatibility between version 2 and version 3, the following should be considered:

- The backward-compatibility across MPEG editions is handled in the second edition of ISO/IEC 23009-1 [4].
- General clarifications and updates are added
- Further restrictions on content authoring compared to version 2 are:
 - forbid time code wrap around
 - the usage of DRM, especially the Content Protection element
 - constraints on trick mode usage
 - additional constraints on the usage of HTTP
 - Adaptation Set constraints
- Relaxations are:
 - Permit usage of additional subtitling format based on CEA-608/708
 - the audio encoding requirements for HE-AACv2
 - permit to not use of the 'lmsg' brand for signaling the last segment
 - the ability to signal bitstream switching set to true
 - the use of remote elements with Xlink

2.3. Use of Key Words

2.3.1. Background

DASH-IF generally does not write specifications, but provides and documents guidelines for implementers to refer to interoperability descriptions. In doing so, the DASH-IF agreed to use key words in order to support readers of the DASH-IF documents to understand better how to interpret the language. The usage of key words in this document is provided below.

2.3.2. Key Words

The key word usage is aligned with the definitions in RFC 2119 [44], namely:

- SHALL: This word means that the definition is an absolute requirement of the specification.
- SHALL NOT: This phrase means that the definition is an absolute prohibition of the specification.
- SHOULD: This word means that there may exist valid reasons in particular circumstances to ignore a particular item, but the full implications must be understood and carefully weighed before choosing a different course.
- SHOULD NOT: This phrase means that there may exist valid reasons in particular circumstances when the particular behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label.

-
- 1 • MAY: This word means that an item is truly optional. One vendor may choose to include
2 the item because a particular marketplace requires it or because the vendor feels that it
3 enhances the product while another vendor may omit the same item.

4 These key words are attempted to be used consistently in this document, but only in small letters.

5 **2.3.3. Mapping to DASH-IF Assets**

6 If an IOP document associates such a key word from above to a content authoring statement then
7 the following applies:

- 8 • SHALL: The conformance software provides a conformance check for this and issues an
9 *error* if the conformance is not fulfilled.
- 10 • SHALL NOT: The conformance software provides a conformance check for this and issues
11 an *error* if the conformance is not fulfilled.
- 12 • SHOULD: The conformance software provides a conformance check for this and issues a
13 *warning* if the conformance is not fulfilled.
- 14 • SHOULD NOT: The conformance software provides a conformance check for this and
15 issues a *warning* if the conformance is not fulfilled.
- 16 • SHOULD and MAY: If present, the feature check of the conformance software documents
17 a feature of the content.

18 If an IOP document associates such a key word from above to a DASH Client then the following
19 applies:

- 20 • SHALL: Test content is necessarily provided with this rule and the reference client imple-
21 ments the feature.
- 22 • SHALL NOT: The reference client does not implement the feature.
- 23 • SHOULD: Test content is provided with this rule and the reference client implements the
24 feature unless there is a justification for not implementing this.
- 25 • SHOULD NOT: The reference client does not implement the feature unless there is a jus-
26 tification for implementing this.
- 27 • MAY: Test content is provided and the reference client implements the feature if there is a
28 justification this.

29 **2.4. Definition and Usage of Interoperability Points**

30 **2.4.1. Profile Definition in ISO/IEC 23009-1**

31 MPEG DASH defines formats for MPDs and Segments. In addition MPEG provides the ability to
32 further restrict the applied formats by the definition of *Profiles* as defined on clause 8 of ISO/IEC
33 23009-1 [4]. Profiles of DASH are defined to enable interoperability and the signaling of the use
34 of features.

35 Such a profile can also be understood as permission for DASH clients that implement the features
36 required by the profile to process the Media Presentation (MPD document and Segments).

37 Furthermore, ISO/IEC 23009-1 permits external organizations or individuals to define restrictions,
38 permissions and extensions by using this profile mechanism. It is recommended that such external
39 definitions be not referred to as profiles, but as *Interoperability Points*. Such an interoperability
40 point may be signalled in the @profiles parameter once a URI is defined. The owner of the

1 URI is responsible to provide sufficient semantics on the restrictions and permission of this in-
2 teroperability point.

3 This document makes use of this feature and provides a set of Interoperability Points. Therefore,
4 based on the interoperability point definition, this document may be understood in two ways:

- 5 • a collection of content conforming points, i.e. as long as the content conforms to the re-
6 strictions as specified by the IOP, clients implementing the features can consume the con-
7 tent.
- 8 • a client capability points that enable content and service providers for flexible service pro-
9 visioning to clients conforming to these client capabilities.

10 This document provides explicit requirements, recommendations and guidelines for content au-
11 thoring that claims conformance to a profile (by adding the `@profiles` attribute to the MPD) as
12 well as for clients that are permitted to consume a media presentation that contains such a profile.

13 **2.4.2. Usage of Profiles**

14 A Media Presentation may conform to one or multiple profiles/interoperability points and con-
15 forms to each of the profiles indicated in the `MPD@profiles` attribute is specified as follows:

16 When `ProfA` is included in the `MPD@profiles` attribute, the MPD is modified into a profile-
17 specific MPD for profile conformance checking using the following ordered steps:

- 18 1. The `MPD@profiles` attribute of the profile-specific MPD contains only `ProfA`.
- 19 2. An `AdaptationSet` element for which `@profiles` does not or is not inferred to in-
20 clude `ProfA` is removed from the profile-specific MPD.
- 21 3. A Representation element for which `@profiles` does not or is not inferred to include
22 `ProfA` is removed from the profile-specific MPD.
- 23 4. All elements or attributes that are either (i) in this Part of ISO/IEC 23009 and explicitly
24 excluded by `ProfA`, or (ii) in an extension namespace and not explicitly included by
25 `ProfA`, are removed from the profile-specific MPD.
- 26 5. All elements and attributes that “may be ignored” according to the specification of `ProfA`
27 are removed from the profile-specific MPD.

28 An MPD is conforming to profile `ProfA` when it satisfies the following:

- 29 1. `ProfA` is included in the `MPD@profiles` attribute.
- 30 2. The profile-specific MPD for `ProfA` conforms to ISO/IEC 23009-1
- 31 3. The profile-specific MPD for `ProfA` conforms to the restrictions specified for `ProfA`.

32 A Media Presentation is conforming to profile `ProfA` when it satisfies the following:

- 33 1. The MPD of the Media Presentation is conforming to profile `ProfA` as specified above.
- 34 2. There is at least one Representation in each Period in the profile-specific MPD for `ProfA`.
- 35 3. The Segments of the Representations of the profile-specific MPD for `ProfA` conform to
36 the restrictions specified for `ProfA`.

37 **2.4.3. Interoperability Points and Extensions**

38 This document defines Interoperability Points and Extensions. Both concepts make use of the pro-
39 file functionality of ISO/IEC 23009-1.

1 Interoperability Points provide a basic collection of tools and features to ensure that content/service
2 providers and client vendors can rely to support a sufficiently good audio-visual experience. Ex-
3 tensions enable content/service providers and client vendors to enhance the audio-visual experi-
4 ence provided by an Interoperability Point in a conforming manner.

5 The only difference between Interoperability Points and Extensions is that Interoperability Points
6 define a full audio-visual experience and Extensions enhance the audio-visual experience in typi-
7 cally only one dimension.

8 Examples for the usage of the @profiles signaling are provided in Annex A of this document.

9 **3. Usage of MPEG-DASH**

10 **3.1. Scope**

11 DASH-IF Interoperability Points use ISO base media file format [7] based encapsulation and pro-
12 vide significant commonality with a superset of the ISO BMFF On-Demand, the ISO BMFF Live
13 profile, and the ISO BMFF Common Profile as defined in ISO/IEC 23009-1 [4], sections 8.3, 8.4
14 and 8.10, respectively. DASH-IF IOPs are intended to provide support for on-demand and live
15 content. The primary constraints imposed by this profile are the requirement that each Represen-
16 tation is provided in one of the following two ways

- 17 • as a single Segment, where Subsegments are aligned across Representations within an Ad-
18 aptation Set. This permits scalable and efficient use of HTTP servers and simplifies seam-
19 less switching. This is mainly for on-demand use cases.
- 20 • as a sequence of Segments where each Segment is addressable by a template-generated
21 URL. Content generated in this way is mainly suitable for dynamic and live services.

22 In both cases (Sub)Segments begin with Stream Access Points (SAPs) of type 1 or 2 [7], i.e. regular
23 IDR frames in case of video. In addition, (Sub)Segments are constrained so that for switching
24 video Representations within one Adaptation Set the boundaries are aligned without gaps or over-
25 laps in the media data. Furthermore, switching is possible by a DASH client that downloads, de-
26 codes and presents the media stream of the come-from Representation and then switches to the go-
27 to Representation by downloading, decoding and presenting the new media stream. No overlap in
28 downloading, decoding and presentation is required for seamless switching of Representations in
29 one Adaptation Set.

30 Additional constraints are documented for bitstream switching set to true as well as special case
31 such as trick modes, etc.

32 **3.2. DASH Formats**

33 **3.2.1. Introduction**

34 This section introduces the detailed constraints of the MPD and the DASH segments in a descrip-
35 tive way referring to ISO/IEC 23009-1 [4]. The DASH-based restrictions have significant com-
36 monality with the ISO BMFF Live and On-Demand profiles from the MPEG-DASH specification.
37 Specifically:

- 38 • Segment formats are based on ISO BMFF with fragmented movie files, i.e. (Sub)Segments
39 are encoded as movie fragments containing a track fragment as defined in ISO/IEC 14496-
40 12 [7], plus the following constraints to make each movie fragment independently de-
41 codable:

-
- 1 • Default values set in the Track Extends ('trex') box may or may not be overridden
2 by the corresponding values set in movie fragments (in 'tfhd' or 'trun' boxes). This
3 permits different deployment models, depending on the headend operation.
4 •
5 • The 'moof' boxes shall not use external data references, the flag 'default-base-
6 is-moof' shall also be set (aka movie-fragment relative addressing) and data-
7 offset shall be used, i.e. base-data-offset-present shall not be used (fol-
8 lows ISO/IEC 23009-1 [4]).
9 • Alignment with ISO BMFF Live & On-Demand Profiles, i.e. within each Adaptation Set
10 the following applies
11 • Fragmented movie files are used for encapsulation of media data
12 • (Sub)Segments are aligned to enable seamless switching

13 Beyond the constraints provided in the ISO BMFF profiles, the following additional restrictions
14 are applied.

- 15 • IDR-like SAPs (i.e., SAPs type 2 or below) at the start of each (Sub)Segment for simple
16 switching.
17 • Segments should have almost equal duration. The maximum tolerance of segment duration
18 shall be $\pm 50\%$ and the maximum accumulated deviation over multiple segments shall be
19 $\pm 50\%$ of the signaled segment duration (i.e. the @duration). Such fluctuations in actual
20 segment duration may be caused by for example ad replacement or specific IDR frame
21 placement. Note that the last segment in a Representation may be shorter according to
22 ISO/IEC 23009-1 [4].

23 Note 1: If accurate seeking to specific time is required and at the same time a fast
24 response is required one may use On-Demand profile for VoD or the **SegmentTimeline** based addressing. Otherwise the offset in segment duration com-
25 pared to the actual media segment duration may result in a less accurate seek position
26 for the download request, resulting in some increased initial start-up.
27

28 Note 2: The maximum tolerance of segment duration is also addressed in ISO/IEC
29 23009-1:2014/Cor.3 [4] and once approved, a reference to the specification may
30 replace the above text.

- 31 • If the **SegmentTimeline** element is used for the signaling of the Segment duration, the
32 timing in the segment timeline shall be media time accurate and no constraints on segment
33 duration deviation are added except the maximum segment duration as specified in the
34 MPD. However, despite the usage of the the **SegmentTimeline**, it is not encouraged to
35 use varying Segment durations. The **SegmentTimeline** element should only be used
36 in order to signal occasional shorter Segments (possibly caused by encoder processes) or
37 to signal gaps in the time line.
38 • only non-multiplexed Representations shall be used, i.e. each Representation only contains
39 a single media component.
40 • Addressing schemes are restricted to
41 • templates with number-based addressing
42 • templates with time-based addressing

-
- 1 • For on-demand profiles, the Indexed Media Segment as defined in ISO/IEC 23009-1
2 [4], clause 6.3.4.4 shall be used. In this case the @indexRange attribute shall be pre-
3 sent. Only a single sidx box shall be present.

4 Note 1: external Representation Indices are considered beneficial, but are only de-
5 fined for MPEG-2 TS at this stage. DASH-IF reached out to ISO/IEC to clarify the
6 applicability to ISO BMFF based media segments and expects to add this feature
7 in a later version of the IOP Guidelines.

8 Note 2: The single sidx restriction was introduced in version 3 of this document
9 based on deployment experience and to enable alignment with DVB DASH.

10 Note 3: If only the movie header is desired, for example for initial capability
11 checks, then downloading at most the first 1000 bytes of the Media Segment is
12 sufficient for all DASH-IF test vectors as available by the end of 2017.

- 13 • In case multiple Video Adaptation Sets as defined in 3.2.13 are offered, exactly one video
14 Adaptation Set shall be signaled as the main one unless different Adaptation Sets contain
15 the same content with different quality or different codecs. In the latter case, all Adaptation
16 Sets with the same content shall be signaled as the main content. Signaling as main content
17 shall be done by using the Role descriptor with @schemeIdUri="urn:mpeg:dash:role:2011" and @value="main".
18
19 • The content offering shall adhere to the presence rules of certain elements and attributes as
20 defined section 3.2.4.

21 It is expected that a client conforming to such a profile is able to process content offered under
22 these constraints. More details on client procedures are provided in section 3.3.

23 **3.2.2. Media Presentation Description constraints for v1 & v2 Clients**

24 **3.2.2.1. Definition according to ISO/IEC 23009-1**

25 This section follows a description according to ISO/IEC 23009-1. In section 3.2.2.2, a restricted
26 content offering is provided that provides a conforming offering.

27 NOTE: The term "ignored" in the following description means, that if an MPD is provided
28 and a client that complies with this interoperability point removes the element that may be
29 ignored, then the MPD is still complying with the constraints of the MPD and segments as
30 defined in ISO/IEC 23001-9, section 7.3.

31 The MPD shall conform to the ISO Base Media File Format Common profile as defined on
32 ISO/IEC 23009-1:2014/Amd.1:2015 [4], section 8.9, except for the following issues:

- 33 • Representations with @mimeType attribute application/xml+ttml shall not be ig-
34 nored.

35 In addition, the Media Presentation Description shall conform to the following constraints:

- 36 — **Representation** elements with a @subsegmentStartsWithSAP value set to 3
37 may be ignored.
38 — **Representation** elements with a @startsWithSAP value set to 3 may be ignored.

1 — If a Period contains multiple Video Adaptation Sets as defined in 3.2.13 then at least one
2 Adaptation Set shall contain a Role element <Role schemeI-
3 dUri="urn:mpeg:dash:role:2011" value="main"> and each Adaptation Set
4 containing such a **Role** element shall provide perceptually equivalent media streams.

5 **3.2.2.2. Simple Restricted Content Offering**

6 A conforming MPD offering based on the ISO BMFF Live Profile shall contain

- 7 • **MPD@type** set to static or set to dynamic.
8 • **MPD@profiles** includes urn:mpeg:dash:profile:isoff-live:2011
9 • One or multiple Periods with each containing one or multiple Adaptation Sets and with
10 each containing one or multiple Representations.
11 • The Representations contain or inherit a **SegmentTemplate** with \$Number\$ or
12 \$Time\$ Identifier.
13 • @segmentAlignment set to true for all Adaptation Sets

14
15 A conforming MPD offering based on the ISO BMFF On-Demand Profile shall contain

- 16 • **MPD@type** set to static.
17 • **MPD@profiles** includes urn:mpeg:dash:profile:isoff-on-demand:2011
18 • One or multiple Periods with each containing one or multiple Adaptation Sets and with
19 each containing one or multiple Representations.
20 • @subSegmentAlignment set to true for all Adaptation Sets

21
22 **3.2.3. Segment format constraints**

23 Representations and Segments referred to by the Representations in the profile-specific MPD for
24 this profile, the following constraints shall be met:

- 25 — Representations shall comply with the formats defined in ISO/IEC 23009-1, section 7.3.
26 — In Media Segments, all Segment Index ('sidx') and Subsegment Index ('ssix') boxes, if
27 present, shall be placed before any Movie Fragment ('moof') boxes.

28 Note: DVB DASH [42] permits only one single Segment Index box ('sidx') for the entire Segment. As this
29 constraints is not severe in the content offering, it is strongly recommended to offer content following this constraint.

- 30 — If the **MPD@type** is equal to "static" and the **MPD@profiles** attribute includes
31 "urn:mpeg:dash:profile:isoff-on-demand:2011", then
32 — Each Representation shall have one Segment that complies with the Indexed Self-Initial-
33 izing Media Segment as defined in section 6.3.5.2 in ISO/IEC 23009-1.

-
- 1 — Time Codes expressing presentation and decode times shall be linearly increasing with in-
2 creasing Segment number in one Representation. In order to minimize the frequency of time
3 code wrap around 64 bit codes may be used or the timescale of the Representation may be
4 chosen as small as possible. In order to support time code wrap around, a new Period may be
5 added in the MPD added that initiates a new Period in order to express a discontinuity.

6 **3.2.4. Presence of Attributes and Elements**

7 Elements and attributes are expected to be present for certain Adaptation Sets and Representations
8 to enable suitable initial selection and switching. Simple rules are provided in this section. A de-
9 tailed description of multi-track content offering is provided in clause 3.9.

10 Specifically the following applies:

- 11 • For any Video Adaptation Sets as defined in 3.2.13 the following attributes shall be present
 - 12 ○ @maxWidth (or @width if all Representations have the same width)
 - 13 ○ @maxHeight (or @height if all Representations have the same height)
 - 14 ○ @maxFrameRate (or @frameRate if all Representations have the same frame
 - 15 ○ @par

17 Note: The attributes @maxWidth and @maxHeight should be used such that they
18 describe the target display size. This means that they may exceed the actual largest size
19 of any coded Representation in one Adaptation Set.

- 20 ○ The attributes @minWidth and @minHeight should not be present. If present,
21 they may be smaller than the smallest @width or smallest @height in the Adap-
22 tation Set.
- 23 • For any Representation within a Video Adaptation Sets as defined in 3.2.13 the following
24 attributes shall be present:
 - 25 ○ @width, if not present in **AdaptationSet** element
 - 26 ○ @height, if not present in **AdaptationSet** element
 - 27 ○ @frameRate, if not present in **AdaptationSet** element
 - 28 ○ @sar

29 Note: @width, @height, and @sar attributes should indicate the vertical and horizontal
30 sample count of encoded and cropped video samples, not the intended display size in pixels.

- 31 • For Adaptation Set or for any Representation within an Video Adaptation Sets as defined
32 in 3.2.13 the attribute @scanType shall either not be present or shall be set to "pro-
33 gressive".
- 34 • For any Audio Adaptation Sets as defined in 3.2.13 the following attributes shall be present
 - 35 ○ @lang
- 36 • For any Representation within an Audio Adaptation Sets as defined in 3.2.13 the following
37 elements and attributes shall be present:

-
- 1 ○ @audioSamplingRate, if not present in **AdaptationSet** element
2 ○ **AudioChannelConfiguration**, if not present in **AdaptationSet** element

3 **3.2.5. MPD Dimension Constraints**

4 No constraints are defined on MPD size, or on the number of elements. However, it should be
5 avoided to create unnecessary large MPDs.

6 Note: DVB DASH [42] adds MPD dimension constraints in section 4.5 of their specification. In order to conform
7 to this specification, it is recommended to obey these constraints.

8 **3.2.6. Generic Metadata**

9 Generic metadata may be added to MPDs based on DASH. For this purpose, the Essential Property
10 Descriptor and the Supplemental Property Descriptor as defined in ISO/IEC 23009-1 [4], clause
11 5.8.4.7 and 5.8.4.8.

12 Metadata identifiers for content properties are provided here: <http://dashif.org/identifiers>.

14 However, it is not expected that DASH-IF clients support all metadata at
15 <http://dashif.org/identifiers> unless explicitly required.

16 **3.2.7. DASH Timing Model**

17 **3.2.7.1. General**

18 According to ISO/IEC 23009-1, DASH defines different timelines. One of the key features in
19 DASH is that encoded versions of different media content components share a common timeline.
20 The presentation time of each access unit within the media content is mapped to the global com-
21 mon presentation timeline for synchronization of different media components and to enable seam-
22 less switching of different coded versions of the same media components. This timeline is referred
23 as Media Presentation timeline. The Media Segments themselves contain accurate Media Presen-
24 tation timing information enabling synchronization of components and seamless switching.

25 A second timeline is used to signal to clients the availability time of Segments at the specified
26 HTTP-URLs. These times are referred to as **Segment availability times** and are provided in wall-
27 clock time. Clients typically compare the wall-clock time to Segment availability times before
28 accessing the Segments at the specified HTTP-URLs in order to avoid erroneous HTTP request
29 responses. For static Media Presentations, the availability times of all Segments are identical. For
30 dynamic Media Presentations, the availability times of segments depend on the position of the
31 Segment in the Media Presentation timeline, i.e. the Segments get available and unavailable over
32 time.

33 Figure 1 provides an overview of the different timelines in DASH and their relation. The diagram
34 shows three Periods, each of the Periods contains multiple Representations (for the discussion it is
35 irrelevant whether these are included in the same Adaptation Set or in different ones).

36 Specifically, the following information is available in the MPD that relates to timing:

- 37 • **MPD@availabilityStartTime**: the start time is the anchor for the MPD in wall-
38 clock time. The value is denoted as *AST*.
39 • **Period@start**: the start time of the Period relative to the MPD availability start time.
40 The value is denoted as *PS*.

- **Representation**@presentationTimeOffset: the presentation time offset of the Representation in the Period, i.e. it provides the media time of the Representation that is supposed to be rendered at the start of the Period. Note that typically this time is either earliest presentation time of the first segment or a value slightly larger in order to ensure synchronization of different media components. If larger, this Representation is presented with short delay with respect to the Period start.

In addition, with the use of the **Representation@duration** or **Representation.SegmentTimeline** the MPD start time for each segment and the MPD duration for each segment can be derived. For details refer to ISO/IEC 23009-1 as well as section 4.

According to Figure 1, the *AST* is a wall-clock time. It provides an anchor to all wall-clock time computation in the MPD. The sum of the **Period@start** of the first Period and the *AST* provides the Period start time *PST1* value in wall-clock time of the first Period.

The media timeline origin for tracks in ISO Media files is defined to be zero. Each Representation is assigned a presentation time offset, either by the value of the attribute **Representation@presentationTimeOffset** or by default set to 0. The value of this attribute is denoted as PTO. It is normally the case that for complete files sequenced as Periods this value is 0. For “partial” files or live streams, @presentationTimeOffset indicates the media composition/presentation time of the samples that synchronize to the start of the Period. @presentationTimeOffset for the live stream will usually not be zero because the encoders are usually started prior to presentation availability, so the media timestamps on the first available Segments will have increased since the encoders started. Encoding timestamps may be set to UTC (as though the encoder was turned on 1/1/1970 at midnight). Representations in Periods of live content typically have the same @presentationTimeOffset as long as the media is continuously encoded, because UTC time and media time increase at the same rate and maintain the same offset.

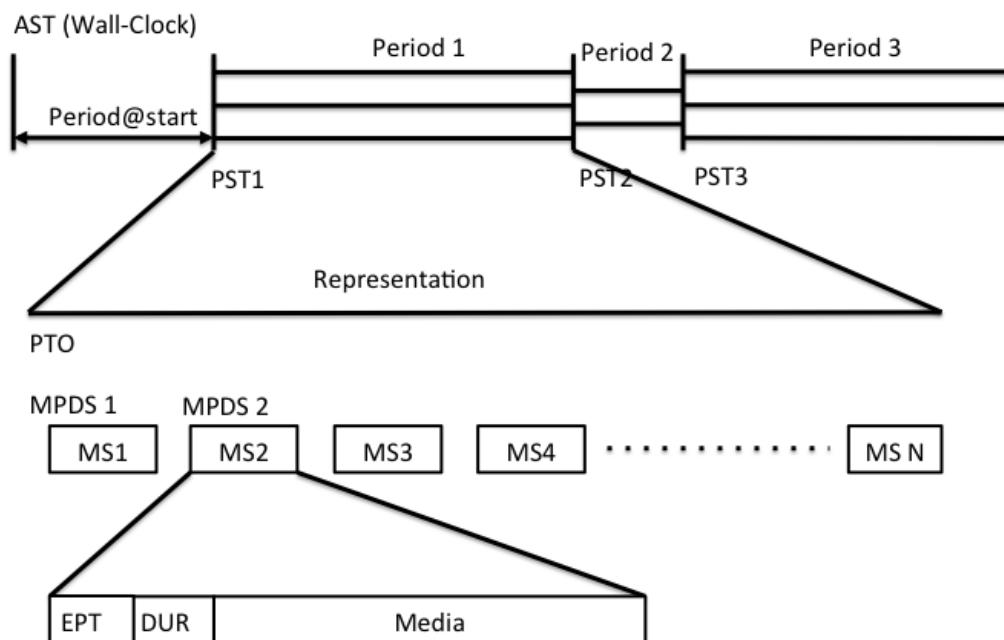


Figure 1 Overview Timing Model

1 Within a Representation, each Segment is assigned an MPD start time and MPD duration according
2 to ISO/IEC 23009-1 (more details for dynamic services are provided in section 4). These two
3 values can be computed from the MPD and provide approximate times for each segment that are
4 in particular useful for random access and seeking.

5 In addition, each segment has an internal sample-accurate presentation time. Therefore, each seg-
6 ment has a media internal earliest presentation time *EPT* and sample-accurate presentation dura-
7 tion *DUR*.

8 For each media segment in each Representation the MPD start time of the segment should approx-
9 imately be *EPT - PTO*. Specifically, the MPD start time shall be in the range of *EPT - PTO -*
10 $0.5*DUR$ and *EPT - PTO + 0.5*DUR* according to the requirement stated above.

11 Each Period is treated independently. Details on processing at Period boundaries are provided in
12 ISO/IEC 23009-1. One example is, that for time code wrap-around a new Period is expected to be
13 added that restarts at presentation time 0.

14 The value of **Period@start** for an ad can be chosen to coincide with an insertion point in the
15 live stream by setting **Period@start** to a presentation time duration equal to the UTC time
16 difference between **@availabilityStartTime** and the scheduled encoding time of the in-
17 sertion point in the live stream.

18 **3.2.7.2. Static Media Presentations**

19 For static media presentations, all Segments shall be available at time *AST*. This means that a
20 DASH client may use the information in the MPD in order to seek to approximate times.

21 **3.2.7.3. Dynamic Media Presentations**

22 For dynamic media presentations, segments get available over time. The latest time they shall be
23 available is at the sum of PST (which is *AST + Period@start*), MPD start time and MPD
24 duration. The latter is added in order to take into account that at the server a segment typically
25 needs to be completed prior to its availability. For details refer to section 4.

26 **3.2.7.4. Usage of XML duration**

27 MPD fields having datatype `xs:duration` shall not use the year and month units.

28 All units expressed in MPD fields of datatype `xs:duration` shall be treated as fixed size:

- 29 - 60S = 1M (minute)
- 30 - 60M = 1H
- 31 - 24H = 1D
- 32 - 30D = 1M (month)
- 33 - 12M = 1Y

34 MPD fields having datatype `xs:duration` should use seconds.

35 **3.2.8. Bandwidth and Minimum Buffer Time**

36 The MPD contains a pair of values for a bandwidth and buffering description, namely the Mini-
37 mum Buffer Time (MBT) expressed by the value of **MPD@minBufferTime** and bandwidth (**BW**)
38 expressed by the value of **Representation@bandwidth**. The following holds:

- 39 • the value of the minimum buffer time does not provide any instructions to the client on
40 how long to buffer the media. The value however describes how much buffer a client

1 should have under ***ideal*** network conditions. As such, MBT is not describing the burstiness
2 or jitter in the network, it is describing the burstiness or jitter in the **content encoding**. To-
3 gether with the BW value, it is a property of the content. Using the "leaky bucket" model,
4 it is the size of the bucket that makes BW true, given the way the content is encoded.

- 5 • The minimum buffer time provides information that for each Stream Access Point (and in
6 the case of DASH-IF therefore each start of the Media Segment), the property of the stream:
7 If the Representation (starting at any segment) is delivered over a constant bitrate channel
8 with bitrate equal to value of the BW attribute then each presentation time PT is available
9 at the client latest at time with a delay of at most $PT + MBT$.
- 10 • In the absence of any other guidance, **the MBT should be set** to the maximum GOP size
11 (coded video sequence) of the content, which quite often is identical **to the maximum**
12 **segment duration** for the live profile or the **maximum subsegment duration** for the On-
13 Demand profile. The MBT may be set to a smaller value than maximum (sub)segment du-
14 ration, but should not be set to a higher value.

15 In a simple and straightforward implementation, a DASH client decides downloading the next
16 segment based on the following status information:

- 17 • the currently available buffer in the media pipeline, $buffer$
- 18 • the currently estimated download rate, $rate$
- 19 • the value of the attribute `@minBufferTime`, MBT
- 20 • the set of values of the `@bandwidth` attribute for each Representation i , $BW[i]$

21 The task of the client is to select a suitable Representation i .

22 The relevant issue is that starting from a SAP on, the DASH client can continue to playout the
23 data. This means that at the current time it does have $buffer$ data in the buffer. Based on this model
24 the client can download a Representation i for which $BW[i] \leq rate * buffer / MBT$ without emptying
25 the buffer.

26 Note that in this model, some idealizations typically do not hold in practice, such as constant bitrate
27 channel, progressive download and playout of Segments, no blocking and congestion of other
28 HTTP requests, etc. Therefore, a DASH client should use these values with care to compensate
29 such practical circumstances; especially variations in download speed, latency, jitter, scheduling
30 of requests of media components, as well as to address other practical circumstances.

31 One example is if the DASH client operates on Segment granularity. As in this case, not only parts
32 of the Segment (i.e., MBT) needs to be downloaded, but the entire Segment, and if the MBT is
33 smaller than the Segment duration, then rather the segment duration needs to be used instead of
34 the MBT for the required buffer size and the download scheduling, i.e. download a Representation
35 i for which $BW[i] \leq rate * buffer / max_segment_duration$.

36 **3.2.9. Trick Mode Support**

37 Trick Modes are used by DASH clients in order to support fast forward, seek, rewind and other
38 operations in which typically the media, especially video, is displayed in a speed other than the
39 normal playout speed. In order to support such operations, it is recommended that the content
40 author adds Representations at lower frame rates in order to support faster playout with the same
41 decoding and rendering capabilities.

1 However, Representations targeted for trick modes are typically not suitable for regular playout.
2 If the content author wants to explicitly signal that a Representation is only suitable for trick mode
3 cases, but not for regular playout, the following is recommended:

- add one or multiple Adaptation Sets that that only contains trick modes Representations
- annotate each Adaptation Set with an **EssentialProperty** descriptor or **SupplementalProperty** descriptor with URL "<http://dashif.org/guidelines/trickmode>" and the @value the value of @id attribute of the Adaptation Set to which these trick mode Representations belong. The trick mode Representations must be time-aligned with the Representations in the main Adaptation Set. The value may also be a white-space separated list of @id values. In this case the trick mode Adaptation Set is associated to all Adaptation Sets with the values of the @id.
- signal the playout capabilities with the attribute @maxPlayoutRate for each Representation in order to indicate the accelerated playout that is enabled by the signaled codec profile and level.
- If the Representation is encoded without any coding dependency on the elementary stream level, i.e. each sample is a SAP type 1, then it is recommended to set the @codingDependency attribute to FALSE.
- If multiple trick mode Adaptation Sets are present for one main Adaptation Set, then sufficient signaling should be provided to differentiate the different trick mode Adaptation Sets. Different Adaptation Sets for example may be provided as thumbnails (low spatial resolution), for fast forward or rewind (no coding dependency with @codingDependency set to false and/or lower frame rates), longer values for @duration to improve download frequencies or different @maxPlayoutRate values. Note also that the @bandwidth value should be carefully documented to support faster than real-time download of Segments.

If an Adaptation Set in annotated with the **EssentialProperty** descriptor with URI "<http://dashif.org/guidelines/trickmode>" then the DASH client shall not select any of the contained Representations for regular playout.

For trick modes for live services, the same annotation should be used. More details on service offerings are provided in section 4.10.

3.2.10. Adaptation Set Constraints

3.2.10.1. Introduction

Content in one Adaptation Set is constrained to enable and simplify switching across different Representations of the same source content. General Adaptation Set constraints allow sequencing of Media Segments from different Representations (“bitrate switching”) prior to a single audio or video decoder, typically requiring the video decoder to be reset to new decoding parameters at the switch point, such as a different encoded resolution or codec profile and level.

Bitstream Switching Adaptation Set constraints allow a switched sequence of Media Segments to be decoded without resetting the decoder at switch points because the resulting Segment stream is a valid track of the source type, so the decoder is not even aware of the switch. In order to signal that the Representations in an Adaptation Set are offered under these constraints, the attribute **Ad-**

1 **aptationSet**@bitstreamSwitching may be set to true. When **Adaptation-**
2 **Set**@bitstreamSwitching attribute is set to TRUE, the decoder can continue decoding with-
3 out re-initialization. When @bitstreamSwitching is set to FALSE, seamless switching
4 across Representations can be achieved with re-initialization of the decoder. Content authors
5 should set **AdaptationSet**@bitstreamSwitching to TRUE only if the content does not
6 need the decoder to be re-initialized.

7 In the following general requirements and recommendations are provided for content in an Adap-
8 tation Set in section 3.2.10.2 and specific requirements when the bitstream switching is set to true
9 in section 3.2.10.3.

10 **3.2.10.2. General**

11 General Adaptation Set constraints require a client to process an Initialization Segment prior to the
12 first Media Segment and prior to each Media Segment selected from a different Representation (a
13 “bitrate switch”).

14 Adaptation Sets shall contain Media Segments compatible with a single decoder that start with
15 SAP type 1 or 2, and in time aligned Representations using the same @timescale, when multiple
16 Representations are present.

17 Edit lists in Initialization Segments intended to synchronize the presentation time of audio and
18 video should be identical for all Representations in an Adaptation Set.

19 Note: Re-initialization of decoders, decryptors, and display processors on some clients during
20 bitrate switches may result in visible or audible artifacts. Other clients may evaluate the differences
21 between Initialization Segments to minimize decoder reconfiguration and maintain seamless
22 presentation equal to the encoded quality.

23 Additional recommendations and constraints may apply for encryption and media coding. For
24 details, please check the relevant sections in this document, in particular section 6.2.5 and 7.7.5.

25 **3.2.10.3. Bitstream Switching**

26 A bitstream switching Adaptation Set is optimized for seamless decoding, and live streams that
27 may change encoding parameters over time. A bitstream switching Adaptation Set may process an
28 Initialization Segment one time from the highest bandwidth Representation in the Adaptation Set,
29 and then process Media Segments from any other Representation in the same Adaptation Set with-
30 out processing another Initialization Segment. The resulting sequence of an Initialization Segment
31 followed by time sequenced Media Segments results in a valid ISO BMFF file with an elementary
32 stream similar to a transport stream.

33 For all Representations within an Adaptation Set with @bitstreamSwitching='true':

- 34 • the Track_ID shall be equal for all Representations
35 • Each movie fragment shall contain one track fragment

36 Note: Multiple Adaptation Sets may be included in an MPD that contain different subsets of the available
37 Representations that are optimized for different decoder and screen limitations. A Representation may be
38 present in more than one Adaptation set, for example a 720p Representation that is present in a 720p Adap-
39 tation Set may also be present in a 1080p Adaptation Set. The 720p Representation uses the same Initializa-
40 tion Segments in each Adaptation Set, but the 1080p Adaptation Set would require decoder and display con-
41 figuration with the 1080p Initialization Segment.

42 Additional recommendation and constraints may apply for encryption and media coding. For de-
43 tails, please see below.

3.2.11. Media Time Information of Segment

The earliest presentation time may be estimated from the MPD using the segment availability start time minus the segment duration announced in the MPD.

The earliest presentation time may be accurately determined from the Segment itself.

If the Segment Index is present than this time is provided in the `earliest_presentation_time` field of the Segment Index. To determine the presentation time in the Period, the value of the attribute `@presentationTimeOffset` needs to be deducted.

If the Segment Index is not present, then the earliest presentation time is deduced from the ISO BMFF parameters, namely the movie fragment header and possibly in combination with the information in the Initialization Segment using the edit list.

The earliest presentation time in the Period for a Segment can be deduced from the decode time taking also into account the composition time offset, edit lists as well as presentation time offsets.

Specifically the following is the case to determine the earliest presentation time assuming that no edit list is present in the Initialization Segment:

- If the SAP type is 1, then the earliest presentation time is identical to the sum of the decode time and the composition offset of the first sample. The decode time of the first sample is determined by the base media decode time of the movie fragment.
- If the SAP type is 2, the first sample may not be the sample with the earliest presentation time. In order to determine the sample with the earliest presentation time, *this sample* is determined as the sample for which the sum of the decode time and the composition offset is the smallest within this Segment. Then the earliest presentation time of the Segment is the sum of the base media decode time and the sum of the decode time and the composition offset for *this sample*. Such an example is shown below.

In addition, if the presentation time needs to be adjusted at the beginning of a period, then the `@presentationTimeOffset` shall be used in order to set the presentation that is mapped to the start of the period. Content authoring shall be such that if edit lists are ignored, then the client can operate without timing and lip sync issues. In the following examples, there is a sequence of I, P, and B frames, each with a decoding time delta of 10. The segmentation, presentation order and storage of the samples is shown in the table below. The samples are stored with the indicated values for their decoding time deltas and composition time offsets (the actual CT and DT are given for reference). The re-ordering occurs because the predicted P frames must be decoded before the bi-directionally predicted B frames. The value of DT for a sample is always the sum of the deltas of the preceding samples. Note that the total of the decoding deltas is the duration of the media in this track.

Example with closed GOP and SAP Type = 1:

Segment	/-	--	--	--	--	--	--	/-	--	--	--	--	--	--\	
	I1	P4	B2	B3	P7	B5	B6	I8	P1	B9	B1	P1	B1	B13	
Presentation Order	==	I1	B2	B3	P4	B5	B6	P7	==	I8	B9	B10	P11	B12	B13
Base media decode time	0								70						

Decode delta	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
DT	0	10	20	30	40	50	60	70	80	90	10	11	12	130		
EPT	10										80					
Composition time offset	10	30	0	0	30	0	0	10	30	0	0	30	0	0		
CT	10	40	20	30	70	50	60	80	11	90	10	14	12	130		

1

2 Example with closed GOP and SAP Type = 2:

3

Segment	/--	--	--	--	--	--\	/-	--	--	--	--	--\				
	I3	P1	P2	P6	B4	B5	I9	P7	P8	P12	B10	B11				
Presentation Order	== P1 P2 I3 B4 B5 P6 == P7 P8 I9 B10 B11 P12 ==															
Base media decode time	0															
Decode Delta	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
DT	0	10	20	30	40	50	60	70	80	90	100	110				
EPT	10										70					
Composition time offset	30	0	0	30	0	0	30	0	0	30	0	0				
CT	30	10	20	60	40	50	90	70	80	120	100	110				

4

5 Example with closed GOP and SAP Type = 2 and negative composition offset:

6

Segment	/--	--	--	--	--	--\	/-	--	--	--	--	--\				
	I3	P1	P2	P6	B4	B5	I9	P7	P8	P12	B10	B11				
Presentation Order	== P1 P2 I3 B4 B5 P6 == P7 P8 I9 B10 B11 P12 ==															
Base media decode time	0															
Decode Delta	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
DT	0	10	20	30	40	50	60	70	80	90	100	110				
EPT	0										60					
Composition offset	20	-10	-10	20	-10	-10	20	-10	-10	20	-10	-10				
CT	20	0	10	50	30	40	80	60	70	110	90	100				

7

1
2 For additional details refer to ISO/IEC 14496-12 [7] and ISO/IEC 23009-1 [1].
3

4 **3.2.12. Content Offering with Periods**

5 Content may be offered with a single Period. If content is offered with a single Period it is suitable
6 to set PSTART to zero, i.e. the initialization segments get available at START on the server.
7 However, other values for PSTART may be chosen.

8 Note: This is aligned with Amd.3 of ISO/IEC 23009-1:2014 [4] and may be referenced in
9 a future version of this document.

10 Content with multiple Periods may be created for different reasons, for example:

- 11 • to enable splicing of content, for example for ad insertion,
12 • to provide synchronization in segment numbering, e.g. compensate non-constant segment
13 durations
14 • to remove or add certain Representations in an Adaptation Set,
15 • to remove or add certain Adaptation Sets,
16 • to remove or add content offering on certain CDNs,
17 • to enable signalling of shorter segments, if produced by the encoder.
18 • for robustness reasons as documented in detail in section 4.8.

19 Periods provide opportunities for resync, for ad insertion, for adding and removing Representations.
20 However, in certain circumstances the content across Period boundaries is continuous and
21 in this case, continuous playout of the client is expected.

22 In certain circumstances the Media Presentation is offered such that the next Period is a continuation
23 of the content in the previous Period, possibly in the immediately following Period or in a
24 later Period (e.g after an advertisement Period had been inserted), in particular that certain media
25 components are continued.

26 The content provider may express that the media components contained in two Adaptation Sets in
27 two different Periods are associated by assigning equivalent Asset Identifiers to both Periods and
28 by identifying both Adaptation Sets with identical value for the attribute @id. Association expresses
29 a logical continuation of the media component in the next Period and may for example be
30 used by the client to continue playing an associated Adaptation Set in the new Period.

31 In addition, two Adaptation Sets in one MPD are period-continuous if all of the following holds:

- 32 • The Adaptation Sets are associated.
33 • The sum of the value of the @presentationTimeOffset and the presentation duration
34 of all Representations in one Adaptation Set are identical to the value of the @presentationTimeOffset
35 of the associated Adaptation Set in the next Period.
36 • If Representations in both Adaptation Sets have the same value for @id, then they shall
37 have functionally equivalent Initialization Segments, i.e. the Initialization Segment may
38 be used to continue the play-out of the Representation. The concatenation of the Initialization
39 Segment of the first Period, if present, and all consecutive Media Segments in the
40 Representation in the first Period and subsequently the concatenation with all consecutive
41 Media Segments in the Representation of the second Period shall represent a conforming

Segment sequence as defined in 4.5.4 conforming to the media type as specified in the @mimeType attribute for the Representation in the first Period. Additionally, the @mimeType attribute for the Representation in the next Period shall be the same as one of the first Period.

Media Presentations should signal period-continuous Adaptation Sets by using a supplemental descriptor on Adaptation Set level with @schemeIdUri set to "urn:mpeg:dash:period-continuity:2015" with

- the @value of the descriptor matching the value of an @id of a Period that is contained in the MPD,
- the value of the **AdaptationSet**@id being the same in both Periods.

MPD should signal period-continuous Adaptation Sets if the MPD contains Periods with identical Asset Identifiers. Period continuity is signaled in the succeeding Period referencing an Adaptation Set in any preceding Period.

There exist special cases, for which the media in one Adaptation Set is a continuation of the previous one, but the timestamps are not continuous. Examples are timestamp wrap around, encoder reset, splicing, or other aspects. Two Adaptation Sets in one MPD are period-connected if all conditions from period-continuity from above hold, except that the timestamps across Period boundaries may be non-continuous, but adjusted by the value of the @presentationTimeOffset at the Period boundary. However, for example the Initialization Segment is equivalent within the two Adaptation Sets. Media Presentations should signal period-connected Adaptation Sets by using a supplemental descriptor on Adaptation Set level with @schemeIdUri set to "urn:mpeg:dash:period-connectivity:2015".

Note that period continuity implies period connectivity.

The content author should use period-continuity signaling or period-connectivity signaling if the content follows the rules. The client should exploit such signals for seamless user experience across Period boundaries.

For details on content offering with multiple Periods, please refer to the requirements and recommendations in section 4 and 5.

3.2.13. Adaptation Set Media Type

In contrast to MPEG-DASH which does not prohibit the use of multiplexed Representations, in the DASH-IF IOPs one Adaptation Set always contains exactly a single media type. The following media types for Adaptation Sets are defined:

- Video Adaptation Set: An Adaptation Set that contains visual information for display to the user. Such an Adaptation Set is identified by @mimeType="video/mp4". For more details on the definition of media type video, refer to RFC 4337 [66]. The DASH-IF IOP restrict the usage of video/mp4 to only @codecs values as defined in this specification.
- Audio Adaptation Set: An Adaptation Set that contains sound information to be rendered to the user. Such an Adaptation Set is identified by @mimeType="audio/mp4". For more details on the definition of media type video, refer to RFC 4337 [66]. The DASH-IF

-
- 1 IOP restrict the usage of audio/mp4 to only @codecs values as defined in this
2 specification.
- 3 - Subtitle Adaptation Set: An Adaptation Set that contains visual overlay information to be
4 rendered as auxiliary or accessibility information. Such an Adaptation Set is identified by
5 @mimeType="application/mp4", a Role descriptor with @schemeIdUri="
6 urn:mpeg:dash:role:2011" and @value="subtitle" and, an @codecs
7 parameter as defined in Table 23 or "application/ttml+xml".
- 8 - Metadata Adaptation Set: An Adaptation Set that contains information that is not
9 expected to be rendered by a specific media handler, but is interpreted by the application.
10 Such an Adaptation Set is identified by @mimeType="application/mp4" and an
11 appropriate sample entry identified by the @codecs parameter.
- 12 The media type is used by the DASH client in order to identify the appropriate handler for
13 rendering. Typically, the DASH client selects at most one Adaptation Set per media type. In
14 addition, the DASH client uses the string included in the @codecs parameter in order to identify
15 if the underlying media playback platform can play the media contained in the Representation.

16 **3.2.14. Seek Preview and Thumbnail Navigation**

17 Seek preview and thumbnail navigation provide DASH clients the possibility to implement thumbnails
18 for UI scrubbing. This may be implemented using a separate video Adaptation Set and using
19 trick mode features as defined in clause 3.2.9. However, this feature may be relatively complex to
20 implement in a player and requires double video decoders. In a simpler approach, a sequence of
21 image tiles may be used, each with multiple thumbnails to provide such thumbnails. An interoperable
22 solution is provided in clause 6.2.6.

23 **3.2.15. MPD Location and Reference Resolution**

24 **3.2.15.1. Signaling and Interoperability**

25 In clause 5.3.1.2 in ISO/IEC 23009-1 [4], the **MPD . Location** element is defined and rules for
26 MPD updates using this element are defined in clause 5.4.1. In clause 5.6.4 of ISO/IEC 23009-1
27 [4], reference resolution is defined, i.e. the usage of MPD URLs as well **BaseUrl** elements on
28 different levels. Clause 5.6.5 in ISO/IEC 23009-1 [4] defines the usage of multiple **BaseUrl**
29 elements.

30 **3.2.15.2. Content Generation and Service Offering guidelines**

31 A content provider may use the **MPD . Location** to redirect the service to a different MPD
32 update server. A content provider may use Base URLs on different levels to easily move and
33 replicate content and offer it for example on multiple CDNs. Redirection may be used to operationally
34 move clients while requesting content.

35 **3.2.15.3. Client Requirements and Recommendations**

36 Based on these descriptions, a DASH client operating on MPD updates and segment requests is
37 expected to operate as follows:

-
- 1 • For MPD update requests (when due) according to clause 5.4, the DASH client requests
2 the MPD according to the following priority
 - 3 • If at least one **MPD . Location** element is present,
 - 4 • any the value of any **MPD . Location** element is used as the MPD request
5 URL
 - 6 • else
 - 7 • If the HTTP request results in an HTTP redirect using a 3xx response code,
8 the redirected URL replaces the original manifest URL,
 - 9 • else
 - 10 • If present and known, the original manifest URL is used for updates
 - 11 • Else updates cannot be done, and client terminates service
 - 12 • In addition, the manifest URL as derived above provides an implicit base URI
 - 13 • Any present **BaseUrl** element does not apply to MPD updates
 - 14 • Based on the knowledge of the MPD request URL according to the previous algorithm, the
15 DASH client requests the MPD according to the following priority
 - 16 • If present, any absolute base URL or an absolute URL is used,
 - 17 • Else
 - 18 • If the base URL is known
 - 19 • the base URL from the above MPD request URL provides the base
20 URL and the relative base URL is used for the segment requests
 - 21 • else
 - 22 • the is service may be terminated.

23 **3.3. Client Implementation Requirements and Guidelines**

24 **3.3.1. Overview**

25 The DASH-related aspects of the interoperability point as defined in section 3.2 can also be un-
26 derstood as permission for DASH clients that only implement the features required by the descrip-
27 tion to process the Media Presentation (MPD document and Segments). The detailed DASH-re-
28 lated client operations are not specified. Therefore, it is also unspecified how a DASH client ex-
29 actly conforms. This document however provides guidelines on what is expected for conformance
30 to this interoperability point. A minimum set of requirements is collected in section 3.3.4.

3.3.2. DASH Client Guidelines

The DASH-related aspects in DASH-IF IOPs as well as for the ISO BMFF based On-Demand and Live profiles of ISO/IEC 23009-1 are designed such that a client implementation can rely on relatively easy processes to provide an adaptive streaming service, namely:

- selection of the appropriate Adaptation Sets based on descriptors and other attributes
- initial selection of one Representation within each adaptation set
- download of (Sub)Segments at the appropriate time
- synchronization of different media components from different Adaptation Sets
- seamless switching of representations within one Adaptation Set

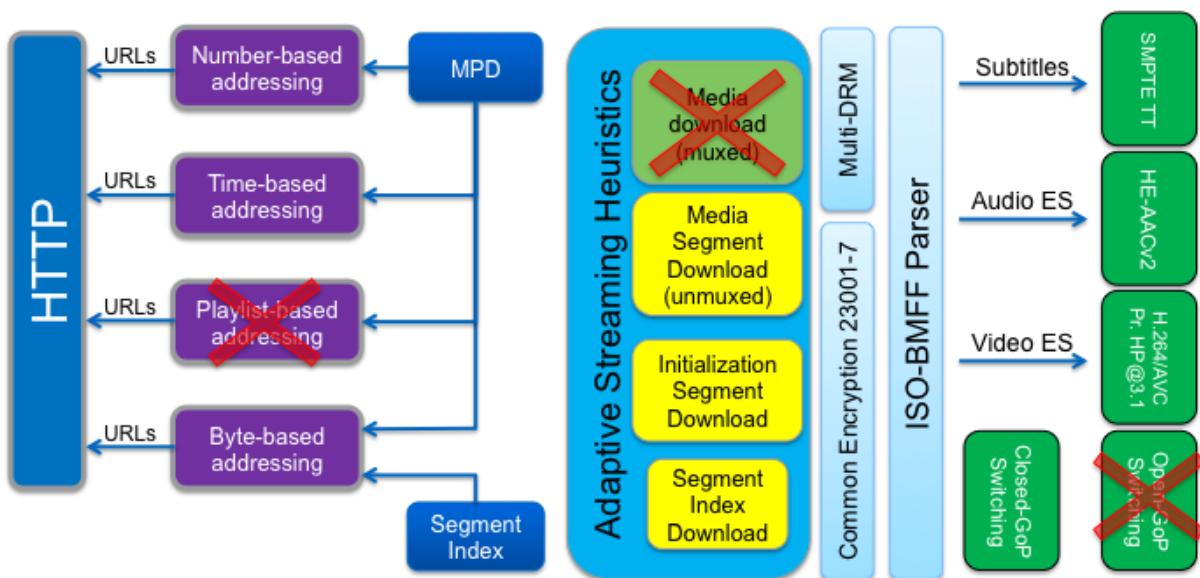


Figure 2 DASH aspects of a DASH-AVC/264 client compared to a client supporting the union of DASH ISO BMFF live and on-demand profile.

Figure 2 shows the DASH aspects of a DASH-AVC/264 client compared to a client supporting all features of the DASH ISO BMFF Live and On-Demand profile. The main supported features are:

- support of HTTP GET and partial GET requests to download Segments and Subsegments
- three different addressing schemes: number and time-based templating as well as byte range based requests.
- support of metadata as provided in the MPD and Segment Index
- download of Media Segments, Initialization Segments and Segment Index
- ISO BMFF parsing
- synchronized presentation of media components from different Adaptation Sets
- switching of video streams at closed GOP boundaries

3.3.3. Seamless switching

The formats defined in section 3.2 are designed for providing good user experience even in case the access bandwidth of the DASH Segment delivery or the cache varies. A key functionality is

the ability that the DASH client can seamlessly switch across different Representations of the same media component. DASH clients should use the common timeline across different Representation representing the same media component to present one Representation up to a certain time t and continue presentation of another Representation from time t onwards. However, in practical implementations, this operation may be complex, as switching at time t may require parallel download and decoding of two Representations. Therefore, providing suitable switching opportunities in regular time intervals simplifies client implementations.

The formats defined in section 3.2 provide suitable switching opportunities at (sub)segment boundaries.

3.3.4. DASH Client Requirements

In order to ensure a minimum level of interoperability, a DASH-IF conforming client shall at least support the following features:

- The DASH client, if it switches, shall provide a seamless experience. A DASH shall be able to switch seamlessly at (sub)segment boundaries according to the definition in ISO/IEC 23009-1 [4], clause 4.5.1.
- If the scheme or the value for the following descriptor elements are not recognized and no equivalent other descriptor is present, the DASH client shall ignore the parent element:
 - **FramePacking**
 - **Rating**
 - **EssentialDescriptor**
 - **ContentProtection**

3.4. Transport and Protocol-Related Issues

3.4.1. General

Servers and clients operating in the context of the interoperability points defined in this document shall support the normative parts of HTTP/1.1 as defined in RFC 7230 [21], RFC 7231 [22], RFC 7232 [23], RFC 7233 [24], and RFC 7234 [25].

Specific requirements and recommendations are provided below.

Note: IETF recently obsoleted RFC 2616 and replaced it with the six RFCs referred above. The changes are generally text clarifications and in some cases, additional constraints to address security or interoperability issues. Each new RFC contains details of the changes compared to RFC2616. The IETF strongly recommends to reference and use the new RFCs that collectively replace RFC2616. This version of DASH-IF IOP addresses this aspect.

MPEG-DASH explicitly permits the use of https as a scheme and hence, HTTP over TLS as a transport protocol as defined in RFC 5246 [65]. For more details refer to section 7.2.

3.4.2. Server Requirements and Guidelines

HTTP Servers serving segments should support suitable responses to byte range requests (partial GETs).

1 If an MPD is offered that contains Representations conforming to the ISO BMFF On-Demand
2 profile, then the HTTP servers offering these Representations shall support suitable responses to
3 byte range requests (partial GETs).

4 HTTP Servers may also support the syntax using Annex E of 23009-1 using the syntax of the
5 second example in Annex E.3,

6 **BaseUrl@byteRange="\$base\$\$query\$&range=\$first\$-\$last\$"**

7 **3.4.3. Client Requirements and Guidelines**

8 Clients shall support byte range requests, i.e. issue partial GETs to subsegments as defined in RFC
9 7233 [24]. Range requests may also be issued by using Annex E of 23009-1 using the syntax of
10 the second example in Annex E.3,

11 **BaseUrl@byteRange="\$base\$\$query\$&range=\$first\$-\$last\$"**

12 Clients shall follow the reaction to HTTP status and error codes as defined in section A.7 of
13 ISO/IEC 23009-1.

14 Clients should support the normative aspects of the HTTP state management mechanisms (also
15 known as Cookies) as defined in RFC 6265 [41] for first-party cookies.

16 **3.4.4. Transforming Proxies and Other Adaptation Middleboxes**

17 A number of video transcoding proxies (aka "middleboxes") are already deployed on the wider
18 Internet may silently transcode Representations. Specifically: a middlebox may see a video/mp4
19 response, transcode that video into a different format (perhaps using a lower bitrate or a different
20 codec), then forward the transcoded video to the DASH client. This will break MPD and/or Seg-
21 ment Index based byte range operations, as those ranges are generally not valid in the transcoded
22 video.

23 If such a threat is possible, one of the following approaches may be considered in order to prevent
24 proxies from transcoding DASH Representations:

- 25 1. serve Media Presentations using encryption (e.g., HTTP over TLS, segment encryption or
26 content protection),
- 27 2. serve Representations with Cache-Control: "no-transform"

28 In all cases the operational impacts on caching and implementations should be considered when
29 using any of the above technologies.

30 In order to prevent middleboxes to manipulate the MPD, e.g. removing certain Representations or
31 Adaptation Sets, the MPD may be securely transported by appropriate means, e.g. HTTPS.

32 **3.5. Synchronization Considerations**

33 In order to properly access MPDs and Segments that are available on DASH servers, DASH serv-
34 ers and clients should synchronize their clocks to a globally accurate time standard. Specifically it
35 is expected that the Segment Availability Times as computed from the MPD according to ISO/IEC
36 23009-1 [4], section 5.3.9.5 and additional details in ISO/IEC 23009-3 [6], section 6.4 are accu-
37 rately announced in the MPD.

38 Options to obtain timing for a DASH client are for example:

- 39 • Usage of NTP or SNTP as defined in RFC5905 [40].

-
- 1 • The Date general-header field in the HTTP header (see RFC 7231 [22] section 7.1.1.2)
2 represents the date and time at which the message was originated, and may be used as an
3 indication of the actual time.

4 Anticipated inaccuracy of the timing source should be taken into account when requesting seg-
5 ments close to their segment availability time boundaries.

6 More details on advanced synchronization support is provided in section 4.7.

7 **3.6. Considerations for Different Services Types**

8 For interoperability aspects of live services, please refer to clause 4.

9 For interoperability aspects of on-demand services, please refer to clause 3.10.

10 **3.7. Considerations on Ad Insertion**

11 For interoperability aspects for ad insertion use cases, please refer to section 5.

12 **3.8. Switching across Adaptation Sets**

13 Note: This technology is expected to be available in ISO/IEC 23009-1:2014/Amd.4:2016 [4], sec-
14 tion 5.3.3.5. Once published by MPEG, this section is expected to be replaced by a reference to
15 the MPEG-DASH standard.

16 Representations in two or more Adaptation Sets may provide the same content. In addition, the
17 content may be time-aligned and may be offered such that seamless switching across Representa-
18 tions in different Adaptation Sets is possible. Typical examples are the offering of the same content
19 with different codecs, for example H.264/AVC and H.265/HEVC and the content author wants to
20 provide such information to the receiver in order to seamlessly switch Representations (as defined
21 in ISO/IEC 23009-1, clause 4.5.1) across different Adaptation Sets. Such switching permission
22 may be used by advanced clients.

23 A content author may signal such seamless switching property across Adaptation Sets by providing
24 a Supplemental Descriptor along with an Adaptation Set with @schemeIdUri set to
25 urn:mpeg:dash:adaptation-set-switching:2016 and the @value is a comma-
26 separated list of Adaptation Set IDs that may be seamlessly switched to from this Adaptation Set.

27 If the content author signals the ability of Adaptation Set switching and as @segmentAlign-
28 ment or @subsegmentAlignment are set to TRUE for one Adaptation Set, the (Sub)Segment
29 alignment shall hold for *all* Representations in *all* Adaptation Sets for which the @id value is
30 included in the @value attribute of the Supplemental descriptor.

31 As an example, a content author may signal that seamless switching across an H.264/AVC Adapt-
32 ation Set with **AdaptationSet**@id="264" and an HEVC Adaptation Set with **Adapta-**
33 **tionSet**@id="265" is possible by adding a Supplemental Descriptor to the H.264/AVC Adaptation
34 Set with @schemeIdUri set to urn:mpeg:dash:adaptation-set-switch-
35 ing:2016 and the @value="265" and by adding a Supplemental Descriptor to the HEVC
36 Adaptation Set with @schemeIdUri set to urn:mpeg:dash:adaptation-set-
37 switching:2016 and the @value="264".

38 In addition, if the content author signals the ability of Adaptation Set switching for

-
- 1 - any Video Adaptation Set as defined in 3.2.13 then the parameters as defined in section
2 3.2.4 for an Adaption Set shall also hold for all Adaptation Sets that are included in the
3 @value attribute.
4 - any Audio Adaptation Set as defined in 3.2.13 then the parameters as defined in section
5 3.2.4 for an Adaption Set shall also hold for all Adaptation Sets that are included in the
6 @value attribute.

7 Note that this constraint may result that the switching may only be signaled with one Adaptation
8 Set, but not with both as for example one Adaptation Set signaling may include all spatial resolu-
9 tions of another one, whereas it is not the case the other way round.

10 **3.9. Annotation and Client Model for Content Selection**

11 **3.9.1. Introduction**

12 Beyond the ability to provide multiple Representations of the same media component in one Ad-
13 aptation Set, DASH MPDs also provide the functionality to annotate Adaptation Sets, such that
14 clients can typically select at most one Adaptation Set for each media type, based on the encoding
15 and description provided in the MPD. The selection is based on client capabilities, client prefer-
16 ences, user preferences and possibly also interactive signalling with the user. Typically, the sig-
17 nalling and selection is independent of the codec in use. This clause provides requirements and
18 recommendations for labelling Adaptation Sets, if multiple tracks are offered. Note that there may
19 be cases that multiple Representations from different Adaptation Sets per media type are chosen
20 for playback, for example if there is a dependency across Representations. In other cases, a DASH
21 client may be asked to select more than one Adaptation Set per media type based on application
22 decisions.

23 Multiple Adaptation Sets may be offered to provide the same content in different encodings, for
24 example different codecs; or different source formats, for example one Adaptation Set encoded
25 from a standard dynamic range master and another encoded from a high dynamic range video
26 master. Alternatively, Adaptation Sets may describe different content, for example different lan-
27 guages, or different camera views of the same event that are provided in a synchronized presenta-
28 tion in one MPD.

29 Proper labelling of Adaptation Sets in MPDs conforming to DASH-IF IOPs is essential in order to
30 enable consistent client implementations. In addition, also a model is needed on how the client
31 makes use of the annotation for a content authors to understand the expected effect of the labelling
32 on playback.

33 **3.9.2. Adaptation Set Labeling Options for Selection**

34 DASH in ISO/IEC 23009-1 [4] provides many options for labelling Adaptation Sets. In order to
35 provide more consistency in the context of DASH-IF, Table 4 provides a restricted subset of labels
36 for which DASH-IF IOPS provide interoperability, i.e. on how they are expected to be used by the
37 content authors and how they are expected to be used by clients. The table provides information
38 specific for each media type.

39 It is expected that DASH clients following the DASH IOPs recognize the descriptors, elements,
40 and attributes as documented in Table 4.

1 Other organizations may define additional descriptors or elements, as well as processing models
2 for clients.

3 **Table 4 Adaptation Set Attributes and Elements and Usage in DASH-IF IOPs (see ISO/IEC 23009-1 [4])**

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
General Attributes and Elements for any media type		
@profiles	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>If not present, it is inherited from the MPD or Period. This may be used for example to signal extensions for new media profiles in the MPD.</p> <p>At least one of the values defined in Table 1 and Table 2 of this document shall be present, or inferred from MPD or Period higher-level.</p>
@group	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>The attribute may be used and shall be different at least for different media type.</p> <p>If present, the value shall be greater than 0. For all Adaptation Sets in the same group, the Group shall be the same.</p> <p>Only one Representation in a Group is intended to be presented at a time. However, two or multiple groups of the same media type may exist, if the content author expects simultaneous presentation of two or more Representation of the same media type.</p>
@selectionPriority	OD default=1	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>This attribute should be used to dis-ambiguate Adaptation Sets within one group for selection and expresses the preference of the MPD author on selecting Adaptation Sets for which the DASH client does make a decision otherwise. Examples include two video codecs providing the same content, but one of the two provides higher compression efficiency and is therefore preferred by the MPD author.</p>
ContentProtection	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>If this element is present, then the content is protected.</p> <p>If not present, no content protection is applied.</p> <p>For details and usage please refer to clause 7.</p>

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
EssentialProperty	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9</p> <p>specifies information about the containing element that is considered essential by the Media Presentation author for processing the containing element.</p> <p>The following schemes are expected to be recognized by a DASH-IF client independent of the media type:</p> <ul style="list-style-type: none"> • http://dashif.org/guidelines/trick-mode (see clause 3.2.9)
SupplementalProperty		<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9</p> <p>specifies information about the containing element that is considered supplemental by the Media Presentation author for processing the containing element. In no case this information is used for differentiation, the information may be used by a DASH client for improved operation.</p> <p>The following schemes are expected to be recognized by a DASH-IF client independent of the media type:</p> <ul style="list-style-type: none"> - urn:mpeg:dash:adaptation-set-switching:2016 (see clause 3.8) - http://dashif.org/guidelines/trick-mode (see clause 3.2.9) - urn:mpeg:dash:period-continuity:2015 (see clause 3.2.12) - urn:mpeg:dash:period-connectivity:2015 (see clause 3.2.12)
Viewpoint	0 ... N	<p>Provides the ability to indicate that media differentiates by a different ViewPoint.</p> <p>If not present, no view point is assigned and no differentiation is taken.</p> <p>For detailed usage of this descriptor, see below.</p>
Label	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>This element enables to provide a textual description of the content. This element should be used if content author expects that clients supports UI for selection. However, this element must not be used as the sole differentiating element as at start-up no user interaction is available.</p>
Attributes and Elements for media type “Video”		
@mimeType	M	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>Shall be set to "video/mp4".</p>
@codecs	M	See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
		<p>This provides the codec that is used for the Adaptation Set. It expresses the codec that is necessary to playback all Representations in one Adaption Set.</p> <p>The following codecs are expected to be recognized by a DASH-IF client:</p> <ul style="list-style-type: none"> • Codecs in Table 19 • Codecs in Table 21
@par	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>shall be present, if the display aspect ratio is a differentiating parameter in the MPD.</p>
@maxWidth	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>This attribute should be present to express the maximum width in samples after decoder sample cropping of any Representation contained in the Adaptation Set.</p> <p>The value should be the maximum horizontal sample count of any SPS in the contained bitstream.</p>
@maxHeight	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>This attribute should be present to express the maximum height in pixel of any Representation contained in the Adaptation Set.</p> <p>The value should be the maximum horizontal sample count of any SPS in the contained bitstream.</p>
@maxFrameRate	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>This attribute should be present to express the maximum frame rate, i.e. the maximum value of any entry in the Decoder configuration record of the signaled frame rate, if constant frame rate is provided. contained in the Adaptation Set.</p>
@scanType	OD Default: progressive	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>This value is expected to be not present. If present, it is expected to be set to "progressive".</p>
EssentialProperty	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>specifies information about the containing element that is considered essential by the Media Presentation author for processing the containing element.</p> <p>The following schemes are expected to be recognized by a DASH-IF client for video:</p> <ul style="list-style-type: none"> • urn:mpeg:mpegB:cicp:<Parameter> as defined in ISO/IEC 23001-8 [49] and <Parameter> one of

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
		the following: ColourPrimaries, TransferCharacteristics, or MatrixCoefficients
Accessibility	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>In DASH-IF IOPs two schemes for accessibility are defined.</p> <ul style="list-style-type: none"> - the Role scheme as defined by MPEG-DASH should be used as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 <p>The DASH role scheme and with the following values is expected to be recognized by a DASH-IF client for media type “video” together with the Accessibility descriptor:</p> <ul style="list-style-type: none"> o sign o captions <ul style="list-style-type: none"> - the scheme when CEA-608 is used as defined in clause 6.4.3.3, with @schemeIdUri set to "urn:scte:dash:cc:cea-608:2015"
Role	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>In DASH-IF IOPs only the Role scheme as defined by MPEG-DASH should be used as defined in ISO/IEC 23009-1 [4], 5.8.5.5, urn:mpeg:dash:role:2011</p> <p>The DASH role scheme and with the following values is expected to be recognized by a DASH-IF client for media type “video” together with the Role descriptor:</p> <ul style="list-style-type: none"> - caption - subtitle - main - alternate - supplementary - sign - emergency <p>If not present, the role is assumed to be main</p>
Rating	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>DASH-IF IOPs do not define a Rating scheme. If present, Adaptation Sets using this descriptor may be ignored by the DASH-IF IOP clients.</p>
Attributes and Elements for media type “Audio”		
@mimeType	M	See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
		Shall set to "audio/mp4".
@codecs	M	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>This provides the codec that is used for the Adaptation Set. It expresses the codec that is necessary to playback all Representations in one Adaption Set.</p> <p>The following codecs are expected to be recognized by a DASH-IF client:</p> <ul style="list-style-type: none"> • Codecs in Table 22 • Codecs in Table 25 • Codecs in Table 26 • Codecs in Table 27 • Codecs in Table 28 • Codecs in Table 29 <p>Note: additional values need to be added with new codecs being added</p>
@lang	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>The language should be present.</p> <p>If not present, the language is unknown or no language applies.</p>
@audioSamplingRate	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>This attribute may be present to support output devices that may only be able to render specific values.</p>
AudioChannelConfiguration	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>specifies information about the Audio channel configuration.The following schemes are expected to be recognized by a DASH-IF client for audio:</p> <ul style="list-style-type: none"> • urn:mpeg:dash:23003:3:audio_channel_configuration:2011 as defined in ISO/IEC 23009-1 [4], 5.8.5.4 • urn:mpeg:mpegB:cicp:ChannelConfiguration as defined in ISO/IEC 23001-8 [49] • tag:dolby.com,2014:dash:audio_channel_configuration:2011 as defined at http://dashif.org/identifiers/audio-source-data/ <p>Note: Annotation may be different for other codecs and may be updated</p>
EssentialProperty	0 ... N	See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
		<p>specifies information about the containing element that is considered essential by the Media Presentation author for processing the containing element.</p> <p>The following schemes are expected to be recognized by a DASH-IF client for audio:</p> <ul style="list-style-type: none"> • urn:mpeg:dash:audio-receiver-mix:2014 as defined in ISO/IEC 23009-1, clause 5.8.5.7.
Accessibility	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>In DASH-IF IOPs only the Role scheme as defined by MPEG-DASH should be used as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011</p> <p>The DASH role scheme and with the following values is expected to be recognized by a DASH-IF client for media type “audio” together with the Accessibility descriptor:</p> <ul style="list-style-type: none"> • description • enhanced-audio-intelligibility
Role	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>In DASH-IF IOPs only the Role scheme as defined by MPEG-DASH should be used as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011</p> <p>The DASH role scheme and with the following values is expected to be recognized by a DASH-IF client for media type “audio” together with the Accessibility descriptor:</p> <ul style="list-style-type: none"> - main - alternate - supplementary - commentary - dub - emergency <p>If not present, the role is assumed to be main</p>
Rating	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>DASH-IF IOPs do not define a Rating scheme. If present, Adaptation Sets using this descriptor may be ignored by the DASH-IF IOP clients.</p>
Attributes and Elements for media type “Subtitle”		
@mimeType	M	See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
		Shall set to "application/mp4" or "application/ttml+xml"
@codecs	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.7.2 Table 9.</p> <p>This provides the codec that is used for the Adaptation Set. It expresses the codec that is necessary to playback all Representations in one Adaption Set.</p> <p>The following codecs are expected to be recognized by a DASH-IF client:</p> <ul style="list-style-type: none"> • Codecs in Table 23 <p>Note: more need to be added with new codecs being added.</p>
@lang	O	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>The language should be present.</p> <p>If not present, the language is unknown or no language applies.</p>
Accessibility	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>In DASH-IF IOPs only the Role scheme as defined by MPEG-DASH should be used as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011</p> <p>The DASH role scheme and with the following values is expected to be recognized by a DASH-IF client for media type "subtitle" together with the Accessibility descriptor:</p> <ul style="list-style-type: none"> - caption - sign
Role	0 ... N	<p>See ISO/IEC 23009-1 [4], clause 5.3.3.2 Table 5.</p> <p>In DASH-IF IOPs only the Role scheme as defined by MPEG-DASH should be used as defined in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011</p> <p>The DASH role scheme and with the following values is expected to be recognized by a DASH-IF client for media type "subtitle" together with the Accessibility descriptor:</p> <ul style="list-style-type: none"> - main - alternate - subtitle - supplementary - commentary - dub - description

Attribute or Element	Use for media type	Detailed Usage in DASH-IF IOPs
		<ul style="list-style-type: none"> - emergency <p>If not present, the role is assumed to be main</p>

3.9.3. Content Model

In order to support the content author in providing content in a consistent manner, Figure 1 provides a conceptual content model for DASH content in one Period of an MPD. The content may be described by an Asset Identifier as a whole and may contain different media types, video, audio, subtitle and application types. Signalling of media types is out of scope for this section, for details refer to section 3.2.12.

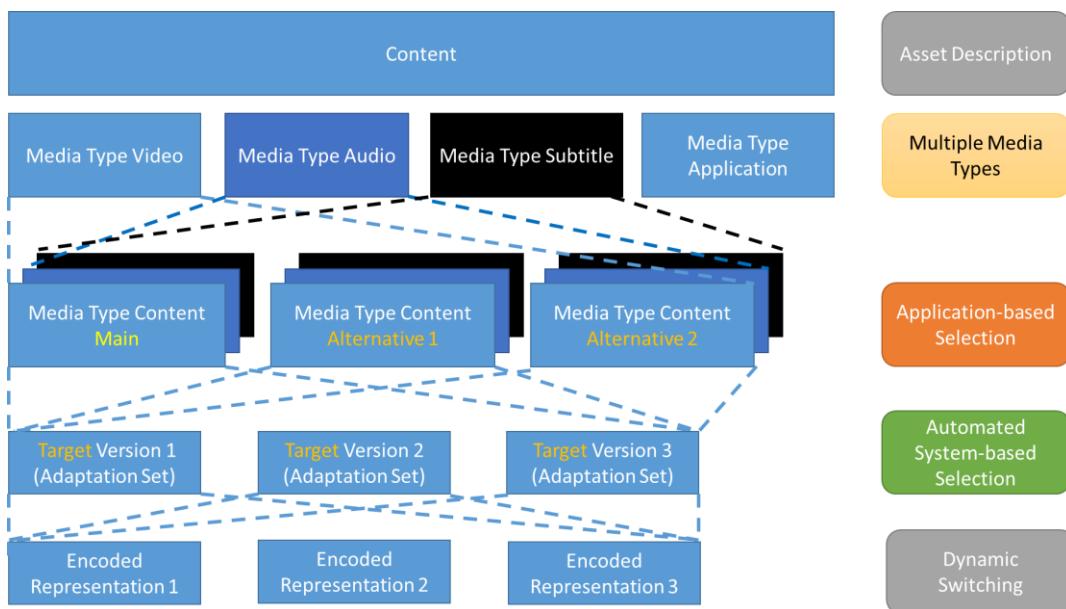


Figure 3 Content Model for DASH Multitrack

Within each media type, the content author may want to offer different alternative content that are time-aligned, but each alternative represents different content. Automatic selection of the alternative content is not expected to be done by the DASH client as the client would not have sufficient information to make such decisions. However, the selection is expected to be done by communication with an application or the user, typically using a user interface appropriate for selection.

However, in the absence of this external communication, or at startup, the DASH client still needs to playback content and therefore benefits from information of what is the default content. Such signalling should be provided by the content author. Such default content is referred to as main content, whereas any content that is not main is referred to as alternative. There may be multiple alternatives which may need to be distinguished. We define *main* and *alternative* content. Examples for such are synchronized camera views of one master content. The main camera view is provided as main content, all other views as alternative content.

1 Furthermore, it may be that content of different media type is linked by the content author, to
2 express that two content of different media type are preferably played together. We define *associated*
3 content for this purpose. As an example, there may be a main commentator associated with
4 the main camera view, but for a different camera view, a different associated commentary is pro-
5 vided.

6 In addition to semantical content level differentiation, each alternative content may be prepared
7 with different target versions, based on content preparation properties (downmix, subsampling,
8 translation, suitable for trick mode, etc.), client preferences (decoding or rendering preferences,
9 e.g. codec), client capabilities (DASH profile support, decoding capabilities, rendering capabili-
10 ties) or user preferences (accessibility, language, etc.). In simple AV playout and in the absence of
11 guidance from an application, a content author expects that the DASH client selects at most one
12 target version for each Group taking into account its capabilities and preferences and the capabili-
13 ties and preferences of the media subsystem. However, an application may obviously select mul-
14 tiple Groups and playout different video Adaptation Sets to support for example picture-in-picture,
15 multi-angle and so on.

16 In addition, the content author may also provide priorities for target versions, if the receivers sup-
17 port multiple of those. Typical examples are that the content is prepared for H.264/AVC and
18 H.265/HEVC capable receivers, and the content author prefers the selection of the H.265/HEVC
19 version as its distribution is more efficient. A device supporting both decoders may then choose
20 the one with higher priority signalled by the content author. In a similar version, the same content
21 may be provided in different languages. In this case, it can still be expected that the language can
22 be automatically selected by the client, so it is assigned to a target version. Again, a content author
23 may express priorities on languages, for example preferring the native language over a dubbed
24 one. Languages may be considered as alternative content as well, but as long as automatic selection
25 can be provided, it may be considered as different target versions. Hence for each content of one
26 media type, different *target versions* may exist and the annotation of the content expressed that it
27 is expected that automated selection can be done. Each target version is preferably accumulated in
28 one Adaptation Set, with exceptions such as scalable codecs.

29 Finally, in the content model, each of the target version typically has multiple Representations that
30 are prepared to enable dynamic switching. This aspect is outside the scope of this section as switch-
31 ing by the client is expected to be done independent of the media type as well as the target version,
32 primarily using the bandwidth and possibly abstract quality information. However, the signalling
33 on the target versions may provide information on how to distribute the available bitrate across
34 different media types.

35 Based on this content model and the available elements, attributes and descriptors from Table 4,
36 requirements and recommendations are provided for Adaptation Set Signalling to address main
37 and alternative content, associated content as well as different target versions. Based on the sig-
38 nalling, a client decision model is developed that may serve a content provider as a reference client
39 to test if the annotation provided in the MPD provides the proper results.

1 **3.9.4. Signalling Requirements and Recommendations**

2 **3.9.4.1. General**

3 Assuming the content author can map its content to the above content model, this section provides
4 signalling requirements and recommendations for such content, such that the content author can
5 expect proper playback of its content for DASH-IF IOP clients

6 In general, if multiple Adaptation Sets for one media types are provided, sufficient information
7 should be provided such that a DASH client make proper selections, either automatically com-
8 municating with its platform or in communication with the application/user.

9 **3.9.4.2. Alternative Content Signalling**

10 If a Period contains alternative content for one media type, then the alternatives should be differ-
11 entiated. In addition, one of the alternatives should be provided as main content. The main content
12 is intended to be selected by the client in the absence of any other information, e.g. at startup or if
13 the annotation of the content cannot be used.

14 *Main content* is signaled by using the Role descriptor with Role scheme as defined by MPEG-
15 DASH in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 with value set to main.
16 *Alternative content* is signaled by using the Role descriptor with Role scheme as defined by
17 MPEG-DASH in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 with value set to
18 alternative. If an Adaptation Set does not include either of the two signals, it is assumed to
19 be *main content*.

20 The alternative content may be selected by the client, if the client does have the capability to select
21 alternatives, typically by either communicating with the application or with the user. If main and
22 alternative content is provided in the Media Presentation, then alternative content shall be signaled
23 by at least one of the two:

- 24 - a **ViewPoint** descriptor. If **ViewPoint** is used for differentiation, then at least each
25 alternative Adaptation Set of the same media type shall include a **ViewPoint** with the
26 same value for @schemeIdUri. The content is differentiated by different values for
27 the @value attribute in the descriptor for different content.
- 28 - a **Label** element. If **Label** is used for differentiation, then at least each alternative Ad-
29 aptation Set shall include a **Label** with the same value for @id. The content is differ-
30 entiated by different values for the **Label** element.

31 A **ViewPoint** descriptor is typically used if a target application (identified by the the value for
32 @schemeIdUri) is expected that can make use of the values in the ViewPoint descriptor. A
33 **Label** element is typically used if the DASH client can provide a user interaction.

34 **3.9.4.3. Associated Content Signalling**

35 For associated content of different media types, the **ViewPoint** descriptor is used. If different
36 media types all belong to one alternative content, they share the same View Point descriptor, i.e.
37 the same value for @schemeIdUri and for @value. Note also that even if the DASH client
38 does not understand the value for @schemeIdUri it would still obey the rules for associated

selection. The DASH client may for example use the labels of different video alternatives for selection, and play the audio according to ViewPoint association.

3.9.4.4. Media-type Independent Target Version Annotation

Adaptation Sets within one media type and alternative content shall differ by at least by one of the following annotation labels

- @profiles,
- **ContentProtection** (need to provide some details on what the options are: present, not-present, different schemes) ➔ work with content protection task force
- **EssentialProperty** (not-present, trickmode, a media type specific value, unknown value, which may be extended)
- Any of those documented in section 3.10.4.5 for media type video, section 3.10.4.6 for media type audio and 3.10.4.7 for media type subtitle.

Adaptation Sets with elements **EssentialProperty** not using any of the permitted values in this document should not be present.

In addition, Adaptation Sets within one media type and alternative content should differ by different values of @selectionPriority. If not present or non-differentiating values are provided, then the content author should expect a random selection of Adaptation Sets in case it is able to handle multiple Adaptation Sets within one media type and alternative content.

3.9.4.5. Video Target Version Annotation

Video Adaptation Sets of one alternative content shall differ by at least by one of the following annotation labels:

- @codecs: specifies the codecs present within the Representation. The codecs parameters shall also include the profile and level information where applicable.
- @maxWidth and @maxHeight specifies the horizontal and vertical visual presentation size of the video media type
- @maxFrameRate specifies the maximum frame rate of the video media type
- **EssentialProperty**: specifies information about the containing element that is considered essential by the Media Presentation author selecting this component.
 - The following different options exist: not-present; generic parameters from above; list in Table 1; unknown value, which may be extended
- Accessibility descriptor with
 - Role scheme as defined by MPEG-DASH in ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 with value set to sign, caption or subtitle. The presence of caption or subtitle signals open (“burned in”) captions or subtitles.

- 1 • the scheme when CEA-608 is used as defined in clause 6.4.3.3, with
2 @schemeIdUri set to "urn:scte:dash:cc:cea-608:2015" indicating
3 the use of CEA-608 captions carried in SEI messages.

4 Adaptation Sets with elements **Rating** and **FramePacking** as well with @scanType not set
5 to progressive should not be present.

6 The content author should use the @selectionPriority attribute in order to express prefer-
7 ence for video selection. If captions are burned in with video Adaptation Set, and other video
8 Adaptation Sets are available as well, the content author should use the @selectionPriority
9 to indicate the selection priority of this Adaptation Set compared to others without burned in cap-
10 tions.

11 **3.9.4.6. Audio Target Version Annotation**

12 Audio Adaptation Sets of one alternative content shall differ by at least by one of the following
13 annotation labels:

- 14 • @codecs: specifies the codecs present within the Representation. The codecs parameters
15 shall also include the profile and level information where applicable.
- 16 • @lang: specifies the dominant language of the audio
 - 17 • If not present, the language is unknown or no language applies
- 18 • @audioSamplingRate specifies the maximum sampling rate of the content
 - 19 • If not present, the audio sampling rate is unknown
- 20 • The **AudioChannelConfiguration** specifies support for output devices that may
21 only be able to render specific values. This element should be present.
 - 22 • If no **AudioChannelConfiguration** is present, then this value is unknown.
 - 23 • If the codec is anyone in Table 22, Table 27, Table 28 or Table 29, then any of the
24 following may be used
 - 25 • urn:mpeg:dash:23003:3:audio_channel_configura-
26 tion:2011 as defined in ISO/IEC 23009-1 [1], 5.8.5.4
 - 27 • urn:mpeg:mpegB:cicp:ChannelConfiguration as defined in
28 ISO/IEC 23001-8 [49]
 - 29 • If the codec is ec-3 or ac-4 according to Table 25, then the following shall be used
 - 30 • tag:dolby.com,2014:dash:audio_channel_configura-
31 tion:2011" as defined at <http://dashif.org/identifiers/audio-source-data/>
32 (see section 9.2.1.2)
 - 33 • If the codec is anyone in Table 26, then refer to DTS specification 9302K62400
34 [39]

-
- 1 • **EssentialProperty**: specifies information about the containing element that is con-
2 sidered essential by the Media Presentation author selecting this component.
3 • The following different options exist: not-present; generic parameters from above;
4 unknown value, which may be extended
5 • Accessibility descriptor with Role scheme as defined by MPEG-DASH in ISO/IEC
6 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 with value set to `description`
7 or `enhanced-audio-intelligibility`.

8 Note that Adaptation Sets with element **Rating** may be ignored by the client and should therefore
9 only be used if the content provider has knowledge that clients can process the applied Rating
10 scheme.

11 **3.9.4.7. Subtitle Target Version Annotation**

12 Subtitle Adaptation Sets of one alternative content shall differ by at least by one of the following
13 annotation labels:

- 14 • `@codecs`: specifies the codecs present within the Representation. The codecs parameters
15 shall also include the profile and level information where applicable.
16 • `@lang`: specifies the language of the subtitle
17 • If not present, the language is unknown or no language applies
18 • **EssentialProperty**: specifies information about the containing element that is con-
19 sidered essential by the Media Presentation author selecting this component.
20 • The following different options exist: not-present; generic parameters from above;
21 unknown value, which may be extended
22 • Accessibility descriptor with Role scheme as defined by MPEG-DASH in ISO/IEC
23 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 with value set to `description`
24 or `caption`.

25 **3.9.4.8. Other Annotation, Auxiliary Data**

26 In addition to selection relevant data, the Adaptation Set may also signal additional auxiliary in-
27 formation. Auxiliary information is expressed by

- 28 - The Role descriptor with the Role scheme as defined by MPEG-DASH as defined in
29 ISO/IEC 23009-1, 5.8.5.5, urn:mpeg:dash:role:2011 with the following values:
30 o `caption`
31 o `subtitle`
32 o `main`
33 o `alternate`
34 o `supplementary`

-
- 1 o sign
 - 2 o emergency
 - 3 o dub
- 4 - The Supplemental descriptor with the @schemeIdUri and @value pairs:
- 5 o Trickmode: @schemeIdUri set to "http://dashif.org/guide-lines/trickmode" and the @value the value of @id attribute of the Adaptation Set to which these trick mode Representations belong.
 - 6 o Period-continuous Adaptation Sets by using Aa @schemeIdUri set to "urn:mpeg:dash:period-continuity:2015" with the @value of the descriptor matching the value of an @id of a Adaptation Set that is contained in the MPD,
 - 7 o Period-connected Adaptation Sets by using Aa @schemeIdUri set to "urn:mpeg:dash:period-connectivity:2015" with the @value of the descriptor matching the value of an @id of a Adaptation Set that is contained in the MPD,
 - 8 o Switching across Adaptation Sets: @schemeIdUri set to urn:mpeg:dash:adaptation-set-switching:2016 and the @value is a comma-separated list of Adaptation Set IDs that may be seamlessly switched to from this Adaptation Set.

20 **3.9.5. Client Processing Reference Model**

21 **3.9.5.1. Introduction**

22 The following client model serves two purposes:

- 23 - In the absence of other information, the following client model may be implemented in a
24 DASH client for the purpose of selection of Adaptation Set for playout
- 25 - A content author may use the model to verify that the annotation is properly done in or-
26 der to get the desired client behaviour.

27 In the model it is assumed that the client can get sufficient information on at least the following
28 properties:

- 29 - For each codec in the @codecs string, the DASH client can get information if the me-
30 dia playback platform can decode the codec as described in the string. The answer
31 should be yes or no.
- 32 - For each DRM system in the **ContentProtection** element string, the DASH client
33 can get information if the media playback platform can handle this Content Protection
34 scheme as described in the string. The answer should be yes or no.

-
- the DASH client can get information on the media playback platform and rendering capabilities in terms of
 - o the maximum spatial resolution for video that can be handled
 - o the maximum frame rate for video that can be handled
 - o the audio channel configuration of the audio system
 - o the audio sampling rate of the audio system
 - the preferred language of the system
 - Accessibility settings for captions, subtitles, audio description, enhanced audio intelligibility,
 - Potentially preferences on media playback and rendering of the platform

Note of any of these functionalities are not fulfilled, then it may still be functional, but it may not result in the full experience as provided by the content author. As an example, if the DASH client cannot determine the preferred language, it may just use the selection priority for language selection.

3.9.5.2. Generic Processing Model

The DASH client uses the MPD and finds the Period that it likes to join, typically the first one for On-Demand content and the one at the live edge for live content. In order to select the media to be played, the DASH client assumes that the content is offered according to the content model above.

1. The DASH client looks for *main content*, i.e. any Adaptation Set with annotation **Role@schemeIdUri="urn:mpeg:dash:role:2011"** and **Role@value="alternative"** is excluded initially for selection. Note that in this model it is assumed that immediate startup is desired. If the DASH client wants to go over the alternatives upfront before starting the service, then the sequence is slightly different, but still follows the remaining principles.

2. DASH Client checks each Adaptation Set for the supported capabilities of the platform
 - Codec support
 - DRM support
 - Rendering capabilities

If any of the capabilities are not supported, then the Adaptation Set is excluded from the selection process.

3. The DASH client checks if it supports for CEA-608 rendering as defined in clause 6.4.3.3. If not supported, any accessibility descriptor with is **@schemeIdUri="urn:scte:dash:cc:cea-608:2015"** removed. Note that the Adaptation Set is maintained as it may used for regular video decoding.

-
- 1 4. DASH Client checks if there are any specific settings for accessibility in the user preferences
- 2
- 3 • If captions are requested by the system, the DASH client extracts
- 4 • all video Adaptation Sets that have an **Accessibility** descriptor assigned with either the @schemeIdUri="urn:mpeg:dash:role:2011" and @value="caption" or @schemeIdUri="urn:scte:dash:cc:cea-608:2015" (burned-in captions and SEI-based), as well as
- 5
- 6 • all subtitle Adaptation Sets that have an **Accessibility** descriptor assigned with either the @schemeIdUri="urn:mpeg:dash:role:2011" and @value="caption"
- 7
- 8 • and makes those available for Adaptation Sets that can be selected by the DASH client for caption support.
- 9
- 10 • If multiple caption Adaptation Sets remain, the DASH client removes all Adaptation Sets from the selection that are not in the preferred language, if language settings are provided in the system. If no language settings in the system are provided, or none of the Adaptation Sets meets the preferred languages, none of the Adaptation Sets are removed from the selection. Any Adaptation Sets that do not contain language annotation are removed, if any of the remaining Adaptation Sets provides proper language settings.
- 11
- 12 • If still multiple caption Adaptation Sets remain, then the ones with the highest value of @selectionPriority is chosen.
- 13
- 14 • If still multiple caption Adaptation Sets remain, then the DASH client makes a random choice on which caption to enable.
- 15
- 16 • else if no captions are requested
- 17 • the **Accessibility** element signaling captions may be removed from the Adaptation Set before continuing the selection.
- 18
- 19 • If sign language is requested
- 20 • all video Adaptation Sets that have an **Accessibility** descriptor assigned with @schemeIdUri="urn:mpeg:dash:role:2011" and @value="sign" are made available for sign language support.
- 21
- 22 • else if no sign language is requested
- 23 • the Adaptation Set signaling sign language with the **Accessibility** element may be removed from the Adaptation Set before continuing the selection.
- 24
- 25 • If audio descriptions are requested

-
- 1 • all video Adaptation Sets that have an **Accessibility** descriptor as-
 - 2 signed with @schemeIdUri="urn:mpeg:dash:role:2011" and
 - 3 @value="description" are made available for audio description
 - 4 support.
 - 5 • else if no audio descriptions are requested
 - 6 • the Adaptation Set signaling audio descriptions with the **Accessibil-**
 - 7 **ity** element may be removed from the Adaptation Set before continuing
 - 8 the selection.
 - 9 • If enhanced audio intelligibility is requested
 - 10 • all audio Adaptation Sets that have an **Accessibility** descriptor as-
 - 11 signed with @schemeIdUri="urn:mpeg:dash:role:2011" and
 - 12 @value="enhanced-audio-intelligibility" are made avail-
 - 13 able for enhanced audio intelligibility support.
 - 14 • else if no enhanced audio intelligibility is requested
 - 15 • the **Accessibility** element may be removed from the Adaptation Set
 - 16 before continuing the selection.
- 17 5. If video rendering is enabled, based on the remaining video Adaptation Sets the client se-
- 18 lects one as follows:
- 19 • Any Adaptation Set for which an Essential Descriptor is present for which the
 - 20 scheme or value is not understood by the DASH client, is excluded from the se-
 - 21 lection
 - 22 • Any Adaptation Set for which an Essential Descriptor is present for which the
 - 23 scheme is <http://dashif.org/guidelines/trickmode>, is excluded
 - 24 from the initial selection
 - 25 • If still multiple video Adaptation Sets remain, then the ones with the highest value
 - 26 of @selectionPriority is chosen.
 - 27 • If still multiple video Adaptation Sets remain, then the DASH client makes a
 - 28 choice for itself, possibly on a random basis.
 - 29 • Note that an Adaptation Set selection may include multiple Adaptation Sets, if
 - 30 Adaptation Set Switching is signaled. However, the selection is done for only one
 - 31 Adaptation Set.
- 32 6. If audio rendering is enabled, based on the remaining audio Adaptation Sets the client se-
- 33 lects one as follows:
- 34 • Any Adaptation Set for which an Essential Descriptor is present for which the
 - 35 scheme or value is not understood by the DASH client, is excluded from the se-
 - 36 lection

-
- 1 • If multiple audio Adaptation Sets remain, the DASH client removes all Adaptation Sets from the selection that are not in the preferred language, if language settings are provided in the system. If no language settings in the system are provided, or none of the Adaptation Sets meets the preferred languages, none of the Adaptation Sets are removed from the selection. Any Adaptation Set that does not contain language annotation are removed, if any of the remaining Adaptation Sets provides proper language settings.
- 2 • If still multiple audio Adaptation Sets remain, then the ones with the highest value of @selectionPriority is chosen.
- 3 • If still multiple audio Adaptation Sets remain, then the DASH client makes a choice for itself, possibly on a random basis.
- 4 • Note that an Adaptation Set may include multiple Adaptation Sets, if Adaptation Set Switching or receiver mix is signaled. However, the selection is done for only one Adaptation Set.
- 5 7. If subtitle rendering is enabled, based on the subtitle Adaptation Sets the client selects one as follows:
- 6 • Any Adaptation Set for which an Essential Descriptor is present for which the scheme or value is not understood by the DASH client, is excluded from the selection
- 7 • If multiple subtitle Adaptation Sets remain, the DASH client removes all Adaptation Sets from the selection that are not in the preferred language, if language settings are provided in the system. If no language settings in the system are provided, or none of the Adaptation Sets meets the preferred languages, none of the Adaptation Sets are removed from the selection. Any Adaptation Set that does not contain language annotation are removed, if any of the remaining Adaptation Sets provides proper language settings.
- 8 • If still multiple subtitle Adaptation Sets remain, then the ones with the highest value of @selectionPriority is chosen.
- 9 • If still multiple subtitle Adaptation Sets remain, then the DASH client makes a choice for itself, possibly on a random basis.
- 10 8. If the DASH client has the ability to possibly switch to alternative content, then alternative content may be selected either through the **Label** function or the **ViewPoint** functionality. This selection may be done dynamically during playout and the DASH client is expected to switch to the alternative content. Once all alternative content is selected, the procedures following from step 2 onwards apply.
- 11 9. At Period boundary a DASH client initially looks for a Period continuity or connectivity, i.e. does the Period include an Adaptation Set that is a continuation of the existing one. If not present it will go back to step 1 and execute the decision logic.

1 **3.10. On-Demand Services**

2 **3.10.1. Introduction**

3 MPEG-DASH [4] provides several tools to support On-Demand streaming services. This section
4 primarily provides requirements and recommendations for both content authoring as well as client
5 implementations.

6 For this purpose, this section

- 7 • clarifies and refines details of interoperability points when used with the features available
8 in MPEG-DASH [4] with respect to different service configurations and client implemen-
9 tations. [4] with respect to different service configurations and client implementations.
- 10 • defines an interoperability point in order to address content authoring and client require-
11 ments to support a broad set of On-Demand services based on the features defined in the
12 third edition (published in 2018) of MPEG-DASH.

13 DASH Media Presentations with **MPD@type** set to "static" indicate that all media announced
14 in the MPD is available at the **MPD@availabilityStartTime** and will cease to be available
15 at the **MPD@availabilityEndTime**. This has two major effects, namely

- 16 1. The content creator announces a DASH Media Presentation which all content is availa-
17 ble.
- 18 2. Clients have flexibility to consume the content at different times, may pause the presenta-
19 tion, may easily apply trick modes such as fast-forward, etc.

20 Static services may be offered using live DASH profiles as defined in clause 3.10.2 as well as on-
21 demand DASH profiles in clause 3.10.3. The use of live profiles is particularly interesting for cases
22 where the content was offered live, but then also made available in a catchup or time shift format
23 as documented in clause 4.6.3.10.2 as well as on-demand DASH profiles in clause.

24 **3.10.2. On-Demand Services with Live profiles**

25 **3.10.2.1. MPD Information**

26 If any of the DASH-IF Live Profiles are used, Table 5 provides an overview on the MPD infor-
27 mation for On-Demand services.

29 **Table 5 -- Information related to Segment Information for Using Live Profiles for On-Demand Services**

MPD Information	Status	Comment
MPD@type	optional, if present set to "static"	the type of the Media Presentation is static, i.e. all Segments get available at the value of the AST.
MPD@availabilityStartTime	optional	the value when all Segments get available. If not present, all Segments are available. The value is denoted as <i>AST</i> in the following.
MPD@mediaPresentationDuration	mandatory (unless the Period duration is present)	provides the duration of the Media Presentation.

MPD@minBufferTime	mandatory	minimum buffer time, used in conjunction with the @bandwidth attribute of each Representation. The value is denoted as <i>MBT</i> . Details on the setting and usage of the parameter is provided in the following.
Period@start	Mandatory for the first Period in the MPD	the start time of the Period relative to the MPD availability start time.
SegmentTemplate@media	mandatory	The template for the Media Segment assigned to a Representation.
SegmentTemplate@startNumber	optional default	number of the first segment in the Period assigned to a Representation
SegmentTemplate@time-scale	optional default	timescale for this Representation.
SegmentTemplate@duration SegmentTemplate.SegmentTimeline	exactly one of SegmentTemplate@duration or SegmentTemplate.SegmentTimeline must be present per Representation.	the duration of each Segment in units of a time.

1

2 **3.10.2.2. Segment Information Derivation**3 **3.10.2.2.1. Introduction**

4 Based on an MPD including information as documented in Table X 5 and available at time *NOW*
 5 on the server, a synchronized DASH client derives the information of the list of Segments for each
 6 Representation in each Period. This section only describes the information that is expressed by the
 7 values in the MPD. The generation of the information on the server and the usage of the infor-
 8 mation in the client is discussed in below.

9 **3.10.2.2.2. MPD Information**

10 For a static DASH services, the following information shall be present and not present in the MPD
 11 (also please refer to Table X 5):

- 12 • The **MPD@type** shall be set to "static".
- 13 • The **MPD@mediaPresentationDuration** shall be present, or the **Period@dura-**
 14 **tion** of the last Period shall be present.
- 15 • The **MPD@minimumUpdatePeriod** shall not be present.
- 16 • **MPD@timeShiftBufferDepth** should not be present. If present, the client is expected
 17 to ignore the value of this attribute.
- 18 • **MPD@suggestedPresentationDelay** should not be present. If present, the client is ex-
 19 pected to ignore the value of this attribute.
- 20 • The **@availabilityTimeOffset** should not be present. If present, the client is ex-
 21 pected to ignore the value of this attribute.

1 **3.10.2.2.3. Period Information**

2 Each Period is documented by a **Period** element in the MPD. An MPD may contain one or more
3 Periods. In order to document the use of multiple Periods, the sequence of Period elements is
4 expressed by an index i with i increasing by 1 for each new Period element.

5 For static Media Presentations, each regular Period i in the MPD is assigned a

- 6 • Period start time $PSwc[i]$ in MPD time,
7 • Period end time $PEwc[i]$, in MPD time.

8 The Period start time $PSwc[i]$ for a regular Period i is determined according to section 5.3.2.1 of
9 ISO/IEC 23009-1 [4]:

- 10 • If the attribute `@start` is present in the **Period**, then $PSwc[i]$ is the value of this attribute
11 minus the value of `@start` of the first Period.
12 • If the `@start` attribute is absent, but the previous **Period** element contains a `@duration`
13 attribute then the start time of the Period is the sum of the start time of the previous
14 Period $PSwc[i]$ and the value of the attribute `@duration` of the previous Period. Note
15 that if both are present, then the `@start` of the new Period takes precedence over the
16 information derived from the `@duration` attribute.

17 The Period end time $PEwc[i]$ for a regular Period i is determined as follows:

- 18 • If the Period is the last one in the MPD, the time $PEwc[i]$ is obtained as
19 o the Media Presentation Duration $MPDur$, with $MPDur$ the value of **MPD@me-**
20 **diaPresentationDuration** if present, or the sum of $PSwc[i]$ of the last Pe-
21 riod and the value of **Period@duration** of the last Period.
22 • else
23 o the time $PEwc[i]$ is obtained as the Period start time of the next Period, i.e. $PEwc[i] = PSwc[i+1]$.

25 **3.10.2.2.4. Representation Information**

26 Based on such an MPD, a list of Segments contained in a Representation in a Period i with Period
27 start time $PSwc[i]$ and Period end time $PEwc[i]$ can be computed.

28 If the **SegmentTemplate.SegmentTimeline** is present and the **SegmentTem-**
29 **plate@duration** is not present, the **SegmentTimeline** element contains N_s **S** elements
30 indexed with $s=1, \dots, N_s$, then let

- 31 • ts the value of the `@timescale` attribute
32 • $t[s]$ be the value of `@t` of the s -th **S** element,
33 • $d[s]$ be the value of `@d` of the s -th **S** element
34 • $r[s]$ be,
35 o if the `@r` value is greater than or equal to zero
36 ■ one more than the value of `@r` of the s -th **S** element. Note that if `@r` is
37 smaller than the end of this segment timeline element, then this Represen-
38 tation contains gaps and no media is present for this gap.
39 o else

-
- 1 • if $t[s+1]$ is present, then $r[s]$ is the ceil of $(t[s+1] - t[s])/d[s]$
 2 • else $r[s]$ is the ceil of $(PEwc[i] - PSwc[i] - t[s]/ts)*ts/d[s])$

3 If the **SegmentTemplate**@duration is present and the **SegmentTemplate.SegmentTimeline** is not present, then

- 5 • $N_s=1$,
 6 • ts the value of the @timescale attribute
 7 • $t[s]$ is 0,
 8 • the $d[s]$ is the value of @duration attribute
 9 • $r[s]$ is the ceil of $(PEwc[i] - PSwc[i] - t[s]/ts)*ts/d[s])$

10 Note that the last segment may not exist and $r[s]$ is one less than this computation provides.

11 3.10.2.2.5. Media Time Information of Segment

12 Each Media Segment at position $k=1,2, \dots$ for each Representation is assigned an earliest media
 13 presentation time $EPT[k,r,i]$ and an accurate segment duration $SDUR[k,r,j]$, all measured in media
 14 presentation time.

15 The earliest presentation time may be estimated from the MPD using the MPD time of the media.

16 The earliest presentation time may be accurately determined from the Segment itself.

17 For details on the derivation of the earliest presentation time, see section 3.2.7.

18 3.10.2.2.6. Segment List Parameters

19 For each Period i with Period start time $PSwc[i]$ and Period end time $PEwc[i]$ and each Representation r in the Period the following information can be computed:

- 21 • the presentation time offset described in the MPD, $o[i,r]$
 22 • the number of the first segment described in the MPD, $k1[i,r]$
 23 • the number of the last segment described in the MPD, $k2[i,r]$
 24 • MPD start time of the initialization segment $MST[0,i,r]$
 25 • MPD end time of the initialization segment $MET[0,i,r]$
 26 • MPD start time of each media segment $MST[k,i,r]$, $k=k1, \dots, k2$
 27 • MPD end time of each media segment $MET[k,i,r]$, $k=k1, \dots, k2$
 28 • segment duration of each media segment $SD[k,i,r]$, $k=k1, \dots, k2$
 29 • the URL of each of the segments, $URL[k,i,r]$

30 Based on the above information, for each Representation r in a Period i , the MPD segment duration
 31 of each segment $MSD[k,i,r]$, and the URL of each of the segments, $URL[k,i,r]$ within one Period i
 32 be derived as follows using the URL Template function $URLTemplate(Replacement-$
 33 $String, Address)$ as documented in clause 4.3.2.2.8:

- 34 • $k=0$
 35 • $MST[0,i,r] = PSwc[i]$
 36 • for $s=1, \dots N_s [i,r]$

-
- ```

1 ○ $k = k + 1$
2 ○ $MST[k,i,r] = PSwc[i] + (t[s,i,r] + d[s,i,r] - o[i,r])/ts$
3 ○ $SD[k,i,r] = d[s,i,r]/ts$
4 ○ $MET[k,i,r] = MST[k,i,r] + d[s,i,r]/ts$
5 ○ if SegmentTemplate@media contains $Number$
6 ▪ Address=@startNumber
7 ▪ URL[k,i,r] = URLTemplate ($Number$, Address)
8 else
9 ▪ Address = t[s,i,r]
10 ▪ URL[k,i,r] = URLTemplate ($Time$, Address)
11 ○ for j = 1, ..., r[s,i,r]
12 ▪ k = k + 1
13 ▪ MST[k,i,r] = MST[k-1,i,r] + d[s,i,r]/ts
14
15 ▪ MET[k,i,r] = MST[k,i,r] + d[s,i,r] /ts
16
17 ▪ SD[k,i,r] = d[s,i,r] /ts
18 ▪ if SegmentTemplate@media contains $Number$
19 ▪ Address = Address + 1
20 ▪ URL[k,i,r] = URLTemplate ($Number$, Address)
21 else
22 ▪ Address = Address + d[s,i,r]
23 ▪ URL[k,i,r] = URLTemplate ($Time$, Address)
24 • k2[i,r] = k
25 • MET[0,i,r] = MET[k2[i,r],i,r]

```

### 3.10.2.2.7. Synchronized Playout and Seamless Switching

In order to achieve synchronized playout across different Representations, typically from different Adaptation Sets, the different Representations are synchronized according to the presentation time in the Period. Specifically, the earliest presentation time of each Segment according to section 5.2.2.5 determines the playout of the Segment in the Period and therefore enables synchronized playout of different media components as well as seamless switching within one media component.

## 3.10.3. On-Demand Services with On-Demand profile

### 3.10.3.1. Motivation and Background

While using the live profile is possible and surely an option in certain cases, the use of On-Demand profiles for On-Demand Services removes the need for the Packager, Origin Server and CDN to manage millions of segment files. Instead, one can have a single ISO BMFF/MP4 conforming file which can then be (sub)indexed to provide at least the same features as the Live Profile for VoD. CDNs may offer capabilities to do this translation in their network, but these requires dedicated origin and possibly cache servers and is not commonly used.

The advantages lie mainly for Content Providers, Service providers and CDNs:

- Easier content management

- 
- 1     • Pre-caching of data so improving server capacity on commodity servers. The cache gen-  
2         erally preloads some byte range following the requested byte range.  
3     • Lower costs (less files, less requests, better caching)  
4     • More efficient encoding (flexible IDR frame placement, constant quality/capped VBR  
5         encoding)  
6     • Easier extensibility with emerging file format features (green metadata, quality metadata,  
7         etc.)

8     However, there are also advantages for client implementations:

- 9         • the clients can be significantly improved adding higher quality (throughput, start-up), ad-  
10          vanced trick mode, and other optimizations in improved user experience, as well as using  
11          the extensions mentioned above. This is as with the segmented approach the maximum  
12          request size is generally one segment. One may use special HTTP methods such as pipe-  
13          lining, but this is not universally available. And it is difficult to vary the request size to  
14          get more accurate bandwidth estimation.

15  
16     From a client perspective assuming that the simple live profile is implemented, then a simple On-  
17         Demand Client just needs to add two new things

- 18         • Request and parse Segment Index to identify byte ranges for the sub-segments  
19         • Issue HTTP requests for these byte ranges, instead of for whole segments

20  
21     In summary, the benefits of an On-Demand profile are significant to achieve a robust, massively  
22         scalable and efficient deployment (which has been recognized by providers that run services,  
23         CDNs and clients).

#### 24     **3.10.3.2. General**

25     Content offered for On-demand services may use on-demand DASH profiles as defined in  
26         ISO/IEC 23009-1 [4].

27  
28     DASH-IF defines a DASH-IF On-Demand Profile "<http://dashif.org/guidelines/dash-if-on-demand>". The following requirements apply for content offered using  
29         the DASH-IF On-Demand Profile:

- 30  
31         - The service offering conforms to the Extended ISO BMFF On-Demand profile as de-  
32          fined in ISO/IEC 23009-1 [4], clause 8.9.  
33         - Each Representation shall have one Segment that complies with the Indexed Self-In-  
34          itializing Media Segment as defined in 6.3.5.2  
35         - In the MPD, the @indexRange attribute shall be present.  
36         - Only a single sidx box shall be present.

37             Note 1: external Representation Indices are considered beneficial, but are only de-  
38          fined for MPEG-2 TS at this stage. DASH-IF reached out to ISO/IEC to clarify the  
39          applicability to ISO BMFF based media segments and expects to add this feature  
40          in a later version of the IOP Guidelines.

41             Note 2: The single sidx restriction was introduced in version 3 of this document  
42          based on deployment experience and to enable alignment with DVB DASH.

1 Note 3: If only the movie header is desired, for example for initial capability  
2 checks, then downloading at most the first 1000 bytes of the Media Segment is  
3 sufficient for all DASH-IF test vectors as available by the end of 2017.  
4

5 Additional Service Offering Requirements and Guidelines are provided in clause 5.3.2.

6 Client Operation, Requirements and Guidelines are provided in clause 5.3.3.

7 **3.10.3.3. MPD Information**

8 For the DASH-IF On-Demand Profile, based on the requirements in ISO/IEC 23009-1 [4], clause  
9 8.3, Table 6 provides an overview on the MPD information.

10 **Table 6 -- Information related to Segment Information for On-Demand Profile**

| MPD Information                               | Status                                            | Comment                                                                                                                                                                                                             |
|-----------------------------------------------|---------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>MPD@type</b>                               | optional, if present set to "static"              | the type of the Media Presentation is static, i.e. all Segments get available at the value of the AST.                                                                                                              |
| <b>MPD@availabilityStartTime</b>              | optional                                          | the value when all Segments get available. If not present, all Segments are available. The value is denoted as <i>AST</i> in the following.                                                                         |
| <b>MPD@mediaPresentationDuration</b>          | mandatory (unless the Period duration is present) | provides the duration of the Media Presentation.                                                                                                                                                                    |
| <b>MPD@minBufferTime</b>                      | mandatory                                         | minimum buffer time, used in conjunction with the @bandwidth attribute of each Representation. The value is denoted as <i>MBT</i> . Details on the setting and usage of the parameter is provided in the following. |
| <b>Period@start</b>                           | Optional, default is 0                            | The first period start time describes an anchor. The remaining ones provide the duration of the Period.                                                                                                             |
| <b>AdaptationSet@subsegmentAlignment</b>      | mandatory to be set to true                       | All subsegments are aligned in an MPD                                                                                                                                                                               |
| <b>AdaptationSet@subsegmentStartsWithSAP</b>  | mandatory to be set to 1 or 2                     | Each subsegment starts with SAP 1 or 2                                                                                                                                                                              |
| <b>Representation@bandwidth</b>               |                                                   | See above                                                                                                                                                                                                           |
| <b>Representation.BaseURL</b>                 |                                                   |                                                                                                                                                                                                                     |
| <b>SegmentTemplate@timescale</b>              | Optional, default is 1                            | timescale for this Representation.                                                                                                                                                                                  |
| <b>SegmentTemplate@presentationTimeOffset</b> | optional default is 0                             | the presentation time offset of the Representation relative to the start of the Period                                                                                                                              |

---

|                                   |           |                                                                                                                                                                                                                                                                                                               |
|-----------------------------------|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>SegmentTemplate@indexRange</b> | mandatory | specifies the byte range that contains the Segment Index in all Media Segments of the Representation.<br><br>The byte range is expressed and formatted as a <code>byte-range-spec</code> as defined in RFC 7233, Clause 2.1. It is restricted to a single expression identifying a contiguous range of bytes. |
|-----------------------------------|-----------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

- 1
- 2     **3.10.3.4. Service Offering Requirements and Guidelines**
- 3         **3.10.3.4.1. MPD Settings**
- 4     The parameters as defined in Table 6 shall be present.
- 5         **3.10.3.4.2. Generation of On-Demand Segments**
- 6     The MPEG-DASH On-Demand Profile provides the ability to structure the Segment Index in a  
7     quite arbitrary manner. However, for many deployment cases, a single Segment Index box de-  
8     scribing all Subsegments (i.e. movie fragments) in the Representations is sufficient to cover a large  
9     amount of use cases. In order to address complexity and interoperability concerns, DASH-IF IOP  
10   restricts the On-Demand profile to a single Segment Index box.
- 11   In this case a single segment index is provided and using the syntax according to ISO/IEC 14496-  
12
- ```

13       aligned(8) class SegmentIndexBox extends FullBox('sidx', version, 0) {
14            unsigned int(32) reference_ID;
15            unsigned int(32) timescale;
16            if (version==0) {
17                unsigned int(32) earliest_presentation_time;
18                unsigned int(32) first_offset;
19            }
20            else {
21                unsigned int(64) earliest_presentation_time;
22                unsigned int(64) first_offset;
23            }
24            unsigned int(16) reserved = 0;
25            unsigned int(16) reference_count;
26            for(i=1; i <= reference_count; i++) {
27                bit (1)                              reference_type;
28                unsigned int(31)                   referenced_size;
29                unsigned int(32)                   subsegment_duration;
30                bit(1)                           starts_with_SAP;
31                unsigned int(3)                   SAP_type;
32                unsigned int(28)                   SAP_delta_time;
33            }
34        }

```
- 35 the following applies:
- 36 • the `reference_ID` is set to the `track_ID` of the movie fragment

-
- 1 • the timescale is the the timescale field of the Media Header Box of the track
2 • the earliest presentation time is typically 0, but if the first sample in the Representation
3 this results from the splitting of a large file or some adjustment for A/V sync is necessary,
4 it may be larger than 0.
5 • first_offset provides the pointer to the first movie fragment header.
6 • reference_count is the total number of subsegments/movie fragments included in
7 the ISO BMFF
8 • reference_type is all set to 0, as reference are only to the movie fragment header
9 box.
10 • referenced_size the distance in bytes from the first byte of the referenced movie
11 fragment to the first byte of the next movie fragment, or in the case of the last entry, the
12 end of the file.
13 • subsegment_duration: this field carries the difference between the earliest presen-
14 tation time of any access unit of the reference stream in the next movie fragment and the
15 earliest presentation time of any access unit of the reference stream in the referenced
16 movie fragment
17 • starts_with_SAP is set to 1, and SAP_type is set to 1 or 2
18 • SAP_delta_time is set to 0

19

20 **3.10.3.5. Content Offering with Multiple Periods**

21 Content may be offered in multiple Periods following the motivation in clause 3.2.12. Content
22 may be offered in a continuous and connected mode.

23 **3.10.3.6. Client Operation, Requirements and Guidelines**

24 **3.10.3.6.1. Basic Operation for Single Period**

25 Given a client supporting the live profile, the above requirements can be implemented as fol-
26 lows:

- 27 • The client already issues requests for the moov box - these needs to be extended to cover
28 the Segment Index (sidx) as well
29 • Parse the Segment Index (<150 LoC in Javascript, open source, DASH-IF reference cli-
30 ent) with the output being a list of (URL, byte range) pairs that would replace the usual
31 list of URLs in the live profile
32 • Create list of requests (URL, byte range) pairs, instead of a list of URLs.

33

34 Only movie fragments are referenced and therefore the entire switching timeline of Representation
35 can be built by downloading the sidx first. However, a smart client may choose to download only
36 the beginning of the Representations's sidx box for a fast start up and download the rest after it
37 starts streaming the initial media subsegments.

38 A DASH client may also determine its request size independent of the segment duration.

39 **3.10.3.6.2. Determining the Subsegment Map**

40 By parsing the sidx for every Representation of interest, the client can create a full subsegment
41 map.

1 **3.10.3.6.3. Multi-Period Content**

2 For details refer to clause 3.2.12.

3 **3.10.4. Mixed On-Demand Content**

4 For on-demand content that is offers a mixture of Periods, DASH-IF defines a DASH-IF Mixed
5 Profile "<http://dashif.org/guidelines/dash-if-mixed-on-demand>".

6 The service is offered as follows:

- 7 - Multiple Periods may be offered
8 - All Representations in one period shall either conform to a DASH Live profile (as defined
9 in clause 3.10.2) or to a DASH On-demand Profile (as defined in clause 3.10.3).
10 - The @profile signalling shall be present in each Period

11 A client implementing the mixed profile shall support

- 12 - Multi-Period offerings
13 - A mixture of on-demand and live periods as defined above

14 Note that this profile does not support dynamic service types.

15

16 **4. Live Services**

17 **4.1. Introduction**

18 MPEG-DASH [1] provides several tools to support live services. This section primarily provides
19 requirements and recommendations for both, content authoring as well as client implementations.

20 For this purpose, this section

- 21 • clarifies and refines details of interoperability points when used with the features available
22 in the 2012 edition of MPEG-DASH with respect to different service configurations and
23 client implementations.
- 24 • defines one new interoperability point in order to address content authoring and client re-
25 quirements to support a broad set of live services based on the features defined in the sec-
26 ond edition (published 2014) of MPEG-DASH as well certain amendments thereof.

27 The main features and differences of these two modes are provided in the following Table 7:

28 **Table 7 Main features and differences of simple and main live services**

Feature	Sim- ple	Main
Support of MPD@type	static, dynamic	static, dynamic
MPD updates	yes	yes
MPD updated triggered	by MPD Event	by In- band

	attrib- ute mini- mum up- date period	mes- sages in the seg- ments.
URL gen- eration	based on MPD	based on MPD and seg- ment infor- mation
Timeline gaps	based on MPD and for entire con- tent	may be sig- naled indi- vidu- ally for each Repre- sen- ta- tion
Segments starts with	closed GOP	closed GOP
Support of Simple Live	Yes	No
Support of Main Live	Yes	Yes

- 1
- 2 To support the definition of the interoperability points, architectures and use cases were collected.
- 3 These are documented in Annex B.

4.2. Overview Dynamic and Live Media Presentations

5 DASH Media Presentations with **MPD@type** set to "dynamic" enable that media is made avail-
6 able over time and its availability may also be removed over time. This has two major effects,
7 namely

- 8 3. The content creator can announce a DASH Media Presentation for which not all content
9 is yet available, but only gets available over time.
- 10 4. Clients are forced into a timed schedule for the playout, such that they follow the sched-
11 ule as desired by the content author.

12 Dynamic services may be used for different types of services:

- 13 1. **Dynamic Distribution of Available Content:** Services, for which content is made avail-
14 able as dynamic content, but the content is entirely generated prior to distribution. In this

1 case the details of the Media Presentation, especially the Segments (duration, URLs) are
2 known and can be announced in a single MPD without MPD updates. This addresses use
3 cases 2 and 3 in Annex B.

- 4 2. **MPD-controlled Live Service:** Services for which the content is typically generated on
5 the fly, and the MPD needs to be updated occasionally to reflect changes in the service
6 offerings. For such a service, the DASH client operates solely on information in the
7 MPD. This addresses the use cases 4 and 5 in Annex B.
- 8 3. **MPD and Segment-controlled Live:** Services for which the content is typically gener-
9 ated on the fly, and the MPD may need to be updated on short notice to reflect changes in
10 the service offerings. For such a service, the DASH client operates on information in the
11 MPD and is expected to parse segments to extract relevant information for proper opera-
12 tion. This addresses the use cases 4 and 5, but also takes into account the advanced use
13 cases.

14 Dynamic and Live services are typically controlled by different client transactions and server-side
15 signaling.

16 For initial access to the service and joining the service, an MPD is required. MPDs may be accessed
17 at join time or may have been provided earlier, for example along with an Electronic Service
18 Guide. The initial MPD or join MPD is accessed and processed by the client and the client having
19 an accurate clock that is synchronized with the server can analyze the MPD and extract suitable
20 information in order to initiate the service. This includes, but is not limited to:

- 21 • identifying the currently active Periods in the service and the Period that expresses the live
22 edge (for more details see below)
- 23 • selecting the suitable media components by selecting one or multiple Adaptation Sets.
24 Within each Adaptation Set selecting an appropriate Representation and identifying the
25 live edge segment in each Representations. The client then issues requests for the Seg-
26 ments.

27 The MPD may be updated on the server based on certain rules and clients consuming the service
28 are expected to update MPDs based on certain triggers. The triggers may be provided by the MPD
29 itself or by information included in Segments. Depending on the service offering, different client
30 operations are required as shown in Figure 4.

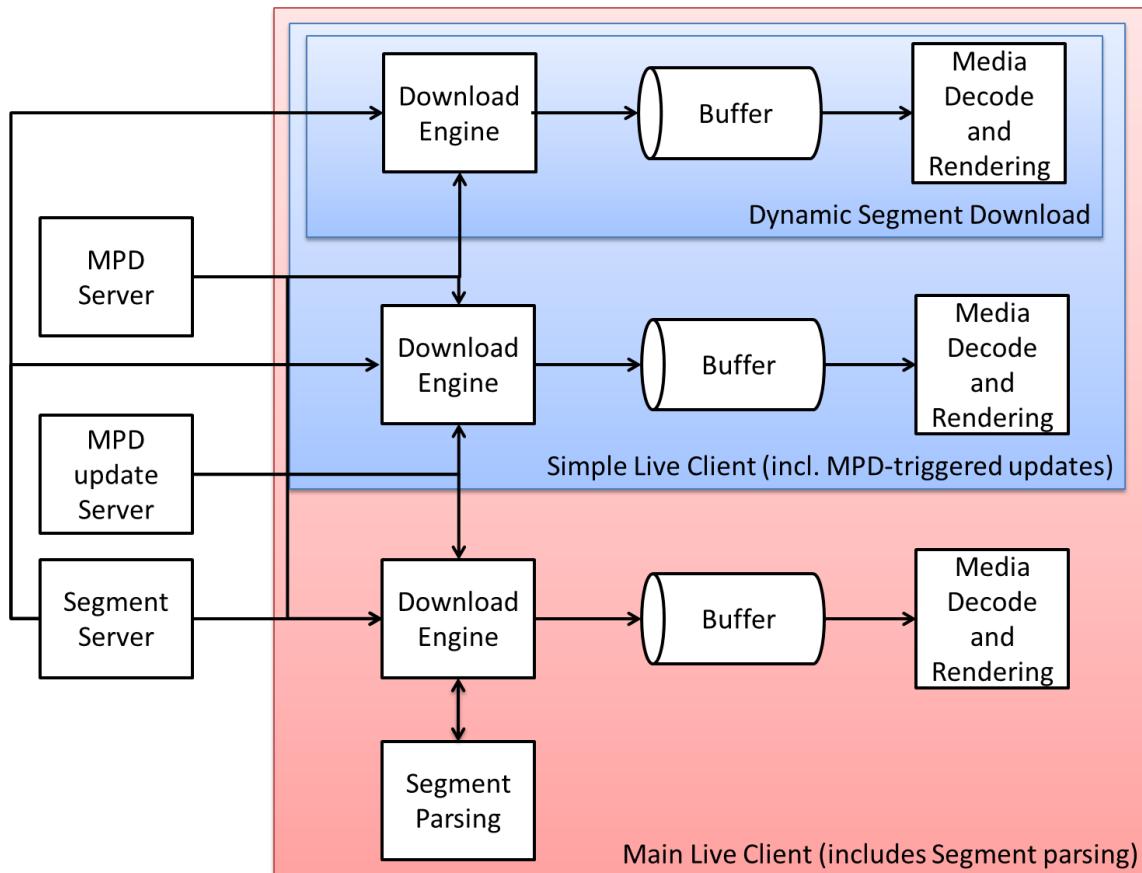


Figure 4 Different Client Models

The basic functions of a live client described in this document are as follows:

1. **Dynamic Segment Download:** This function creates a list of available Segments based on a single MPD and joins the service by downloading Segments at the live edge or may use the Segments that are available in the time shift buffer.
2. **Simple Live Client:** This client includes the dynamic segment download function and enables updates of the MPD based on information in the MPD in order to extend the Segment list at the live edge. MPDs are refetched and revalidated when the currently available MPD expires, i.e. an expired MPD can no longer be used for Segment URL generation.
3. **Main Live Client:** This client includes all features of the simple Live DASH client. In addition it generates Segment URLs and it updates the MPD based on information in the

-
- 1 Segments if the service offering provides this feature. MPDs are refetched and revali-
2 dated when the currently available MPD expires based on expiry information in the Seg-
3 ments.
- 4 Requirements and recommendations for the dynamic segment download functions are defined in
5 in section 4.3.
- 6 Requirements and recommendations for simple live service offerings and corresponding clients
7 are defined in section 4.4.
- 8 Requirements and recommendations for main live service offerings and corresponding clients are
9 defined in section 4.5.
- 10 Requirements and recommendations when offering live services as on-demand are provided in
11 section 4.6.
- 12 Requirements and recommendations for client-server timing synchronization are defined in sec-
13 tion 4.7.
- 14 Requirements and recommendations for robust service offerings and corresponding clients are
15 defined in section 4.8.
- 16 Interoperability Aspects are defined in section 4.9.

4.3. Dynamic Segment Download

4.3.1. Background and Assumptions

The dynamic segment download function is a key component of live services. In addition, the dynamic segment download function may also be used for scheduling a playout. In the remainder of this subsection, it is assumed that the client has access to a single instance of an MPD and all information of the entire Media Presentation is contained in the MPD.

We refer to this service as dynamic service as the main feature is that the Segments are made available over time following the schedule of the media timeline.

Dynamic services are primarily documented in order to provide insight into the timing model of Segment availabilities. This forms the basis for live services and explains the key concepts and rules for Segment availabilities.

4.3.2. Preliminaries

4.3.2.1. MPD Information

If the Media Presentation is of type dynamic, then Segments have different Segment availability times, i.e. the earliest time for which the service provider permits the DASH client to issue a request to the Segment and guarantees, under regular operation modes, that the client gets a 200 OK response for the Segment. The Segment availability times for each Representation can be computed based on the information in an MPD.

For a dynamic service the MPD should at least contain information as available in Table 8. Information included there may be used to compute a list of announced Segments, Segment Availability Times and URLs.

- 1 Assume that an MPD is available to the DASH client at a specific wall-clock time *NOW*. It is
 2 assumed that the client and the DASH server providing the Segments are synchronized to wall-
 3 clock, either through external means or through a specific client-server synchronization. Details
 4 on synchronization are discussed in section 4.7.
- 5 Assuming synchronization, the information in the MPD can then be used by the client at time *NOW*
 6 to derive the availability (or non-availability) of Segments on the server.

7 **Table 8 -- Information related to Segment Information and Availability Times for a dynamic service**

MPD Information	Status	Comment
MPD@type	mandatory, set to "dynamic"	the type of the Media Presentation is dynamic, i.e. Segments get available over time.
MPD@availabilityStartTime	mandatory	the start time is the anchor for the MPD in wall-clock time. The value is denoted as <i>AST</i> in the following.
MPD@mediaPresentationDuration	mandatory (for the considered use cases)	provides the duration of the Media Presentation.
MPD@suggestedPresentationDelay	optional, but recommended	suggested presentation delay as delta to segment availability start time. The value is denoted as <i>SPD</i> . Details on the setting and usage of the parameter is provided in the following.
MPD@minBufferTime	mandatory	minimum buffer time, used in conjunction with the @bandwidth attribute of each Representation. The value is denoted as <i>MBT</i> . Details on the setting and usage of the parameter is provided in the following.
MPD@timeShiftBufferDepth	optional, but recommended	time shift buffer depth of the media presentation. The value is denoted as <i>TSB</i> . Details on the setting and usage of the parameter is provided in the following.
Period@start	Mandatory for the first Period in the MPD	the start time of the Period relative to the MPD availability start time.
Representation@availabilityTimeOffset	Optional default	The offset in availability time for this Representation. The same attribute may also be available on a Base URL or default. For processing model and more details refer to section 4.3.2.2.5. NOTE: the value of "INF" implies availability of all segments starts at MPD@availabilityStartTime
SegmentTemplate@media	mandatory	The template for the Media Segment assigned to a Representation.

SegmentTemplate@startNumber	optional default	number of the first segment in the Period assigned to a Representation
SegmentTemplate@time-scale	optional default	timescale for this Representation.
SegmentTemplate@duration SegmentTemplate.SegmentTimeline	exactly one of SegmentTemplate@duration or SegmentTemplate.SegmentTimeline must be present per Representation.	the duration of each Segment in units of a time.

1 **4.3.2.2. Segment Information Derivation**

2 **4.3.2.2.1. Introduction**

3 Based on an MPD including information as documented in Table 8 and available at time *NOW* on
4 the server, a synchronized DASH client derives the information of the list of Segments for each
5 Representation in each Period. This section only describes the information that is expressed by the
6 values in the MPD. The generation of the information on the server and the usage of the infor-
7 mation in the client is discussed in section 4.3.3 and 4.3.4, respectively.

8 MPD information is provided in subsection 4.3.2.2.3. The Period based information is documented
9 in sub-section 4.3.2.2.4, and the Representation information is documented in sub-section
10 4.3.2.2.5.

11 **4.3.2.2.2. Definitions**

12 The following definitions are relevant and aligned with ISO/IEC 23009-1:

- 13 • available Segment is a Segment that is accessible at its assigned HTTP-URL. This means
14 that a request with an HTTP GET to the URL of the Segment results in a reply of the
15 Segment and 2xx status code.
- 16 • valid Segment URL is an HTTP-URL that is promised to reference a Segment during its
17 Segment availability period.
- 18 • *NOW* is a time that is expressing the time on the content server as wall-clock time. All
19 information in the MPD related to wall-clock is expressed as a reference to the time *NOW*.

20 **4.3.2.2.3. MPD Information**

21 For a dynamic service without MPD updates, the following information shall be present and not
22 present in the MPD (also please refer to Table 8):

- 23 • The **MPD@type** shall be set to "dynamic".
- 24 • The **MPD@mediaPresentationDuration** shall be present, or the **Period@dura-**
25 **tion** of the last Period shall be present.
- 26 • The **MPD@minimumUpdatePeriod** shall not be present.

27 Furthermore, it is recommended to provide a value for **MPD@timeShiftBufferDepth** and
28 **MPD@suggestedPresentationDelay**.

1 **4.3.2.2.4. Period Information**

2 Each Period is documented by a **Period** element in the MPD. An MPD may contain one or more
3 Periods. In order to document the use of multiple Periods, the sequence of Period elements is
4 expressed by an index i with i increasing by 1 for each new Period element.

5 Each regular Period i in the MPD is assigned a

- 6 • Period start time $PSwc[i]$ in wall-clock time,
- 7 • Period end time $PEwc[i]$, in wall-clock time.

8 Note: An MPD update may extend the Period end time of the last regular Period. For details refer to section 4.4.

9 The Period start time $PSwc[i]$ for a regular Period i is determined according to section 5.3.2.1 of
10 ISO/IEC 23009-1:

- 11 • If the attribute `@start` is present in the **Period**, then $PSwc[i]$ is the sum of AST and the
12 value of this attribute.
- 13 • If the `@start` attribute is absent, but the previous **Period** element contains a `@duration`
14 attribute then the start time of the Period is the sum of the start time of the previous
15 Period $PSwc[i]$ and the value of the attribute `@duration` of the previous Period. Note
16 that if both are present, then the `@start` of the new Period takes precedence over the
17 information derived from the `@duration` attribute.

18 The Period end time $PEwc[i]$ for a regular Period i is determined as follows:

- 19 • If the Period is the last one in the MPD, the time $PEwc[i]$ is obtained as
 - 20 ○ the sum of AST and Media Presentation Duration $MPDur$, with $MPDur$ the value
21 of **MPD@mediaPresentationDuration** if present, or the sum of $PSwc[i]$ of
22 the last Period and the value of **Period@duration** of the last Period.
- 23 • else
 - 24 ○ the time $PEwc[i]$ is obtained as the Period start time of the next Period, i.e. $PEwc[i] = PSwc[i+1]$.

26 **4.3.2.2.5. Representation Information**

27 Based on such an MPD at a specific time NOW , a list of Segments contained in a Representation
28 in a Period i with Period start time $PSwc[i]$ and Period end time $PEwc[i]$ can be computed.

29 If the **SegmentTemplate**.**SegmentTimeline** is present and the **SegmentTemplate@duration**
30 is not present, the **SegmentTimeline** element contains N_s **S** elements
31 indexed with $s=1, \dots, N_s$, then let

- 32 • ts the value of the `@timescale` attribute
- 33 • ato is the sum of the values of the `@availabilityTimeOffset` attribute present on
34 on all levels that are processed in determining the URL for the corresponding segment, if
35 present at all. If not present, it is zero.
- 36 • $t[s]$ be the value of `@t` of the s -th **S** element,
- 37 • $d[s]$ be the value of `@d` of the s -th **S** element
- 38 • $r[s]$ be,
 - 39 ○ if the `@r` value is greater than or equal to zero

-
- 1 ▪ one more than the value of @r of the s -th **S** element. Note that if @r is
 2 smaller than the end of this segment timeline element, then this Representation
 3 contains gaps and no media is present for this gap.
 4 ○ else
 5 ▪ if $t[s+1]$ is present, then $r[s]$ is the ceil of $(t[s+1] - t[s])/d[s]$
 6 ▪ else $r[s]$ is the ceil of $(PEwc[i] - PSwc[i] - t[s]/ts)*ts/d[s])$

7 If the **SegmentTemplate** @duration is present and the **SegmentTemplate**.**SegmentTimeline** is not present, then
 8

- 9 • $N_s=1$,
 10 • ato is the sum of the values of the **@availabilityTimeOffset** attribute present on
 11 on all levels that are processed in determining the URL for the corresponding segment, if
 12 present at all. If not present, it is zero.
 13 • ts the value of the **@timescale** attribute
 14 • $t[s]$ is 0,
 15 • the $d[s]$ is the value of **@duration** attribute
 16 • $r[s]$ is the ceil of $(PEwc[i] - PSwc[i] - t[s]/ts)*ts/d[s])$

17 Note that the last segment may not exist and $r[s]$ is one less than this computation provides.
 18 For more details, refer to clause 4.4.3.6.

4.3.2.2.6. Media Time Information of Segment

20 Each Media Segment at position $k=1,2, \dots$ for each Representation has assigned an earliest media
 21 presentation time $EPT[k,r,i]$ and an accurate segment duration $SDUR[k,r,j]$, all measured in media
 22 presentation time.

23 The earliest presentation time may be estimated from the MPD using the segment availability start
 24 time minus the segment duration announced in the MPD.

25 The earliest presentation time may be accurately determined from the Segment itself.
 26

27 For details on the derivation of the earliest presentation time, see section 3.2.11.
 28

4.3.2.2.7. Segment List Parameters

29 For each Period i with Period start time $PSwc[i]$ and Period end time $PEwc[i]$ and each Representation r in the Period the following information can be computed:
 30

- 31 • the presentation time offset described in the MPD, $o[i,r]$
 32 • the availability time offset of this Representation, $ato[r]$
 33 • the number of the first segment described in the MPD, $k1[i,r]$
 34 • the number of the last segment described in the MPD, $k2[i,r]$
 35 • segment availability start time of the initialization segment $SAST[0,i,r]$
 36 • segment availability end time of the initialization segment $SAET[0,i,r]$
 37 • segment availability start time of each media segment $SAST[k,i,r]$, $k=k1, \dots, k2$
 38 • segment availability end time of each media segment $SAET[k,i,r]$, $k=k1, \dots, k2$

-
- 1 • adjusted segment availability start time ASAST[0,i,r], k=0, k1, ..., k2
 2 • segment duration of each media segment SD[k,i,r], k=k1, ..., k2
 3 • the URL of each of the segments, URL[k,i,r]
- 4 In addition,
- 5 • the latest available Period $i[NOW]$ and the latest segment available at the server $k[NOW]$
 6 can be computed. This segment is also referred to as *live edge segment*.
 7 • the earliest available Period $i^*[NOW]$ and the earliest segment available at the server
 8 $k^*[NOW]$ can be computed.
- 9 Based on the above information, for each Representation r in a Period i , the segment availability
 10 start time $SAST[k,i,r]$, the segment availability end time of each segment $SAET[k,i,r]$, the segment
 11 duration of each segment $SD[k,i,r]$, and the URL of each of the segments, $URL[k,i,r]$ within one
 12 Period i be derived as follows using the URL Template function $URLTemplate(Replace-$
 13 $mentString, Address)$ as documented in subsection 4.3.2.2.8:
- 14 • $k=0$
 15 • $SAST[0,i,r] = PSwc[i]$
 16 • $ASAST[0,i,r] = PSwc[i] - ato$
 17 • for $s=1, \dots N_s [i,r]$
 18 ○ $k = k + 1$
 19 ○ $SAST[k,i,r] = PSwc[i] + (t[s,i,r] + d[s,i,r] - o[i,r])/ts$
 20 ○ $ASAST[k,i,r] = SAST[k,i,r] - ato$
 21 ○ $SD[k,i,r] = d[s,i,r]/ts$
 22 ○ $SAET[k,i,r] = SAST[k,i,r] + TSB + d[s,i,r]/ts$
 23 ○ if **SegmentTemplate**@media contains \$Number\$
 24 ■ *Address*=@startNumber
 25 ■ $URL[k,i,r] = URLTemplate (\$Number\$, Address)$
 26 else
 27 ■ $Address = t[s,i,r]$
 28 ■ $URL[k,i,r] = URLTemplate (\$Time\$, Address)$
 29 ○ for $j = 1, \dots, r[s,i,r]$
 30 ■ $k = k + 1$
 31 ■ $SAST[k,i,r] = SAST[k-1,i,r] + d[s,i,r]/ts$
 32 ■ $ASAST[k,i,r] = SAST[k,i,r] - ato$
 33 ■ $SAET[k,i,r] = SAST[k,i,r] + TSB + d[s,i,r] /ts$
 34 ■ $SD[k,i,r] = d[s,i,r] /ts$
 35 ■ if **SegmentTemplate**@media contains \$Number\$
 36 ■ $Address = Address + 1$
 37 ■ $URL[k,i,r] = URLTemplate (\$Number\$, Address)$
 38 else
 39 ■ $Address = Address + d[s,i,r]$

-
- 1 ■ $URL[k,i,r] = \text{URLTemplate}(\$Time\$, Address)$
 2 • $k2[i,r] = k$
 3 • $SAET[0,i,r] = SAET[k2[i,r],i,r]$
- 4 Note that not all segments documented above may necessarily be accessible at time *NOW*, but only
 5 those that are within the segment availability time window.
 6 Hence, the number of the first media segment described in the MPD for this Period, $k1[i,r]$, is the
 7 smallest $k=1, 2, \dots$ for which $SAST[k,i,r] \geq NOW$.
 8 The latest available Period $i[NOW]$ is the Period i with the largest $PEwc[i]$ and $PEwc[i]$ is smaller
 9 than or equal to *NOW*.
 10 The latest available segment $k[NOW]$ available for a Representation of Period $i[NOW]$ (also the
 11 live edge segment) is the segment with the largest $k=0,1,2,\dots$ such that $SAST[k,i,r]$ is smaller than
 12 or equal to *NOW*. Note that this contains the Initialization Segment with $k=0$ as not necessarily any
 13 media segment may yet be available for Period $i[NOW]$. In this case, last media segment
 14 $k2[i[NOW]-1,r]$, i.e., the last media segment of the previous Period is the latest accessible media
 15 Segment.
 16 However, if the `@availabilityTimeOffset` is present, then the segments for this Represen-
 17 tation are available earlier than the nominal segment availability start time, namely at $ASAST[k,i,r]$.
- 4.3.2.8. URL Generation with Segment Template**
- 19 The function URL Template function `URLTemplate(ReplacementString, Address)` gen-
 20 erates a URL. For details refer to ISO/IEC 23009-1 [1], section 5.3.9.4. Once the Segment is gen-
 21 erated, processing of the Base URLs that apply on this segment level is done as defined in ISO/IEC
 22 23009-1, section 5.6.
 23 For the avoidance of doubt, only `%0 [width]d` is permitted and no other identifiers. The reason
 24 is that such a string replacement can be easily implemented without requiring a specific library.
- 4.3.2.9. Synchronized Playout and Seamless Switching**
- 26 In order to achieve synchronized playout across different Representations, typically from different
 27 Adaptation Sets, the different Representations are synchronized according to the presentation time
 28 in the Period. Specifically, the earliest presentation time of each Segment according to section
 29 4.3.2.2.6 determines the playout of the Segment in the Period and therefore enables synchronized
 30 playout of different media components as well as seamless switching within one media component.
- 4.3.3. Service Offering Requirements and Guidelines**
- 4.3.3.1. General Service Offering Requirements**
- 33 For dynamic service offerings, the MPD shall conform to DASH-IF IOP as defined in section 3
 34 and shall at least contain the mandatory information as documented in Table 8.
 35 If such an MPD is accessible at time *NOW* at the location **MPD . Location**, then
- 36 • all Segments for all Representations in all Periods as announced in an MPD shall be
 37 available latest at the announced segment availability start time $SAST[k,i,r]$ at all
 38 $URL[k,i,r]$ as derived in section 4.3.2.2;

- all Segments for all Representations in all Periods as announced in an MPD shall at least be available until the announced segment availability end time $SAET[k,i,r]$ at all $URL[k,i,r]$ as derived in section 4.3.2.2;
- for all Media Segments for all Representations in all Periods as announced in an MPD the Segment in this Period is available prior to the sum of Period start, earliest presentation time and segment duration, i.e. $SAST[k,i,r] \leq PSw[i] + SD[k,r,i] + EPT[k,r,i]$;
- if a Media Segments with segment number k is delivered over a constant bitrate channel with bitrate equal to value of the @bandwidth attribute then each presentation time PT is available at the client latest at time with a delay of at most $PT + MBT$.

4.3.3.2. Dynamic Service Offering Guidelines

4.3.3.2.1. Introduction

In order to offer a simple dynamic service for which the following details are known in advance,

- start at wall-clock time START,
- exact duration of media presentation PDURATION,
- location of the segments for each Representation at "`http://example.com/$RepresentationID$/$Number$`",

a service provider may offer an MPD as follows:

Table 9 – Basic Service Offering

MPD Information	Value
MPD@type	dynamic
MPD@availabilityStartTime	START
MPD@mediaPresentationDuration	PDURATION
MPD@suggestedPresentationDelay	SPD
MPD@minBufferTime	MBT
MPD@timeShiftBufferDepth	TSB
MPD.BaseURL	" <code>http://example.com/</code> "
Period@start	PSTART
Representation@bandwidth	BW
SegmentTemplate@media	" <code>\$RepresentationID\$/\$Number\$</code> "
SegmentTemplate@startNumber	1
SegmentTemplate@duration	SDURATION

Note that the setting of capitalized parameters is discussed in section 4.3.3.2.2.

According to the work-flow shown in Annex B:

- the MPD is generated and published prior to time START such that DASH clients may access it prior to the start of the Media Presentation.
- no redundant tools are considered.

-
- 1 • the encoder and the segmenter generate segments of duration SDURATION and publish
2 those on the origin server, such that they are available at URL[k] latest at their announced
3 segment availability start time SAST[k].

4 Based on the details in section 4.3.2.2, the Segment Information is derived as:

- 5 • k1 = 1
6 • k2 = ceil (PDURATION / SDURATION)
7 • for k = 1, ..., k2
8 ○ SAST[k] = START + PSTART + k * SDURATION
9 ○ SAET[k] = SAST[k] + TSB + SDURATION
10 ○ SD[k] = SDURATION
11 ○ URL[k] = http://example.com/\$RepresentationID\$/k
12 • The segment availability times of the Initialization Segment are as follows:
13 ○ SAST[0] = START + PSTART
14 ○ SAET[0] = SAET[k2]

15 **4.3.3.2.2. Basic Parameter Settings**

16 In the following recommendations are provided for the

- 17 • Time Shift Buffer Depth (TSB):
18 ○ If the content should be consumed at the live edge, then the time shift buffer depth
19 should be set short. However, the TSB should not be smaller than the recommended
20 value of 4 * SDURATION and 6 seconds in media time in order for the client to do
21 some prebuffering in more difficult network conditions.
22 ○ If no restrictions on the accessibility of the content are provided, then the TSB may
23 be set to a large value that even exceeds PDURATION.
- 24 • Suggested Presentation Delay (SPD)
25 ○ If synchronized play-out with other devices adhering to the same rule is desired
26 and/or the service provider wants to define the typical live edge of the program,
27 then this value should be provided. The service provider should set the value taking
28 into account at least the following:
29 ■ the desired end-to-end latency
30 ■ the typical required buffering in the client, for example based on the net-
31 work condition
32 ■ the segment duration SDURATION
33 ■ the time shift buffer depth TSB
34 ○ A reasonable value may be 2 to 4 times of the segment duration SDURATION, but
35 the time should not be smaller than 4 seconds in order for the client to maintain
36 some buffering.
- 37 • Segment Duration (SDURATION)
38 ○ The segment duration typically influences the end-to-end latency, but also the
39 switching and random access granularity as in DASH-264/AVC each segment
40 starts with a stream access point which can also be used as a switch point. The
41 service provider should set the value taking into account at least the following:
42 ■ the desired end-to-end latency

- the desired compression efficiency
 - the start-up latency
 - the desired switching granularity
 - the desired amount of HTTP requests per second
 - the variability of the expected network conditions
 - Reasonable values for segment durations are between 1 second and 10 seconds.
 - Minimum Buffer Time (MBT) and bandwidth (BW)
 - the value of the minimum buffer time **does not provide any instructions to the client on how long to buffer the media**. This aspect is covered in 4.3.4.4. The value describes how much buffer a client should have under **ideal** network conditions. As such, MBT is not describing the burstiness or jitter in the network, it is describing the burstiness or jitter in the **content encoding**. Together with the BW value, it is a property of the content. Using the "leaky bucket" model, it is the size of the bucket that makes BW true, given the way the content is encoded.
 - The minimum buffer time provides information that for each Stream Access Point (and in the case of DASH-IF therefore each start of the Media Segment), the property of the stream: If the Representation (starting at any segment) is delivered over a constant bitrate channel with bitrate equal to value of the BW attribute then each presentation time *PT* is available at the client latest at time with a delay of at most $PT + MBT$.
 - In the absence of any other guidance, **the MBT should be set** to the maximum GOP size (coded video sequence) of the content, which quite often is identical **to the maximum segment duration**. The *MBT* may be set to a smaller value than maximum segment duration, but should not be set to a higher value.

In a simple and straightforward implementation, a DASH client decides downloading the next segment based on the following status information:

- the currently available buffer in the media pipeline, $buffer$
 - the currently estimated download rate, $rate$
 - the value of the attribute `@minBufferTime`, MBT
 - the set of values of the `@bandwidth` attribute for each Representation i. $BW[i]$

The task of the client is to select a suitable Representation i .

The relevant issue is that starting from a SAP on, the DASH client can continue to playout the data. This means that at the current time it does have *buffer* data in the buffer. Based on this model the client can download a Representation i for which $BW[i] \leq rate * buffer / MBT$ without emptying the buffer.

Note that in this model, some idealizations typically do not hold in practice, such as constant bitrate channel, progressive download and playout of Segments, no blocking and congestion of other HTTP requests, etc. Therefore, a DASH client should use these values with care to compensate such practical circumstances; especially variations in download speed, latency, jitter, scheduling of requests of media components, as well as to address other practical circumstances.

One example is if the DASH client operates on Segment granularity. As in this case, not only parts of the Segment (i.e., MBT) needs to be downloaded, but the entire Segment, and if the MBT is smaller than the Segment duration, then rather the segment duration needs to be used instead of

1 the MBT for the required buffer size and the download scheduling, i.e. download a Representation
2 i for which $BW[i] \leq rate * buffer / max_segment_duration$.

3 For low latency cases, the above parameters may be different.

4 **4.3.3.2.3. Example**

5 Assume a simple example according to Table 12.

6 **Table 10 – Basic Service Offering**

MPD Information	Value
MPD@type	dynamic
MPD@availabilityStartTime	START
MPD@mediaPresentationDuration	43sec
MPD@suggestedPresentationDelay	15sec
MPD@minBufferTime	5sec
MPD@timeShiftBufferDepth	25sec
MPD.BaseURL	"http://example.com/"
Period@start	0
SegmentTemplate@media	"\$RepresentationID\$/\$Number\$"
SegmentTemplate@startNumber	1
SegmentTemplate@duration	5sec

7 Based on the derivation in section 4.3.3.2.1, the following holds:

- 8 • $k_1 = 1, k_2 = 9$
- 9 • for $k = 1, \dots, k_2$
 - 10 ○ $SAST[k] = START + k * 5\text{sec}$
 - 11 ○ $SAET[k] = SAST[k] + 30\text{sec}$
 - 12 ○ $URL[k] = http://example.com/1/k$
- 13 • The segment availability times of the Initialization Segment are as follows:
 - 14 ○ $SAST[0] = START$
 - 15 ○ $SAET[0] = START + 75 \text{ sec}$

16 Figure 5 shows the availability of segments on the *server* for different times *NOW*. In particular,
17 before START no segment is available, but the segment URLs are valid. With time *NOW* advanc-
18 ing, segments get available.



Figure 5 Segment Availability on the Server for different time NOW (blue = valid but not yet available segment, green = available Segment, red = unavailable Segment)

4.3.3.3. Content Offering with Periods

4.3.3.3.1. General

For content offered within a Period, and especially when offered in multiple Periods, then the content provider should offer the content such that actual media presentation time is as close as possible to the actual Period duration. It is recommended that the Period duration is the maximum of the presentation duration of all Representations contained in the Period.

A typical Multi-Period Offering is shown in Table 11. This may for example represent a service offering where main content provided in Period 1 and Period 3 are interrupted by an inserted Period 2.

Table 11 Multi-Period Service Offering

MPD Information	Value
MPD@type	dynamic
MPD@availabilityStartTime	START
MPD@mediaPresentationDuration	PDURATION
MPD@suggestedPresentationDelay	SPD
MPD@minBufferTime	MBT
MPD@timeShiftBufferDepth	TSB
MPD .BaseUrl	"http://example.com/"
Period@start	PSTART
SegmentTemplate@media	"1/\$RepresentationID\$/\$Number\$"

SegmentTemplate@startNumber	1
SegmentTemplate@duration	SDURATION1
Period@start	PSTART2
Representation@availabilityTimeOffset	ATO2
SegmentTemplate@media	"2/\$RepresentationID\$/Number\$"
SegmentTemplate@startNumber	1
SegmentTemplate@duration	SDURATION2
Period@start	PSTART3
SegmentTemplate@media	"1/\$RepresentationID\$/Number\$"
SegmentTemplate@startNumber	STARTNUMBER2
SegmentTemplate@duration	SDURATION1
SegmentTemplate@presentationTimeOffset	PTO

- 1 The work flow for such a service offering is expected to be similar to the one in section 4.3.2.2.1.
 2 Based on the details in section 4.3.2.2, the Segment Information is derived as:

- 3 • Period 1
 - 4 ○ PSwc[1] = START + PSTART
 - 5 ○ PEwc[1] = START + PSTART2
 - 6 ○ k1 = 1
 - 7 ○ k2 = ceil ((PSTART2 - PSTART1) / SDURATION)
 - 8 ○ for k = 1, ..., k2
 - 9 ▪ SAST[k] = PSwc[1] + k*SDURATION
 - 10 ▪ SAET[k] = SAST[k] + TSB + SDURATION
 - 11 ▪ SD[k] = SDURATION
 - 12 ▪ URL[k] = http://example.com/1/\$RepresentationID\$/k
 - 13 ○ SAST[0] = PSwc[1]
 - 14 ○ SAET[0] = SAET[k2]
- 15 • Period 2
 - 16 ○ PSwc[2] = START + PSTART2
 - 17 ○ PEwc[2] = START + PSTART3
 - 18 ○ k1 = 1
 - 19 ○ k2 = ceil ((PSTART3 - PSTART2) / SDURATION2)
 - 20 ○ for k = 1, ..., k2
 - 21 ▪ SAST[k] = PSwc[2] + k*SDURATION2
 - 22 ▪ ASAST[k] = SAST[k] - ATO2
 - 23 ▪ SAET[k] = SAST[k] + TSB + SDURATION2
 - 24 ▪ SD[k] = SDURATION2
 - 25 ▪ URL[k] = http://example.com/2/\$RepresentationID\$/k
 - 26 ○ SAST[0] = PSwc[2]
 - 27 ○ SAET[0] = SAET[k2]

-
- 1 • Period 3
- 2 ○ PSwc[3] = START + PSTART3
- 3 ○ PEwc[3] = START + PDURATION
- 4 ○ k1 = 1
- 5 ○ k2 = ceil((PDURATION - PSTART3) / SDURATION1)
- 6 ○ for k = 1, ..., k2
- 7 ■ SAST[k] = PSwc[3] + k * SDURATION1
- 8 ■ SAET[k] = SAST[k] + TSB + SDURATION1
- 9 ■ SD[k] = SDURATION1
- 10 ■ URL[k] = "http://example.com/1/\$RepresentationID\$/(k+STARTNUMBER2-1)"
- 11
- 12 ○ SAST[0] = PSwc[3]
- 13 ○ SAET[0] = SAET[k2]

14 Note that the number k describes position in the Period. The actual number used in the segment
15 template increased by the one less than the actual start number.

16 In order to ensure that the attribute **Period@start** can accurately document the duration of the
17 previous Period and to avoid that the player may fall into a loop searching for a Segment in the
18 wrong Period, it is recommended to accurately document the **Period@start** time. In order to
19 fulfill this, it is recommended to use video track time scale to document the exact duration of the
20 Period. A media time scale of at most 90 kHz is recommended and may be represented by the
21 xs:duration type of **Period@start**.

22 **4.3.3.3.2. Continuous Period Offering**

23 Continuous Period offering as defined in section 3.2.12 may be used. If multiple periods are of-
24 fered primarily for robustness or MPD changes, the continuous period should be used by the con-
25 tent author to provide seamless experiences for the user. If the condition of continuous timelines
26 is not fulfilled, but all other conditions, then period-connectivity may be used as defined in section
27 3.2.12.

28 **4.3.3.4. Content Offering with Segment Timeline**

29 **4.3.3.4.1. Basic Operation**

30 In order to offer a dynamic service that takes into account

- 31 • variable segment durations
- 32 • gaps in the segment timeline of one Representation,

33 the Segment timeline as defined in ISO/IEC 23009-1, section 5.3.9.6 may be used as an alternative
34 to the @duration attribute as shown in section 4.3.3.2.

35 **Table 12 – Service Offering with Segment Timeline**

MPD Information	Value
MPD@type	dynamic
MPD@availabilityStartTime	START
MPD@mediaPresentationDuration	PDURATION

MPD@suggestedPresentationDelay	SPD
MPD@minBufferTime	MBT
MPD@timeShiftBufferDepth	TSB
MPD.BaseURL	"http://example.com/"
Period@start	PSTART
SegmentTemplate@media	"\$RepresentationID\$/Number\$"
SegmentTemplate@startNumber	1
SegmentTemplate.SegmentTimeline	t[i], n[i], d[i], r[i]

1 According to the work-flow shown in Annex B:

- 2 the MPD is generated and published prior to time START such that DASH clients may
3 access it prior to the start of the Media Presentation.
- 4 no redundant tools are considered.
- 5 the encoder and the segmenter generally should generate segments of constant duration
6 SDURATION and publish those on the origin server, such that they are available at URL[k]
7 latest at their announced segment availability start time SAST[k]. However, the server may
8 offer occasional shorter segments for encoding optimizations, e.g. at scene changes, or
9 segment gaps (for details see section 6). If such an irregular segment is published the MPD
10 needs to document this by a new **S** element in the segment timeline.

11 If the segment timeline is used and the \$Time\$ template is used, then the times in the MPD shall
12 accurately present media internal presentation times.

13 If the segment timeline is and the \$Number\$ template is used, then the MPD times shall at most
14 deviate from the earliest presentation time documented in the MPD by 0.5sec.

15 Based on these considerations, it is not feasible to operate with a single MPD if the content is not
16 yet known in advance. However, pre-prepared content based on the segment timeline may be of-
17 fered in a dynamic fashion. The use of the Segment Timeline is most suitable for the case where
18 the MPD can be updated. For details refer to section 4.4.

19 **4.3.3.4.2. Basic Parameter Settings**

20 The parameters for TSB and SPD should be set according to section 4.3.3.2.2. The segment dura-
21 tion SDURATION may be set according to section 4.3.3.2.2, but it should be considered that the
22 service provider can offer shorter segments occasionally.

23 **4.3.3.5. Joining Recommendation**

24 By default, an MPD with **MPD@type="dynamic"** suggests that the client would want to join
25 the stream at the live edge, therefore to download the latest available segment (or close to, depend-
26 ing on the buffering model), and then start playing from that segment onwards.

27 However there are circumstances where a dynamic MPD might be used with content intended for
28 playback from the start, or from another position. For example, when a content provider offers
29 ‘start again’ functionality for a live program, the intention is to make the content available as an
30 on-demand program, but not all the segments will be available immediately.

- 1 This may be signalled to the DASH client by including an MPD Anchor, with either
 2 • the `t` parameter, or
 3 • both the `period` and `t` parameter, in the MPD URL provided to the DASH client, or
 4 • the `POSIX` parameter, for details refer to Amd.3 of ISO/IEC 23009-1:2014 [4][4].

5 The format and behaviour of MPD Anchors is defined in section C.4 of ISO/IEC 23009-1. Specifically the `POSIX` parameter is defined in Amd.3 of ISO/IEC 23009-1:2014 [4].

7 For example to start from the beginning of the MPD the following would be added to the end of
 8 the MPD URL provided to the DASH client:

9 `#t=0`

10 Or to start from somewhere other than the start, in this case 50 minutes from the beginning of the
 11 period with Period ID “`program_part_2`”:

12 `#period=program_part_2&t=50:00`

13 Starting from a given UTC time can be achieved using the `POSIX` clock with `t` parameter. For
 14 example, starting playback from Wed, 08 Jun 2016 17:29:06 GMT would be expressed as

15 `#t=posix:1465406946`

16 `#t=posix:now` stands for “live edge”

17 Notes:

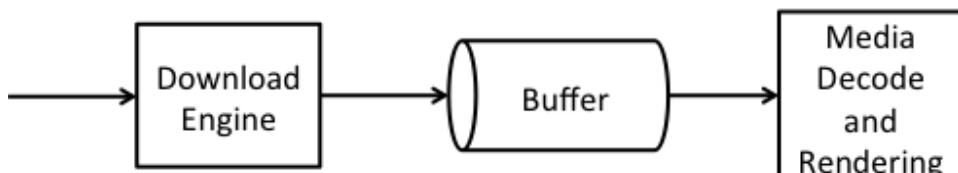
- 18 • as per section C.4 of ISO/IEC 23009-1 the time indicated using the `t` parameter is as per
 19 the field definition of the W3C Media Fragments Recommendation v1.0 section 4.2.1.
- 20 • the period ID has to be URL encoded/decoded as necessary and needs to match one of the
 21 `Period@id` fields in the MPD.

22 Where an MPD Anchor is used it should refer to a time for which segments are currently available
 23 in the MPD.

24 **4.3.4. Client Operation, Requirements and Guidelines**

25 **4.3.4.1. Basic Operation for Single Period**

26 A DASH client is guided by the information provided in the MPD. A simple client model is shown
 27 in Figure 6.



28
 29 **Figure 6 Simple Client Model**

30 Assume that the client has access to an MPD and the MPD contains the parameters in Table 8, i.e.
 31 it consumes a dynamic service with fixed media presentation duration.

1 In addition in the following for simplicity it is assumed that the MPD only contains a single Period
2 with period start time $PSwc[i]$ and the MPD-URL does not include any fragment parameters ac-
3 cording to section 4.3.3.5.

4 The following example client behavior may provide a continuous streaming experience to the user:

- 5 1) The client parses the MPD, selects a collection of Adaptation Sets suitable for its environ-
6 ment based on information provided in each of the **AdaptationSet** elements.
- 7 2) Within each Adaptation Set it selects one Representation, typically based on the value of
8 the @bandwidth attribute, but also taking into account client decoding and rendering
9 capabilities.
- 10 3) The client creates a list of accessible Segments at least for each selected Representation
11 taking into account the information in the MPD as documented in Table 8 and the current
12 time *JOIN* in the client and in particular the segment closest to the live edge referred to the
13 *live edge segment*. For details refer to section 4.3.4.2.
- 14 4) The client downloads the initialization segment of the selected Representations and then
15 accesses the content by requesting entire Segments or byte ranges of Segments. Typically
16 at any time the client downloads the next segment at the larger of the two: (i) completion
17 of download of current segment or (ii) the Segment Availability Start Time of the next
18 segment. Note that if the @availabilityTimeOffset is present, then the segments
19 may be downloaded earlier, namely at the adjusted segment availability start time. Based
20 on the buffer fullness and other criteria, rate adaptation is considered. Typically the first
21 media segment that is downloaded is the *live edge segment*, but other decisions may be
22 taken in order to minimize start-up latency. For details on initial buffering, refer to section
23 4.3.4.4.
- 24 5) According to Figure 6 media is fed into buffer and at some point in time, the decoding and
25 rendering of the media is kicked off. The downloading and presentation is done for the
26 selected Representation of each selected Adaptation. The synchronization is done using
27 the presentation time in the Period as documented in section 4.3.2.2.9. For synchronized
28 playout, the exact presentation times in the media shall be used.
29 Once presentation has started, the playout process is continuous. The playout process ex-
30 pects media to be present in the buffer continuously. If the **MPD**@suggestedPresen-
31 tationDelay is present, then this value may be used as the presentation delay *PD*. If
32 the **MPD**@suggestedPresentationDelay is not present, but the client is expected
33 to consume the service at the live edge, then a suitable presentation delay should be se-
34 lected, typically between the value of @minBufferTime and the value of
35 @timeShiftBufferDepth. It is recommended that the client starts rendering the first
36 sample of the downloaded media segment *k* with earliest presentation time *EPT(k)* at
37 $PSwc[i] + (EPT(k) - o[r,i]) + PD$. For details on selecting and minimizing end-to-end la-
38 tency as well as the start-up latency, see section 4.3.4.4.
- 39 6) The client may request Media Segments of the selected Representations by using the gen-
40 erated Segment list during the availability time window.

-
- 1 7) Once the presentation has started, the client continues consuming the media content by
 2 continuously requesting Media Segments or parts of Media Segments and playing content
 3 that according to the media presentation timeline. The client may switch Representations
 4 taking into updated information from its environment, e.g. change of observed throughput.
 5 In a straight-forward implementation, with any request for a Media Segment starting with
 6 a stream access point, the client may switch to a different Representation. If switching at
 7 a stream access point, the client shall switch seamlessly at such a stream access point.
 8 8) With the wall-clock time *NOW* advancing, the client consumes the available Segments. As
 9 *NOW* advances the client possibly expands the list of *available* Segments for each Repre-
 10 sentation in the Period according to the procedures specified in 4.3.4.2.
 11 9) Once the client is consuming media contained in the Segments towards the end of the
 12 announced media in the Representation, then either the Media Presentation is terminated,
 13 a new Period is started (see subsection 4.3.4.3) or the MPD needs to be refetched. For
 14 details on MPD updates and refetching, please refer to section 4.4.

15 **4.3.4.2. Determining the Segment List**

16 For a single Period content the client determines the available Segment List at time *NOW* according
 17 to section 4.3.2.2.7 taking into account the simplified offering in Table 9 as

- 18 • $k_1 = 1$
- 19 • $k_2 = \text{ceil}(\text{PDURATION}/\text{SDURATION})$
- 20 • $\text{SAST}[k] = \text{START} + \text{PSTART} + k * \text{SDURATION}$ for $k = 0, 1, \dots, k_2$
- 21 • $\text{ASAST}[k] = \text{SAST}[k] - \text{ATO}$
- 22 • $\text{SAET}[k] = \text{SAST}[k] + \text{TSB} + \text{SDURATION}$ for $k = 1, \dots, k_2$
- 23 • $\text{SAET}[0] = \text{SAET}[k_2]$
- 24 • $\text{SD}[k] = \text{SDURATION}$
- 25 • $\text{URL}[k] = \text{http://example.com/$RepresentationID$}/k$
- 26 • $k[\text{NOW}] = \text{MIN}(\text{floor}((\text{NOW} - \text{START} - \text{PSTART})/\text{SDURATION}), k_2)$
- 27 • $k^*[\text{NOW}] = \text{MAX}(k_1, \text{floor}((\text{NOW} - \text{START} - \text{PSTART} - \text{TSB})/\text{SDURATION}))$

28 Note that if $k[\text{NOW}]$ is 0, then only the Initialization Segment is available. The live edge segment
 29 if provided as $k[\text{NOW}]$. If the `@availabilityTimeOffset` is present, then the segments for
 30 this Representation may be downloaded from $\text{ASAST}[k]$ onwards.

31 **4.3.4.3. Multi-Period Content**

32 In an extension to the description in section 4.3.4.1 assume now that the client has access to an
 33 MPD and the MPD contains content with multiple Periods, for example following the parameters
 34 in Table 11. The start time of each Period is computed as period start time $\text{PSwC}[i]$. and the MPD-
 35 URL does not include any fragment parameters according to section 4.3.3.5.

36 In an extension of bullet 3 in section 4.3.4.1,

37 the client creates a list of accessible Segments at least for each selected Representation
 38 taking into account the information in the MPD as documented in Table 8 and the current
 39 time *NOW* in the client and in particular the segment closest to the live edge referred to the
 40 *live edge segment*.

For this it needs to take into account the latest Period $i[NOW]$. The latest Period and the latest segment are obtained as follows with i^* the index of the last Period.:

- if $NOW \leq PSwc[1]$
 - no segment is yet available
- else if $NOW > PSwc[i^*]$
 - the last one and the latest segment is available is $k2[i^*]$
- else if $NOW > PSwc[i^*] + TSB$
 - no segment is available any more
- else if $PSwc[1] < NOW \leq PEwc[i^*]$
 - i' the such that $PSwc[i'] < NOW \leq PEwc[i']$
 - $k[NOW] = \text{MIN}(\text{fLOOR}((NOW - PEwc[i'] - PSwc[i']) / SDURATION[i']), k2)$
- Note again that if $k[NOW]$ is 0, then only the Initialization Segment is available. If the Period is not the first one, then the last available Media Segment is the last Media Segment of the previous Period.

In an extension of bullet 9 in section 4.3.4.1,

the client consumes media in one Period. Once the client is consuming media contained in the Segments towards the end of the announced media in the Representation, and the Representation is contained not in the last Period, then the DASH clients generally needs to reselect the Adaptation Sets and a Representation in same manner as described in bullet 1 and 2 in section 4.3.4.1. Also steps 3, 4, 5 and 6 need to be carried out at the transition of a Period. Generally, audio/video switching across period boundaries may not be seamless. According to ISO/IEC 23009-1, section 7.2.1, at the start of a new Period, the playout procedure of the media content components may need to be adjusted at the end of the preceding Period to match the *PeriodStart* time of the new Period as there may be small overlaps or gaps with the Representation at the end of the preceding Period. Overlaps (respectively gaps) may result from Media Segments with actual presentation duration of the media stream longer (respectively shorter) than indicated by the Period duration. Also in the beginning of a Period, if the earliest presentation time of any access unit of a Representation is not equal to the presentation time offset signalled in the *@presentation-TimeOffset*, then the playout procedures need to be adjusted accordingly.

The client should play the content continuously across Periods, but there may be implications in terms of implementation to provide fully continuous and seamless playout. It may be the case that at Period boundaries, the presentation engine needs to be reinitialized, for example due to changes in formats, codecs or other properties. This may result in a re-initialization delay. Such a re-initialization delay should be minimized. If the Media Presentation is of type dynamic, the addition of the re-initialisation delay to the playout may result in drift between the encoder and the presentation engine. Therefore, the playout should be adjusted at the end of each Period to provide a continuous presentation without adding drift between the time documented in the MPD and the actual playout, i.e. the difference between the actual playout time and the Period start time should remain constant.

If the client presents media components of a certain Adaptation Set in one Period, and if the following Period has assigned an identical Asset Identifier, then the client should identify an associated Period and, in the absence of other information, continue playing the content in the associated Adaptation Set.

If furthermore the Adaptation Sets are period-continuous, i.e. the presentation times are continuous and this is signalled in the MPD, then the client shall seamlessly play the content across the Period boundary under the constraints in section 4.3.3.3.2. The presentation time offset should be ignored. Most suitably the client may continue playing the Representation in the Adaptation Set with the same @id, but there is no guarantee that this Representation is available. In this case the client shall seamlessly switch to any other Representation in the Adaptation Set.

If otherwise the Adaptation Sets are period-connected and this is signaled in the MPD, then the client should avoid re-initializing media decoders. The client should inform the media decoder on a timeline discontinuity obeying the value of @presentationTimeOffset attribute, but it may continue processing the incoming Segments without re-initializing the media decoder. The presentation time offset should be used to seamlessly play the content across the Period boundary under the constraints in section 4.3.3.3.2.

4.3.4.4. Joining, Initial Buffering and Playout Recommendations

4.3.4.4.1. General

A DASH client should start playout from:

- The time indicated by the MPD Anchor, if one is present
- The live edge, if there is no MPD Anchor and **MPD@type="dynamic"**.

4.3.4.4.2. Joining at the live edge

For joining at the live edge there are basically two high-level strategies:

- Every client participating in the service commits to the same presentation delay (PD) relative to the announced segment availability start time at start-up and in continuous presentation, possibly using one suggested by the Content Provider and then attempts to minimize start-up latency and maintain the buffer. The content provider may have provided the **MPD@suggestedPresentationDelay** (SPD) or may have provided this value by other means outside the DASH formats. The content author should be aware that the client may ignore the presence of **MPD@suggestedPresentationDelay** and may choose its own suitable playout scheduling.
- The client individually picks the presentation delay (PD) in order to maximize stable quality and does this dependent on its access, user preferences and other considerations.

In both cases the client needs to decide, which segment to download first and when to schedule the playout of the segment based on the committed PD.

A DASH client would download an available segment and typically render the earliest presentation time EPT(k) of the segment at $PSwc[i] + (EPT(k) - o[r,i]) + PD$. As PD may be quite large, for example in order to provision for downloading in varying bitrate conditions, and if a segment is downloaded that was just made available it may result in larger start up delay.

1 Therefore, a couple of strategies may be considered as a tradeoff of for start-up delay, presentation
2 delay and sufficient buffer at the beginning of the service, when joining at the live edge:

- 3 1. The client downloads the next available segment and schedules playout with delay PD.
4 This maximizes the initial buffer prior to playout, but typically results in undesired long
5 start-up delay.
- 6 2. The client downloads the latest available segment and schedules playout with delay PD.
7 This provides large initial buffer prior to playout, but typically results in undesired long
8 start-up delay.
- 9 3. The client downloads the earliest available segment that can be downloaded to schedules
10 playout with delay PD. This provides a smaller initial prior to playout, but results in rea-
11 sonable start-up delay. The buffer may be filled gradually by downloading later segments
12 faster than their media playout rate, i.e. by initially choosing Representations that have
13 lower bitrate than the access bandwidth.

14 In advanced strategies the client may apply also one or more of the following:

- 15 1. Actual rendering may start not with the sample of the earliest presentation time, but the
16 one that matches as closely as possible $PSwc[i] + (PT - o[r,i]) + PD$ equal to NOW.
- 17 2. The client may start rendering even if only a segment is downloaded partially.

18 Also if the @availabilityTimeOffset is present and the segment has an adjusted segment
19 availability start time, then the segments may be downloaded earlier.

20 **4.3.4.5. Requirements and Recommendations**

21 In summary, a client that access a dynamic MPD shall at least obey the following rules:

- 22 • The client shall be able to consume single Period and multi-Period content
- 23 • If multi-period content is offered in a seamless manner, the client shall play seamlessly
24 across Period boundaries.

25 **4.3.5. Additional DVB-DASH alignment aspects**

26 For alignment with DVB-DASH [42], the following should be considered:

- 27 • Reasonable requirements on players around responding to response codes are provided in
28 DVB DASH in section 10.8.6.
- 29 • Further guidelines on live edge aspects are provided in DVB DASH section 10.9.2.

30 DVB DASH also provides recommendations in order to apply weights and priorities to different
31 networks in a multi Base URL offering in section 10.8.2.1.

32 **4.3.6. Considerations on live edge**

33 Detecting the live edge segment in DASH as well as providing a sanity check for the MPD author
34 on the correctness of the offering may be achieved for example by the following means:

- 35 • If the MPD contains a @publishTime attribute with value PUBT, then at the publication
36 of the MPD all Segments according to the computation in section 4.3.4.2 and 4.3.4.3 with
37 NOW set to PUBT shall be available.

-
- 1 • If the MPD contains a `@publishTime` attribute with value `PUBT` and a Representation
2 contains a Segment timeline with the `@r` attributed of the last `s` element being non-nega-
3 tive, then the last Segment describe in this Segment timeline shall have a Segment availa-
4 bility start time smaller than `PUBT` and the sum of the segment duration and the segment
5 availability start time shall be larger than `PUBT`.

6 A DASH client should avoid being too aggressive in requesting segments exactly at the computed
7 segment availability start time, especially if it is uncertain to be fully synchronized with the server.
8 If the DASH client observes issues, such as 404 responses, it should back up slightly in the re-
9 quests.

10 In addition, for a content authoring to avoid too aggressive requests and possible 404 responses,
11 the content author may schedule the segment availability start time in the MPD with a small safety
12 delay compared to the actual publish time. This also provides the content author a certain amount
13 of flexibility in the publishing of Segments. However, note that such safety margins may lead to
14 slightly increased end-to-end latencies, so it is a balance to be taken into account.

15 **4.4. Simple Live Service Offering including MPD Updates**

16 **4.4.1. Background and Assumptions**

17 If many cases, the service provider cannot predict that an MPD that is once offered, may be used
18 for the entire Media Presentations. Examples for such MPD changes are:

- 19 • The duration of the Media Presentation is unknown
20 • The Media Presentation may be interrupted for advertisements which requires proper splicing
21 of data, for example by adding a Period
22 • Operational issues require changes, for example the addition or removal of Representations
23 or Adaptation Sets.
24 • Operational problems in the backend, for example as discussed in section 4.8.
25 • Changes of segment durations, etc.

26 In this case the MPD typically only can describe a limited time into the future. Once the MPD
27 expires, the service provider expects the client to recheck and get an updated MPD in order to
28 continue the Media Presentation.

29 The main tool in MPEG-DASH is Media Presentation Description update feature as described in
30 section 5.4 of ISO/IEC 23009-1. The MPD is updated at the server and the client is expected to
31 obtain the new MPD information once the determined Segment List gets to an end.

32 If the MPD contains the attribute `MPD@minimumUpdatePeriod`, then the MPD in hand will
33 be updated.

34 According to the clustering in section 4.2, we distinguish two different types of live service offer-
35 ings:

- 36 • MPD controlled live service offering: In this case the DASH client typically frequently
37 polls the MPD update server whether an MPD update is available or the existing MPD can
38 still be used. The update frequency is controlled by MPD based on the attribute `MPD@min-`
39 `imumUpdatePeriod`. Such a service offering along with the client procedures is shown
40 in section 4.4.2.

- 1 • MPD and segment controlled offerings. In this case the DASH client needs to parse seg-
 2 ments in order to identify MPD validity expirations and updates on the MPD update server.
 3 MPD expiry events as described in section 5.10 of ISO/IEC 23009-1 "are pushed" to the
 4 DASH client as parts of downloaded media segments. This offering along with the client
 5 procedures is shown in section 4.5.

6 This section describes the first type of offering. In section 4.5 the MPD and segment controlled
 7 offerings are described. Under certain circumstances a service offering may be provided to both
 8 types of clients. An overview how such a service offering may be generated is shown in Annex A.

9 **4.4.2. Preliminaries**

10 **4.4.2.1. MPD Information**

11 As the MPD is typically updated over time on the server, the MPD that is accessed when joining
 12 the service as well as the changes of the MPD are referred to as MPD instances in the following.

13 This expresses that for the same service, different MPDs exist depending on the time when the
 14 service is consumed.

15 Assume that an MPD instance is present on the DASH server at a specific wall-clock time *NOW*.
 16 For an MPD-based Live Service Offering, the MPD instance may among others contain infor-
 17 mation as available in Table 13. Information included there may be used to compute a list of an-
 18 nounced Segments, Segment Availability Times and URLs.

19 **Table 13 – Information related to Live Service Offering with MPD-controlled MPD Updates**

MPD Information	Status	Comment
MPD@type	mandatory, set to "dynamic"	the type of the Media Presen-tation is dynamic, i.e. Seg-ments get available over time.
MPD@availabilityStartTime	mandatory	the start time is the anchor for the MPD in wall-clock time. The value is denoted as <i>AST</i> .
MPD@minimumUpdatePeriod	mandatory	this field is mandatory except for the case where the MPD@mediaPresenta-tionDuration is present. However, such an MPD falls then in an instance as docu-mented in section 4.3.
Period@start	mandatory	the start time of the Period relative to the MPD availabil-ity start time. The value is de-noted as <i>PS</i> .
SegmentTemplate@media	mandatory	the template for the Media Segment

SegmentTemplate@startNumber	optional default	the number of the first segment in the Period. The value is denoted as <i>SSN</i> .
SegmentTemplate@duration	exactly one of SegmentTemplate@duration or SegmentTemplate.SegmentTimeline	the duration of each Segment in units of a time. The value divided by the value of @timescale is denoted as <i>MD[k]</i> with <i>k</i> =1, 2, ... The segment timeline may contain some gaps.

1 **4.4.2.2. Segment Information Derivation**

2 Based on an MPD instance including information as documented in Table 13 and available at time
3 *NOW* on the server, a DASH client may derive the information of the list of Segments for each
4 Representation in each Period.

5 If the Period is the last one in the MPD and the **MPD@minimumUpdatePeriod** is present, then
6 the time *PEwc[i]* is obtained as the sum of *NOW* and the value of **MPD@minimumUpdate-**
7 **Period**.

8 Note that with the MPD present on the server and *NOW* progressing, the Period end time is ex-
9 tended. This issue is the only change compared to the segment information generation in section
10 4.3.2.2.

11 **4.4.2.3. Some Special Cases**

12 If the **MPD@minimumUpdatePeriod** is set to 0, then the MPD documents all available seg-
13 ments on the server. In this case the @r count may be set accurately as the server knows all avail-
14 able information.

15 **4.4.3. Service Offering Requirements and Guidelines**

16 **4.4.3.1. General**

17 The same service requirements as in section 4.3.3.1 hold for any time *NOW* the MPD is present
18 on the server with the interpretation that the Period end time *PEwc[i]* of the last Period is obtained
19 as the sum of *NOW* and the value of **MPD@minimumUpdatePeriod**.

20 In order to offer a simple live service with unknown presentation end time, but only a single Period
21 and the following details are known in advance,

- 22 • start at wall-clock time START,
- 23 • location of the segments for each Representation at " [http://example.com/\\$RepresentationID\\$/\\$Number\\$](http://example.com/$RepresentationID$/$Number$) ",

25 a service provider may offer an MPD with values according to Table 14.

26 **Table 14 – Basic Service Offering with MPD Updates**

MPD Information	Value
MPD@type	dynamic

MPD@availabilityStartTime	START
MPD@publishTime	PUBTIME1
MPD@minimumUpdatePeriod	MUP
MPD . BaseURL	"http://example.com/"
Period@start	PSTART
SegmentTemplate@media	"\$RepresentationID\$/Number\$"
SegmentTemplate@startNumber	1
SegmentTemplate@duration	SDURATION

1 According to the work-flow shown in Annex B,

- 2 the MPD is generated and published prior to time START such that DASH clients may
3 access it prior to the start of the Media Presentation. The MPD gets assigned a publish time
4 PUBTIME1, typically a value that is prior to START + PSTART
- 5 no redundant tools are considered.
- 6 the encoder and the segmenter generate segments of duration SDURATION and publish
7 those on the origin server, such that they are available at URL[k] latest at their announced
8 segment availability start time SAST[k].

9 Based on the details in section 4.3.2.2 and 4.4.2.2, the Segment Information can be derived at each
10 time NOW by determining the end time of the Period $PE_{wc}[1] = NOW + MUP$.

11 The service provider may leave the MPD unchanged on the server. If this is the case the Media
12 Presentation may be terminated with an updated MPD that

- 13 adds the attribute **MPD@mediaPresentationDuration** with value PDURATION
- 14 removes the attribute **MPD@minimumUpdatePeriod**
- 15 changes the **MPD@publishTime** attribute to PUBTIME2

16 The MPD must be published latest at the end of the Media Presentation minus the value of MUP,
17 i.e. $PUBTIME2 \leqslant START + PSTART + PDURATION - MUP$. For details to convert such a termi-
18 nated live service into an on-demand service, refer to clause 4.6.

19 The minimum update period may also be changed during an ongoing Media Presentation. Note
20 that as with any other change to the MPD, this will only be effective with a delay in media time of
21 the value of the previous MUP.

22 The principles in this document also holds for multi-period content, for which an MPD update may
23 add a new Period. In the same way as for signalling the end of the Media Presentation, the publish
24 time of the updated MPD with the new period needs to be done latest at the start of the new Period
25 minus the value of the **MPD@minimumUpdatePeriod** attribute of the previous MPD.

26 Track fragment decode times should not roll over and should not exceed 2^{53} (due to observed
27 limitations in ECMAScript). Two options may be considered:

- 28 the timescale value should be selected that the above mentioned issues are avoided. 32 bit
29 timescales are preferable for installed-base of browsers.

-
- 1 • if large track timescale values are required and/or long-lasting live sessions are setup, this
2 likely requires the use of 64 bit values. Content authors should use 64 bit values for track
3 fragment decode times in these cases, but should not exceed to 2^{53} to avoid truncation
4 issues.

5 **4.4.3.2. Setting the Minimum Update Period Value**

6 Setting the value of the minimum update period primarily affects two main service provider aspects:
7 A short minimum update period results in the ability to change and announce new content
8 in the MPD on shorter notice. However, by offering the MPD with a small minimum update period,
9 the client requests an update of the MPD more frequently, potentially resulting in increased uplink
10 and downlink traffic.

11 A special value for the minimum update period is 0. In this case, the end time of the period is the
12 current time *NOW*. This implies that all segments that are announced in the MPD are actually
13 available at any point in time. This also allows changing the service provider to offer changes in
14 the MPD that are instantaneous on the media timeline, as the client, prior for asking for a new
15 segment, has to revalidate the MPD.

16 **4.4.3.3. Permitted Updates in an MPD**

17 According to clause 5.4 of ISO/IEC 23009-1, when the MPD is updated

- 18 • the value of **MPD@id**, if present, shall be the same in the original and the updated MPD;
19 • the values of any **Period@id** attributes shall be the same in the original and the updated
20 MPD, unless the containing **Period** element has been removed;
21 • the values of any **AdaptationSet@id** attributes shall be the same in the original and
22 the updated MPD unless the containing **Period** element has been removed;
23 • any Representation with the same @id and within the same Period as a Representation
24 appearing in the previous MPD shall provide functionally equivalent attributes and
25 elements, and shall provide functionally identical Segments with the same indices in the cor-
26 responding Representation in the new MPD.

27 In addition, updates in the MPD only extend the timeline. This means that information provided
28 in a previous version of the MPD shall not be invalidated in an updated MPD. For failover cases,
29 refer to section 4.8.

30 **MPD@availabilityStartTime** and **Period@start** shall not be changed over MPD up-
31 dates.

32 If Representations and Adaptations Sets are added or removed or the location of the Segments is
33 changed, it is recommended to update the MPD and provide Adaptation Sets in a period-continu-
34 ous manner as defined in clause 4.3.3.3.2.

35 DASH clients operating in real-time playout are expected to use the **Period@id** for consistency
36 across MPD updates in order to find the respective playing Period.

37 **4.4.3.4. Usage of Segment Timeline**

38 If the Segment Timeline is used and @minimumUpdatePeriod greater than 0, then

- 39 • the operation as described in section 4.3.3.4 applies, and for all Representations that use
40 the Segment Timeline:

- 1 ○ the @r value of the last **S** element of the last regular Period shall be a negative
 2 value,
 3 ○ only \$Number\$ template shall be used,
- 4 • an MPD may be published for which the additional **S** elements are added at the end. An
 5 addition of such **S** element shall be such that clients that have not updated the MPD can
 6 still generate the Segment Information based on the previous MPD up to the Period end
 7 time. Note that this may lead that such clients have a different segment availability time,
 8 but the availability time may be corrected once the MPD is updated.

9 An example for such an offering is shown in Table 15 where the RVALUE needs to be increased
 10 by 1 for each newly published segment.

11 **Table 15 – Service Offering with Segment Timeline and MUP greater than 0**

MPD Information	Value
MPD@type	dynamic
MPD@availabilityStartTime	START
MPD@publishTime	PUBTIME1
MPD@minimumUpdatePeriod	MUP > 0
MPD . BaseURL	"http://example.com/"
Period@start	PSTART
SegmentTemplate@media	"\$RepresentationID\$/\$Time\$"
SegmentTemplate@d	SDURATION
SegmentTemplate.SegmentTimeline.S@r	-1

12 **4.4.3.5. Last Segment Message**

13 The content author may signal the last segment of a Representation by using the lmsg brand in
 14 the segment. If lmsg is signaled in the Representation, the @segmentProfiles attribute for
 15 this Representation should signal the 'lmsg' brand for this Representation. If the @segmentPro-
 16 files includes the 'lmsg' brand for a Representation, then the 'lmsg' brand shall be included
 17 for the last segment of the Representation in a Period.

18 For non-live MPDs, i.e. @minimumUpdatePeriod not present, and if the lmsg is signaled in
 19 the MPD, the DASH client should search for the lmsg brand at least the last two Segments of
 20 a Period, and not request Segments that are later than the one for which the lmsg brand was
 21 provided. The player may also parse every Segment for lmsg.

22 For live MPDs, i.e. @minimumUpdatePeriod is present, if the @segmentProfiles con-
 23 tains the 'lmsg' brand for a certain Representation, then the 'lmsg' brand for signaling the last
 24 segment shall be applied for any content with **MPD@minimumUpdatePeriod** present and the
 25 **MPD@type="dynamic"**.

DASH clients operating based on such an MPD and consuming the service at the live edge typically need to request a new MPD prior to downloading a new segment. However, in order to minimise MPD requests and resulting traffic load, the client may use one or more of the following optimisations:

- If the client fetches the MPD using HTTP, the client should use conditional GET methods as specified in RFC 7232 [23] to reduce unnecessary network usage in the downlink.
- If the `@segmentProfiles` contains the '`lmsg`' brand clients may also rely on the '`lmsg`' message and request a new MPD only in case a segment is received with an '`lmsg`' brand. Otherwise the client may use template constructions to continue determining the URL and the segment availability start time of segments.

If the attribute `MPD@minimumUpdatePeriod` is set to a value greater than 0 then all Segments with availability start time less than the sum of the request time and the value of the `MPD@minimumUpdatePeriod` will eventually get available at the advertised position at their computed segment availability start time. Note that by providing a `MPD@minimumUpdatePeriod` is set to a value greater than 0, DASH servers reduce the polling frequency of clients, but at the same time cannot expect that clients will request an updated MPD to be informed on changes in the segment URL constructions, e.g. at the start of a new Period.

4.4.3.6. Signalling the last segment number in Period

As indicated in clause 4.3.2.2, the content provider may not offer the last segment that is signaled in the MPD. If this is the case, the content provider should signal that the last segment is not the one indicated in the MPD.

At least the following three options may be considered:

- Use the `lmsg` signalling as defined in clause 4.4.3.5.
- Use the Segment Timeline with `@r` value greater or equal to 0.

Add a Supplemental Descriptor with `@schemeIdUri` set to <http://dashif.org/guidelines/last-segment-number> with the `@value` set to the last segment number.

4.4.4. MPD-based Live Client Operation based on MPD

In an extension to the description in section 4.3.4.1 and section 4.3.4.3, the client now has access to an MPD and the MPD contains the `MPD@minimumUpdatePeriod`, for example following the parameters in Table 14. The start time of each Period is computed as period start time $PSwC[i]$ and the MPD-URL does not include any fragment parameters according to section 4.3.3.5.

The client fetches an MPD with parameters in Table 13 access to the MPD at time `FetchTime`, at its initial location if no `MPD.Location` element is present, or at a location specified in any present `MPD.Location` element. `FetchTime` is defined as the time at which the server processes the request for the MPD from the client. The client typically should not use the time at which it actually successfully received the MPD, but should take into account delay due to MPD delivery and processing. The fetch is considered successful either if the client obtains an updated MPD or the client verifies that the MPD has not been updated since the previous fetching.

1 If the client fetches the MPD using HTTP, the client should use conditional GET methods as specified in RFC 7232 [23] to reduce unnecessary network usage in the downlink.

3 In an extension of bullet 3 in section 4.3.4.1 and section 4.3.4.3

4 the client creates a list of accessible Segments at least for each selected Representation
5 taking into account the information in the MPD as documented in Table 13 and the current
6 time *NOW* by using the Period end time of the last Period as `FetchTime + MUP`.

7 In an extension of bullet 9 in section 4.3.4.1 and section 4.3.4.3,

8 the client consumes media in last announced Period. Once the client is consuming media
9 contained in the Segments towards the end of the announced Period, i.e. requesting seg-
10 ments with segment availability start time close to the validity time of the MPD defined as
11 `FetchTime + MUP`, them, then the DASH client needs to fetch an MPD at its initial
12 location if no **MPD . Location** element is present, or at a location specified in any present
13 **MPD . Location** element.

14 If the client fetches the updated MPD using HTTP, the client should use conditional GET
15 methods as specified in in RFC 7232 [23] to reduce unnecessary network usage in the
16 downlink.

17 The client parses the MPD and generates a new segment list based on the new `FetchTime`
18 and `MUP` of the updated MPD. The client searches for the currently consumed Adaptation
19 Sets and Representations and continues the process of downloading segments based on the
20 updated Segment List.

21 **4.5. MPD and Segment-based Live Service Offering**

22 **4.5.1. Preliminaries**

23 **4.5.1.1. MPD Information**

24 In order to offer a service that relies on both, information in the MPD and in Segments, the Service
25 Provider may announce that Segments contains inband information. An MPD as shown in Table
26 9 provides the relevant information. In contrast to the offering in Table 6, the following infor-
27 mation is different:

- 28 • The **MPD@minimumUpdatePeriod** is present but is recommended to be set to 0 in order
29 to announce instantaneous segment updates.
- 30 • The **MPD@publishTime** is present in order to identify different versions of MPD in-
31 stances.
- 32 • all Representations of all audio Adaptation Sets or if audio is not present, of all video Ad-
33 aptation Sets, shall contain an **InbandEventStream** element with `@scheme_id_uri`
34 = "urn:mpeg:dash:event:2012" and `@value` either set to 1 or set to 3. The **In-**
35 **bandEventStream** element with `@scheme_id_uri` =
36 "urn:mpeg:dash:event:2012" and `@value` either set to 1 or set to 3 may be pre-
37 sent in all Representations of all Adaptation Sets.
- 38 • **InbandEventStream** element with `@scheme_id_uri` =
39 "urn:mpeg:dash:event:2012" and `@value` either set to 1 or set to 3 shall only be
40 signaled on Adaptation Set level.

1 The information included there may be used to compute a list of announced Segments, Segment
2 Availability Times and URLs.

3 **Table 16 – Service Offering with MPD and Segment-based Live Services**

MPD Information	Status	Comment
MPD@type	mandatory, set to "dynamic"	the type of the Media Presentation is dynamic, i.e. Segments get available over time.
MPD@publishTime	mandatory	specifies the wall-clock time when the MPD was generated and published at the origin server. MPDs with a later value of @publishTime shall be an update as defined in 5.4 to MPDs with earlier @publishTime.
MPD@availabilityStartTime	mandatory	the start time is the anchor for the MPD in wall-clock time. The value is denoted as AST.
MPD@minimumUpdatePeriod	mandatory	recommended/mandate to be set to 0 to indicate that frequent DASH events may occur
Period@start	mandatory	the start time of the Period relative to the MPD availability start time. The value is denoted as PS.
AdaptationSet.InbandEvent-Stream	mandatory	if the @schemeIdUri is urn:mpeg:dash:event:2014 and the @value is 1, 2 or 3, then this described an Event Stream that supports extending the validity of the MPD.
SegmentTemplate@media	mandatory	the template for the Media Segment
SegmentTemplate@startNumber	optional default	The number of the first segment in the Period. The value is denoted as SSN.
SegmentTemplate@duration		

SegmentTemplate.SegmentTimeline	exactly one of SegmentTemplate@duration or SegmentTemplate.SegmentTimeline must be present	the duration of each Segment in units of a time. The value divided by the value of @timescale is denoted as $MD[k]$ with $k=1, 2, \dots$. The segment timeline may contain some gaps.
--	--	--

1 **4.5.1.2. Segment Information Derivation**

2 Based on an MPD instance including information as documented in Table 13 and available at time
3 *NOW* on the server, a DASH client may derive the information of the list of Segments for each
4 Representation in each Period.

5 If the Period is the last one in the MPD and the **MPD@minimumUpdatePeriod** is present, then
6 the time $PEwc[i]$ is obtained as the sum of *NOW* and the value of **MPD@minimumUpdate-**
7 **Period**.

8 Note that with the MPD present on the server and *NOW* progressing, the Period end time is ex-
9 tended. This issue is the only change compared to the segment information generation in section
10 4.3.2.2.

11 If the **MPD@minimumUpdatePeriod** is set to 0, then the MPD documents all available seg-
12 ments on the server. In this case the @r count may be set accurately as the server knows all avail-
13 able information.

14 **4.5.2. Service Offering Requirements and Guidelines**

15 **4.5.2.1. Background**

16 In section 5.10 of ISO/IEC 23009-1, section 5.10, DASH events are defined. For service offerings
17 based on the MPD and segment controlled services, the DASH events specified in section 5.10.4
18 may be used. Background is provided in the following.

19 DASH specific events that are of relevance for the DASH client are signalled in the MPD. The
20 URN "urn:mpeg:dash:event:2012" is defined to identify the event scheme defined in Ta-
21 ble 10.

1 **Table 17 InbandEventStream@value** attribute for scheme with a value "urn:mpeg:dash:event:2012"

@value	Description
1	indicates that MPD validity expiration events as defined in 5.10.4.2 are signalled in the Representation. MPD validity expiration is signalled in the event stream as defined in 5.10.4.2 at least in the last segment with earliest presentation time smaller than the event time.
2	indicates that MPD validity expiration events as defined in 5.10.4.3 are signalled in the Representation. MPD validity expiration is signalled in the event stream as defined in 5.10.4.2 at least in the last segment with earliest presentation time smaller than the event time. In addition the message includes an MPD Patch as defined in 5.10.4.3 in the message_data field.
3	indicates that MPD validity expiration events as defined in 5.10.4.3 are signalled in the Representation. MPD validity expiration is signalled in the event stream as defined in 5.10.4.2 at least in the last segment with earliest presentation time smaller than the event time. In addition the message includes a full MPD as defined in 5.10.4.4 in the message_data field.

2 Note: DVB DASH specification [42] does not include the value 3.

3 MPD validity expiration events provide the ability to signal to the client that the MPD with a
4 specific publish time can only be used up to a certain media presentation time.

5 Figure 4 shows an example for MPD validity expiration method. An MPD signals the presence of
6 the scheme in one or several Representations. Once a new MPD gets available, that adds new
7 information not present in the MPD with @publishTime="2012-11-01T09:06:31.6",
8 the expiration time of the current MPD is added to the segment by using the emsg box. The infor-
9 mation may be present in multiple segments.

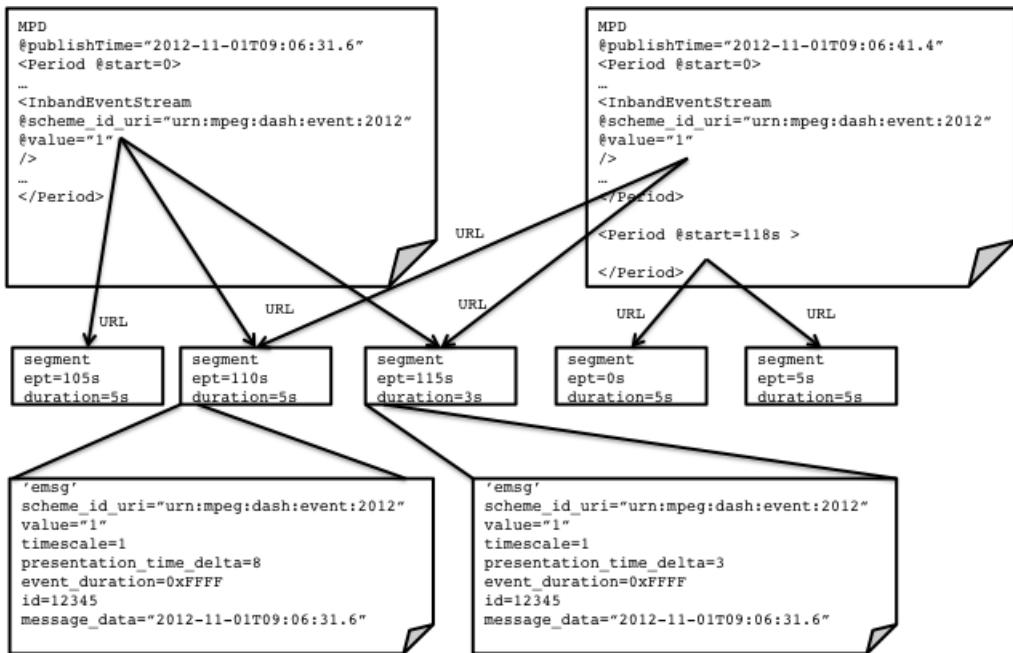


Figure 4 Example for MPD validity expiration to signal new Period

If the scheme_id_uri is set to "urn:mpeg:dash:event:2012" and the value is set to 1, then the fields in the event message box document the following:

- the message_data field contains the publish time of an MPD, i.e. the value of the **MPD@publishTime**.
- The media presentation time beyond the event time (indicated time by presentation_time_delta) is correctly described only by MPDs with publish time greater than indicated value in the message_data field.
- the event duration expresses the remaining duration of Media Presentation from the event time. If the event duration is 0, Media Presentation ends at the event time. If 0xFFFF, the media presentation duration is unknown. In the case in which both presentation_time_delta and event_duration are zero, then the Media Presentation is ended.

This implies that clients attempting to process the Media Presentation at the event time or later are expected to operate on an MPD with a publish time that is later than the indicated publish time in this box.

Note that event boxes in different segments may have identical id fields, but different values for presentation_time_delta if the earliest presentation time is different across segments.

4.5.2.2. Service Offering

A typical service offering with an Inband event stream is provided in Table 11. In this case the MPD contains information that one or multiple or all Representations contain information that the Representation contains an event message box flow in order to signal MPD validity expirations. The **MPD@publishTime** shall be present.

Table 18 – Basic Service Offering with Inband Events

MPD Information	Value
MPD@type	dynamic
MPD@availabilityStartTime	START
MPD@publishTime	PUBTIME1
MPD@minimumUpdatePeriod	MUP
MPD.BaseURL	"http://example.com/"
Period@start	PSTART
InbandEventStream@scheme_id_URI	urn:mpeg:dash:event:2012
InbandEventStream@value	1 or 3
SegmentTemplate@duration	SDURATION

2 For a service offering based on MPD and segment-based controls, the DASH events shall be used
 3 to signal MPD validity expirations.

4 In this case the following shall apply:

- 5 • at least all Representations of all audio Adaptation Sets shall contain an **InbandEvent-**
 6 **Stream** element with **scheme_id_uri** = "urn:mpeg:dash:event:2014" and
 7 **@value** either set to 1 or set to 3.
- 8 • for each newly published MPD, that includes changes that are not restricted to any of the
 9 following (e.g. a new Period):
 - 10 ○ The value of the **MPD@minimumUpdatePeriod** is changed,
 - 11 ○ The value of a **SegmentTimeline.S@r** has changed,
 - 12 ○ A new **SegmentTimeline.S** element is added
 - 13 ○ Changes that do not modify the semantics of the MPD, e.g. data falling out of the
 14 timeshift buffer can be removed, changes to service offerings that do not affect the
 15 client, etc.

16 the following shall be done

- 17 • a new MPD shall be published with a new publish time **MPD@publishTime**
- 18 • an 'emsg' box shall be added to each segment of each Representation that con-
 19 tains an **InbandEventStream** element with
 - 20 ○ **scheme_id_uri** = "urn:mpeg:dash:event:2012"
 - 21 ○ **@value** either set to 1 or set to 3
 - 22 ○ If **@value** set to 1 or 3
 - 23 ▪ the value of the **MPD@publishTime** of the previous MPD as the
 24 **message_data**

25 In addition, the following recommendations should be taken into account: All Representations of
 26 at least one media type/group contain an **InbandEventStream** element with
 27 **scheme_id_uri** = "urn:mpeg:dash:event:2012" and **@value** either set to 1 or set
 28 to 3.

4.5.3. Client Requirements and Guidelines

4.5.3.1. Introduction

A DASH client is guided by the information provided in the MPD. An advanced client model is shown in Figure 7. In contrast to the client in section 4.4.3.5, the advanced client requires parsing of segments in order to determine the following information:

- to expand the Segment List, i.e. to generate the Segment Availability Start Time as well as the URL of the next Segment by parsing the Segment Index.
- to update the MPD based on Inband Event Messages using the 'emsg' box with scheme_id_uri="urn:mpeg:dash:event:2012" and @value either set to 1 or set to 3.

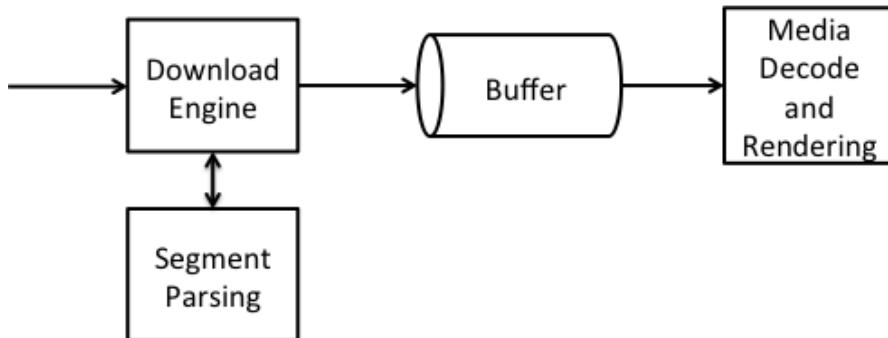


Figure 7 Advanced Client Model

Assumes that the client has access to an MPD and the MPD contains the mandatory parameters in Table 9, i.e., it contains the following information:

- MPD@minimumUpdatePeriod is set to 0
- MPD@publishTime is included and the value is set to PUBTIME
- At least one Representation is present that contains an **InbandEventStream** element with scheme_id_uri="urn:mpeg:dash:event:2012" and @value either set to 1 or set to 3.
- Either the @duration or **SegmentTimeline** for the Representation is present.

In an extension of bullet 7, 8 and 9 in section 4.3.4.1 and section 4.3.4.3, the following example client behaviour may provide a continuous streaming experience to the user as documented in the following.

4.5.3.2. MPD Validity expiration and Updates

The DASH client shall download at least one Representation that contains **InbandEventStream** element with scheme_id_uri = "urn:mpeg:dash:event:2012" and @value either set to 1 or set to 3. It shall parse the segment at least up to the first 'moof' box. The DASH client shall parse the segment information and extract the following values:

- ept the earliest presentation time of the media segment
- dur the media presentation duration of the media segment

1 If an 'emsg' is detected `scheme_id_uri = "urn:mpeg:dash:event:2012"` and
2 @value either set to 1 or set to 3, the DASH client shall parse the segment information and extract
3 the following values:

- 4 • `emsg.publish_time` the publish time documented in the message data of the emsg,
5 either directly or from the patch.
- 6 • `emsg.ptd` the presentation time delta as documented in the emsg.
- 7 • `emsg.ed` the event duration as documented in the emsg

8 After parsing, the Segment is typically forwarded to the media pipeline if it also used for rendering,
9 but it may either be dumped (if the Representation is only used to access the DASH event, such as
10 muted audio).

11 If no 'emsg' validity expiration event is included, then

- 12 • the current MPD can at least be used up to a media presentation time `ept + dur`

13 else if an 'emsg' validity expiration event is included, then

- 14 • the MPD with publish time equal to `emsg.publish_time` can only be used up to a
15 media presentation time `ept + emsg.ptd`. Note that if `dur > emsg.ptd`, then the
16 Period is terminated at `ept + emsg.ptd`.
- 17 • any MPD with publish time greater than `emsg.publish_time` can at least be used up
18 to a media presentation time `ept + emsg.ptd`
- 19 • prior to generating a segment request with earliest presentation time greater than `ept +`
20 `emsg.ptd`, the MPD shall either
 - 21 ◦ be refetched and updated by the client.
 - 22 ◦ or if `@value=3`, it may be used as included in the message.

23
24 NOTE: The DVB DASH profile [42] explicitly forbids downloading a Representation solely to gain access
25 to an Inband Event Stream contained within it. For reference, the relevant part of the DVB DASH specifica-
26 tion is section 9.1.6.

27 **4.5.3.3. Extended Segment Information**

28 The DASH client shall download the selected Representation and shall parse the segment at least
29 up to the first 'moof' box. The DASH client shall parse the segment information and extract the
30 following values:

- 31 • `ept` the earliest presentation time of the media segment
 - 32 ◦ if the Segment Index is present use the Segments Index
 - 33 ◦ if not use the `baseMediaDecodeTime` in 'tfdt' of the first movie frag-
34 ment as the earliest presentation time
- 35 • `dur` the media presentation duration of the media segment
 - 36 ◦ if the Segment Index is present use the Segments Index
 - 37 ◦ if not use aggregated sample durations of the first movie fragment as the dura-
38 tion

1 Using this information, the DASH client should extend the Segment information and, if present
2 the Segment Timeline with the information provided in the Segment. This information can then be
3 used to generate the URL of the next Segment of this Representation. This avoids that the client
4 fetches the MPD, but uses the information of the Segment Timeline. However, in any doubt of the
5 information, for example if a new Adaptation Set is selected, or if Segments are lost, or in case of
6 other operational issues, the DASH client may refetch the MPD in order to obtain the complete
7 information from the MPD.

8 **4.6. Provisioning of Live Content in On-Demand Mode**

9 **4.6.1. Scenario**

10 A common scenario for DASH distribution results that a live generated service is also made available
11 for On-Demand offering after the live program is completed. The typical scenario is as follows:

- 13 - The Segments as generated for the live service are also used for the On-Demand case. This
14 avoids reformatting and also permits to reuse the Segments that are already cached.
15 - The MPD is modified to reflect that the content is available as On-Demand now.
16 - Problems that result from live delivery may be solved, e.g. variable segment durations, or
17 issues of segment unavailability.
18 - The content may be augmented with ads.
19 - The content may be trimmed from a longer, e.g. 24/7 stream, at the beginning and/or end.

20 In an extension to this scenario, the same MPD URL is used for live and on-demand content. This
21 transition scenario is discussed in clause 4.6.4.

22 **4.6.2. Content Offering Requirements and Recommendations**

23 In order to provide live content as On-Demand in the above scenario, the following is
24 recommended:

- 25 - The same Segments as generated for the live (i.e. **MPD@type** is set to `dynamic`)
26 distribution are reused also for static distribution (i.e. **MPD@type** is set to `static`).
27 - Typically, the Segments also will have the same URL in order to exploit caching
28 advantages.
29 - An MPD should be generated latest at the end of the live session, but also may be
30 created during an ongoing live session to document a certain window of the program that
31 is offered for On-Demand. For providing a transition from a live service into an On-
32 Demand service, refer to clause 4.6.4.
33 - A new MPD is generated that should contain the following information
34 ○ The **MPD@type** is set to `static`.
35 ○ The **MPD@availabilityStartTime** should be removed or be maintained
36 from the live MPD since all resources referenced in the MPD are available. If the
37 **MPD@availabilityStartTime** is maintained for a portion of the live
38 program that is offered in the static MPD the **Period@start** value (including

the presentation time offset and the start number) and the presentation duration shall be set accordingly. The relationship to the wall-clock time should be maintained by offsetting the **Period**@start without changing the **MPD**@availabilityStartTime.

- As profile, the simple live profile may be used
- The attributes @timeShiftBufferDepth and @minimumUpdatePeriod shall not be present (in contrast to the live MPD), i.e. it is expected that such attributes are removed. Note that according to ISO/IEC 23009-1, that if present, a client is expected to ignore these attributes for **MPD**@type set to static.
- The presentation duration is determined through either the @mediaPresentationDuration attribute or, if not present, through the sum of the *PeriodStart* and the **Period**@duration attribute of the last Period in the MPD.
- Content may be offered in the same Period structure as for live or in a different one.
 - If Periods are continuous, it is preferable to remove the Period structure.
 - If new Periods are added for Ad Insertion, the Periods preferably be added in a way that they are at Segment boundaries.
- Independent whether the @duration attribute or the **SegmentTimeline** element was used for the dynamic distribution, the static distribution version may have a **SegmentTimeline** with accurate timing to support seeking and to possibly also signal any gaps in the Segment timeline. To obtain the accurate timeline, the segments may have to be parsed (at least up to the *tfdt*) to extract the duration of each Segment.
- The same templating mode as used in the live service should also be used for static distribution.
- MPD validity expiration events should not be present in the MPD. However, it is not expected that ‘emsg’ boxes are removed from Segments.

4.6.3. Client Behavior

For a DASH client, there is basically no difference on whether the content was generated from a live service or the content is provided as On-Demand. However, there are some aspects that may be “left-overs” from a live service distribution that a DASH client should be aware of:

- The Representations may show gaps in the Segment Timeline. Such gaps should be recognized and properly handled. For example a DASH client may find a gap only in one Representation of the content and therefore switches to another Representation that has no gap.
- The DASH client shall ignore any possibly present DASH Event boxes ‘emsg’ (e.g., MPD validity expirations) for which no Inband Event Stream is present in the MPD.

4.6.4. Transition Phase between Live and On-Demand

In the scenario for which the same MPD URL is used for live and On-Demand content, once the URL and publish time of the last Segment is known for the live service, and the duration of the service is known as well, the service provider acts as defined in clause 4.4.3.1, i.e.,

- adds the attribute **MPD@mediaPresentationDuration**
- removes the attribute **MPD@minimumUpdatePeriod**

Once the last segment is published, the service provider may also replace the **MPD@type** from dynamic to static and perform all the processes defined in clause 4.6.2. DASH clients will no longer update the MPD in this case.

DASH clients should support the transition from **MPD@type** being dynamic to static in the case when the @minimumUpdatePeriod is no longer present in the MPD.

4.7. Availability Time Synchronization between Client and Server

4.7.1. Background

According to ISO/IEC 23009-1 [1] and section 4.3, in order to properly access MPDs and Segments that are available on origin servers or get available over time, DASH servers and clients should synchronize their clocks to a globally accurate time standard.

Specifically Segment Availability Times are expected to be wall-clock accurately announced in the MPD and the client needs to have access to the same time base as the MPD generation in order to enable a proper service. In order to ensure this, this section provides server and client requirements to ensure proper operation of a live service.

4.7.2. Service Provider Requirements and Guidelines

If the Media Presentation is dynamic or if the **MPD@availabilityStartTime** is present then the service shall provide a Media Presentation as follows:

- The segment availability times announced in the MPD should be generated from a device that is synchronized to a globally accurate timing source, preferably using NTP.
- The MPD should contain at least one **UTCTiming** element with @schemeIdUri set to one of the following:
 - urn:mpeg:dash:utc:http-xsdate:2014
 - urn:mpeg:dash:utc:http-iso:2014
 - urn:mpeg:dash:utc:http-ntp:2014
 - urn:mpeg:dash:utc:ntp:2014
 - urn:mpeg:dash:utc:http-head:2014
 - urn:mpeg:dash:utc:direct:2014
- If the MPD does not contain any element **UTCTiming** then the segments shall be available latest at the announced segment availability time using a globally accurate timing source.
- If the MPD contains an element **UTCTiming** then
 - the announced timing information in the **UTCTiming** shall be accessible to the DASH client, and

-
- 1 ○ the segments shall be available latest at the announced segment availability time
2 in the MPD for any device that uses one of announced time synchronization meth-
3 ods at the same time.

4 Despite the latter three technologies may save one or several HTTP transactions, the usage of them
5 should be considered carefully by the MPD author, and rather not be used if the MPD author is not
6 controlling the entire distribution system:

- 7 • If `urn:mpeg:dash:utc:ntp:2014` is used, client and server need to implement an
8 NTP client, which may be non-trivial, especially in browser-based clients.
9 • If `urn:mpeg:dash:utc:http-head:2014` is used, then the server specified in the
10 @value attribute of the **UTCTiming** element should be the server hosting the DASH
11 segments such that with each request the Date general-header field in the HTTP header
12 (see in RFC 7231 [22], section 7.1.12) can be used by the client to maintain synchroniza-
13 tion. Also the MPD generator should be aware that caching infrastructures may add inac-
14 curacies to the Date header if the edge caches are not wall-clock synchronized. Therefore,
15 it should not use this method, if they cannot verify that the Date header is set accurately by
16 the edge cache from where each Segment is served. CORS
17 • If `urn:mpeg:dash:utc:direct:2014` is used, then the MPD generator is ex-
18 pected to write the wall-clock time into the MPD. This basically requires a customized
19 MPD for each request and the MPD should be offered such that it is *not* cached as otherwise
20 the timing is flawed and inaccurate.

21
22 Note that in practical deployments segment availability may be an issue due to failures, losses,
23 outages and so on. In this case the Server should use methods as defined in section 4.8 to inform
24 DASH clients about potential issues on making segments available.

25 A leap second is added to UTC every 18 months on average. A service provider should take into
26 account the considerations in RFC 7164 [50]. The MPD time does not track leap seconds. If these
27 occur during a live service they may advance or retard the media against the real time.

28 **4.7.3. Client Requirements and Guidelines**

29 If the Media Presentation is dynamic or if the **MPD@availabilityStartTime** is present then
30 client should do the following:

- 31 • If the MPD does not contain any element **UTCTiming** it should acquire an accurate
32 wall-clock time from its system. The anticipated inaccuracy of the timing source should
33 be taken into account when requesting segments close to their segment availability time
34 boundaries.
35 • If the MPD contains one or several elements **UTCTiming** then the client should at least
36 use one of the announced timing information in the **UTCTiming** to synchronize its
37 clock. The client must not request segments prior to the segment availability start time
38 with reference to any of the chosen **UTCTiming** methods.

39 Note: The DVB DASH [42] spec requires support for `http-xsdate` and `http-head` but al-
40 lows content providers to include others in addition, and allows clients to choose others in prefer-
41 ence if they wish. For details, refer to section 4.7 of the DVB DASH specification.

-
- 1 • The client may take into account the accuracy of the timing source as well as any trans-
2 mission delays if it makes segment requests.
3 • Clients shall observe any difference between their time zone and the one identified in the
4 MPD, as MPDs may indicate a time which is not in the same timezone as the client.
5 • If the client observes that segments are not available at their segment availability start
6 time, the client should use the recovery methods defined in section 4.8.
7 • Clients should not access the **UTCTiming** server more frequently than necessary.

8

4.8. Robust Operation

9

4.8.1. Background

10 In order to support some of the advanced use cases documented in section 2, robust service offer-
11 ings and clients are relevant. This document lists the relevant ones.

12

4.8.2. Tools for Robust Operations

13

4.8.2.1. General Robustness

14 General Guidelines in ISO/IEC 23009-1 [1] DASH spec in A.7:

- 15 • The DASH access client provides a streaming service to the user by issuing HTTP requests
16 for Segments at appropriate times. The DASH access client may also update the MPD by
17 using HTTP requests. In regular operation mode, the server typically responds to such re-
18 quests with status code 200 OK (for regular GET) or status code 206 Partial Content (for
19 partial GET) and the entity corresponding to the requested resource. Other Successful 2xx
20 or Redirection 3xx status codes may be returned.
21 • HTTP requests may result in a Client Error 4xx or Server Error 5xx status code. Some
22 guidelines are provided in this subclause as to how an HTTP client may react to such error
23 codes.
24 • If the DASH access client receives an HTTP client or server error (i.e. messages with 4xx
25 or 5xx error code), the client should respond appropriately (e.g. as indicated in RFC
26 7231 [22]) to the error code. In particular, clients should handle redirections (such as 301
27 and 307) as these may be used as part of normal operation.
28 • If the DASH access client receives a repeated HTTP error for the request of an MPD, the
29 appropriate response may involve terminating the streaming service.
30 • If the DASH access client receives an HTTP client error (i.e. messages with 4xx error code)
31 for the request of an Initialization Segment, the Period containing the Initialization Seg-
32 ment may not be available anymore or may not be available yet.
33 • Similarly, if the DASH access client receives an HTTP client error (i.e. messages with 4xx
34 error code) for the request of a Media Segment, the requested Media Segment may not be
35 available anymore or may not be available yet. In both these case the client should check
36 if the precision of the time synchronization to a globally accurate time standard or to the
37 time offered in the MPD is sufficiently accurate. If the clock is believed accurate, or the
38 error re-occurs after any correction, the client should check for an update of the MPD. . If
39 multiple **BaseUrl** elements are available, the client may also check for alternative in-
40 stances of the same content that are hosted on a different server.

-
- 1 • Upon receiving server errors (i.e. messages with 5xx error code), the client should check
2 for an update of the MPD. If multiple **BaseURL** elements are available, the client may also
3 check for alternative instances of the same content that are hosted on a different server.

4 **4.8.3. Synchronization Loss of Segmenter**

5 In order to address synchronization loss issues at the segmenter, the following options from the
6 DASH standard should be considered with preference according to the order below:

- 7 1. The server is required to always offer a conforming media stream. In case the input
8 stream or encoder is lost, the content author may always add dummy content. This may
9 be done using a separate Period structure and is possible without any modifications of the
10 standard.
- 11 2. Usage of the Segment timeline: In general, with every generated segment on the server,
12 the DASH content generator writes the Segment Timeline following the rules in 4.5.2.2.
13 Only if changes are done beyond the ones excluded, an MPD validity expiration is added.
14 Specifically, when an encoder fails for one or more specific Representations to generate
15 the next Segment, then the DASH content generator may add a segment in the Segment
16 timeline.
- 17 3. Early Terminated Periods as included Cor.1 of the second edition of ISO/IEC 23009-1 [4].
18 Early Terminated Periods may be added that contain both **Period@start** and **Pe-**
19 **riod@duration**. For Early Terminated Periods, the value of the **Period@duration**
20 is the presentation duration in Media Presentation time of the media content represented
21 by the Representations in this Period. The MPD is updated using the **@minimumUpdate-**
22 **Period**, i.e. the timeline is progressing. This permits server to signal that there is an out-
23 age of media generation, but that the service is continuing. It is then up to the client to take
24 appropriate actions.
- 25 4. Gap filling segments: If the decoder or the playback system supports the ability to over-
26 come losses by providing codec independent gap filling segments (such as a null seg-
27 ment), such an approach may be used as well. No standardized way to support this func-
28 tionality exists as of today, but it is expected that MPEG will define such functionalities
29 in the near future.

30 **4.8.4. Encoder Clock Drift**

31 In order to support robust offering even under encoder drift circumstances, the segmenter should
32 avoid being synced to the encoder clock. In order to improve robustness, in the case of an MPD-
33 based offering Periods should be added in a period continuous manner. In the case of MPD and
34 segment-based control, the producer reference box should be added to media streams in order for
35 the media pipeline to be aware of such drifts. In this case the client should parse the segment to
36 obtain this information.

37 **4.8.5. Segment Unavailability**

38 To address signaling of segment unavailability between the client and server and to indicate the
39 reason for this, it is recommended to use regular 404s. In addition, unless a UTC Timing has been
40 defined prior in the MPD, the Date-Header specifying the time of the server should be used. In this

1 case, the DASH client, when receiving a 404, knows that if its time is matching the Date Header,
2 then the loss is due to a segment loss.

3 **4.8.6. Swapping across Redundant Tools**

4 To enable swapping across redundant tools doing hot and warm swaps, the following should be
5 considered

- 6 1. the content author is offering the service redundant to the client (for example using multi-
7 ple Base URLs) and the client determines the availability of one or the other. This may be
8 possible under certain circumstances
- 9 2. Periods may be inserted at a swap instance in order to provide the new information after
10 swap. If possible, the offering may be period-continuous or period-connected, but the of-
11 ferring may also be non-continuous from a media time perspective.
- 12 3. A completely new MPD is sent that removes all information that was available before
13 any only maintains some time continuity. However, this tool is not fully supported yet in
14 any DASH standard and not even considered.

15 There is a clear preference for the bullets above in their order 1, 2 and 3 as the service continuity
16 is expected to be smoother with higher up in the bullet list. At the same time, it may be the case
17 that the failure and outages are severe and only the third option may be used.

18 **4.8.7. Service Provider Requirements and Guidelines**

19 The requirements and guidelines in subsections 8.2 to 8.6 shall be followed.

20 **4.8.8. Client Requirements and Guidelines**

21 The client shall implement proper methods to deal with service offerings provided in section 8.2
22 to 8.6.

23 **4.9. Interoperability Aspects**

24 **4.9.1. Introduction**

25 In order to provide interoperability based on the tools introduce in this section a restricted set of
26 interoperability points are defined.

27 **4.9.2. Simple Live Operation**

28 **4.9.2.1. Definition**

29 The simple live interoperability point permits service offerings with formats defined in the first
30 edition of ISO/IEC 23009-1 [4] as well as in DASH-IF IOPs up to version 2. The DASH client *is*
31 ***not required to parse media segments for proper operation***, but can rely exclusively on the infor-
32 mation in the MPD.

33 **4.9.2.2. Service Requirements and Recommendations**

34 Service offerings conforming to this operation shall follow

- The general requirements and guidelines in section 4.3.3
- the MPD Update requirements and guidelines in section 4.4.3

-
- 1 • the requirements and guidelines for service offering of live content in on-demand mode in
2 section 4.6.2
3 • the synchronization requirements and guidelines in section 4.7.2
4 • the robustness requirements and guidelines in section 4.8.7

5 **4.9.2.3. Client Requirements and Recommendations**

6 Clients claiming conformance to this operation shall follow

- 7 • The general requirements and guidelines in section 4.3.4
8 • the MPD Update requirements and guidelines in section 4.4.3.5
9 • the requirements and guidelines for service offering of live content in on-demand mode in
10 section 4.6.3.
11 • the synchronization requirements and guidelines in section 4.7.3,
12 • the robustness requirements and guidelines in section 4.8.8,

13 **4.9.3. Main Live Operation**

14 **4.9.3.1. Definition**

15 The main live operation permits service offerings with formats defined in the second edition of
16 ISO/IEC 23009-1 [1]. In this case the DASH client **may be required to parse media segments for**
17 **proper operation.**

18 **4.9.3.2. Service Requirements and Recommendations**

19 Service offerings claiming conformance to main live shall follow

- 20 • the requirements and guidelines in section 4.3.3
21 • either
22 ○ the requirements and guidelines in section 4.4.3. Note that in this case no profile
23 identifier needs to be added.
24 • or
25 ○ the segment-based MPD update requirements and guidelines in section 4.5.2. In
26 this case the profile identifier shall be added.
27 • the requirements and guidelines for service offering of live content in on-demand mode in
28 section 4.6.2
29 • the synchronization requirements and guidelines in section 4.7.2
30 • the robustness requirements and guidelines in section 4.8.7

31 **4.9.3.3. Client Requirements and Recommendations**

32 Clients claiming conformance to main live shall follow

- 33 • the requirements and guidelines in section 4.3.4,
34 • the MPD-update requirements and guidelines in section 4.4.3.5,
35 • the segment-based MPD update requirements and guidelines in section 4.5.3,
36 • the requirements and guidelines for service offering of live content in on-demand mode in
37 section 4.6.3.
38 • the synchronization requirements and guidelines in section 4.7.3,
39 • the robustness requirements and guidelines in section 4.8.8.

1 **4.10. Trick Mode for Live Services**

2 **4.10.1. Introduction**

3 In certain use cases, along with the offering of the main content, a content author also wants to
4 provide a trick mode version primarily of the video Adaptation Set along with the live content that
5 can be used for rewind and fast forward in the time shift buffer of the Media Presentation. In
6 section 3.2.9 signalling is introduced to flag and customize Adaptation Sets for Trick Modes. This
7 clause provides additional service offering requirements and recommendations for trick modes in
8 case of a live service. Typically, a reduced frame rate Representation or an I-frame only version is
9 provided for supporting such trick mode operations.

10 **4.10.2. Service Offering Requirements and Recommendations**

11 If trick mode is to be supported for live services, the trick mode Representations should be offered
12 using the same segment duration as in the main Adaptation Set or each segment duration should
13 aggregate an integer multiple of the segments in the main Adaptation Set. The content author needs
14 to find a balance between the segment duration affecting the amount of requests in fast forward
15 or fast rewind and the availability of trick mode segments at the live edge.

16 However, longer segment durations for the trick mode Representation delay the Segment availa-
17 bility time of such Segments by the duration of the Segment, i.e. for the live edge the trick mode
18 may not be fully supported. Based on this it is a content author's decision to provide one or more
19 of the following alternatives for trick mode for live services:

- 20 - Provide one trick mode Adaptation Set that generates a Segment for every Segment in
21 the main Adaptation Set. Note that if this Adaptation Set is used, it may result in in-
22 creased amount of HTTP requests when the player does a fast forward or fast rewind.
- 23 - Provide one trick mode Adaptation Set that generates a Segment only after several Seg-
24 ments in the main Adaptation Set have been generated and aggregate the trick mode
25 samples in a single Segment of longer duration. This will results that possibly no trick
26 mode samples are available at the live edge
- 27 - Provide multiple trick mode Adaptation Sets with different segment durations. If done it
28 is recommended that the `@ti`
- 29 - `meShiftBuffer` for short trick mode segment Adaptation Sets is kept small and in the
30 full timeshift buffer, only trick mode Representations with longer segment durations are
31 maintained. The content author should offer the trick mode Adaptation Sets such that
32 those with longer segment durations can switch to those with shorter duration.
- 33 - Provide trick mode Adaptation sets with a single Indexed Media Segment per Period and
34 use Period boundaries with Period connectivity for both, the main Adaptation Set as
35 well as the trick mode Adaptation Set. This means that only for Periods which are not
36 the live Period, trick mode Adaptation Sets are available, or combinations with the above
37 are possible.

38 **4.10.3. Client Implementation Guidelines**

39 If a client wants to access a trick mode Adaptation Set in a live service, it is recommended to
40 minimize the amount of requests to the network, i.e. it should fetch segments with longer segment
41 duration.

4.10.4. Conversion for Live-to-VoD for Trick Mode Adaptation Sets

If the service is converted from live to VoD as described in clause 4.6, it is recommended that trick mode Adaptation Sets are offered with a single Indexed Media Segment per Period.

4.11. Deployment Scenarios

4.11.1. Introduction

This section addresses specifically considered deployment scenarios and provides proposed service configurations based on the technologies introduced in section 4.

4.11.2. Reliable and Consistent-Delay Live Service

4.11.2.1. Scenario Overview

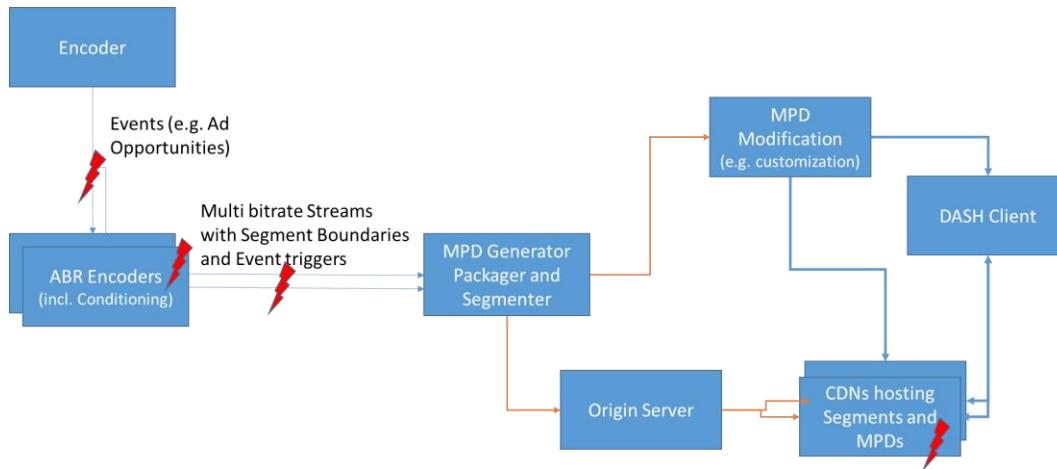
A service provider wants to run a live DASH service according to the below Figure 8. As an example, a generic encoder for a 24/7 linear program or a scheduled live event provides a production encoded stream. Such streams typically include inband events to signal program changes, ad insertion opportunities and other program changes. An example for such signalling are SCTE-35 [54] messages. The stream is then provided to one or more Adaptive Bitrate (ABR) encoders, which transcodes the incoming stream into multiple bitrates and also conditions the stream for segmentation and program changes. These multiple encoders may be used for increased ABR stream density and/or then distributed downstream for redundancy purposes. The resultant streams are received by the DASH generation engines that include: MPD generator, packager and segmenter. Typically the following functions are applied by the MPD packager:

- Segmentation based on in-band information in the streams produced by the ABR encoders
- Encapsulation into ISO BMFF container to generate DASH segments
- Dynamic MPD generation with proper customization options downstream
- Event handling of messages
- Any other other DASH related adaptation

Downstream, the segments may be hosted on a single origin server, or in one or multiple CDNs. The MPD may even be further customized downstream, for example to address specific receivers. Customization may include the removal of certain Adaptation Sets that are not suitable for the capabilities of downstream clients. Specific content may be spliced based on regional services, targeted ad insertion, media blackouts or other information. Events carried from the main encoder may be interpreted and removed by the MPD packager, or they may be carried through for downstream usage. Events may also be added as MPD events to the MPD.

In different stages of the encoding and distribution, errors may occur (as indicated by lightning symbols in the diagram), that for itself need to be handled by the MPD Generator and packager, the DASH client, or both of them. The key issue for this section is the ability for the DASH Media Presentation Generator as shown in to generate services that can handle the incoming streams and provide offerings such that DASH clients following DASH-IF IOPs can support.

1 Hence this section primarily serves to provide guidelines for implementation on MPD Generators
2 and Packagers.



3
4 **Figure 8 Example Deployment Architecture**

5 More detailed service requirements and recommendations are provided in the following.

6 **4.11.2.2. Service Considerations**

7 The following scenarios are considered in the service setup:

8 1. The distribution latency should be consistent, typically what is observed for broadcast TV
9 services. This means that the MPD Generator should add minimum delay, and the service
10 should be setup such that the delay between MPD generator and DASH client playout is
11 consistent, and preferably small.

12 2. Program events may occur for different reasons, for example Program changes, switches
13 from Programs to Advertisements or vice versa, media blackouts or other program
14 changes. Such changes are typically anticipated only on short notice, i.e. within a few
15 seconds. In the following we refer to the time that changes are announced as change lead
16 time. The service should also provide a minimum change lead time, i.e. the smallest time
17 in media time between the change being announced in the stream and the time between
18 the change occurs. Changes may for example include one or more of the following:

- 19 a. Number of source audio languages or formats can change. For example:
20 i. Programming with English and Spanish to other content with only English
21 ii. Descriptive audio may disappear / reappear
22 iii. Programming with 5.1 E-AC-3 and AAC Stereo content to other content
23 with only Stereo AAC

- 24 b. Resolution or format of source video content can change, e.g. HD to/from SD,
25 HDR to/from SDR, etc.
26 c. Codecs may change, or at least the profile or level of the codecs
27 d. The number of Representations in an Adaptation Set may change

1 e. A distribution network may be changed, added or removed.

2 As an example, at broadcast origination points if MPEG-2 TS is used, then the Program
3 Map Table (PMT) typically indicates changes such changes. Typically, these changes also
4 result in discontinuities for in the media timeline.

5 3. The segmentation is determined by the ABR encoders. This encoding may result in occa-
6 sional slight variations in segment durations during a period (as compared to the last seg-
7 ment in a period) due to encoding optimizations around scene changes near the segment
8 duration point (for example: making a segment slightly shorter or longer to align segment
9 IDR to a scene change).

10 4. Unanticipated losses and operational failures or outages, possibly happen just for a single
11 encoding (typically at the input of the encoder, but also possibly also downstream pack-
12 aging). Examples are

13 a. An encoder for one or more Representations or the output of an encoder fails for
14 some time and does not produce content.

15 b. An encoder or the input to the encoder or the output of the encoder fails for a me-
16 dia component/Adaptation Set for some time and do not produce content, e.g. re-
17 ferring to issues as documented in Annex B.4.

18 c. All encoding or the input to the encoder fails for some time e.g. referring to issues
19 as documented in Annex B.4.

20 In all cases an MPD can still be written and the MPD is up and running.

21 Also in the distribution, single Segments may be lost for different reasons and the client
22 typically gets 404.

23 5. MPD updates should be minimized, whereby MPD updates includes the following as-
24 pects for every MPD request

25 a. Client sending uplink requests for MPDs

26 b. Sending full MPD with every request from the server to the client

27 c. Parsing and processing of MPD at the client

28 d. Writing a new MPD on the server if the MPD is changed

29 All factors are relevant to some extent, but primarily the issues a and b should be mini-
30 mized.

31 **4.11.3. Relevant DASH-IF IOP Technologies**

32 **4.11.3.1. Introduction**

33 This document includes technologies that permit to solve the problems addressed above. We re-
34 view the available technologies and justify the selection of the technology for the considered sce-
35 nario. A proposed service configuration is provided in clause 4.11.4.

1 **4.11.3.2. Consistent Latency**

2 The scenario as introduced in clause 4.11.2 does not ask for very low latency, but for consistent
3 latency. In DASH-IF IOP, latency can primarily be controlled by the following means:

- 4 - segment duration: the segment duration typically directly impacts the end-to-end la-
5 tency. Smaller segment sizes provide improved latency and segments of 1-2 seconds
6 may be chosen, if latency is an important aspect. However, too small segments may re-
7 sult in issues, as compression efficiency decreases due to more frequent closed GOPs in
8 the elementary stream. In addition, the number of files/requests to be handled is higher,
9 and finally, with shorter segments, TCP throughput may be such that not the full availa-
10 ble capacity on the link can be exploited. Annex B.4 and clause 4.3.3.2.2 provide some
11 guidelines on this.
- 12 - If files are available in chunks on the origin, for example due to specific encoding or de-
13 livery matters, chunked delivery may be supported. If this feature is offered, then the
14 @availabilityTimeOffset attribute may be provided to announce how much
15 earlier than the nominal segment availability the segment can be accessed.
- 16 - In order to provide tight synchronization between client and server, and therefore
17 providing the receiver the ability to request the segment at the actual segment availabil-
18 ity time, the availability time synchronization as defined in clause 4.7 should be pro-
19 vided and signalled in the MPD. Typically support for http-xsd:date is sufficient for
20 consistent latency support. Accurate NTP synchronization is recommended, but not re-
21 quired for the MPD packager or the DASH client as long as the time synchronization
22 API is provided.
- 23 - It is proposed that a client consistently implements and joins at a segment that is slightly
24 offset (e.g. 4 segments earlier) from the live edge segment. The exact number depends
25 on the distribution system (for example in a fully managed environment, the offset may
26 be smaller in contrast to best effort networks). The MPD author may support consistency
27 by providing a suggested presentation delay in the service offering. For details on join-
28 ing at the live edge, please refer to clause 4.3.4.4.2.

29 **4.11.3.3. Unanticipated New Periods**

30 To avoid that the clients take future segment existence for granted even if a sudden change on the
31 service offering is necessary, the MPD service provider must set to the **MPD@minimumUpdate-**
32 **Period** to a low value. All Segments with availability start time less than the sum of the request
33 time and the value of the **MPD@minimumUpdatePeriod** will eventually get available at the
34 advertised position at their computed segment availability start time.

35 In the most conservative case, the MPD author sets the **MPD@minimumUpdatePeriod** to 0.
36 Then only Segments with availability start time less than the request time are available, i.e. no
37 promise for future segments is provided. The DASH client is forced to revalidate the MPD prior
38 to any new Segment request. For this purpose, basically two options exists:

- 39 - Option 1) Client revalidates MPD with every Segment request according to clause 4.4.4,
40 preferably using a conditional GET in order to avoid unnecessary downlink traffic and
41 processing in the client.

-
- 1 - Option 2) Client relies on MPD validity expiration events in event messages, if content
2 provider announces those in the MPD and by this, it can revalidate.

3 Note that the two methods are not mutually exclusive. More details are discussed further below.

4 In case of option 1 using MPD level validation, with every generated segment on the server, the
5 DASH content generator checks the validity of the MPD offering. If still valid, no changes to the
6 MPD are done. Only if changes are done that are no longer valid, a new MPD is written.

7 **4.11.3.4. Segment Duration Variations**

8 Variable segment durations impact the accuracy of the MPD times of the Segments. MPD times
9 are used for the computation of the segment availability time. With variable segment durations,
10 the segment availability times vary and can impact the DASH-IF IOPs basically provide to options
11 to deal with variable segment durations

- 12 - Option 1)

- 13 ○ Signalling of constant segment duration using `@duration`, permitting a varia-
14 tion of +/- 50% of the segment duration. According to clause 3.2.7.1, for each
15 media segment in each Representation the MPD start time of the segment should
16 approximately be $EPT - PTO$. Specifically, the MPD start time shall be in the
17 range of $EPT - PTO - 0.5*DUR$ and $EPT - PTO + 0.5*DUR$ according to the re-
18 quirement stated above.

19 Note that the encoder should provide segments of a virtual segmentation that ad-
20 heres to this rule. However, there may be reasons that the encoder does break this
21 rule occasionally.

- 22 ○ If the DASH packager receives a segment stream such that the drift can no
23 longer be compensated, then a new Period should be added, that adjusts the pa-
24 rameters for the segment availability computation, but also signals that the Pe-
25 riod is continuous as defined in 4.3.3.3.2. Note that this is done for all Represen-
26 tations in the MPD and a change of the MPD is happening, i.e. this needs to be
27 announced. However, no segment parsing is necessary for the client.

- 28 - Option 2) Following the rules in 4.5.2.2 and using the Segment Timeline to accurately
29 signal the different segment durations. If the segment duration changes, then the `@x` at-
30 tribute of the last `S` element in the Segment timeline is terminated and a new `S` element
31 is added to the MPD with the new segment duration. Note that this results in a change of
32 the MPD. The client should determine such changes independent of MPD updates by de-
33 tailed segment parsing to obtain the earliest presentation time of the segment and the
34 segment duration.

35 **4.11.3.5. Losses and Operational Failures**

36 One of the most complex aspects are occasional operational issues, such as losses, outages, failo-
37 vers of input streams, encoders, packagers and distribution links. Section 4.8 provides detailed
38 overview on available tools that should be used by network service offering and clients in order to
39 deal with operational issues. Several types of losses may occur as shown in Figure 9:

Seg	Seg	Seg	Seg	Seg	Seg	Seg	Seg	No or concealed Loss
Seg	Seg	Seg <small>loss</small>	Seg	Seg	Seg	Seg	Seg	Loss w/in a Segment
Seg	Seg	Seg	Seg	Seg <small>loss</small>	Seg	Seg	Seg	Loss at end of a Segment
Seg	Seg	Seg	<small>loss</small>	Seg	Seg	Seg	Seg	Loss at anticipated Segment Boundary. Next segment does not start at predicted time
Seg	Seg	Seg	<small>loss</small>		Seg	Seg	Seg	Loss of Segment Boundary
Seg	Seg	Seg	<small>loss</small>			Seg	Seg	Loss across multiple segments

1

2

Figure 9 Loss scenarios

3 Losses may occur in the middle of a Segment, at the end of a Segment, at the start of a new Segment.
 4 At the elementary stream level, losses may be within a compressed access unit (AU), pro-
 5 ducing a syntactically corrupt bitstream, or may be the result of the ABR encoder simply not en-
 6 coding a source frame in which case the duration of the prior AU is extended producing a con-
 7 forming bitstreams. Losses may impact an entire Segment or may just impact a part of the Seg-
 8 ment. Typically, service oriented losses will occur until the next Random access point, i.e. a loss
 9 is to be signaled from the start of the lost sample up to the next random access point, typically
 10 coinciding with the start of a new Segment.

11 In order to deal with this, the MPD packager basically has the following options that are not mu-
 12 tually exclusive:

- 13 • Option 1: Operation with Main live profile expecting main profile clients. In this case the
 14 Segment Timeline is used to signal gaps in the media. For details see below.
- 15 • Option 2: Operation with simple live profile expecting simple profile clients. In this case
 16 the early terminated Period is used to signal gaps in the media. For details see below.
- 17 • Option 3: Gap filling segments. Note that Gap filling segments are not part of the IOP
 18 recommendations as the DASH client would be unaware that such gaps are happening.
 19 The DASH client can not take proper measures, such as switching to a Representation
 20 that is not under loss and does provides the actual content. Therefore, gap filling seg-
 21 ments should be used by the content author carefully taking this into account.

1 Note: At the time of writing we are aware that MPEG is addressing the issue of gap filling
2 segments. It is expected that at foreseeable time an update on this issue will be provided.

3 In addition to the above, the content provider may offer to provide the same content on different
4 Base URLs. In this case, the temporary non-availability may be signaled as well through the MPD.

5 **4.11.3.6. Minimizing MPD Updates**

6 MPD updates, the frequency of MPD updates and the actions included in MPD updates are differ-
7 ent ones, and their effects may have different impacts on deployments. To avoid confusion on the
8 generally overloaded term, some more details are discussed in the following section. In non-DASH
9 adaptive streaming solutions, MPD updates result in the following additional processing and de-
10 livery overhead:

- 11 1. The client sends an uplink requests for the MPD. At least from a CDN perspective, this is
12 issue is considered less critical, typically the bounds of operation are reached by through-
13 out, not by the number of requests.
- 14 2. The server needs to send a full MPD with every request, which for itself causes overhead
15 from all the way of the origin server to the client. This is in particular relevant if the man-
16 ifest contains a list of URLs, and some timeshift buffer is maintained.
- 17 3. Yet another aspect is the regular parsing and processing of the manifest in the client.
18 Whereas the processing is likely less of a burden, the consistency across two parsing in-
19 stances is relevant and requires to keep state.
- 20 4. MPD updates may also result in writing a new MPD on the server. This may be less prob-
21 lematic for certain cases, especially for unicast, but it results in significant overhead if
22 DASH formats are used for broadcast.

23 DASH-IF IOP provides different means to avoid one or the more of the above issues. Assuming
24 that the **MPD@minimumUpdatePeriod** is set to a low value for reasons documented above,
25 then issues mentioned above can be addressed by the following means in DASH-IF IOP

- 26 1. Client Requests: can be avoided by signalling inband that an MPD is has expired. The
27 most obvious tool is the use of Inband Events with MPD expiry. However, this requires
28 inband events being added during packaging.
- 29 2. Sending Full MPD: Instead of requesting the full MPD, the client can support this opera-
30 tion by issuing a conditional GET. If the MPD has not changed, no MPD needs to be sent
31 and the downlink rate is small. However, this requires the usage of @duration or
32 **SegmentTimeline** with @r=-1.
- 33 3. MPD Parsing and Processing: This can be avoided by using either of the solutions docu-
34 mented above.
- 35 4. MPD writing on server: This goes hand-in-hand with 2, i.e. the usage of @duration or
36 **SegmentTimeline** with @r=-1.

37 Generally, DASH-IF IOP provide several tools to address different aspects of minimizing MPD
38 updates. Based on the deployment scenario, the appropriate tools should be used. However, it is

1 preferable that DASH clients support different tools in order to provide choices for the service
2 offering.

3 **4.11.4. Proposed Service Configuration and Generation of the MPD and Seg-**
4 **ments based on a “Segment Stream”**

5 **4.11.4.1. Introduction and Assumptions**

6 The core concept is the availability of a segment stream at the input to a packager. The segment
7 stream may be made available as individual segments or as boundary markers in a continuous
8 stream. In addition, the stream may contain information that is relevant for the packager, such as
9 program changes. The segment stream determines for each segment the earliest presentation time,
10 the presentation duration, as well as boundaries in program offerings.

11 Furthermore, it is assumed that multiple bitrates may exist that are switchable. In the following we
12 focus on one segment stream, but assume that in the general case multiple bitrates are available
13 and the encoding and segment streams are generated such that they can be switched.

14 The high-level assumptions for the service are summarized in 4.11.2. Based on these assumptions,
15 a more detailed model is provided.

- 16 - A segment stream is provided for each Representation. The segmentation is the same for
17 Representations that are included in one Adaptation Set. Each segment i has assigned a
18 duration $d[i]$ and an earliest presentation time $ept[i]$. In addition, the segment stream has
19 a nominal segment duration $d0$ that the ABR encoders attempts to maintain. However,
20 variation may occur for different reasons, documented above.
- 21 - Losses may occur in the segment stream, spanning a part of a segment, multiple seg-
22 ments, a full segment and so on. The loss may be in one Representation or in multiple
23 Representations at the same time (see above for more discussions).
- 24 - The latency of the time that the segment is made available to the DASH packager and
25 that it is offered as an available segment in the MPD should be small, i.e. the segment
26 availability time should be shortly after the time when the full segment is received in the
27 DASH packager. Any permitted delay by the MPD Packager can be view as additive to
28 change lead time and may therefore improve efficiency and robustness, but may at the
29 same time increase the end-to-end latency.
- 30 - Changes in the program setup may occur, that signal changes as discussed in 4.11.2. A
31 change is possibly announced with a time referred to as change lead time. Note that sig-
32 nal changes such as SCTE-35 only indicate where a change may occur, it does not indi-
33 cate what type of change will occur.

34 The different scenarios are summarized in Figure 10. For the third part, it shows the notion of the
35 change lead time. Segment of Period with index j are provided. In this case, at the start of segment
36 ($j, i+1$) (i.e. its earliest presentation time) an indication is provided that the media will change after
37 segment ($j, i+2$), i.e. the change lead time is $d[j, i+1] + d[j, i+2]$. A new Period $j+1$ is
38 generated that starts with a new segment numbering.

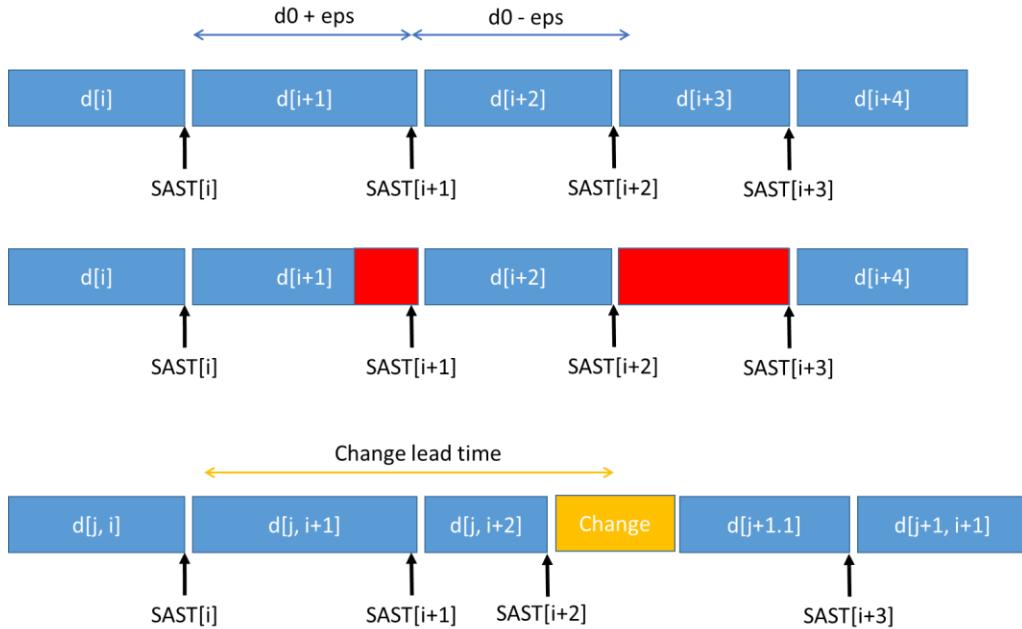


Figure 10: Different properties of a segment stream

Based on the discussions in 4.11.2, proposed service configuration for such a service are proposed.
The service configuration differentiates two deployment scenarios:

- 1) Clients implementing the simple live client, i.e. no emsg support and no segment parsing is implemented.
- 2) Clients implementing the main client, i.e. emsg is supported and segment parsing is implemented.

In the following, reference is made to technologies in section 4.11.3.

4.11.4.2. Service Configuration for Simple Live

Assuming that the input stream is a segment stream with the properties documented above is received by the DASH packager.

The DASH packager may operate as follows:

- The @minimumUpdatePeriod is set to a value that is equal or smaller than the change lead time provided by the segment stream.
- The @timescale of the Adaptation Set is set to the timescale of the included media
- The @duration attribute is set such that the nominal duration d0 is documented in the MPD for this Adaptation Set.
- \$Number\$ is used of segment templating.
- With incoming segments of the segment stream, a new segment is generated by the DASH packager and the DASH packager checks the validity of the MPD offering. If

1 still valid, no changes to MPD are done. Only if changes are done that are no longer
2 valid, a new MPD is written. Specifically,

- 3 ○ The MPD start time of the next segment must be in the range of $EPT - PTO -$
4 $0.5*DUR$ and $EPT - PTO + 0.5*DUR$ with DUR the value of `@duration`.
- 5 ○ If this is not fulfilled a new Period is written that includes the following:
 - 6 ■ The `Period@start` is set such that the MPD start time is correct.
 - 7 ■ The `@presentationTimeOffset` is set to the EPT of the first seg-
8 ment
 - 9 ■ The `@startNumber` is set to the first segment in the new Period.
 - 10 ■ The Adaptation Sets are continued by providing Period continuity sig-
11 nalling with each Adaptation Set.

- 12 - when an encoder fails for one or more specific Representations to generate the next seg-
13 ment, then the DASH content generator

- 14 ○ terminates the Segment with the last sample in the segment, (which is possibly
15 corrupted)
- 16 ○ generates a new MPD as follows:
 - 17 ■ The `@minimumUpdatePeriod` is set to 0.
 - 18 ■ If all or at least many Representations fail, the `Period@duration` is set
19 to the value of the media time in the Period that is still available.
 - 20 ■ If only a subset of the Representations fail, the `@presentationDu-`
21 ration for the last segment is set to the value of the last presentation
22 time in the Representation that is still available.
 - 23 ■ By doing so, the content provider basically informs the DASH client that
24 for the duration of the Segment as announced, no media is available. The
25 DASH client revalidates this after every Segment duration. The MPD is
26 not changed on the server until either the decoder resumes or the Media
27 Presentation is terminated.
 - 28 ■ If the `@minimumUpdatePeriod` is long, then the client may request
29 non-existent segments, which itself may then trigger that the DASH cli-
30 ent revalidates the MPD. If the DASH client has the possibility, it should
31 add the ‘lmsg’ brand as a compatibility brand to the last generated seg-
32 ment. In addition, when the segment is distributed over HTTP, the HTTP
33 header should signal the content type of the segment including the com-
34 patibility brand ‘lmsg’. If the DASH client can identify this, it is ex-
35 pected to refetch the MDP and may by this means observe the early ter-
36 minated Period or Representations.

-
- 1 ○ Only after the encoder resumes, a new MPD is written as follows:
 - 2 ■ A new Period is provided with **Period@start** according to the value
3 of the new Period. The **@presentationTimeoffset** of the Repre-
4 sentation of the Period shall match the the earliest presentation time of
5 the newly generated Segment. If appropriate, Period connectivity should
6 be signaled.
 - 7 ■ The **@minimumUpdatePeriod** is set again to the minimum change
8 lead time.
 - 9 - when a program change is announced, generates a new MPD as follows:
 - 10 ○ The **@minimumUpdatePeriod** is set to 0.
 - 11 - When the program change occurs
 - 12 ○ Write a new MPD with all the parameters
 - 13 ○ Reset the **@minimumUpdatePeriod** is set to a value that is equal or smaller
14 than the change lead time provided
- 15 **4.11.4.3. Service Configuration for Main Live**
- 16 Assuming that the input stream is a segment stream with the properties documented above is re-
17 ceived by the DASH packager.
- 18 The DASH packager may operate as follows:
- 19 - The **@minimumUpdatePeriod** is set to 0.
 - 20 - The **@timescale** of the Adaptation Set is set to the timescale of the included media
 - 21 - The segment timeline is used. Addressing may used: **\$Number\$** or **\$Time\$**.
 - 22 - The MPD is assigned an **MPD@publishTime**
 - 23 - With incoming segments of the segment stream, following the rules in 4.5.2.2 the DASH
24 Packager uses the Segment Timeline to accurately signal the different segment dura-
25 tions. If the segment duration changes, then the **@r** attribute of the last **S** element in the
26 Segment timeline is terminated and a new S element is added to the MPD with the new
27 segment duration. The values **@t** and **@d** need to be set correctly:
 - 28 ○ **@r** of the last segment element may be set to -1. In this case a new MPD is only
29 written if the segment duration changes
 - 30 ○ **@r** of the last segment element may be set to the actual published number of seg-
31 ments. In this case a new MPD is written for each new segment
 - 32 - Whenever a new MPD is written, the **MPD@publishTime** is updated.

-
- 1 - when an encoder fails for one or more specific Representations to generate the next seg-
2 ment, then the DASH packager
- 3 ○ terminates the Segment with the last sample in the segment (may be corrupt)
- 4 ○ adds emsg to this last generated segment. The MPD validity expiration is set to
5 the duration of the current segment or smaller. This emsg may be added to all
6 Representation that have observed this failure, to all Representations in the Ad-
7 aptation Set or to all Representations in the MPD. The content author should be
8 aware that if the emsg is not signaled with all Representations, then there exist
9 cases that a switch *to* the erroneous Representation causes a request to a non-ex-
10 isting Segment. That loss would be signaled in the MPD, but the client is not
11 aware that an update of the MPD is necessary.
- 12 ○ The emsg shall be added to all Representations that announce that they carry the
13 message as an inband stream.
- 14 ○ The MPD is updated on the server such that the last generated segment is docu-
15 mented in the Segment timeline and no new **S** element is added to the timeline.
- 16 ○ Only after the Representation(s) under loss resumes, a new **S** element is written
17 with **s@t** matching the earliest presentation time of the newly generated Seg-
18 ment. The DASH client with its next update will resume and possibly take into
19 account again this Representation.
- 20 ○ If the encoder does not resume for a specific Representation over a longer time,
21 it is recommended to terminate this Period and remove this Representation at
22 least temporarily until the encoder resumes again. Period continuity should be
23 signaled.
- 24 - when the program change occurs
- 25 ○ adds emsg to this last generated segment. The MPD validity expiration is set to
26 the duration of the current segment or smaller. This emsg shall be added to all
27 Representations that announce the Inband Event stream for the MPD validity ex-
28 piration.
- 29 ○ Write a new MPD with all the parameters
- 30 - Whenever a new MPD is written, the **MPD@publishTime** is updated.

31 **4.11.5. Client Support Considerations**

32 **4.11.5.1. Introduction**

33 Generally the client should support the rules in this section for the specific clients.

34 **4.11.5.2. Client Requirements for Simple Live**

35 The client shall follow the details in clause 4.3.4 and 4.4.4. In addition, the DASH client is ex-
36 pected to handle any losses signalled through early terminated Periods.

1 **4.11.5.3. Client Requirements for Main Live**

2 The client shall follow the details in clause 4.3.4 and 4.5.3. In addition, the DASH client is ex-
3 pected to handle any losses signalled through gaps in the segment timeline.

4 The DASH client having received an MPD that signals gaps is expected to either look for alterna-
5 tive Representations that are not affected by the loss, or if not possible, do some appropriate error
6 concealment. The DASH client also should go back regularly to check for MPD updates whether
7 the Representation gets available again.

8 **5. Ad Insertion in DASH**

9 **5.1. Introduction**

10 **5.1.1. General**

11 This section provides recommendations for implementing ad insertion in DASH. Specifically, it
12 defines the reference architecture and interoperability points for a DASH-based ad insertion solu-
13 tion.

14 The baseline reference architecture addresses both server-based and app-based scenarios. The for-
15 mer approach is what is typically used for Apple HLS, while the latter is typically used with Mi-
16 crosoft SmoothStreaming and Adobe HDS.

17 **5.1.2. Definitions**

18 The following definitions are used in this section:

19 **Ad Break:** A location or point in time where one or more ads may be scheduled for delivery; same as
20 *avail* and *placement opportunity*.

21 **Ad Decision Service:** functional entity that decides which ad(s) will be shown to the user. It
22 interfaces deployment-specific and are out of scope for this document.

23 **Ad Management Module:** logical service that, given cue data, communicates with the ad decision
24 service and determines which advertisement content (if at all) should be presented during the ad
25 break described in the cue data.

26 **Cue:** indication of time and parameters of the upcoming ad break. Note that cues can indicate a
27 pending switch to an ad break, pending switch to the next ad within an ad break, and pending switch
28 from an ad break to the main content.

29 **CDN node:** functional entity returning a segment on request from DASH client. There are no
30 assumptions on location of the node.

31 **Packager:** functional entity that processes conditioned content and produces media segments suitable
32 for consumption by a DASH client. This entity is also known as fragmenter, encapsulator, or
33 segmenter. Packager does not communicate directly with the origin server – its output is written to
34 the origin server’s storage.

35 **Origin:** functional entity that contains all media segments indicated in the MPD, and is the fallback if
36 CDN nodes are unable to provide a cached version of the segment on client request.

37 **Splice Point:** point in media content where its stream may be switched to the stream of another
38 content, e.g. to an ad.

39 **MPD Generator:** functional entity returning an MPD on request from DASH client. It may be
40 generating an MPD on the fly or returning a cached one.

1 **XLink resolver:** functional entity which returns one or more remote elements on request from
2 DASH client.

3 **5.1.3. DASH Concepts**

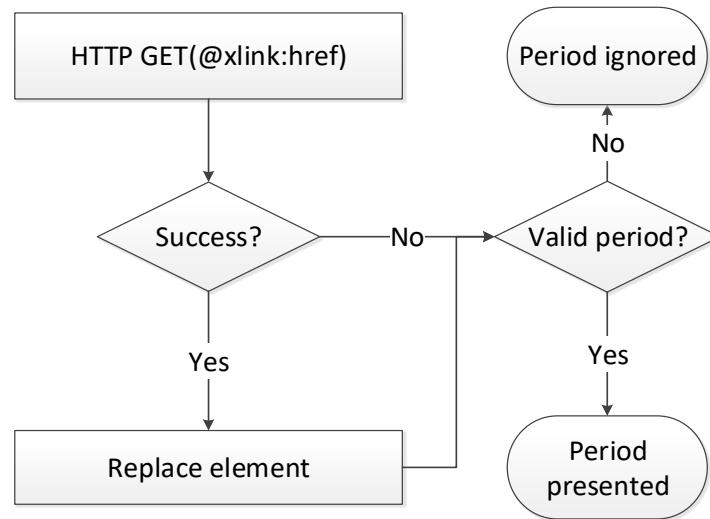
4 **5.1.3.1. Introduction**

5 DASH ad insertion relies on several DASH tools defined in the second edition of ISO/IEC 23009-
6 1 [4], which are introduced in this section. The correspondence between these tools and ad inser-
7 tion concepts are explained below.

8 **5.1.3.2. Remote Elements**

9 *Remote elements* are elements that are not fully contained in the MPD document but are referenced
10 in the MPD with an HTTP-URL using a simplified profile of XLink.

11 A remote element has two attributes, @xlink:href and @xlink:actuate. @xlink:href
12 contains the URL for the complete element, while @xlink:actuate specifies the resolution
13 model. The value "onLoad" requires immediate resolution at MPD parse time, while "onRe-
14 quest" allows deferred resolution at a time when an XML parser accesses the remote element.
15 In this text we assume deferred resolution of remote elements, unless explicitly stated otherwise.
16 While there is no explicit timing model for earliest time when deferred resolution can occur, the
17 specification strongly suggests it should be close to the expected playout time of the corresponding
18 Period. A reasonable approach is to choose the resolution at the nominal download time of the
19 Segment.



20
21 **Figure 11: XLink resolution**

22 Resolution (a.k.a. dereferencing) consists of two steps. Firstly, a DASH client issues an HTTP
23 GET request to the URL contained in the @xlink:href, attribute of the *in-MPD element*, and
24 the XLink resolver responds with a *remote element entity* in the response content. In case of error
25 response or syntactically invalid remote element entity, the @xlink:href and
26 @xlink:actuate attributes the client shall remove the *in-MPD element*.

27 If the value of the @xlink:href attribute is urn:mpeg:dash:resolve-to-zero:2013,
28 HTTP GET request is not issued, and the in-MPD element shall be removed from the MPD. This
29 special case is used when a remote element can be accessed (and resolved) only once during the
30 time at which a given version of MPD is valid.

1 If a syntactically valid remote element entity was received, the DASH client will replace in-MPD
2 element with remote period entity.

3 Once a remote element entity is resolved into a fully specified element, it may contain an
4 @xlink:href attribute with @xlink:actuate set to 'onRequest', which contains a new
5 XLink URL allowing repeated resolution.

6 Note that the only information passed from the DASH client to the XLink resolver is encoded
7 within the URL. Hence there may be a need to incorporate parameters into it, such as splice time
8 (i.e., *PeriodStart* for the remote period) or cue message.

9
10 Note: In ISO/IEC 23009-1:2014/Cor.3 it is clarified that if multiple top-level remote elements
11 are included, the remote element entity is not a valid XML document.

12
13 **5.1.3.3. Periods**
14 **5.1.3.3.1. Timing**

15 Periods are time-delimited parts of a DASH Media Presentation. The value of *PeriodStart* can be
16 explicitly stated using the **Period@start** attribute or indirectly computed using **Period@du-**
17 **ration** of the previous Periods.

18 Precise period duration of period i is given by $PeriodStart(i+1) - PeriodStart(i)$. This can accom-
19 modate the case where media duration of period i is slightly longer than the period itself, in which
20 case a client will schedule the start of media presentation for period $i+1$ at time *PeriodStart*($i+1$).
21 **Representation@presentationTimeOffset** specifies the value of the presentation
22 time at *PeriodStart*(i) .

23 **5.1.3.3.2. Segment Availability**
24 In case of dynamic MPDs, Period-level **BaseURL@availabilityTimeOffset** allow earlier
25 availability start times. A shorthand notation @availabilityTimeOffset="INF" at a Pe-
26 riod-level **BaseURL** indicates that the segments within this period are available at least as long as
27 the current MPD is valid. This is the case with stored ad content. Note that DASH also allows
28 specification of @availabilityTimeOffset at Adaptation Set and Representation level.

29 **5.1.3.3.3. Seamless transition**
30 The DASH specification says nothing about Period transitions – i.e., there are no guarantees for
31 seamless continuation of playout across the period boundaries. Content conditioning and receiver
32 capability requirements should be defined for applications relying on this functionality. However,
33 Period continuity or connectivity should be used and signaled as defined in section 3.2.12 and
34 ISO/IEC 23009-1:2014/Amd.3 [4].

35 **5.1.3.3.4. Period labeling**
36 Period-level **AssetIdentifier** descriptors identify the asset to which a given Period belongs.
37 Beyond identification, this can be used for implementation of client functionality that depends on
38 distinguishing between ads and main content (e.g. progress bar and random access).

39 **5.1.3.4. DASH events**
40 DASH events are messages having type, timing and optional payload. They can appear either in
41 MPD (as period-level event stream) or inband, as ISO-BMFF boxes of type `emsg`. The `emsg`

1 boxes shall be placed at the very beginning of the Segment, i.e. prior to any media data, so that
2 DASH client needs a minimal amount of parsing to detect them.

3 DASH defines three events that are processed directly by a DASH client: MPD Validity Expiration,
4 MPD Patch and MPD Update. All signal to the client that the MPD needs to be updated – by
5 providing the publish time of the MPD that should be used, by providing an XML patch that can
6 be applied to the client’s in-memory representation of MPD, or by providing a complete new MPD.
7 For details please see section 4.5.

8 User-defined events are also possible. The DASH client does not deal with them directly – they
9 are passed to an application, or discarded if there is no application willing or registered to process
10 these events. A possible client API would allow an application to register callbacks for specific
11 event types. Such callback will be triggered when the DASH client parses the `emsg` box in a
12 Segment, or when it parses the **Event** element in the MPD.

13 In the ad insertion context, user-defined events can be used to signal information, such as cue
14 messages (e.g. SCTE 35 [54])

15 **5.1.3.5. MPD Updates**

16 If **MPD@minimumUpdatePeriod** is present, the MPD can be periodically updated. These up-
17 dates can be *synchronous*, in which case their frequency is limited by **MPD@minimumUpdate-**
18 **Period**. In case of the main live profiles MPD updates may be triggered by DASH events. Fir
19 details refer to section 4.5.

20 When new period containing stored ads is inserted into a linear program, and there is a need to
21 unexpectedly alter this period the inserted media will not carry the `emsg` boxes – these will need
22 to be inserted on-the-fly by proxies. In this case use of synchronous MPD updates may prove
23 simpler.

24 **MPD@publishTime** provides versioning functionality: MPD with later publication times in-
25 clude all information that was included all MPDs with earlier publication times.

26 **5.1.3.6. Session information**

27 In order to allow fine-grain targeting and personalization, the identity of the client/viewer, should
28 be known i.e. maintain a notion of a session.

29 HTTP is a stateless protocol, however state can be preserved by the client and communicated to
30 the server.

31 The simplest way of achieving this is use of cookies. According to RFC 6265 [41], cookies set via
32 2xx, 4xx, and 5xx responses must be processed and have explicit timing and security model.

33 **5.1.3.7. Tracking and reporting**

34 The simplest tracking mechanism is server-side logging of HTTP GET requests. Knowing request
35 times and correspondence of segment names to content constitutes an indication that a certain part
36 of the content was requested. If MPDs (or remote element entities) are generated on the fly and
37 identity of the requester is known, it is possible to provide more precise logging. Unfortunately
38 this is a non-trivial operation, as same user may be requesting parts of content from different CDN
39 nodes (or even different CDNs), hence log aggregation and processing will be needed.

40 Another approach is communicating with existing tracking server infrastructure using existing ex-
41 ternal standards. An IAB VAST-based implementation is shown in section 5.3.3.7.

1 DASH Callback events are defined in ISO/IEC 23009-1:2014 AMD3 [4], are a simple native im-
2 plementation of time-based impression reporting (e.g., quartiles). A callback event is a promise by
3 the DASH client to issue an HTTP GET request to a provided URL at a given offset from *Period-*
4 *Start*. The body of HTTP response is ignored. Callback events can be both, MPD and inband
5 events.

6 **5.2. Architectures**

7 The possible architectures can be classified based on the location of component that communicates
8 with the ad decision service: a *server-based* approach assumes a generic DASH client and all com-
9 munication with ad decision services done at the server side (even if this communication is trig-
10 gered by a client request for a segment, remote element, or an MPD. The *app-based* approach
11 assumes an application running on the end device and controlling one or more generic DASH
12 clients.

13 Yet another classification dimension is amount of media engines needed for a presentation – i.e.,
14 whether parallel decoding needs to be done to allow seamless transition between the main and the
15 inserted content, or content is conditioned well enough to make such transition possible with a
16 single decoder.

17 Workflows can be roughly classified into *linear* and *elastic*. Linear workflows (e.g., live feed from
18 an event) has ad breaks of known durations which have to be taken: main content will only resume
19 after the end of the break and the programmer / operator needs to fill them with some inserted
20 content. Elastic workflows assume that the duration of an ad break at a given cue location not
21 fixed, thus the effective break length can vary (and can be zero if a break is not taken).

5.3. Server-based Architecture

2 5.3.1. Introduction

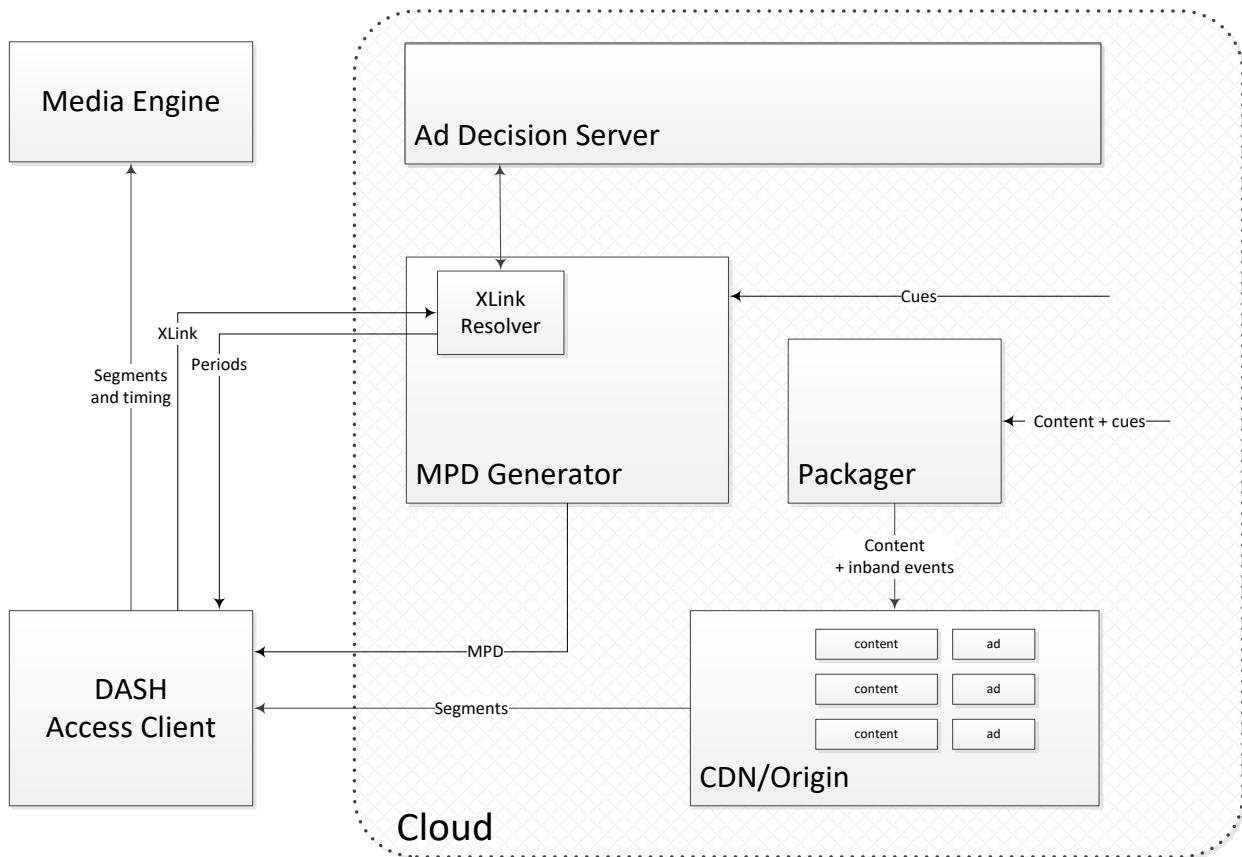


Figure 12: Server-based architecture

5 In the server-based model, all ad-related information is expressed via MPD and segments, and ad
6 decisions are triggered by client requests for MPDs and for resources described in them (Segments,
7 remote periods).

8 The server-based model is inherently MPD-centric – all data needed to trigger ad decision is con-
9 centrated in the MPD. In case where ad break location (i.e., its start time) is unknown at the MPD
10 generation time, it is necessary to rely on MPD update functionality. The two possible ways of
11 achieving these are described in 5.1.3.5.

12 In the live case, packager receives feed containing inband cues, such as MPEG-2 TS with SCTE
13 35 cue messages [54]. The packager ingests content segments into the CDN. In the on demand
14 case, cues can be provided out of band.

15 Ad management is located at the server side (i.e., in the cloud), thus all manifest and content con-
16 ditioning is done at the server side.

17 5.3.2. Mapping to DASH

18 5.3.2.1. Period elements

19 5.3.2.1.1. General

20 A single ad is expressed as a single **Period** element.

1 Periods with content that is expected to be interrupted as a result of ad insertion should contain
2 explicit start times (**Period@start**), rather than durations. This allows insertion of new periods
3 without modifying the existing periods. If a period has media duration longer than the distance
4 between the start of this period and the start of next period, use of start times implies that a client
5 will start the playout of the next period at the time stated in the MPD, rather than after finishing
6 the playout of the last segment.

7 An upcoming ad break is expressed as Period element(s), possibly remote.

8 5.3.2.1.2. Remote Period elements.

9 Remote Periods are resolved on demand into one or more than one Period elements. It is possible
10 to embed parameters from the cue message into the XLink URL of the corresponding remote pe-
11 riod, in order to have them passed to the ad decision system via XLink resolver at resolution time.
12 In an elastic workflow, when an ad break is not taken, the remote period will be resolved into a
13 period with zero duration. This period element will contain no adaptation sets.

14 If a just-in-time remote Period dereferencing is required by use of `@xlink:actuate="onRe-`
15 `quest"`, MPD update containing a remote period should be triggered close enough to the intended
16 splice time. This can be achieved using MPD Validity events and full-fledged MPD update, or
17 using MPD Patch and MPD Update events (see sec. 5.1.3.5 and 5.1.3.4). However, due to security
18 reasons MPD Patch and MPD Update events should only be used with great care.

19 In case of `Period@xlink:actuate="onRequest"`, MPD update and XLink resolution
20 should be done sufficiently early to ensure that there are no artefacts due to insufficient time given
21 to download the inserted content. Care needs to be taken so that the client is given a sufficient
22 amount of time to (a) request and receive MPD update, and (b) dereference the upcoming remote
23 period.

24 NOTE: It may be operationally simpler to avoid use of `Period@xlink:actu-`
25 `ate="onRequest"`, dereferencing in case of live content.

26 5.3.2.1.3. Timing and dereferencing

27 The only interface between DASH client and the XLink resolver is the XLink URL (i.e., the
28 `Period@xlink:href` attribute). After resolution, the complete remote Period element is
29 replaced with Period element(s) from the remote entity (body of HTTP response coming from
30 XLink resolver). This means that the XLink resolver is (in the general case) unaware of the
31 exact start time of the ad period.

32 In case of linear content, start of the ad period is only known a short time before the playback.
33 The recommended implementation is to update the MPD at the moment the start of the ad
34 period is known to the MPD generator.

35 The simplest approach for maintaining time consistency across dereferencing is to have the
36 MPD update adding a `Period@duration` attribute to the latest (i.e., the currently playing)
37 main content period. This means that the MPD resolver needs to include the `Period@du-`
38 `ration` attribute into each of the Period elements returned in the remote entity. The down-
39 side of this approach is that the DASH client needs to be able to update the currently playing
40 period.

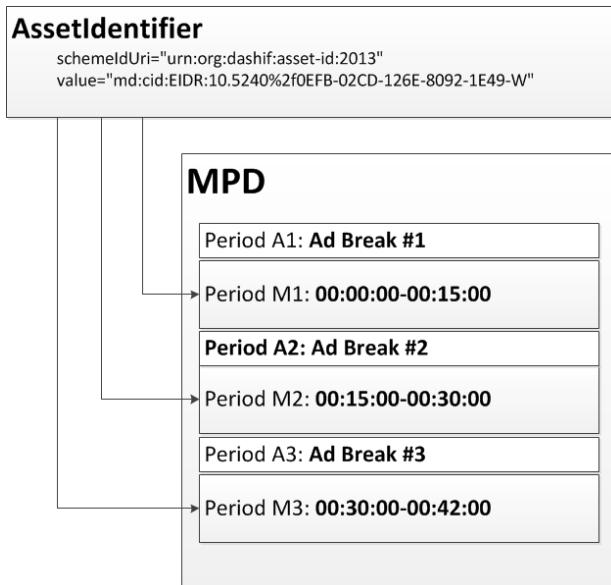
1 An alternative approach is to embed the desired value of **Period@start** of the first period
2 of the remote entity in the XLink URL (e.g., using URL query parameters). This approach is
3 described in clause 5.3.5. The downside of this alternative approach is that the DASH specifi-
4 cation does not constrain XLink URLs in any way, hence the XLink resolver needs to be aware
5 of this URL query parameter interface defined in clause 5.3.5.

6

7 5.3.2.2. Asset Identifiers

8 **AssetIdentifier** descriptors identify the asset to which a Period belongs. This can be used
9 for implementation of client functionality that depends on distinguishing between ads and main
10 content (e.g. progress bar).

11 Periods with same **AssetIdentifier** should have identical Adaptation Sets, Initialization
12 Segments and same DRM information (i.e., DRM systems, licenses). This allows reuse of at least
13 some initialization data across periods of the same asset, and ensures seamless continuation of
14 playback if inserted periods have zero duration. Period continuity or connectivity should be sig-
15 naled, if the content obeys the rules.



16
17 **Figure 13 Using an Asset Identifier**

18

19 5.3.2.3. MPD updates

20 MPD updates are used to implement dynamic behavior. An updated MPD may have additional
21 (possibly – remote) periods. Hence, MPD update should be triggered by the arrival of the first cue
22 message for an upcoming ad break. Ad breaks can also be canceled prior to their start, and such
23 cancellation will also trigger an MPD update.

24 Frequent regular MPD updates are sufficient for implementing dynamic ad insertion. Unfortu-
25 nately they create an overhead of unnecessary MPD traffic – ad breaks are rare events, while MPD
26 updates need to be frequent enough if a cue message is expected to arrive only several seconds
27 before the splice point. Use of HTTP conditional GET requests (i.e., allowing the server to respond

1 with "304 Not Modified" if MPD is unchanged) is helpful in reducing this overhead, but asynchronous MPD updates avoid this overhead entirely.

3 DASH events with scheme "urn:mpeg:dash:event:2013" are used to trigger asynchronous MPD updates.

5 The simple mapping of live inband cues in live content into DASH events is translating a single
6 cue into an MPD Validity expiration event (which will cause an MPD update prior to the splice
7 time). MPD Validity expiration events need to be sent early enough to allow the client request a
8 new MPD, resolve XLink (which may entail communication between the resolver and ADS), and,
9 finally, download the first segment of the upcoming ad in time to prevent disruption of service at
10 the splice point.

11 If several `emsg` boxes are present in a segment and one of them is the MPD Validity Expiration
12 event, `emsg` carrying it shall always appear first.

13 **5.3.2.4. MPD events**

14 In addition to tracking events (ad starts, quartile tracking, etc.) the server may also need to signal
15 additional metadata to the video application. For example, an ad unit may contain not only inline
16 linear ad content (that is to be played before, during, or after the main presentation), it may also
17 contain a companion display ad that is to be shown at the same time as the video ad. It is important
18 that the server be able to signal both the presence of the companion ad and the additional tracking
19 and click-through metadata associated with the companion.

20 With that said, there is no need to have a generic DASH client implement this functionality – it is
21 enough to provide opaque information that the client would pass to an external module. Event
22 @schemeIdUri provides us with such addressing functionality, while MPD events allow us to
23 put opaque payloads into the MPD.

24 **5.3.3. Workflows**

25 **5.3.3.1. General**

26 In the workflows below we assume that our inputs are MPEG-2 transport streams with embedded
27 SCTE 35 cue messages [54]. In our opinion this will be a frequently encountered deployment,
28 however any other in-band or out-of-band method of getting cue messages and any other input
29 format lend themselves into the same model.

30 **5.3.3.2. Linear**

31 A real-time MPEG-2 TS feed arrives at both packager and MPD generator. While real-time multi-
32 cast feeds are a very frequently encountered case, the same workflow can apply to cases such as
33 ad replacement in a pre-recorded content (e.g., in time-shifting or PVR scenarios).

34 MPD generator generates dynamic MPDs. Packager creates DASH segments out of the arriving
35 feed and writes them into the origin server. Client periodically requests the MPDs so that it has
36 enough time to transition seamlessly into the ad period.

37 Packager and MPD generator may be tightly coupled (e.g. co-located on the same physical ma-
38 chine), or loosely coupled as they both are synchronized only to the clock of the feed.

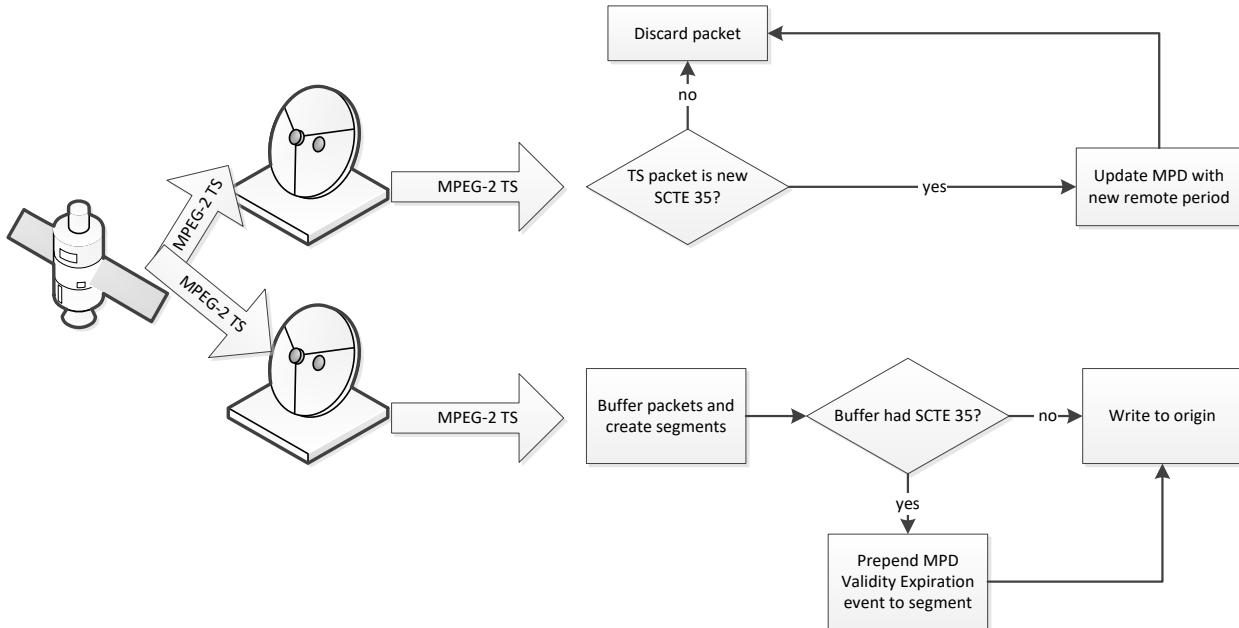


Figure 14: Live Workflow

5.3.3.2.1. Cue Interpretation by the MPD generator

When an SCTE 35 cue message indicating an upcoming splice point is encountered by the MPD generator, the latter creates a new MPD for the same program, adding a remote period to it.

The `Period@start` attribute of the inserted period has `splice_time()` translated into the presentation timeline. Parameters derived from the cue message are inserted into the `Period@xlink:href` attribute of the inserted period. Examples below show architectures that allow finer targeting.

5.3.3.2.1.1. Example 1: Immediate ad decision

MPD generator keeps an up-to-date template of an MPD. At each cue message arrival, the generator updates its template. At each MPD request, the generator customizes the request based on the information known to it about the requesting client. The generator contacts ad decision server and produces one or more non-remote ad periods. In this case XLink is not needed.

5.3.3.2.1.2. Example 2: Stateful cue translation

MPD generator keeps an up-to-date template of an MPD. At each cue message arrival, the generator updates its template. At each MPD request, the generator customizes the request based on the information known to it about the requesting client.

The operator targets separately male and female audiences. Hence, the generator derives this from the information it has regarding the requesting client (see 5.1.3.6), and inserts an XLink URL with the query parameter `?gender=male` for male viewers, and `?gender=female` for the female viewers.

Note that this example also showcases poor privacy practices – would such approach be implemented, both parameter name and value should be encrypted or TLS-based communication should be used

1 **5.3.3.2.1.3. Example 3: Stateless cue translation**

2 At cue message arrival, the MPD generator extracts the entire SCTE 35 `splice_info_section` (starting at the `table_id` and ending with the `CRC_32`) into a buffer. The buffer is then
3 encoded into URL-safe base64url format according to RFC 4648 [60], and inserted into the XLink
4 URL of a new remote Period element. `splice_time` is translated into `Period@start` attrib-
5 ute. The new MPD is pushed to the origin.

6
7 Note: this example is a straightforward port of the technique defined for SCTE 67 [55], but uses base64url
8 and not base64 encoding as the section is included in a URI.

9 **5.3.3.2.2. Cue Interpretation by the packager**

10 Cue interpretation by the packager is optional and is an optimization, rather than core functionality.
11 On reception of an SCTE 35 cue message signaling an upcoming splice, an ``emsg`` with MPD
12 Validity Expiration event is inserted into the first available segment. This event triggers an MPD
13 update, and not an ad decision, hence the sum of the earliest presentation time of the ``emsg``-
14 bearing segment and the ``emsg`.presentation_time_delta` should be sufficiently ear-
15 lier than the splice time. This provides the client with sufficient time to both fetch the MPD and
16 resolve XLink.

17 `splice_time()` of the cue message is translated into the media timeline, and last segment be-
18 fore the splice point is identified. If needed, the packager can also finish the segment at the splice
19 point and thus having a segment shorter than its target duration.

20 **5.3.3.2.3. Multiple cue messages**

21 There is a practice of sending several SCTE 35 cue messages for the same splice point (e.g., the
22 first message announces a splice in 6 seconds, the second arrives 2 seconds later and warns about
23 the same splice in 4 seconds, etc.). Both the packager and the MPD generator react on the same
24 first message (the 6-sec warning in the example above), and do nothing about the following mes-
25 sages.

26 **5.3.3.2.4. Cancelation**

27 It is possible that the upcoming (and announced) insertion will be canceled (e.g., ad break needed
28 to be postponed due to overtime). Cancelation is announced in a SCTE 35 cue message.

29 When cancelation is announced, the packager will insert the corresponding ``emsg`` event and the
30 MPD generator will create a newer version of the MPD that does not contain the inserted period
31 or sets its duration to zero. This implementation maintains a simpler less-coupled server side sys-
32 tem at the price of an increase in traffic.

33 **5.3.3.2.5. Early termination**

34 It is also possible that a planned ad break will need to be cut short – e.g., an ad will be cut short
35 and there will be a switch to breaking news. The DASH translation of this would be creating an
36 ``emsg`` at the packager and updating the MPD appropriately. Treatment of early termination here
37 would be same as treatment of a switch from main content to an ad break.

38 It is easier to manipulate durations when `Period@duration` is absent and only `Pe-`
39 `riod@start` is used – this way attributes already known to the DASH client don't change.

1 5.3.3.2.6. Informational cue messages

SCTE 35 can be used for purposes unrelated to signaling of placement opportunities. Examples of such use are content identification and time-of-day signaling. Triggering MPD validity expiration and possibly XLink resolution in this case may be an overreaction.

5.3.3.2.7. Ad decision

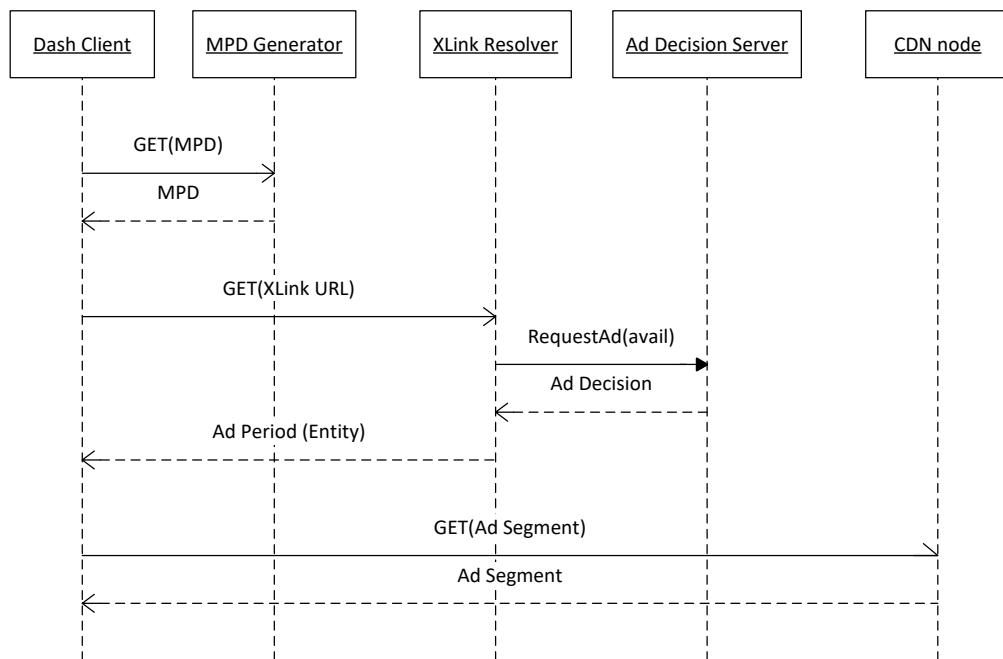


Figure 15: Ad Decision

8 A client will attempt to dereference a remote period element by issuing an HTTP GET for the URL
9 that appears in **Period**@xlink:href. The HTTP server responding to this request (XLink re-
10 solver) will contact the ad decision service, possibly passing it parameters known from the request
11 URL and from client information available to it from the connection context. In case described in
12 5.3.3.2.1.3, the XLink resolver has access to a complete SCTE 35 message that triggered the splice.
13 The ad decision service response identifies the content that needs to be presented, and given this
14 information the XLink resolver can generate one or more Period elements that would be then re-
15 turned to the requesting DASH client.

16 A possible optimization is that resolved periods are cached – e.g. in case of 5.3.3.2.1.1 "male" and
17 "female" versions of the content are only generated once in T seconds, with HTTP caching used
18 to expire the cached periods after T seconds.

19 5.3.3.3. On Demand

20 In a VoD scenario, cue locations are known ahead of time. They may be available multiplexed into
21 the mezzanine file as SCTE 35 or SCTE 104, or may be provided via an out-of-band EDL.

22 In VoD workflows both cue locations and break durations are known, hence there is no need for a
23 dynamic MPD. Thus cue interpretation (which is same as in 5.3.3.2) can occur only once and result

1 in a static MPD that contains all remote elements with all Period elements having **Pe-**
2 **riod@start** attribute present in the MPD.

3 In elastic workflows ad durations are unknown, thus despite our knowledge of cue locations within
4 the main content it is impossible to build a complete presentation timeline. **Period@duration**
5 needs to be used. Remote periods should be dereferenced only when needed for playout. In case
6 of a “jump” – random access into an arbitrary point in the asset – it is a better practice not to
7 dereference Period elements when it is possible to determine the period from which the playout
8 starts using **Period@duration** and asset identifiers. The functionality described in 5.3.3.2 is
9 sufficient to address on-demand cases, with the only difference that a client should be able to
10 handle zero-duration periods that are a result of avails that are not taken.

11 **5.3.3.4. Capture to VoD**

12 Capture to VoD use case is a hybrid between pure linear and on demand scenarios: linear content
13 is recorded as it is broadcast, and is then accessible on demand. A typical requirement is to have
14 the content available with the original ad for some time, after which ads can be replaced

15 There are two possible ways of implementing the capture-to-VoD workflow.

16 The simplest is treating capture-to-VoD content as plain VoD, and having the replacement policy
17 implemented on the XLink resolver side. This way the same Period element(s) will be always
18 returned to the same requester within the window where ad replacement is disallowed; while after
19 this window the behavior will be same as for any on-demand content. An alternative implementa-
20 tion is described in 5.3.3.5 below.

21 **5.3.3.5. Slates and ad replacement**

22 A content provider (e.g., OTT) provides content with ad breaks filled with its own ads. An ISP is
23 allowed to replace some of these with their own ads. Conceptually there is content with slates in
24 place of ads, but all slates can be shown and only some can be replaced.

25 An ad break with a slate can be implemented as a valid in-MPD Period element that also has XLink
26 attributes. If a slate is replaceable, XLink resolution will result in new Period element(s), if not –
27 the slate is played out.

28 **5.3.3.6. Blackouts and Alternative content**

29 In many cases broadcast content cannot be shown to a part of the audience due to contractual
30 limitations (e.g., viewers located close to an MLB game will not be allowed to watch it, and will
31 be shown some alternative content). While unrelated to ad insertion per se, this use case can be
32 solved using the same “default content” approach, where the in-MPD content is the game and the
33 alternative content will be returned by the XLink resolver if the latter determines (in some unspec-
34 ified way) that the requester is in the blackout zone.

35 **5.3.3.7. Tracking and reporting**

36 A Period, either local or a remote entity, may contain an EventStream element with an event con-
37 taining IAB VAST 3.0 Ad element [53]. DASH client does not need to parse the information and
38 act accordingly – if there is a listener to events of this type, this listener can use the VAST 3.0 Ad
39 element to implement reporting, tracking and companion ads. The processing done by this listener
40 does not have any influence on the DASH client, and same content would be presented to both
41 “vanilla” DASH client and the player in which a VAST module registers with a DASH client a
42 listener to the VAST 3.0 events. VAST 3.0 response can be carried in an **Event** element where

1 **EventStream**@schemeIdUri value is http://dashif.org/identifiers/vast30.

3 An alternative implementation uses DASH Callback events to point to the same tracking URLs.
4 While DASH specification permits both inband and MPD Callback events, inband callback events
5 shall not be used.

6 **5.3.4. Examples**

7 **5.3.4.1. MPD with mid-roll ad breaks and default content**

8 In this example, a movie (“Top Gun”) is shown on a linear channel and has two mid-roll ad breaks.
9 Both breaks have default content that will be played if the XLink resolver chooses not to return
10 new Period element(s) or fails.

11 In case of the first ad break, SCTE 35 cue message is passed completely to the XLink resolver,
12 together with the corresponding presentation time.

13 In case of the second ad break, proprietary parameters u and z describe the main content and the
14 publishing site.

15

```
<?xml version="1.0"?>
<MPD xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xmlns="urn:mpeg:dash:schema:mpd:2011"
      xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011 DASH-MPD.xsd"
      type="dynamic"
      minimumUpdatePeriod="PT2S"
      timeShiftBufferDepth="PT600S"
      minBufferTime="PT2S"
      profiles="urn:mpeg:dash:profile:isoff-live:2011"
      availabilityStartTime="2012-12-25T15:17:50">
  <BaseURL>http://cdn1.example.com/</BaseURL>
  <BaseURL>http://cdn2.example.com/</BaseURL>

  <!-- Movie -->
  <Period start="PT0.00S" duration="PT600.6S" id="movie period #1">
    <AssetIdentifier schemeIdUri="urn:org:dashif:asset-id:2013"
                     value="md:cid:EIDR:10.5240%2f0EFB-02CD-126E-8092-1E49-W"/>
    <AdaptationSet mimeType="video/mp4" codecs="avc1.640828"
                    frameRate="24000/1001" segmentAlignment="true" startWithSAP="1">
      <BaseURL>video_1/</BaseURL>
      <SegmentTemplate timescale="90000" initialization="$Band-
width%/init.mp4v"
                        media="$Bandwidth$/Number%05d$.mp4v"/>
      <Representation id="v0" width="320" height="240" bandwidth="250000"/>
      <Representation id="v1" width="640" height="480" bandwidth="500000"/>
      <Representation id="v2" width="960" height="720" bandwidth="1000000"/>
    </AdaptationSet>
  </Period>

  <!-- Mid-roll advertisement, passing base64url-coded SCTE 35 to XLink resolver ->
  <Period duration="PT60.6S" id="ad break #1"
          xlink:href="https://adserv.com/avail.mpd?scte35-time=PT600.6S&
                      scte35-cue=DAIAAAAAAAAAQAAZ_I0VniQAQAgBDVUVJQAAAAH+cAAAAAA%3D%3D"
          xlink:actuate="onRequest" >

  <!-- Default content, replaced by elements from remote entity -->
  <AdaptationSet mimeType="video/mp4" codecs="avc1.640828">
```

```

        frameRate="30000/1001"
        segmentAlignment="true" startWithSAP="1">
      <BaseURL availabilityTimeOffset="INF">default_ad/</BaseURL>
      <SegmentTemplate timescale="90000" initialization="$Band-
width%/init.mp4v"
                        media="$Bandwidth%/$Time$.mp4v"/>
      <Representation id="v0" width="320" height="240" bandwidth="250000"/>
      <Representation id="v1" width="640" height="480" bandwidth="500000"/>
      <Representation id="v2" width="960" height="720" bandwidth="1000000"/>
    </AdaptationSet>
  </Period>

<!--Movie, cont'd -->
<Period duration="PT600.6S" id="movie period #2">
  <AssetIdentifier schemeIdUri="urn:org:dashif:asset-id:2013"
    value="md:cid:EIDR:10.5240%2f0EFB-02CD-126E-8092-1E49-W"/>
  <AdaptationSet mimeType="video/mp4" codecs="avc1.640828"
    frameRate="24000/1001"
    segmentAlignment="true" startWithSAP="1">
    <BaseURL>video_2/</BaseURL>
    <SegmentTemplate timescale="90000" initialization="$Band-
width%/init.mp4v"
                      media="$Bandwidth%/$Time$.mp4v"/>
    <Representation id="v0" width="320" height="240" bandwidth="250000"/>
    <Representation id="v1" width="640" height="480" bandwidth="500000"/>
    <Representation id="v2" width="960" height="720" bandwidth="1000000"/>
  </AdaptationSet>
  </Period>

<!-- Mid-roll advertisement, using proprietary parameters -->
<Period duration="PT60.6S" id="ad break #2"
  xlink:href="https://adserv.com/avail.mpd?u=0EFB-02CD-126E-8092-1E49-
W&z=spam"
  xlink:actuate="onRequest" >

  <!-- Default content, replaced by elements from remote entity -->
  <AdaptationSet mimeType="video/mp4" codecs="avc1.640828"
    frameRate="30000/1001"
    segmentAlignment="true" startWithSAP="1">
    <BaseURL availabilityTimeOffset="INF">default_ad2/</BaseURL>
    <SegmentTemplate timescale="90000" initialization="$Band-
width%/init.mp4v"
                      media="$Bandwidth%/$Time$.mp4v"/>
    <Representation id="v0" width="320" height="240" bandwidth="250000"/>
    <Representation id="v1" width="640" height="480" bandwidth="500000"/>
    <Representation id="v2" width="960" height="720" bandwidth="1000000"/>
  </AdaptationSet>
  </Period>
</MPD>
```

1

Figure 16: Example of MPD for "Top Gun" movie

2

5.3.5. Use of query parameters

3 Parameters can be passed into the XLink resolver as a part of the XLink URL. Clause
 4 5.3.3.2.1.3 shows an example of this approach when an SCTE 35 cue message is embedded
 5 into the XLink URL.

This approach can be generalized and several parameters (i.e., name-value pairs) can be defined. SCTE 214-1 2016 [56] takes this approach and defines parameters expressing splice time (i.e., **Period**@start of the earliest ad period), SCTE 35 cue message, and syscode (a geolocation identifier used in US cable industry). The first two parameters are also shown in example in clause 5.3.4.1 of this document.

Note 1: Effectively this creates a RESTful API for XLink dereferencing. While discussion above implies that these parameters are embedded by the MPD generator into the XLink URL, the parameter values may as well be calculated by the client or the embedded values may be modified by the client.

Note 2: The same RESTful API approach can be used with MPD URLs as well.

Note 3: More parameters may be defined in the future version of these guidelines.

5.4. App-based Architecture

5.4.1. Introduction

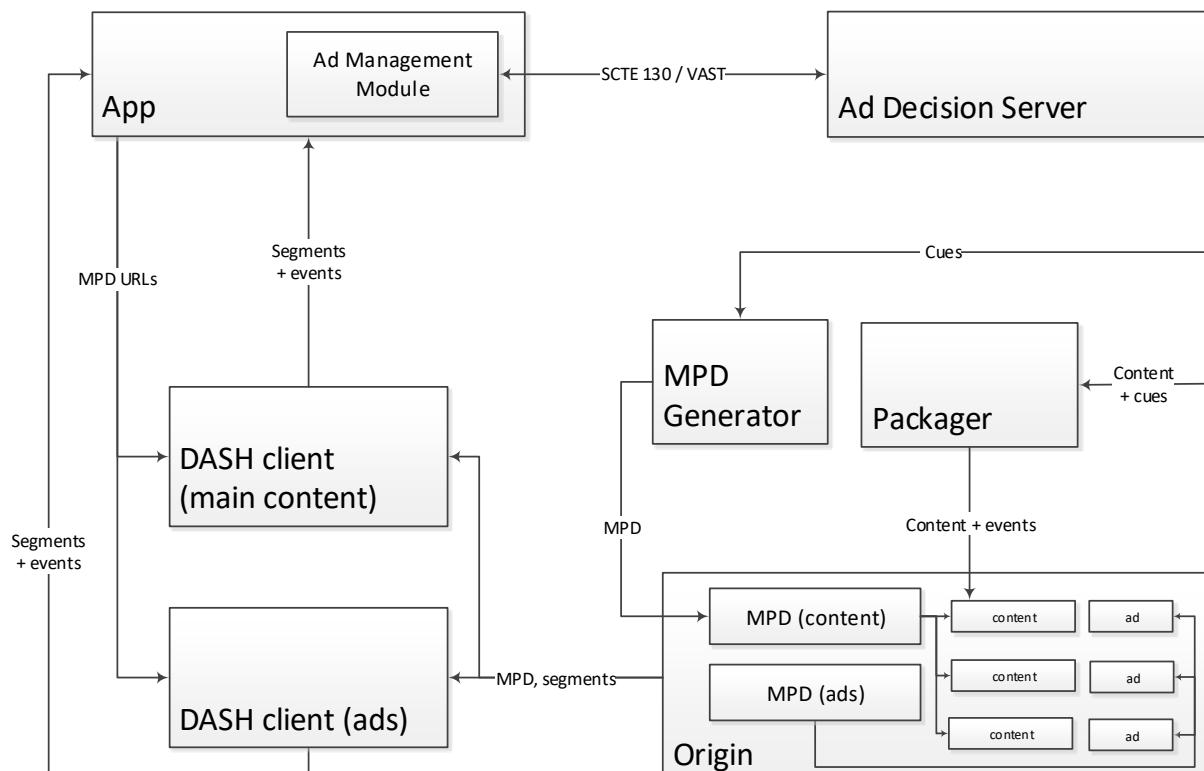


Figure 17: App-based architecture

Inputs in this use case are same as the ones described in sec. 5.3. At the packaging stage, cues are translated into a format readable by the app or/and DASH client and are embedded into media segments or/and into the manifest

Ad management module is located at the client side. The DASH client receives manifest and segments, with cues embedded in either one of them or in both.

1 Cue data is passed to the ad management module, which contacts the ad decision service and re-
2 ceives information on content to be played. This results in an MPD for an inserted content and a
3 splice time at which presentation of main content is paused and presentation of the inserted content
4 starts.

5 Note that this architecture does not assume multiple decoders – with careful conditioning it is
6 possible to do traditional splicing where inserted content is passed to the same decoder. In this case
7 it is necessary to keep a player state and be able to initialize a player into this state.

8 **5.4.2. Mapping to DASH**

9 This section details mapping of elements of the reference architecture into DASH concepts per the
10 2nd edition of the specification (i.e., ISO/IEC 23009-1:2014).

11 **5.4.2.1. MPD**

12 Each ad decision results in a separate MPD. A single MPD contains either main content or inserted
13 content; existence of multiple periods or/and remote periods is possible but not essential.

14 **5.4.2.2. SCTE 35 events**

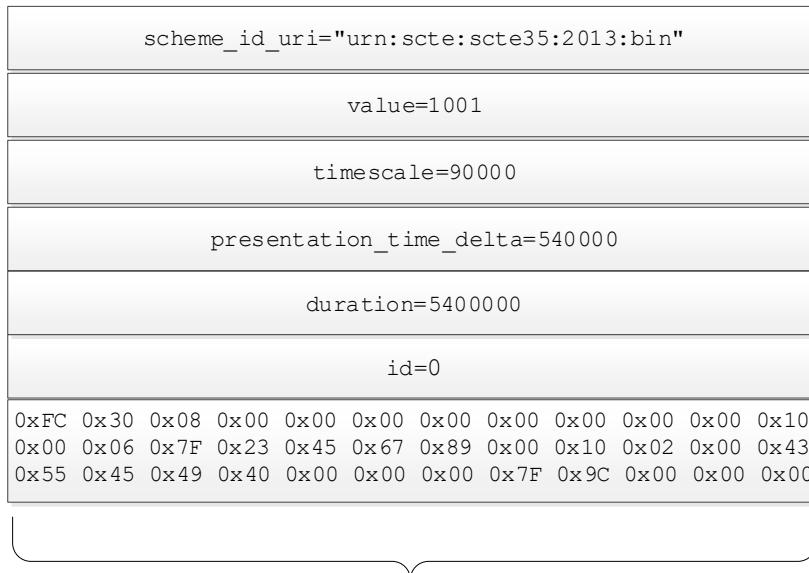
15 **5.4.2.2.1. General**

16 Cue messages are mapped into DASH events, using inband '`emsg`' boxes and/or in-MPD events.
17 Note that SCTE 35 cue message may not be sufficient by itself.

18 The examples below show use of SCTE 35 in user-defined events, and presentation time indicates
19 the timing in within the Period.

20 Figure 18 below shows the content of an '`emsg`' box at the beginning of a segment with earliest
21 presentation time T . There is a 6-sec warning of an upcoming splice – delta to splice time is indi-
22 cated as 6 seconds – and duration is given as 1 minute. This means that an ad will start playing at
23 time $T + 6$ till $T + 66$. This example follows a practice defined in SCTE 214-3 [57].

24



Binary SCTE 35 cue message

Figure 18 Inband carriage of SCTE 35 cue message

Figure 19 below shows the same example with an in-MPD SCTE35 cue message. The difference is in the in-MPD event the splice time is relative to the Period start, rather than to the start of the event-carrying segment. This figure shows a one-minute ad break 10 minutes into the period.

```
<EventStream schemeIdUri="urn:scte:scte35:2014:xml+bin">
  <Event timescale="90000" presentationTime="54054000" duration="5400000" id="1">
    <scte35:Signal>
      <scte35:Binary>
        /DAIAAAAAAAAAAAQAAZ/I0VniQAQAgBDVUVJQAAAAH+cAAAAAA==
      </scte35:Binary>
    </scte35:Signal>
  </Event>
</EventStream>
```

Figure 19: In-MPD carriage of SCTE 35 cue message

Note: for brevity purposes SCTE 35 2014 allows use of base64-encoded section in **Signal.Binary** element as an alternative to carriage of a completely parsed cue message.

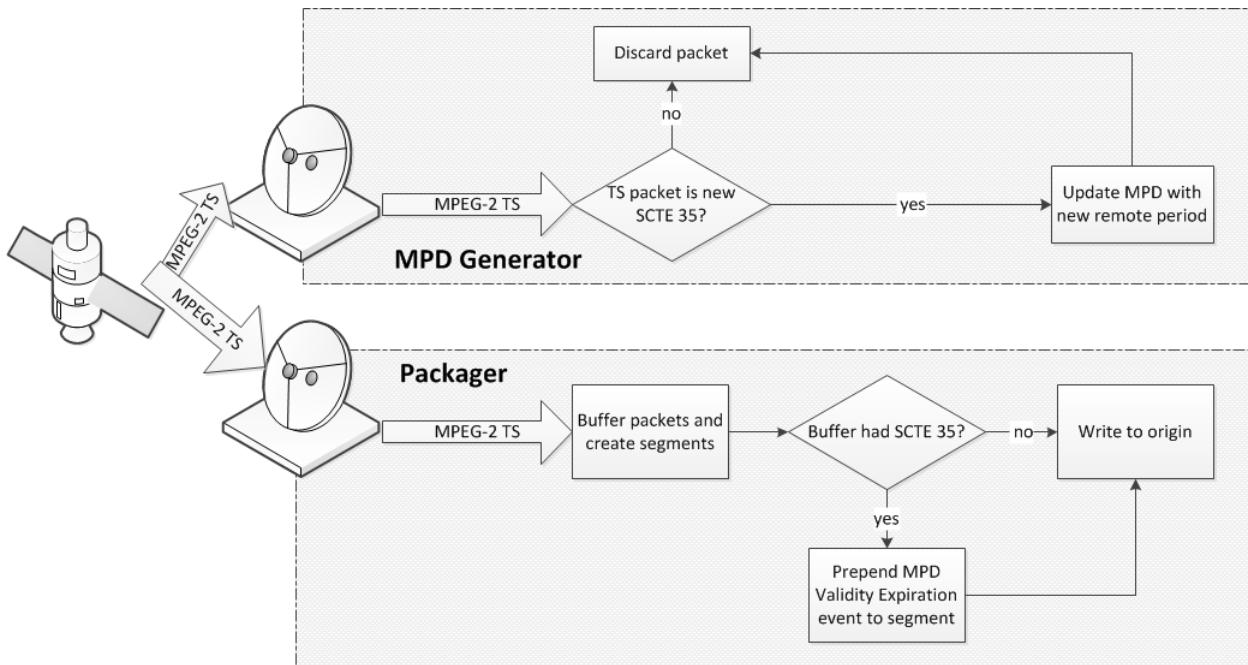
Normative definitions of carriage of SCTE 35 cue messages are in ANSI/SCTE 214-1 [56] sec 6.8.4 (MPD) and SCTE 214-3 [57] sec 8.3.3.

5.4.2.3. Asset Identifiers

See sec. 5.3.2.2 for details.

1 5.4.3. Workflows

2 5.4.3.1. Linear



3 **Figure 20: Linear workflow for app-driven architecture**

4 A real-time MPEG-2 TS feed arrives at a packager. While real-time multicast feeds are a very
5 frequently encountered case, the same workflow can apply to cases such as ad replacement in a
6 pre-recorded content (e.g., in time-shifting or PVR scenarios).

7 Packager creates DASH segments out of the arriving feed and writes them into the origin server.
8 The packager translates SCTE 35 cue messages into inband DASH events, which are inserted into
9 media segments.

10 MPD generator is unaware of ad insertion functionality and the packager does the translation of
11 SCTE 35 cue messages into inband user-defined DASH events. On reception of an SCTE 35 cue
12 message signaling an upcoming splice, a `emsg` with a translation of the cue message in its
13 `emsg`.message_data [] field is inserted into the most recent Segment. This event triggers
14 client interaction with an ad decision server, hence the sum of the earliest presentation time of the
15 `emsg`-bearing segment and the `emsg`.presentation_time_delta should be a trans-
16 lation of splice_time() into the media timeline.

17 An alternative implementation which is more compatible with server-based architecture in section
18 5.3, an MPD generator can generate separate MPDs for both server-based and app-based architec-
19 tures creating remote periods for server-based and in-MPD SCTE 35 events for app-based archi-
20 tectures, while a packager can insert inband MPD validity expiration events.

21 A DASH client will pass the event to the app controlling it (e.g., via a callback registered by the
22 app). The app will interpret the event and communicate with the ad decision server using some
23 interface (e.g., VAST). This interface is out of the scope of this document.

24 The communication with ad decision service will result in an MPD URL. An app will pause the
25 presentation of the main content and start presentation of the inserted content. After presenting the

1 inserted content the client will resume presentation of the main content. This assumes either proper
2 conditioning of the main and inserted content or existence of separate client and decoder for in-
3 inserted content. The way pause/resume is implemented is internal to the API of the DASH client.
4 Interoperability may be achieved by using the DASH MPD fragment interface, see ISO/IEC
5 23009-1 [4], Annex C.4.

6 **5.4.3.2. On Demand**

7 As in the server-based case, functionality defined for the live case is sufficient. Moreover, the fact
8 that that app-based implementation relies heavily on app's ability to pause and resume the DASH
9 client, support for elastic workflows is provided out of the box.

10 In the on demand case, as cue locations are well-known, it is advantageous to provide a static MPD
11 with SCTE 35 events than run a dynamic service that relies on inband events.

12 **5.5. Extensions for ad insertion**

13 **5.5.1. Asset Identifiers**

14 **AssetIdentifier** descriptor shall be used for distinguishing parts of the same asset within a
15 multi-period MPD, hence it shall be used for main content and may be used for inserted content.

16 In order to enable better tracking and reporting, unique IDs should be used for different assets.

17 Use of EIDR and Ad-ID identification schemes is recommended. The value of @schemeIdUri
18 set to "urn:eidr" signals use of EIDR. The value of @value attribute shall be a valid canon-
19 ical EIDR entry as defined in [67].

20 Use of Ad-ID for asset identification is signaled by setting the value of @schemeIdUri to
21 "urn:smpte:ul:060E2B34.01040101.01200900.00000000" ("designator"
22 URN defined in SMPTE 2092-1 [68]). The value of @value attribute shall be a canonical full
23 Ad-ID identifier as defined in SMPTE 2092-1 [68].

24 Other schemes may be used, including user private schemes, by using appropriately unique values
25 of @schemeIdUri.

26 In the absence of other asset identifier schemes, a DASH-IF defined scheme may be used with the
27 value of @schemeIdUri set to "urn:org:dashif:asset-id:2014". If used, the value
28 of @value attribute descriptor shall be a MovieLabs ContentID URN ([58], 2.2.1) for the content.
29 It shall be the same for all parts of an asset. Preferred schemes are EIDR (main content) and Ad-
30 ID (advertising).

31 If a Period has one-off semantics (i.e., an asset is completely contained in a single period, and its
32 continuation is not expected in the future), the author shall not use asset identifier on these assets.

33 Periods that do not contain non-remote **AdaptationSet** elements, as well as zero-length peri-
34 ods shall not contain the **AssetIdentifier** descriptor.

35 **5.5.2. Remote Periods**

36 An MPD may contain remote periods, some of which may have default content. Some of which
37 are resolved into multiple Period elements.

38 After dereferencing MPD may contain zero-length periods or/and remote Periods.

1 In case of **Period**@xlink:actuate="onRequest", MPD update and XLink resolution
2 should be done sufficiently early to ensure that there are no artefacts due to insufficient time given
3 to download the inserted content.

4 **Period**@xlink:actuate="onRequest" shall not be used if **MPD**@type = "dynamic"

5 **5.5.3. User-defined events**

6 **5.5.3.1. Cue message**

7 Cue messages used in app-driven architecture shall be SCTE 35 events [54]. SCTE 35 event carriage
8 is defined in ANSI/SCTE 214-1 (MPD) and ANSI/SCTE 214-3 (inband). For MPD events,
9 the XML schema is defined in SCTE 35 2014 [54] and allows either XML representation or concise
10 base64-coded representation.

11 NOTE: PTS offset appearing in SCTE 35 shall be ignored, and only DASH event timing mechanism may be
12 used to determine splice points.

13 **5.5.3.2. Reporting**

14 MPD events with embedded IAB VAST 3.0 [53] response may be used for reporting purposes.

15 If only time-based reporting is required (e.g., reporting at start, completion, and quartiles), use of
16 DASH callback event may be a simpler native way of implementing tracking. Callback events are
17 defined in ISO/IEC 23009-1:2014 AMD3 [4].

18 **5.5.3.3. Ad Insertion Event Streams**

19 Recommended Event Stream schemes along with their scheme identifier for app-driven ad insertion
20 are:

21 1. "urn:scte:scte35:2013:bin" for inband SCTE 35 events containing a complete
22 SCTE 35 section in binary form, as defined in ANSI/SCTE 214-3.

23 2. "urn:scte:scte35:2014:xml+bin" for SCTE 35 MPD events containing only
24 base64 cue message representation, as defined in ANSI/SCTE 214-1.

25 NOTE: the content of Event element is an XML representation of the complete SCTE 35 cue message,
26 that contains **Signal.Binary** element rather than the **Signal.SpliceInfoSection** element,
27 both defined in SCTE 35 2014.

28 3. "http://dashif.org/identifiers/vast30" for MPD events containing
29 VAST3.0 responses [53].

30 4. urn:mpeg:dash:event:callback:2015 for DASH callback events.

31 **5.6. Interoperability Aspects**

32 **5.6.1. Server-based Ad insertion**

33 For server-based ad insertion, the following aspects needs to be taken into account:

- 34 • Service offerings claiming conformance to server-based ad insertion shall follow the requirements
35 and guidelines for service offerings in sections 5.3.2, 5.5.1, and 5.5.2..
- 36 • Clients claiming conformance to server-based ad insertion shall follow shall follow the requirements
37 and guidelines for clients in section 5.3.2, 5.5.1, and 5.5.2. .

1 **5.6.2. App-based Ad Insertion**

2 For app-based ad insertion, the logic for ad insertion is outside the scope of the DASH client. The
3 tools defined in section 5.4 and 5.5 may be used to create an interoperable system that includes
4 DASH-based delivery and ad insertion logic.

5 **6. Media Coding Technologies**

6 **6.1. Introduction**

7 In addition to DASH-specific constraints, DASH-IF IOPs also adds restrictions on media codecs
8 and other technologies. This section provides an overview on technologies for different media
9 components and how they fit into the DASH-related aspects of DASH-IF IOPs.

10 **6.2. Video**

11 **6.2.1. General**

12 The codec considered for basic video support up to 1280 x 720p at 30 fps is H.264 (AVC) Pro-
13 gressive High Profile Level 3.1 decoder [8]. This choice is based on the tradeoff between content
14 availability, support in existing devices and compression efficiency.

15 Further, it is recognized that certain clients may only be capable to operate with H.264/AVC "Pro-
16 gressive" Main Profile Level 3.0 and therefore content authors may provide and signal a specific
17 subset of DASH-IF IOP.

18 Notes

- 19 • H.264 (AVC) Progressive High Profile Level 3.1 decoder [8] can also decode any content
20 that conforms to
 - 21 ○ H.264 (AVC) Constrained Baseline Profile up to Level 3.1
 - 22 ○ H.264 (AVC) "Progressive" Main Profile up to Level 3.1.
- 23 • H.264 (AVC) H.264/AVC "Progressive" Main Profile Level 3.0 decoder [8] can also de-
24 code any content that conforms to H.264 (AVC) Constrained Baseline Profile up to Level
25 3.0.

26 Further, the choice for HD extensions up to 1920 x 1080p and 30 fps is H.264 (AVC) Progressive
27 High Profile Level 4.0 decoder [8].

28 The High Efficiency Video Coding (HEVC) resulted from a joint video coding standardization
29 project of the ITU-T Video Coding Experts Group (ITU-T Q.6/SG 16) and ISO/IEC Moving Pic-
30 ture Experts Group (ISO/IEC JTC 1/SC 29/WG 11). The final specification is available here [19].
31 Additional background information may be found at <http://hevc.info>.

32 The DASH-IF is interested in providing Interoperability Points and Extensions for established co-
33 decs configurations. It is not the intent of the DASH-IF to define typically deployed HEVC pro-
34 files/levels or the associated source formats. However, at the same time it is considered to provide
35 implementation guidelines supported by test material for DASH-based delivery as soon as the
36 industry has converged to profile/level combinations in order to support a dedicated format. For
37 this version of this document the following is considered:

- 38 • For HEVC-based video, it is expected that the minimum supported format is 720p. The
39 codec considered to support up to 1280 x 720p at 30 fps is HEVC Main Profile Main Tier
40 Level 3.1 [19].

- The choice for 8-bit HD extensions based on HEVC to support up to 2048 x 1080 and 60 fps is HEVC Main Profile Main Tier Level 4.1 [19].
 - The choice for 10-bit HD extensions based on HEVC to support up to 2048 x 1080 and 60 fps and 10 bit frame depth is HEVC Main10 Profile Main Tier Level 4.1 [19].
 - For UHD extensions refer to section 10.

Other profile/level combinations will be considered in updated versions of this document.

This document uses the terms SD, HD and UHD to the best knowledge of the community, but does not provide requirements for the formats, but primarily specifies the signaling and required receiver capabilities.

6.2.2. DASH-specific aspects for H.264/AVC video

For the integration of the above-referred codecs in the context of DASH, the following applies for H.264 (AVC):

- The encapsulation of H.264/MPEG-4 AVC video data is based on the ISO BMFF as defined in ISO/IEC 14496-15 [9].
 - Clients shall support H.264/AVC sample entries when SPS/PPS is provided in the Initialization Segment only according to ISO/IEC 14496-15, [9], i.e. sample entry 'avc1'.
 - Clients shall support Inband Storage for SPS/PPS based ISO/IEC 14496-15, [9], i.e. sample entry 'avc3'.
 - Service offerings using H.264/AVC may use sample entry 'avc1' or 'avc3'.
 - SAP types 1 and 2 correspond to IDR-frames in [8].
 - The signaling of the different video codec profile and levels for the codecs parameters according to RFC6381 [10] is documented in Table 19. Note that any of the codecs present in Table 19 conforms to the profile level combination that is supported in DASH-AVC/264. Other codecs strings may be used and conform as well.
 - Additional constraints within one Adaptation Set are provided in section 6.2.5.

Note: For a detailed description on how to derive the signaling for the codec profile for H.264/AVC, please refer to DVB DASH section 5.1.3.

Table 19 H.264 (AVC) Codecs parameter according to RFC6381 [10]



e
r

E 3 a
. . v
2 0 c
6 [
4 1
(,
A 3
V]
C .
) 4
" D
P Y
r 0
o 1
g E
r
e
s
s
i
v
e"
M
a
i
n
P
r
o
f
i
l
e

E 3 a
. . v
2 1 c
6 [
4 1
(,
A 3
V]
C .

```

) 6
P 4
r Y
o 0
g 1
r F
e 4 a
s . v
s 0 c
i [
v 1
e 1
E ,
i 3
g ]
h .
P 6
r 4
o Y
f 0
i 2
l 8
e

```

6.2.3. DASH-specific aspects for H.265/HEVC video

For the integration in the context of DASH, the following applies for HEVC

- The encapsulation of HEVC video data in ISO BMFF is defined in ISO/IEC 14496-15 [9]. Clients shall support both sample entries ' using 'hvc1' and 'hev1', i.e.. inband Storage for VPS/SPS/PPS.
- Additional constraints within one Adaptation Set are provided in section 6.2.5.
- For the signaling of HEVC IRAP Pictures in the ISOBMFF and in DASH, in particular the use of the sync sample table and of the SAP sample group, please refer to Table 20.

Table 20 Signaling of HEVC IRAP Pictures in the ISOBMFF and in DASH

NAL Unit Type	ISOBMFF sync status	DASH SAP type
IDR_N_LP	true	1
IDR_W_RADL	true	2 (if the IRAP has associated RADL pictures) 1 (if the IRAP has no associated RADL pictures)
BLA_N_LP	true	1
BLA_W_RADL	true	2 (if the IRAP has associated RADL pictures)

			1 (if the IRAP has no associated RADL pictures)
BLA_W_LP	false		3 (if the IRAP has associated RASL pictures)
	true		2 (if the IRAP has no associated RASL pictures but has associated RADL pictures)
	true		1 (if the IRAP has no associated leading pictures)
CRA	false		3 (if the IRAP has associated RASL pictures)
	true		2 (if the IRAP has no associated RASL pictures but has associated RADL pictures)
	true		1 (if the IRAP has no associated leading pictures)

In the above table, when there are multiple possible values for a given NAL Unit Type, if the entity creating the signaling is not able to determine correctly which signaling to use, it shall use the values in the first row of this table associated to the NAL Unit Type.

- The signaling of the different video codec profile and levels for the codecs parameters is according to ISO/IEC 14496-15 [9] Annex E. Note that any of the codecs present in Table 1 conforms to the profile level combination that is supported in DASH-HEVC.

NOTE: For a detailed description on how to derive the signaling for the codec profile for H.264/AVC, please refer to DVB DASH, section 5.2.2.

Table 21 Codecs parameter according to ISO/IEC 14496-15 [9]

Profile	Level	Tier	Codec Parameter
HEVC Main	3.1	Main	hev1.1.2.L93.B0 hvc1.1.2.L93.B0
	4.1	Main	hev1.1.2.L123.B0 hvc1.12.L123.B0
		Main	hev1.2.4.L123.B0 hvc1.2.4.L123.B0
HEVC Main-10	4.1	Main	hev1.2.4.L123.B0 hvc1.2.4.L123.B0

6.2.4. Video Metadata

The provisioning of video metadata in the MPD is discussed in section 3.2.4.

6.2.5. Adaptation Sets Constraints

6.2.5.1. General

Video Adaptation Sets shall contain Representations that are alternative encodings of the same source content. Video Adaptation Sets may contain Representations encoded at lower resolutions that are exactly divisible subsamples of the source image size. As a result, the cropped vertical and horizontal sample counts of all Representations can be scaled to a common display size without position shift or aspect ratio distortion that would be visible during adaptive switching. Subsample ratios must result in integer values for the resulting encoded sample counts (without rounding or truncation). The encoded sample count shall scale to the source video's exact active image aspect ratio when combined with the encoded sample aspect ratio value `aspect_ratio_idc` stored in the video Sequence Parameter Set NAL. Only the active video area shall be encoded so that devices can frame the height and width of the encoded video to the size and shape of their currently selected display area without extraneous padding in the decoded video, such as "letterbox bars" or "pillarbox bars".

All decoding parameter sets referenced by NALs in a Representation using 'avc1' or 'hvc1' sample description shall be indexed to that track's sample description table and decoder configuration record in the 'avCC' or 'hvCC' box contained in its Initialization Segment. All decoding parameter sets referenced by NALs in a Representation using 'avc3' or 'hev1' sample description shall be indexed to a Sequence Parameter NAL (SPS) and Picture Parameter NAL (PPS) stored prior to the first video sample in that Media Segment. For 'avc3' and 'hev1' sample description Representations, the SPS and PPS NALs stored in 'avCC' or 'hvCC' in the Initialization Segment shall only be used for decoder and display initialization, and shall equal the highest Tier, Profile, and Level of any SPS in the Representation. SPS and PPS stored in each Segment shall be used for decoding and display scaling.

For all Representations within an Adaptation Set with the following parameters shall apply.

- All the Initialization Segments for Representations within an Adaptation Set shall have the same sample description `codingname`. For example the inclusion of 'avc1' and 'avc3' based Representations within an Adaptation Set or the inclusion 'avc1' and 'hev1' based Representations within an Adaptation Set is not permitted.
- All Representations shall have equal timescale values in all `@timescale` attributes and '`tkhd`' `timescale` fields in Initialization Segments.
- If 'avc1' or 'hvc1' sample description is signaled in the `AdaptationSet@codecs` attribute, an edit list may be used to synchronize all Representations to the presentation timeline, and the edit offset value shall be equal for all Representations.
- Representations in one Adaptation Set shall not differ in any of the following parameters: Color Primaries, Transfer Characteristics and Matrix Coefficients. If Adaptation Sets differ in any of the above parameters, these parameters should be signaled on Adaptation Set level. If signaled, a Supplemental or Essential Property descriptor shall be used, with the `@schemeIdUri` set to `urn:mpeg:mpegB:cicp:<Parameter>` as defined in ISO/IEC 23001-8 [49] and `<Parameter>` one of the following: `ColourPrimaries`, `TransferCharacteristics`, **or** `MatrixCoefficients`. The `@value` attribute shall be set as defined in ISO/IEC 23001-8 [49].

1 **6.2.5.2. Bitstream Switching**

2 For AVC and HEVC video data, if the `@bitstreamswitching` flag is set to true, then the
3 following additional constraints shall apply:

- 4 • All Representations shall be encoded using 'avc3' sample description for AVC or
5 'hev1' for HEVC, and all IDR pictures shall be preceded by any SPS and PPS NAL
6 decoding parameter referenced by a video NAL in that codec video sequence.

7
8 Note: NAL parameter indexes in a Media Segment are scoped to that Segment. NALs and
9 indexes in the Initialization Segment may be different, and are only used for decoder initia-
10 lization, not Segment decoding.

- 11 • All Representations within a video Adaptation Set shall include an Initialization Segment
12 containing an 'avcC' or 'hvcC' Box containing a Decoder Configuration Record con-
13 taining SPS and PPS NALs that equal the highest Tier, Profile, Level, vertical and hori-
14 zontal sample count of any Media Segment in the Representation. HEVC Decoder Con-
15 figuration Records shall also include a VPS NAL.
- 16 • The **AdaptationSet@codecs** attribute shall be present and equal the maximum pro-
17 file and level of any Representation contained in the Adaptation Set.
- 18 • The **Representation@codecs** attribute may be present and in that case shall equal
19 the maximum profile and level of any Segment in the Representation.
- 20 • Edit lists shall not be present in video Adaptation Sets unless they are offered in On-Demand profile.
- 21 • If present in any Representation in the Adaptation Set, the edit list shall be identical in all
22 Representations.
- 23 • Video Media Segments shall set the first presented sample's composition time equal to
24 the first decoded sample's decode time, which equals the `baseMediaDecodeTime` in
25 the Track Fragment Decode Time Box ('tfdt').

26
27 Note: This requires the use of negative composition offsets in a v1 Track Run Box
28 ('trun') for video samples, otherwise video sample reordering will result in a delay of
29 video relative to audio.

- 30
31 • The `@presentationTimeOffset` attribute shall be sufficient to align audio video,
32 subtitle, and presentation timelines at presentation a Period's presentation start time. Any
33 edit lists present in Initialization Segments shall be ignored. It is strongly recommended
34 that the Presentation Time Offset at the start of each Period coincide with the first frame
35 of a Segment to improve decoding continuity at the start of Periods.

36 NOTE: An Adaptation Set with the attribute `AdaptationSet@bitstreamSwitching="true"` fulfills
37 the requirements of the DVB DASH specification [42].

38 See section 7.7 for additional Adaptation Set constraints related to content protection.

39 **6.2.6. Tiles of thumbnail images**

40 For providing easily accessible thumbnails with timing, Adaptation Sets with the new `@con-`
41 `tentType="image"` may be used in the MPD. A typical use case is for enhancing a scrub bar

1 with visual cues. The actual asset referred to is a rectangular tile of temporally equidistant thumb-
2 nails combined into one jpeg or png image. A tile, therefore is very similar to a video segment
3 from MPD timing point of view, but is typically much longer. As for video, different spatial reso-
4 lutions can be collected into one Adapation Set. To limit the implementation effort, only SegmentTemplate
5 with \$Number\$ is used to described the thumbnail tiles and their timing.

6 It is typically expected that the DASH client is able to process such Adaptation Sets by download-
7 ing the images and using browser-based processing to assign the thumbnails to the Media Presen-
8 tation timeline.

9 A lot of parameters are the same as for video, but the ones which are new for thumbnail tiles, the
10 rectangular grid dimensions are given as the value of the **EssentialProperty** with
11 @schemeIdUri set to "http://dashif.org/guidelines/thumbnail_tile".

- 12 • If the **EssentialProperty** descriptor with @schemeIdUri set to
13 http://dashif.org/guidelines/thumbnail_tile is present, the following
14 attributes and elements of the Adaptation Set shall be used to describe the tiling as fol-
15 lows: The value of the descriptor provides the horizontal and vertical number of the tiles
16 as unsigned integer, separated by an 'x'. The two values are referred to as `htiles` and
17 `vtiles` in the following.
- 18 • `@duration` expresses the duration of one tile in the media presentation timeline in the
19 timescale provided by the value of the `@timescale` attribute, if present, otherwise in
20 seconds. The value is referred to as `tduration` in the following.
- 21 • `@bandwidth` expresses the maximum tile size in bits divided by the duration of one tile
22 as provided by the `tduration` value.
- 23 • `@width` and `@height` expresses the spatial resolution of the tile. Note that the maxi-
24 mum dimension of a JPEG image is 64k in width and height.
- 25 • Each tile has assigned a number starting with 1. The tile number is referred as `tnumber`.
- 26 • The `@startNumber` may be present to indicate the number of the first tile in the Period
27 in the Period. If not present the first number is defaulted to 1. The value of the start num-
28 ber is referred to as `startnumber`.
- 29 • The `@presentationTimeOffset` may be present to indicate the presentation time
30 of the thumbnail sequence at the start of the period. If not present, the first number is de-
31 faulted to 0. The timescale is provided by the value of the `@timescale` attribute, if pre-
32 sent, otherwise in seconds. The value of the presentation time offset is referred to as `pto`.
- 33 Based on this information, the following information can be derived:

-
- 1 • information on how many thumbnails are included in one tile by multiplying the
 2 htiles with vtiles. This number is referred to as ttiles.
 3 • The first tile that can be requested in the Period is referred to as startnumber which is
 4 used in the segment template.
 5 • The presentation time in the period for each tile is defined as (tnumber - 1)*tdura-
 6 tion - pto.
 7 • The duration of each thumbnail is defined as tduration/ttiles. This value is re-
 8 ferred to as thduration.
 9 • Thumbnail ordering is from left to right, row by row, starting from the top row. The last
 10 tile in a Period may have thumbnails outside the time interval. The content provider may
 11 use any padding pixel to such a tile, e.g. add black thumbnails. The number of a thumb-
 12 nail in a tile is referred to as thnumber.
 13 • The presentation time of the thumbnail within the tile is defined as (thnumber -
 14 1)*thduration, i.e. the presentation within a period is defined as (tnumber-1)
 15)*tduration + (thnumber - 1)*thduration - pto.
 16 • The vertical size of each thumbnail is defined as @height/vtiles and the horizontal
 17 size of each thumbnail is defined as @width/htiles.
 18 • The total number of thumbnails in a Period, referred to ttumbnails is the ceiling of the
 19 period duration divided by the thduration.

20 An example Adaptation Set for tile-based thumbnails is provided below:

```
21 <AdaptationSet id="3" mimeType="image/jpeg" contentType="image">
22   <SegmentTemplate media="$RepresentationID$/tile$Number$.jpg" duration="125"
23   startNumber="1"/>
24   <Representation bandwidth="10000" id="thumbnails" width="6400" height="180">
25     <EssentialProperty schemeIdUri="http://dashif.org/guidelines/thumbtile"
26     value="25x1"/>
27   </Representation>
28 </AdaptationSet>
```

29

30 Here

- 31 • htiles is 25 and vtiles is set to 1. ttiles is derived as 25.
 32 • tduration is 125 seconds and thduration is derived as 5 seconds.
 33 • Assuming a Period duration of 899 seconds, the total number of thumbnails is 180.
 34 • The vertical size of each thumbnail is 256 and the horizontal size is 180.
 35 • The maximum bandwidth for each tile is 10 kbit/s.

1 6.3. Audio

2 6.3.1. General

3 Content offered according to DASH-IF IOP is expected to contain an audio component in most
4 cases. Therefore, clients consuming DASH-IF IOP-based content are expected to support stereo
5 audio. Multichannel audio support and support for additional codecs is defined in extensions in
6 section 9 of this document.

7 The codec for basic stereo audio support is MPEG-4 High Efficiency AAC v2 Profile, level 2 [11].

8 Notes

- 9 • HE-AACv2 is also standardized as Enhanced aacPlus in 3GPP TS 26.401 [13].
- 10 • HE-AACv2 Profile decoder [8] can also decode any content that conforms to
 - 11 ○ MPEG-4 AAC Profile [11]
 - 12 ○ MPEG-4 HE-AAC Profile [11]

13 Therefore, Broadcasters and service providers encoding DASH-AVC/264 content are free to use
14 any AAC version. It is expected that clients supporting the DASH-IF IOP interoperability point
15 will be able to play AAC-LC, HE-AAC and HE-AACv2 encoded content.

16 For all HE-AAC and HE-AACv2 bitstreams, explicit backwards compatible signaling should be
17 used to indicate the use of the SBR and PS coding tools.

18 Note: To conform to the DVB DASH profile [42], explicit backwards compatible signaling
19 shall be used to indicate the use of the SBR and PS coding tools.

20 For advanced audio technologies, please refer to section 9.

21 6.3.2. DASH-specific aspects for HE-AACv2 audio

22 In the context of DASH, the following applies for the High Efficiency AAC v2 Profile

- 23 • The content should be prepared according to the MPEG-DASH Implementation Guidelines
24 [6] to make sure each (Sub)Segment starts with a SAP of type 1.
- 25 • The signaling of MPEG-4 High Efficiency AAC v2 for the codecs parameters is according
26 to IETF RFC6381 [10] and is documented in Table 22. Table 22 also provides information
27 on the ISO BMFF encapsulation.
- 28 • For content with SBR, i.e. @codecs=mp4a.40.5 or @codecs=mp4a.40.29, @au-
29 dioSamplingRate signals the resulting sampling rate after SBR is applied, e.g. 48 kHz
30 even if the AAC-LC core operates at 24 kHz. For content with PS, i.e. @co-
31 decs=mp4a.40.29, **AudioChannelConfiguration** signals the resulting channel
32 configuration after PS is applied, e.g. stereo even if the AAC-LC core operates at mono.

33 34 Table 22 HE-AACv2 Codecs parameter according to RFC6381 [10]

Codec	Codec Parameter	ISO BMFF Encapsulation	SAP type
MPEG-4 AAC Profile [11]	mp4a.40.2	ISO/IEC 14496-14 [12]	1

MPEG-4 HE-AAC Profile [11]	mp4a.40.5	ISO/IEC 14496-14 [12]	1
MPEG-4 HE-AAC v2 Profile [11]	mp4a.40.29	ISO/IEC 14496-14 [12]	1

1 **Note:** Since both, HE-AAC and HE-AACv2 are based on AAC-LC, for the above-mentioned “Co-
2 dec Parameter” the following is implied:

- 3 • mp4a.40.5 = mp4a.40.2 + mp4a.40.5
- 4 • mp4a.40.29 = mp4a.40.2 + mp4a.40.5 + mp4a.40.29

5 **6.3.3. Audio Metadata**

6 **6.3.3.1. General**

7 Metadata for audio services is defined in ISO/IEC 23009-1.

8 **6.3.3.2. ISO/IEC 23009-1 audio data**

9 With respect to the audio metadata, the following elements and attributes from ISO/IEC 23009-1
10 are relevant:

- 11 • the @audioSamplingRate attribute for signaling the sampling rate of the audio media
12 component type in section 5.3.7 of ISO/IEC 23009-1
- 13 • the **AudioChannelConfiguration** element for signaling audio channel configura-
14 tion of the audio media component type.in section 5.3.7 of ISO/IEC 23009-1. For this ele-
15 ment the scheme and values defined in ISO/IEC 23001-8 for the **ChannelConfigura-**
16 **tion** should be used.

17 **6.4. Auxiliary Components**

18 **6.4.1. Introduction**

19 Beyond regular audio and video support, TV programs typically also require support for auxiliary
20 components such as subtitles and closed captioning, often due to regulatory requirements. DASH-
21 IF IOP provides tools to addresses these requirements.

22 **6.4.2. Subtitles and Closed Captioning**

23 Technologies for subtitles are as follows:

- 24 • CEA-608/708 Digital Television (DTV) Closed Captioning [14]
- 25 • IMSC1 [61] conformant profiles of TTML, packaged as Segments conforming to MPEG-
26 4, Part 30 [29], including subsets such as:
 - 27 ○ W3C TTML [16]
 - 28 ○ SMPTE Timed Text [17] (including image-based subtitles and closed captioning)
 - 29 ○ EBU-TT [20]
- 30 • 3GPP Timed Text [15]
- 31 • Web VTT [18]

32 For simple use cases, CEA-608/708 based signaling as defined in section 6.4.3 may be used.

1 For any other use cases, IMSC1 [61] should be used as defined in section 6.4.4. It is expected that
2 most subset profiles of IMSC1 would be reasonably decodable.

3 TTML and WebVTT Media Segments shall be referenced by Representation elements in MPDs,
4 downloaded, initialized, and synchronized for multimedia presentation the same as audio and
5 video Segments.

6 Note: DASH playback applications such as Web pages can download TTML or WebVTT text files, initialize
7 renderers, and synchronize rendering and composition. This specification does not specify interoperable play-
8 back of these “sidecar” subtitle files in combination with a DASH audio visual presentation. However, sec-
9 tion 6.4.5 provides guidelines on how to synchronize side car files at Period boundaries.

10 **6.4.3. CEA-608/708 in SEI messages**

11 **6.4.3.1. Background**

12 In order to provide the signaling of the presence of SEI-based data streams and closed captioning
13 services on MPD level, descriptors on DASH level are defined. This section provides some back-
14 ground.

15 Note: This method is compatible with draft SCTE specification DVS 1208 and therefore
16 SCTE URNs are used for the descriptor @schemeIdUri. In an updated version of this
17 document more details on the exact relation to the SCTE specification will be provided.

18 The presence of captions and their carriage within the SEI message of a video track is defined in
19 ANSI/SCTE 128-1 2013 [43], section 8.1 Encoding and transport of caption, active format de-
20 scription (AFD) and bar data.

21 Based on this it is enabled that a video track carries SEI message that carry CEA-608/708 CC. The
22 SEI message payload_type=4 is used to indicates that Rec. ITU-T T.35 based SEI messages
23 are in use.

24 In summary the following is included in ANSI/SCTE 128-1 2013 to signal CEA-608/708 CC:

- 25 • SEI payloadType is set to 4
- 26 • itu_t_t35_country_code – A fixed 8-bit field, the value of which shall be 0xB5.
- 27 • itu_t_35_provider_code – A fixed 16-bit field registered by the ATSC. The value
shall be 0x0031.
- 28 • user_identifier – This is a 32 bit code that indicates the contents of the
user_structure() and is 0x47413934 (“GA94”).
- 29 • user_structure() – This is a variable length data structure ATSC1_data() defined
in section 8.2 of ANSI/SCTE 128 2013-a.
- 30 • user_data_type_code is set to 0x03 for indicating captioning data in the
user_data_type_structure()
- 31 • user_data_type_structure() is defined in section 8.2.2 of ANSI/SCTE 128-1
2013 for Closed Captioning and defines the details on how to encapsulate the captioning
data.

32 The semantics of relevant Caption Service Metadata is provided in CEA-708 [14], section 4.5:

- 33 • the total number of caption services (1-16) present over some transport-specific period.
- 34 • For each service:

-
- The type of the service, i.e. being 608 or 708. According to CEA-708 [14], section 4.5, there shall be at most one CEA-608 data stream signaled. The CEA-608 datastream itself signals the individual CEA-608-E caption channels.
 - When the type of the service is 708, then the following 708-related metadata should be conveyed:
 - SERVICE NUMBER: the service number as found on the 708 caption service block header (1-31). This field provides the linkage of the remaining metadata to a specific 708 caption service
 - LANGUAGE: the dominant language of the caption service, recommended to be encoded from ISO 639.2/B [45].
 - DISPLAY ASPECT RATIO {4:3, 16:9}: The display aspect ratio assumed by the caption authoring in formatting the caption windows and contents.
 - EASY READER: this metadata item, when present, indicates that the service contains text tailored to the needs of beginning readers.

6.4.3.2. MPD-based Signaling of SEI-based CEA-608/708 Closed Caption services

This subsection provides methods MPD-based Signaling of SEI-based CEA-608/708 Closed Caption services, i.e.

- The presence of one or several SEI-based closed caption services in a Representation.
- The signaling of the relevant Caption Service Metadata as defined in CEA-708 [14], section 4.5.

The descriptor mechanism in DASH is used for this purpose.

Signaling is provided by including **Accessibility** descriptors, one each for CEA 608 and CEA 708 and is described in sections 6.4.3.3 and 6.4.3.4, respectively. The **Accessibility** descriptor is included for the **AdaptationSet** and all included Representations shall provide equivalent captions.

The @value attribute of each descriptor can be either list of languages or a complete map of services (or CC channels, in CEA-608 terminology). Listing languages without service or channel information is strongly discouraged if more than one caption service is present.

These definitions are equivalent to SCTE 214-1 [56].

6.4.3.3. Signaling CEA-608 caption service metadata

The **Accessibility** descriptor shall be provided with @schemeIdUri set to urn:scte:dash:cc:cea-608:2015, and an optional @value attribute to describe the captions. If the @value attribute is not present, the Representation contains a CEA-608 based closed captioning service.

If present, the @value attribute shall contain a description of caption service(s) provided in the stream as a list of channel-language pairs. Alternatively, a simple list of language codes may be provided, but this is strongly discouraged as it will not provide sufficient information to map the language with the appropriate caption channel.

The @value syntax shall be as described in the ABNF below.

```
1 @value          = (channel *3 [";" channel]) / (language *3[";" language])
2 channel         = channel-number "=" language
3 channel-number = CC1 | CC2 | CC3 | CC4
4 language        = 3ALPHA ; language code per ISO 639.2/B [45]
```

5 **6.4.3.4. Signaling CEA-708 caption service metadata**

6 DASH-IF IOPs do not provide any interoperability guidelines for CEA-708.

7 Note: Caption Service Metadata is provided in SCTE 214-1 [56], section 4.5.

8 **6.4.3.5. Examples**

9 Simple signaling of presence of CEA-608 based closed caption service (Note: Not signaling lan-
10 guages is a discouraged practice)

```
11 <Accessibility
12   schemeIdUri="urn:scte:dash:cc:cea-608:2015"/>
```

13 Signaling of presence of CEA-608 closed caption service languages in English and German

```
14 <Accessibility
15   schemeIdUri="urn:scte:dash:cc:cea-608:2015"
16   value="eng;deu"/>
```

17 Signaling of presence of CEA-608 closed caption service in English and German, with channel
18 assignments

```
19 <Accessibility
20   schemeIdUri="urn:scte:dash:cc:cea-608:2015"
21   value="CC1=eng;CC3=deu"/>
```

22 Signaling of presence of CEA-708 closed caption service in English and German

```
23 <Accessibility
24   schemeIdUri="urn:scte:dash:cc:cea-708:2015"
25   value="1=lang:eng;2=lang:deu"/>
```

26 Signaling of presence of CEA-708 closed caption service in English and easy reader English

```
27 <Accessibility
28   schemeIdUri="urn:scte:dash:cc:cea-708:2015"
29   value="1=lang:eng;2=lang:eng,war:1,er:1"/>
```

30 **6.4.4. Timed Text (IMSC1)**

31 W3C TTML [16] and its various profiles - W3C IMSC1 (text and image profiles) [61], SMPTE
32 Timed Text [17], and EBU Timed Text [20] - provide a rich feature set for subtitles. Beyond basic
33 subtitles and closed captioning, for example, graphics-based subtitles and closed captioning are
34 also supported by IMSC1. Conversion of CEA-608 and CEA-708 into IMSC1 may be done ac-
35 cording to SMPTE 2052-10 [27] and SMPTE-2052-11 [28], respectively. The Timed Text track
36 shall conform to IMSC1 [61]. Note that by the choice of IMSC1 as the supported format at the
37 client, other formats such as EBU TT [20] are also supported because they are subset profiles.

38 In the context of DASH, the following applies for text/subtitling:

- 39 • All graphics type samples shall be SAP type 1. The signalling of the different text/subtitling
40 codecs for the codecs parameters is according to W3C TTML Profile Registry [62] and is
41 documented in Table 23.
- 42 • Table 23 also provides information on ISO BMFF encapsulation.

1
2

Table 23 Subtitle MIME type and codecs parameter according to IANA and W3C registries

Codec	MIME type	Codecs Parameter @codecs	ISO BMFF En- capsulation
IMSC1 Timed Text [61] without encapsulation	application/ttml+xml ^(1,3)	See [62]	n/a
IMSC1 Timed Text [61] with ISO BMFF encapsulation	application/mp4	See [62]	ISO/IEC 14496-12 [7] ISO/IEC 14496-30 [29]

Notes:

(1) DVB DASH only supports ISO BMFF encapsulated TT, but not XML-based.

3

6.4.5. Guidelines for side-loaded TTML and WebVTT files

Side-loaded TTML or WebVTT subtitles or caption files can be used by some players including dash.js. Such files can be indicated in the manifest like:

```
7 <AdaptationSet contentType="text" mimeType="application/ttml+xml" lang="swe">
8   <Role schemeIdUri="urn:mpeg:dash:role:2011" value="subtitle"/>
9   <Representation id="xml_swe" bandwidth="1000">
10    <BaseUrl>sub_swe_short.xml</BaseUrl>
11   </Representation>
12 </AdaptationSet>
```

Only one file for the full period is permitted, practically limiting this use case to non-live content.

Such external files are assumed to have a timeline aligned with the Period, so that TTML time 00:00:00.000 corresponds to the start of the Period. The presentation time offset is expected to be not presented, and if present, expected to be ignored by the DASH client.

The same applies to side-loaded WebVTT files. In that case, the @mimeType is text/vtt.

If segmented subtitles are needed, such as for live sources, segments shall be packaged in ISOBMFF containers, in TTML or WebVTT format, with timing according to [29]. In particular, this means that the TTML timing inside the segments is with respect to the media timeline. ".

6.4.6. Annotation of Subtitles

Subtitles should be annotated properly using descriptors available in ISO/IEC 23009-1, Specifically Role, Accessibility, Essential Property and Supplemental Property descriptors and the DASH

1 role scheme may be used. Guidelines for annotation are for example provided in DVB DASH,
2 section 7.1.2 or SCTE 214-1 [56], section 7.2.

3 **7. Content Protection and Security**

4 **7.1. Introduction**

5 DASH-IF IOPs do not intend to specify a full end-to-end DRM system. However DASH-IF IOP
6 provides a framework for multiple DRMs to protect DASH content by adding instructions or *Pro-
7 tection System Specific*, proprietary information in predetermined locations in MPDs, or DASH
8 content that is encrypted with Common Encryption as defined in ISO/IEC 23001-7 [30].

9 The Common Encryption ('cenc') protection scheme specifies encryption parameters that can be
10 applied by a scrambling system and key mapping methods using a common key identifier (KID)
11 to be used by different DRM systems such that the same encrypted version of a file can be com-
12 bined with different DRM systems that can store proprietary information for licensing and key
13 retrieval in the Protection System Specific Header Box ('pssh'), or in **ContentProtection**
14 Descriptors in an MPD. The DRM scheme for each pssh is identified by a DRM specific Sys-
15 temID.

16 The recommendations in this document reduce the encryption parameters and use of the encryption
17 metadata to specific use cases for VOD and live content with key rotation.

18 The base technologies are introduced first followed by informative chapter on standardized ele-
19 ments. Additional Content Protection Constraints are then listed that are specific to conformance
20 to DASH-264/AVC IOP.

21 **7.2. HTTPS and DASH**

22 Transport security in HTTP-based delivery may be achieved by using HTTP over TLS (HTTPS)
23 as specified in RFC 5246. HTTPS is a protocol for secure communication which is widely used
24 on the Internet and also increasingly used for content streaming, mainly for the following purposes:

- 25 • Protecting the privacy of the exchanged data from eavesdropping by providing encryp-
26 tion of bidirectional communications between a client and a server, and
- 27 • Ensuring integrity of the exchanged data against man-in-the-middle attacks against tam-
28 pering with and/or forging the contents of the communication.

29 As a MPD carries links to media resources, web browsers follow the W3C recommendation for
30 mixed content (<https://www.w3.org/TR/mixed-content/>). To ensure that HTTPS benefits are main-
31 tained once the MPD is delivered, it is recommended that if the MPD is delivered with HTTPS,
32 then the media also be delivered with HTTPS.

33 In addition, MPEG-DASH explicitly permits the use of https as a scheme and hence, HTTP over
34 TLS as a transport protocol. When using HTTPS in DASH, one can for instance specify that all
35 media segments are delivered over HTTPS, by declaring that all the <BaseURL>'s are HTTPS based,
36 as follow:

37 <BaseURL><https://cdn1.example.com/></BaseURL>
38 <BaseURL><https://cdn2.example.com/></BaseURL>

1 One can also use HTTPS for retrieving other types of data carried with an MPD that are HTTP-
2 URL based, such as, for example, DRM licenses specified within the <ContentProtection> de-
3 scriptor:

4 <ContentProtection schemeIdUri="http://example.net/052011/drm">
5 <drm:License>https://MoviesSP.example.com/protect?license=kljklksdfiowek</drm:License>
6 </ContentProtection>

7 It is recommended to adopt HTTPS for delivering DASH content. It must be noted nevertheless,
8 that HTTPS does hurt proxies that attempt to intercept, cache and/or modify content between the
9 client and the CDN that holds the delivery certs. Since the HTTPS traffic is opaque to these inter-
10 mediate nodes, they can loose much of their intended functionality when faced with HTTPS traffic.

11 While using HTTPS in DASH provides good levels of trust and authenticity for data exchanged
12 between DASH servers and clients connected over HTTPS, it should be pointed out that HTTPS
13 only protects the transport link, but not the access to streaming content and the usage of streamed
14 content. HTTPS itself does not imply user authentication and content authorization (or access con-
15 trol). This is especially the case that HTTPS provides no protection to any streamed content cached
16 in a local buffer at a client for playback. HTTPS does not replace a DRM.

17 **7.3. Base Technologies Summary**

18 The normative standard that defines common encryption in combination with ISO BMFF is
19 ISO/IEC 23001-7 [30]. It includes:

- 20 • Common ENCryption (CENC) of NAL structure video and other media data with AES-
21 128 CTR mode
- 22 • Support for decryption of a single Representation by multiple DRM systems
- 23 • Key rotation (changing media keys over time)
- 24 • XML syntax for expressing a default KID attribute and pssh element in MPDs

25 The main DRM components are:

- 26 1. The **ContentProtection** descriptors in the MPD (see [4], 5.3.7.2-Table 9, 5.8.5.2 and
27 [4] 5.8.4.1) that contains the URI for signaling of the use of Common Encryption or the
28 specific DRM being used.
- 29 2. ‘tenc’ parameters that specify encryption parameters and `default_KID` (see [30] 8.2).
30 The ‘tenc’ information is in the Initialization Segment. Any KIDs in Movie Fragment
31 sample group description boxes override the ‘tenc’ parameter of the `default_KID`, as
32 well as the ‘not encrypted’ parameter. Keys referenced by KID in sample group descrip-
33 tions must be available when samples are available for decryption, and may be stored in a
34 protection system specific header box (‘pssh’) in each movie fragment box (‘moof’).
35 The `default_KID` information may also appear in the MPD (see [30] 11).
- 36 3. ‘senc’ parameters that may store initialization vectors and subsample encryption ranges.
37 The ‘senc’ box is stored in each track fragment box (‘traf’) of an encrypted track (see
38 [30] 7.1), and the stored parameters accessed using the sample auxiliary information offset
39 box (‘saio’)) and the sample auxiliary information size box (‘saiz’) (see [4] 8.7.8 and
40 8.7.9).

1 4. ‘pssh’ license acquisition data or keys for each DRM in a format that is “Protection System
2 Specific”. ‘pssh’ refers to the Protection System Specific Header box described in
3 [30], 8.1.2. ‘pssh’ boxes may be stored in Initialization or Media Segments (see [31]
4 8.1 and 8.2). It may also be present in a cenc:pssh element in the MPD (see [4] 5.8.4.1,
5 [30] 11.2.1). cenc:pssh information in the MPD allows faster parsing, earlier access,
6 identification of duplicate license requests, and addition of DRMs without content modifi-
7 cation. ‘pssh’ boxes in Initialization Segments are not recommended because they trigger
8 a license request each time an Initialization Segment is processed in a Web browser for
9 each Representation and bitrate switch.

10 *Note: The duplication of the pssh information in the Initialization Segment may cause
11 difficulties in playback with HTML5 - EME based players. I.e. content will fail unless play-
12 ers build complex DRM specific license handling.*

13 5. Key rotation is mainly used to allow changes in entitlement for continuous live content. It
14 is used as defined in [30] with the following requirements:

- 15 • Sample To Group Box ('sbgp') and Sample Group Description Box ('sgpd')
16 of type 'seig' are used to indicate the KID applied to each sample, and changes
17 to KIDs over time (i.e. “key rotation”). (see [4] 8.9.4) KIDs referenced by sample
18 groups must have the keys corresponding to those KIDs available when the sam-
19 ples in a Segment are available for decryption. Keys referenced by sample groups
20 in a Segment may be stored in that Segment in Protection System Specific Header
21 Boxes ('pssh') stored in the Movie Fragment Box ('moof'). A version 1
22 ‘pssh’ box may be used to list the KID values stored to enable removal of
23 duplicate boxes if a file is defragmented.
- 24 • Keys stored in Media Segment ‘pssh’ boxes must be stored in the same DRM
25 format for all users so that the same Media Segments can be shared by all users.
26 User-specific information must be delivered “out of band”, as in a “root” license
27 associated with the default_KID, which can be individualized for each DRM
28 client, and control access to the shared ‘pssh’ information stored in Media Seg-
29 ments, e.g. by encrypting the keys stored in Segment ‘pssh’ boxes with a “root
30 key” provided by the user-specific DRM root license. Common Encryption speci-
31 fies ‘pssh’ to enable key storage in movie fragments/Segments; but it does not
32 preclude other methods of key delivery that satisfy KID indexing and availability
33 requirements.
- 34 • For details see Section 7.5.

35 **7.4. ISO BMFF Support for Common Encryption and DRM**

36 **7.4.1. Box Hierarchy**

37 The ISO Media Format carries content protection information in different locations. Their hierar-
38 chy is explained in the informational chapter below, followed by a reference on where these ele-
39 ments are standardized.

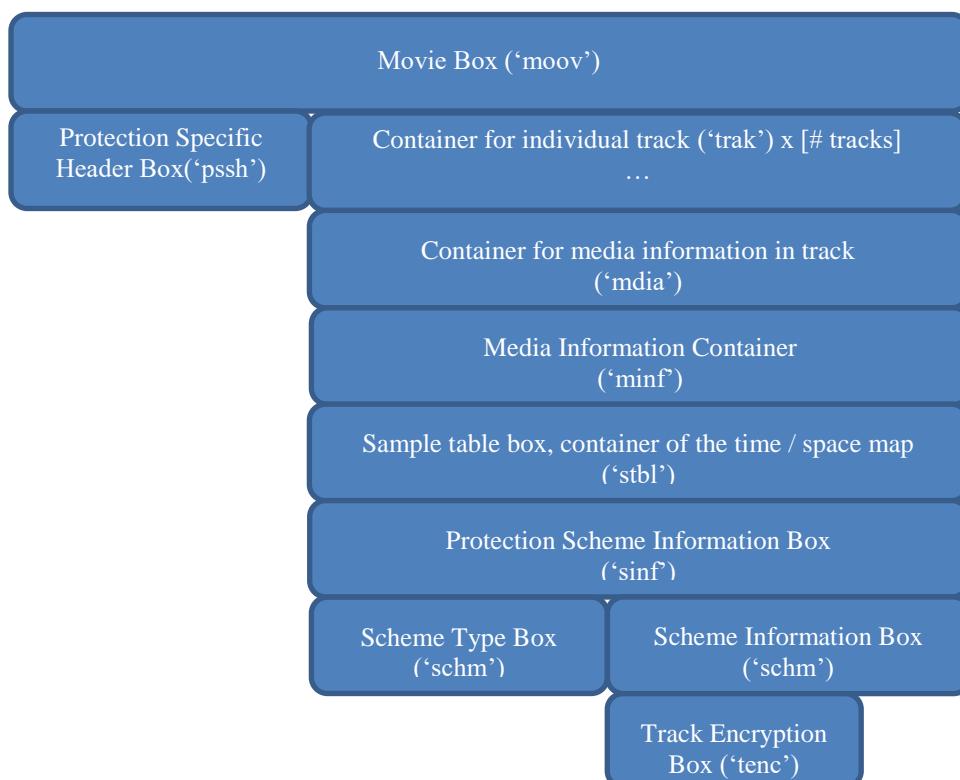
1 The following shows the box hierarchy and composition for relevant boxes, when using common
2 encryption:

- 3 • moov/pssh (zero or one per system ID)
- 4 • moov/trak/mdia/minf/stbl/stsd/sinf/schm (one, if encrypted)
- 5 • moov/trak/mdia/minf/stbl/stsd/sinf/schi/tenc (one, if encrypted)
- 6 • moof/traf/saiz (one, if encrypted)
- 7 • moof/traf/saio (one, if encrypted)
- 8 • moof/traf/senc (one, if encrypted)

9 for key rotation

- 10 • moof/traf/sbgp (one per sample group)
- 11 • moof/traf/sgpd 'seig' (sample group entry) (one per sample group)
- 12 • moof/pssh (zero or one per system ID)

13 Graphical overviews of above structure for VOD content and live content are shown in Figure 21
14 and Figure 22 respectively.



15

16

Figure 21: Visualization of box structure for single key content

17

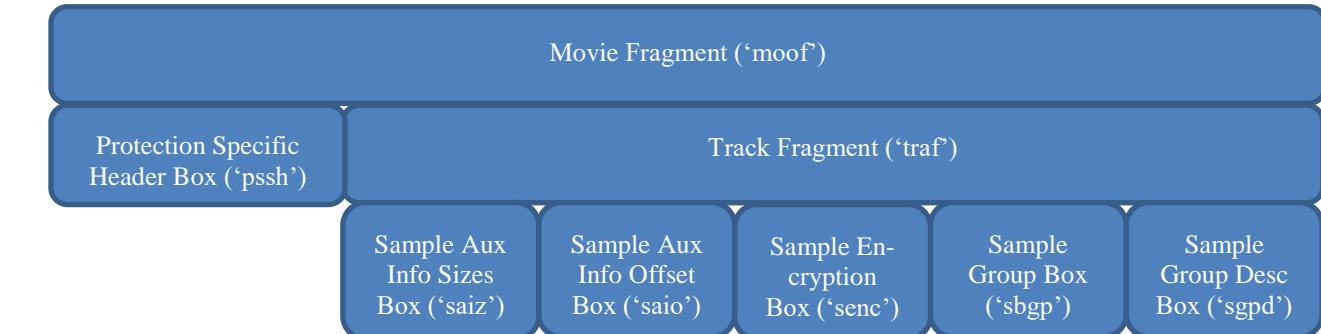


Figure 22: Visualization of box structure with key rotation

7.4.2. ISO BMFF Structure Overview

Table 24 provides pointers to relevant information in the specifications to understand the standard DRM components and if the main description is located in the ISO base media file format ([7]), or the Common Encryption specification ([30]).

Table 24 Boxes relevant for DRM systems

Box	Full Name / Usage	Reference
moof	movie fragment header <i>One 'moof' box for each fragment, i.e. Media Segment/Subsegment.</i>	[7] 8.32 + [4]
moov	movie header, container for metadata <i>One 'moov' box per file.</i>	[7] 8.1
pssh	Protection System Specific Header Box <i>Contains DRM specific data. pssh box version 1 (specified in Common Encryption 2nd edition) contains a list of KIDs to allow removing duplicate 'pssh' boxes when defragmenting a file by comparing their KIDs</i>	[30] 8.1.1
saio	Sample Auxiliary Information Offsets Box <i>Contains the offset to the IVs & subsample encryption byte ranges.</i>	[7] 8.7.9
saiz	Sample Auxiliary Information Sizes Box <i>Contains the size of the IVs & subsample encryption byte ranges.</i>	[7] 8.7.8
senc	Sample Encryption Box <i>Contains Initialization Vectors; and subsample ranges for a Media Segment</i>	[30] 7.1
schi	Scheme Information Box <i>Container boxes used by that protection scheme type.</i>	[7] 8.12.6 + [30] 4
schm	Scheme Type Box <i>Contains the encryption scheme, identified by a 4 character code, e.g. 'cenc'</i>	[7], 8.12.5 + [30] 4
seig	Cenc Sample Encryption Information Video Group Entry <i>A sample description containing KIDs describing sample groups in this segment, for key rotation.</i>	[30] 6

sbgp	Sample to Group Box <i>lists a group of samples</i>	[7] +[30] 5
sgpd	Sample Group Description Box <i>Describes properties of a sample group</i>	[7] 8.9.3 + [30] 5
sinf	Protection Scheme Information Box <i>Signals that the stream is encrypted</i>	[7] 8.12.1 + [30] 4
sts _d	Sample description table (codec type, initialization parameters, stream layout, etc.)	[7] 8.16
tenc	Track Encryption Box <i>Contains default encryption parameters for the entire track, e.g. default_KID</i>	[30] 8.2.1

- 1
- 2 **7.5. Periodic Re-Authorization**
- 3 **7.5.1. Introduction**
- 4 This section explains different options and tradeoffs to enable change in keys (aka key rotation),
5 considering different use cases, application scenarios, content encoding variants and signaling re-
6 quirements.
- 7 **7.5.2. Use Cases and Requirements**
- 8 The main use case in this context is to enable service changes at program boundaries, not to in-
9 crease security of CENC by preventing e.g. key factoring or key redistribution. In order to clarify
10 this application, the term *periodic re-authorization* is used instead of the term *key rotation*.
- 11 In addition, this is one of the ways to implement counting of active streams as they are periodically
12 requesting keys from a license server.
- 13 The following use cases and requirements have been considered:
- 14
 - Ability to force a client device to re-authorize to verify that it is still authorized for con-
15 tent consumption.
 - Support for distribution models such as: Live content, PVR, PPV, VOD, SVOD, live to
16 VOD, network DVR. This includes where live content is converted into another con-
17 sumption license for e.g. catch up TV.
 - Uninterrupted playback when keys are rotated.
 - Preventing of client storm: Requests from client should be distributed where pos-
21 sible to prevent spiking loads at isolated times.
 - Quick recovery: If the server or many client devices fail, the service should be
22 able to resume quickly.
 - Player visibility into the key rotation signal
 - Regional blackout: Device location may be taken into account to enable de-activation of
26 content in a geographical area.

-
- 1 • Hybrid broadcast/unicast networks in which receivers operating in broadcast-only mode at
2 least some of the time, i.e. unable to always download licenses on-demand through unicast.
3 • No required changes to the standard process and validity of MPDs.

4 **7.5.3. Implementation Options**

5 **7.5.3.1. General**

6 This section describes approaches for *periodic re-authorization*; recommended because they best
7 cover the use cases and allow interoperable implementation. Other approaches are possible and
8 may be considered by individual implementers.

9 One of those is explicit signaling using e.g. `esmg` messages, using a custom key rotation signal to
10 indicate future KIDs.

11 To prevent the initial client storm to retrieve the first keys, before they are rotated, the initial `pssh`
12 parameters SHOULD be included in the MPD as described in 7.4.1.

14 **7.5.3.2. Period Boundaries**

15 One possibility is to use a DASH Period as minimum key duration interval and existing MPD level
16 signaling for KID.

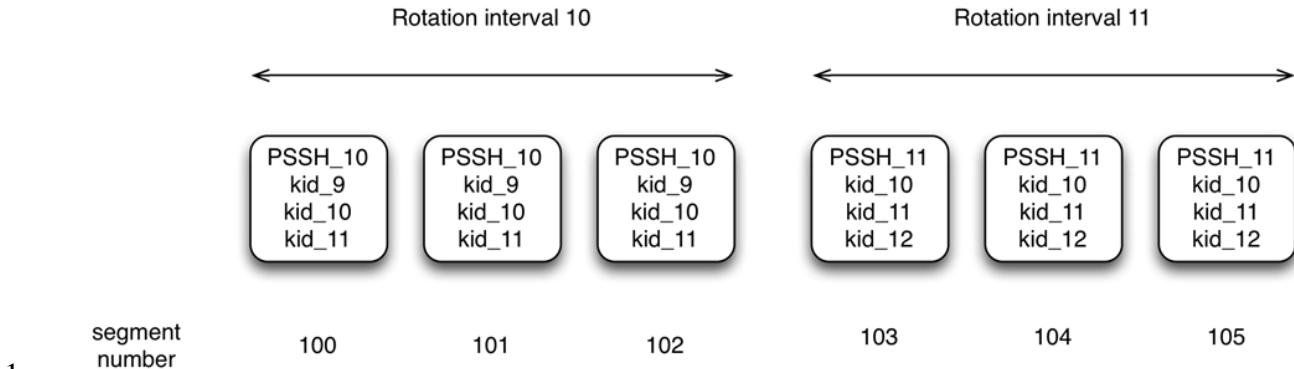
17 This is a simple implementation and a possible alternative but has limitations in the flexibility:

- 18 • The signal does not allow for early warning and time to switch the encryption keys and
19 context.
- 20 • The logic of the periods is decided by content creation not DRM. Boundaries may not be
21 suited and period may be longer than desired key interval

22 **7.5.3.3. Future Keys in pssh**

23 This approach considers the protection system to be responsible to manage notification and key
24 retrieval that prevents a client storm. The `pssh` information is used for signaling in a *content protected system*
25 proprietary form. No additional signaling mechanism is created and the DRM is
26 managing key rotation by providing extra information in the Protection System Specific Header
27 Box ('`pssh`') (see [4]). To prevent a client storm on key change boundaries the following imple-
28 mentation options can be considered. They are listed for informational purpose and do not affect
29 the guidelines on content formatting.

30 Current and future keys or access information and validity times are provided in a proprietary
31 format in the `pssh` (see example in figure below). The client can chose a random time to use the
32 access information to request licenses so that requests are distributed over time.



1 scheme indicated in the @value attribute. The file structure of content protection schemes is
2 specified in [7], 5.8.5.2, and the @value = ‘cenc’ for the Common Encryption scheme, as spec-
3 ified in [30]. Although the **ContentProtection** Descriptor for UUID Scheme described be-
4 low is usually used for license acquisition, the **ContentProtection** Descriptor with
5 @schemeIdUri="urn:mpeg:dash:mp4protection:2011" and with @cenc:de-
6 fault_KID may be sufficient to acquire a license or identify a previously acquired license that
7 can be used to decrypt the Adaptation Set. It may also be sufficient to identify encrypted content
8 in the MPD when combined with license acquisition information stored in ‘pssh’ boxes in Ini-
9 tialization Segments.

10 A **ContentProtection** Descriptor for the mp4 Protection Scheme shall be used to identify
11 the default KID, as specified by the ‘tenc’ box, using the @cenc:default_KID attribute de-
12 fined in [30], section 11.1. The value of the attribute is the KID expressed in UUID string notation.
13

```
<ContentProtection schemeIdUri="urn:mpeg:dash:mp4protection:2011"  
value="cenc" cenc:default_KID="34e5db32-8625-47cd-ba06-68fca0655a72"/>
```

14 When starting playback of any Adaptation Set, the client should interact with the DRM system to
15 verify that the media key identified by the adaptation set’s default KID is available and should not
16 assume that a media key is available for decrypting content unless so signaled by the DRM system.
17

18 When the default_KID is present on each Adaptation Set, it allows a player to determine if a
19 new license needs to be acquired for each Adaptation Set by comparing their default_KIDs
20 with each other, and with the default_KIDs of stored licenses. A player can simply compare
21 these KID strings and determine what unique licenses are necessary without interpreting license
22 information specific to each DRM system.

23 7.6.2.2. ContentProtection Descriptor for UUID Scheme

24 A UUID **ContentProtection** descriptor in the MPD may indicate the availability of a par-
25 ticular DRM scheme for license acquisition. An example is provided below:
26

```
<ContentProtection  
schemeIdUri="urn:uuid:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx"  
value="DRMNAME version"/>
```

27 The schemeIdUri uses a UUID URN with the UUID string equal to the registered SystemID for a
28 particular DRM system. A list of known DRM SystemIDs can be found in the DASH identifier
29 repository available here: <http://www.dashif.org/identifiers/protection>.
30

31 This is specified in [7], 5.8.5.2 and is referred to as “**ContentProtection** Descriptor for
32 UUID Scheme” in the following.

1 **7.6.2.3. Protection System Specific Header Box cenc:pssh element in MPD**

2 A 'pssh' box is defined by each DRM system for use with their registered SystemID, and the
3 same box can be stored in the MPD within a **ContentProtection** Descriptor for UUID
4 scheme using an extension element in the "cenc:" namespace. Examples are provided in [6] and
5 in [30] sec. 11.2.

6 Carrying cenc:default_KID attribute and a cenc:pssh element in the MPD is useful to
7 allow key identification, license evaluation, and license retrieval before live availability of initial-
8 ization segments. This allows clients to spread license requests and avoid simultaneous requests
9 from all viewers at the instant that an Initialization Segments containing license acquisition infor-
10 mation in 'pssh' becomes available. With cenc:default_KID indicated in the mp4protection
11 **ContentProtection** Descriptor on each Adaptation Set, clients can determine if that key and
12 this presentation is not available to the viewer (e.g. without purchase or subscription), if the key is
13 already downloaded, or which licenses the client SHOULD download before the @availabil-
14 ityStartTime of the presentation based on the default_KID of each **AdaptationSet**
15 element selected.

16 **7.6.2.4 Use of W3C Clear Key with DASH**

17 When using Clear Key [69] with MPEG DASH, Clear Key management availability is signaled in
18 the MPD with a **ContentProtection** element that has the following format.

19 The Clear Key ContentProtection element attributes take the following values:

- 20 • The UUID e2719d58-a985-b3c9-781a-b030af78d30e is used for the
21 @schemeIdUri attribute.
22 • The @value attribute is equal to the string "ClearKey1.0"

23 The following element MAY be added under the **ContentProtection** element:

- 24 • **Laurl** element that contains the URL for a Clear Key license server allowing to receive
25 a Clear Key license in the format defined in [69] section 9.1.4. It has the attribute
26 @Lic_type that is a string describing the license type served by this license server. Pos-
27 sible value is "EME-1.0" when the license served by the Clear Key license server is in
28 the format defined in [69] section 9.1.4.

29 The name space for the **Laurl** element is <http://dashif.org/guidelines/clearKey>

30 An example of a Clear Key **ContentProtection** element is as follows

```
<xss:schema xmlns:ck="http://dashif.org/guidelines/clearKey">

<ContentProtection
    schemeIdUri="urn:uuid:xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx"
    value="ClearKey1.0">
    <ck:Laurl
        Lic_type="EME-1.0">
        https://clearKeyServer.foocompany.com</ck:Laurl>
</ContentProtection>
```

1 W3C also specifies the use of the SystemID="1077efec-c0b2-4d02-ace3-
2 3c1e52e2fb4b" in [70] section 4 to indicate that tracks are encrypted with Common Encryption
3 [33], and list the KID key identifiers of keys used to encrypt the track in a version 1 'pssh' box
4 with that SystemID. However, the presence of this Common PSSH box does not indicate
5 whether keys are managed by DRM systems or Clear Key management specified in this section.
6 Browsers are expected to provide decryption in the case where Clear Key management is used,
7 and a DRM system where a DRM key management system is used.

8 Therefore, clients SHALL NOT use the signalling of SystemID 1077efec-c0b2-4d02-
9 ace3-3c1e52e2fb4b as an indication that the Clear Key mechanism is to be used.

10 W3C specifies that in order to activate the Clear Key mechanism, the client must provide Clear
11 Key initialization data to the browser. The Clear Key initialization data consists of a listing of the
12 default KIDs required to decrypt the content.

13 The MPD SHALL NOT contain Clear Key initialization data. Instead, clients SHALL construct
14 Clear Key initialization data at runtime, based on the default KIDs signaled in the MPD using
15 ContentProtection elements with the urn:mpeg:dash:mp4protection:2011 scheme.

16 When requesting a Clear Key license to the license server, it is recommended to use a secure
17 connection as described in Section 7.2.

18 When used with a license type equal to "EME-1.0":

- 19 • The GET request for the license includes in the body the JSON license request format de-
20 fined in [69] section 9.1.3. The license request MAY also include additional authentica-
21 tion elements such as access token, device or user ID.
- 22 • The response from the license server includes in the body the Clear Key license in the
23 format defined in [69] section 9.1.4 if the device is entitled to receive the Content Keys.

24 Clear Key licenses SHALL NOT be used to manage a key and KID that is also used by a DRM
25 system. The use of an unprotected DRM key risks the security of DRM systems using that key,
26 and violates the terms of use of most DRM systems.

27 **7.7. Additional Content Protection Constraints**

28 The following describes additional constraints for presentations to be conformant with DASH-
29 264/AVC, for both MPD and ISO Media files.

30 **7.7.1. ISO BMFF Content Protection Constraints**

- 31 • There SHALL be identical values of default_KID in the Track Encryption Box
32 ('tenc') of all Representation referenced by one Adaptation Set. Different Adaptation
33 Sets may have equal or different values of default_KID.
- 34 • If a W3C Common 'pssh' box [69] is used with encrypted content, its list of KIDs SHALL
35 contain only the default_KID from the 'tenc' box.
- 36 • 'pssh' boxes SHOULD NOT be present in Initialization Segments, and cenc:pssh el-
37 ements in **ContentProtection** Descriptors used instead. If 'pssh' boxes are present
38 in Initialization Segments, each Initialization Segment within one Adaptation Set SHALL
39 contain an equivalent pssh box for each SystemID, i.e. license acquisition from any

1 Representation is sufficient to allow switching between Representations within the Adap-
2 tation Set without acquiring a new license.

3 Note: ‘pssh’ boxes in Initialization Segments may result in playback failure during
4 browser playback when a license request is initiated each time an Initialization Segment is
5 processed, such as the start of each protected Representation, each track selection, and each
6 bitrate switch. This content requires DASH clients that can parse the ‘pssh’ box contents
7 to determine the duplicate license requests and block them.

8 A cenc:pssh element is parsed at most once per Adaptation Set by a client’s MPD par-
9 ser, and the potential need for a new license request is identified by a new cenc:de-
10 fault_KID value. In this case, only the DASH client initiates license requests, and may
11 do so per Period, if cenc:default_KID is a new value and the DRM system does not
12 already have the key available for use.

13 7.7.2. MPD Content Protections Constraints

14 For an encrypted Adaptation Set, **ContentProtection** Descriptors shall always be present in
15 the **AdaptationSet** element and apply to all contained Representations.

16 A **ContentProtection** Descriptor for the mp4protection Scheme (@schemeIdUri equals
17 to "urn:mpeg:dash:mp4protection:2011" and @value='cenc') SHALL be pre-
18 sent in the **AdaptationSet** element if content represented by the contained Representations is
19 encrypted. This allows clients to recognize that content is encrypted with a common encryp-
20 tion scheme without the need to understand any DRM system specific element. This element
21 SHALL contain the attribute @cenc:default_KID. The ‘tenc’ box that specifies the en-
22 coded track encryption parameters shall be considered the source of truth for the default key ID
23 value since it contains the default_KID field. The MPD @cenc:default_KID attribute
24 SHALL match the ‘tenc’ default_KID value. This allows general purpose clients to identify
25 the default KID from the MPD using a standard location and format without the need to understand
26 any DRM specific element.

27 For each **ContentProtection** element with UUID Scheme, the @cenc:pssh element
28 SHOULD be present. The base64 encoded contents of the element SHALL be equivalent to a
29 ‘pssh’ box including its header. The information SHOULD be sufficient to allow for license
30 acquisition. The @value attribute SHOULD contain the DRM system name and version in a
31 human readable form.

32 *Note: A player such as DASH.js hosted by a browser may pass the contents of this pssh
33 element through the Encrypted Media Extension (EME) API to the DRM system Content
34 Decryption Module (CDM) with a SystemID equal to the element’s UUID value. This al-
35 lows clients to acquire a license using only information in the MPD, prior to downloading
36 Segments.*

37 Below is an example of the recommended format for a hypothetical acme DRM service:

```
38 <ContentProtection schemeIdUri="urn:uuid:d0ee2730-09b5-459f-8452-200e52b37567"  
39   value="Acme DRM 2.0">  
40   <!-- base64 encoded 'pssh' box with SystemID matching the containing Con-  
41   tentProtection Descriptor -->  
42   <cenc:pssh>  
43     YmFzZTY0IGVuY29kZWQgY29udGVudHMgb2YgkXB  
44     zc2iSIGJveCB3aXRoIHRoaXMgU3lzdGVtSUQ=
```

```
1      </cenc:pssh>
2      </ContentProtection>
```

5 **7.7.3. Other Content Protections Constraints**

6 In the case where the ‘pssh’ box element is present in the MPD and in the Initialization Segment,
7 the ‘pssh’ box element in the MPD SHALL take precedence, because the parameters in the MPD
8 will be processed first, are easier to update, and can be assumed to be up to date at the time the
9 MPD is fetched.

10 Recommended scheduling of License and key delivery:

- 11 • Request licenses on initial processing of an MPD if **ContentProtection** Descriptors
12 or Initialization Segments are available with license acquisition information. This is in-
13 tended to avoid a large number of synchronized license requests at **MPD@availabil-**
14 **ityStartTime**.
- 15 • Prefetch licenses for a new Period in advance of its presentation time to allow license
16 download and processing time, and prevent interruption of continuous decryption and play-
17 back. Advanced requests will also help prevent a large number of synchronized license
18 requests during a live presentation at **Period@start** time.

19 **7.7.4. Additional Constraints for Periodic Re-Authorization**

- 20 • Key rotation should not occur within individual segments, as their duration is typically short
21 enough to enable the intended use cases.
- 22 • Each Movie Fragment SHOULD contain one ‘pssh’ in each ‘moof’ box per Sys-
23 temID that contains sufficient information for the DRM system with matching Sys-
24 temID to obtain protected keys for this movie fragment, when combined with:
 - 25 ◦ Information from ‘pssh’ in ‘moov’ or **cenc:pssh** in MPD.
 - 26 ◦ KID associated with each sample from ‘seig’ sample group description box.
 - 27 ◦ Sample to group boxes that list all the samples that use a particular KID.
- 28 • The KID should be observable by the player by reading the clear key_ids in PSSH defini-
29 tion v1.
- 30 • If the key is does not need to be retrieved, a pssh update may not result in a license request.
- 31 • If **key_id** cannot be observed, the player may perform binary comparison of pssh seg-
32 ments to understand updates.

33 **7.7.5. Encryption of Different Representations**

34 Representations contained in one Adaptation Set SHALL be protected by the same license for each
35 protection system (“DRM”), and SHALL have the same value of ‘default_KID’ in their
36 ‘tenc’ boxes in their Initialization Segments. This is to enable seamless switching within Adap-
37 tation Sets, which is generally not possible if a new DRM license needs to be authorized, client
38 bound, generated, downloaded, and processed for each new Representation.

39 In the case of key rotation, if root licenses are used, the same requirement applies to the root li-
40 censes (one license per Adaptation Set for each DRM), and also means all Representations SHALL

have the same value of ‘`default_KID`’ in their ‘`tenc`’ boxes in their Initialization Segments. The use of root and leaf licenses is optional and DRM specific, but leaf licenses are typically delivered in band to allow real time license acquisition, and do not require repeating client authentication, authorization, and rebuilding the security context with each key change in order to enable continuous playback without interruption cause by key acquisition or license processing.

In cases where SD and HD and UHD Representations are contained in one presentation, different license rights may be required for each quality level and may be sold separately. If different licenses are required for different quality levels, then it is necessary to create separate Adaptation Sets for each quality level, each with a different license and value of ‘`default_KID`’.

Representations that are equivalent resolution and bitrate but encrypted with different keys may be included in different Adaptation Sets. Seamless switching between UHD, HD and SD Representations is difficult because these quality levels typically use different decryption licenses and keys, use different DRM output rules (prohibit analog interfaces, require resolution down-scaling, require HDCP encryption on output, etc.), and use different decoding parameters for e.g. subsampling, codec, profile, bit depth, aspect ratios and color spaces.

If any Representation is encrypted in an Adaptation Set, then all must be encrypted using the same `default_KID` in the Track Encryption Box (‘`tenc`’) to avoid realtime changes to the DRM licenses and security context. KID values may change over time (“key rotation”) as specified in Common Encryption and a particular DRM system.

For all Representations within an Adaptation Set with `@bitstreamSwitching="false"` (default), the following parameters shall apply.

- ‘`tenc`’ `default_KID` shall be equal for all Representations

7.7.6. Encryption of Multiple Periods

If a new license is needed and `cenc:default_KID` is to be changed, it SHALL be at the beginning of a Period. A different file is indicated by a different `default_KID` signaled in the ‘`tenc`’ box in the Initialization Segment.

A file associated with a single license may be continued over multiple Periods by being referenced by multiple Representations over multiple Periods (for instance, a program interspersed with ad Periods). A client can recognize the same `cenc:default_KID` value and avoid having to download the same license again; but the DRM system may require a complete erase and rebuild of the security context, including all key material, samples in process, etc., between Periods with different licenses or no license (between protected and clear Periods).

7.7.7. DRM System Identification

The DRM system is signaled in the MPD and ‘`pssh`’ boxes with a `SystemID`. A list of known DRMs can be found in the DASH identifier repository available here: <http://www.dashif.org/identifiers/protection>.

7.7.8. Protection of Media Presentations that Include SD, HD and UHD Adaptation Sets

Per DASH IF interop points, Representations with separate keys, licenses, and license policy are contained in different Adaptation Sets.

1 Adaptive bitrate switching can function automatically within an Adaptation Set without changing
2 keys, licenses, robustness and output rules, etc.

3 A player may download licenses for multiple Adaptation Sets in a Group, and seamlessly switch
4 between them if it is able. Seamless switching between Adaptation Sets is allowed, but not re-
5 quired. DASH may need to signal which Adaptation Sets are intended for seamless switching, i.e.
6 have identical source content, same picture aspect ratio, same exact rescaled pixel registration,
7 same sample description (e.g. ‘avc3’), same initialization behavior (@bitstreamSwitch-
8 ing = true/false), same Timescale and @timescale, and are mutually time-aligned.

9 The DASH-IF interop points are intended to make bitrate switching within an Adaptation Set sim-
10 ple and automatic, whether Representations are encrypted or not. Placement of Representations
11 in different Adaptation Sets informs players that those Representations need to be initialized with
12 different parameters, such as a different key and license. The full initialization process is repeated
13 per Period. Adaptation Sets with @bitstreamSwitching = “true” only need to be initial-
14 ized once per Period. Adaptation Sets with @bitstreamSwitching = “false” need to be
15 partially re-initialized on each Representation switch (to change the SPS parameter sets referenced
16 from NALs to those stored in the containing track’s ‘avcC’), but most initialized parameters
17 such as timescale, codec Profile/Level, display buffer size, colorspace, etc.; and licenses and the
18 DRM system ... do not need to be changed.

19 Fetching and resetting keys and licenses during adaptive switching requires processing Initializa-
20 tion Segments with different ‘tenc’ default_KID and possibly ‘pssh’ boxes. That may not be
21 seamless, especially in browser playback where the decoders are only aware of player switching
22 when an Initialization Segment flows through the MSE buffer and a needKey() event is raised
23 via EME.

24 Note that switching between Adaptation Sets with different Media Profiles could be restricted by
25 key and license policy, e.g. the user only purchased SD rights, the player only has analog output
26 and HD content requires a protected digital output, UHD content requires hardware protected
27 DRM, etc.

28 Implementations that seamlessly switch between Representations with different keys and policies
29 generally require a standardized presentation ID or content ID system that associates multiple keys
30 and licenses to that ID and presentation, then downloads only the keys/licenses authorized for that
31 user and device (e.g. SD or HD+SD). The player must then install those licenses and use player
32 logic to select only Representations in an Adaptation Set for which a license is installed and output
33 controls, display configuration, etc. allow playback (e.g. only Representations keyed for an in-
34 stalled SD license). Players and license servers without this pre-configuration protocol and adap-
35 tive switching logic will encounter key/license requests in the process of adaptive switching, and
36 may find output blocked by different license policies, user rights, etc.

37 **7.7.9. Client Interactions with DRM Systems**

38 The client interacts with one or more DRM systems during playback in order to control the de-
39 cryption of content. Some of the most important interactions are:

- 40 1) Determining the availability of media keys.
41 2) Requesting the DRM system to acquire media keys.

42 In both of these interactions, the client and DRM system use the default_KID as an abstract mech-
43 anism to communicate information regarding the capability to decrypt adaptation sets that use a

1 particular default_KID. A DRM system may also make use of other media keys in addition to the
2 one signalled by default_KID (e.g. in key derivation or sample variant schemes) but this SHALL
3 be transparent to the client, with only the default_KID being used in communications between the
4 client and DRM system.

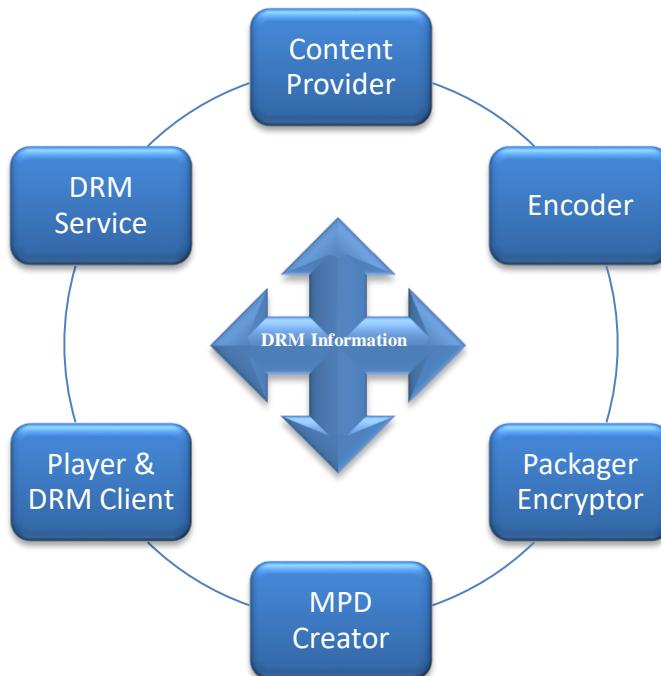
5 A client SHALL determine the required set of media keys based on the default KIDs signalled in
6 the manifest for the adaptation sets selected for playback.

7 Upon determining that one or more required media keys signalled by default KIDs are not available
8 the client SHOULD interact with the DRM system and request the missing media keys. The client
9 MAY also request media keys that are known to be usable. Clients SHALL explicitly request all
10 required media keys signaled by default KIDs and SHALL NOT assume that requesting one key
11 from this set will implicitly make others available.

12 The client and/or DRM system MAY batch multiple key requests (and the respective responses)
13 into a single transaction (for example, to reduce the chattiness of license acquisition traffic).

14

15 **7.8. Workflow Overview**



16
17 **Figure 24 Logical Roles that Exchange DRM Information and Media**

18 Figure 24 shows logical entities that may send or receive DRM information such as media keys,
19 asset identifiers, licenses, and license acquisition information. A physical entity may combine
20 multiple logical roles, and the point of origin for information, such as media keys and asset iden-
21 tifiers, can differ; so various information flows are possible. This is an informative example of
22 how the roles are distributed to facilitate the description of workflow and use cases. Alternative
23 roles and functions can be applied to create conformant content.

24 **Description of Logical Roles:**

1 **Content Provider** – A publisher who provides the rights and rules for delivering protected media,
2 also possibly source media (mezzanine format, for transcoding), asset identifiers, key identifiers
3 (KID), key values, encoding instructions, and content description metadata.

4 **Encoder** – A service provider who encodes Adaptation Sets in a specified media format, number
5 of streams, range of bitrates and resolutions, seamless switching constraints, etc., possibly deter-
6 mined by the publisher. An asset identifier needs to be assigned to each encoded track in order to
7 associate a key identifier, a Representation element in an MPD, a possible ‘pssh’ box in the file
8 header, and a DRM license separately downloaded.

9 **Packager / Encryptor** – A service provider who encrypts and packages media files, inserting
10 default_KID in the file header ‘tenc’ box, initialization vectors and subsample byte ranges
11 in track fragments indexed by ‘saio’ and ‘saiz’ boxes, and possibly packages ‘pssh’ boxes
12 containing license acquisition information (from the DRM Provider) in the file header. Tracks that
13 are partially encrypted or encrypted with multiple keys require sample to group boxes and sample
14 group description boxes in each track fragment to associate different KIDs to groups of samples.
15 The Packager could originate values for KIDs, media keys, encryption layout, etc., then send that
16 information to other entities that need it, including the DRM Provider and Streamer, and probably
17 the Content Provider. However, the Packager could receive that information from a different point
18 of origin, such as the Content Provider or DRM Provider.

19 **MPD Creator** – The MPD Creator is assumed to create one or more types of DASH MPD, and
20 provide indexing of Segments and/or ‘sidx’ indexes for download so that players can byte range
21 index Subsegments. The MPD must include descriptors for Common Encryption and DRM key
22 management systems, and SHOULD include identification of the default_KID for each Ad-
23 aptationSet element, and sufficient information in UUID ContentProtection De-
24 scriptor elements to acquire a DRM license. The default_KID is available from the Packager
25 and any other role that created it, and the DRM specific information is available from the DRM
26 Provider.

27 **Player / DRM Client** – Gets information from different sources: MPD, Media files and DRM
28 License.

29 **DRM Service** – The DRM Provider creates licenses containing a protected media key that can
30 only be decrypted by a trusted client.

31 The DRM Provider needs to know the default_KID and DRM SystemID and possibly other
32 information like asset ID and player domain ID in order to create and download one or more li-
33 censes required for a Presentation on a particular device. Each DRM system has different license
34 acquisition information, a slightly different license acquisition protocol, and a different license
35 format with different playback rules, output rules, revocation and renewal system, etc. The DRM
36 Provider typically must supply the Streamer and the Packager license acquisition information for
37 each UUID ContentProtection Descriptor element or ‘pssh’ box, respectively.

38 The DRM Service may also provide logic to manage key rotation, DRM domain management,
39 revocation and renewal and other content protection related features.

40

41 Figure 25 shows a simple workflow with pssh information in the Initialization Segment for in-
42 formational purpose.

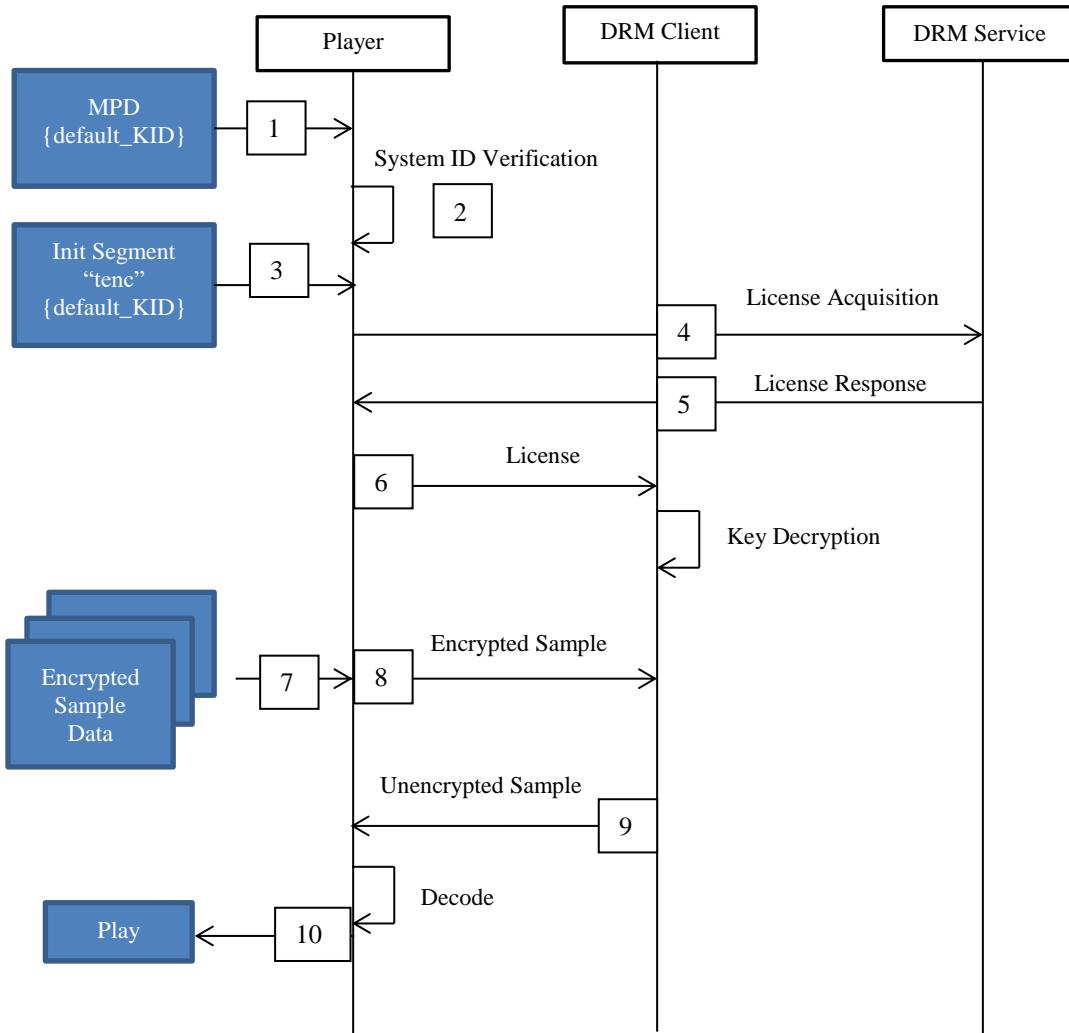


Figure 25 Example of Information flow for DRM license retrieval

[1] A MPD may include **ContentProtection** Descriptors to indicate that the ‘cenc’ scheme is used to encrypt the referenced media, and to provide license acquisition information for one (or more) DRM system(s) with the indicated SystemID.

[2] The Player verifies if a specified DRM is supported using System ID value(s) from the MPD. With unique KIDs, a license request using the cenc:default_KID attribute value is sufficient to identify a DRM license containing that key that will enable playback of the Components, Representations, Adaptation Sets, or Periods that the ContentProtection Descriptor element and default_KID describe.

[3] The TrackEncryptionBox (‘tenc’) contains default values for the IsEncrypted, IV_size, and KID for the entire track. These values are used as the encryption parameters for the samples in this track unless over-ridden by the sample group description associated with a group of samples. The license acquisition information could be also present in ‘pssh’ boxes in the initialization segment.

-
- 1 [4] Decryption Key acquisition can be performed either by the Player or the DRM Client.
2 [5] DRM License / Decryption Key response includes the required material for enabling access.
3 [6] DRM licenses/rights need not be stored in order to look up a key using KID values stored in
4 the file and decrypt media samples using the encryption parameters stored in each track.
5 [7] The Player requesting encrypted sample data.
6 [8] The Player provides encrypted sample data to the DRM Client for decryption using the decryp-
7 tion key. How the DRM system locates the identified decryption key is left to the DRM.
8 [9] The Player received unencrypted sample data from the DRM Client.

9 **8. DASH-IF Interoperability Points**

10 **8.1. Introduction**

11 This version of the document defines Interoperability Points in this section. Earlier versions of this
12 document, especially version 2 [2] defines legacy IOPs.

13 **8.2. DASH-AVC/264 main**

14 **8.2.1. Introduction**

15 The scope of the DASH-AVC/264 main interoperability point is basic support of high-quality
16 video distribution over the top based on H.264/AVC up to 1080p. Both, live and on-demand ser-
17 vices are supported.

18 The compliance to DASH-AVC/264 main may be signaled by a @profiles attribute with the
19 value "http://dashif.org/guidelines/dash264main"

20 **8.2.2. Definition**

21 A DASH client conforms to the IOP by supporting at least the following features:

- 22 • All DASH-related features as defined in section 3 of this document.
- 23 • The requirements and guidelines in section 4.9.2 for simple live operation.
- 24 • The requirements and guidelines in section 5.6.1 for server-based ad insertion.
- 25 • H.264/MPEG AVC Progressive High Profile at level 4.0 as defined in section 6.2 together
26 with all AVC-related requirements and recommendation in section 6.2.
- 27 • MPEG-4 HE-AAC v2 level 2 profile audio codec as defined in section 6.3. Dynamic Range
28 Control is not expected to be supported.
- 29 • subtitle and closed captioning support
 - 30 ○ using SMPTE-TT as defined in section 6.4.2
 - 31 ▪ For On-Demand single file download is sufficient.
 - 32 ▪ For live services and/or if key rotation is to be supported, the encapsulation
33 into ISO BMFF is necessary.
 - 34 ○ Using CEA-608/708 as defined in section 6.4.3.

-
- 1 • content protection based on common encryption and key rotation as defined in section 7.
2 And specifically, the client supports MPD-based parsing and movie box based parsing of
3 DRM related parameters for common encryption.

4 Content shall only be authored claiming conformance to this IOP if such a client can properly play
5 the content. In addition, the content shall follow the mandatory aspects and should take into ac-
6 count the recommendations and guidelines for content authoring documented in section 3 (DASH
7 features), section 4.9.2 (simple live operation), section 5.6.1 (server-based ad insertion), AVC-
8 related issues in section 6.2, section 6.3 (audio), section 6.4.2 (SMPTE-TT), section 6.4.3 (CEA-
9 608/708), and section 7 (Content Protection).

10 If content is offered claiming conformance to this IOP, the content author is encouraged to use the
11 HTTP-URL construction as defined in [6], section 5.1.4.

12 **8.3. DASH-AVC/264 high**

13 **8.3.1. Introduction**

14 The scope of the DASH-AVC/264 interoperability point is support of high-quality video distribu-
15 tion over the top based on H.264/AVC up to 1080p. Both, live and on-demand services are sup-
16 ported as well as features for main live and advanced ad insertion.

17 The compliance to DASH-AVC/264 may be signaled by a @profiles attribute with the value
18 "http://dashif.org/guidelines/dash264high"

19 **8.3.2. Definition**

20 A client that attempts to consume content generated conforming to this IOP shall support the fol-
21 lowing features:

- 22 • All features required for DASH-264/AVC main as defined in section 8.2.
23 • The client requirements and recommendations for the main live operation as defined in
24 section 4.9.3.

25 Content shall only be authored claiming conformance to this IOP if such a client can properly play
26 the content. In addition, the content shall follow the mandatory aspects and should take into ac-
27 count the recommendations and guidelines for content authoring documented in section 8.2
28 (DASH-264/AVC main), section 4.9.3 (main live operation), and section 5.6.2 (app-based ad in-
29 sertion).

30 If content is offered claiming conformance to this IOP, the content author is encouraged to use the
31 HTTP-URL construction as defined in [6], section 5.1.4.

32 **8.4. DASH-IF IOP simple**

33 **8.4.1. Introduction**

34 The scope of the DASH-IF IOP simple interoperability point is the basic support of efficient high-
35 quality video distribution over the top with HD video up to 1080p including support for HEVC 8
36 bit.

37 The compliance to *DASH-IF IOP simple* may be signaled by a @profiles attribute with the
38 value "http://dashif.org/guidelines/dash-if-simple"

8.4.2. Definition

A DASH client conforms to the IOP by supporting at least the following features:

- All DASH-related features as defined in section 3 of this document.
- The requirements and guidelines in section 4.9.2 for simple live operation.
- The requirements and guidelines in section 5.6.1 for server-based ad insertion.
- H.264/MPEG AVC Progressive High Profile at level 4.0 as defined in section 6.2 together with all AVC-related requirements and recommendation in section 6.2.
- H.265/MPEG-H HEVC Main Profile Main Tier at level 4.1 as defined in section 6.2 together with all HEVC-related requirements and recommendation in section 6.2.
- MPEG-4 HE-AAC v2 level 2 profile audio codec as defined in section 6.3. Dynamic Range Control is not expected to be supported.
- subtitle and closed captioning support
 - using SMPTE-TT as defined in section 6.4.2
 - For On-Demand single file download is sufficient.
 - For live services and/or if key rotation is to be supported, the encapsulation into ISO BMFF is necessary.
 - Using CEA-608/708 as defined in section 6.4.3.
- content protection based on common encryption and key rotation as defined in section 7. And specifically, the client supports MPD-based parsing and movie box based parsing of DRM related parameters for common encryption.

Content shall only be authored claiming conformance to this IOP if such a client can properly play the content. In addition, the content shall follow the mandatory aspects and should take into account the recommendations and guidelines for content authoring documented in section 3 (DASH features), section 4.9.2 (simple live operation), section 5.6.1 (server-based ad insertion), section 6.2 (video), section 6.3 (audio), section 6.4.2 (SMPTE-TT), section 6.4.3 (CEA-608/708), and section 7 (Content Protection).

If content is offered claiming conformance to this IOP, the content author is encouraged to use the HTTP-URL construction as defined in [6], section 5.1.4.

8.5. DASH-IF IOP Main

8.5.1. Introduction

For the support of broad set of use cases the DASH-IF IOP Main Interoperability Point is defined. In addition the features of DASH-264/AVC main as defined in section 8.2 the interoperability point requires DASH clients for real-time segment parsing and 10-bit HEVC.

The compliance to *DASH-IF IOP main* may be signalled by a @profile attribute with the value "http://dashif.org/guidelines/dash-if-main"

8.5.2. Definition

A client that attempts to consume content generated conforming to this IOP shall support the following features:

- All features required for DASH-264/AVC high as defined in section 8.3.

-
- 1 • H.265/MPEG-H HEVC Main Profile Main Tier at level 4.1 as defined in section 6.2 to-
2 gether with all HEVC-related requirements and recommendation in section 6.2.
3 • H.265/MPEG-H HEVC Main 10 Profile Main Tier at level 4.1 as defined in section 6.2
4 together with all HEVC-related requirements and recommendation in section 6.2.

5 Content shall only be authored claiming conformance to this IOP if such a client can properly play
6 the content. In addition, the content shall follow the mandatory aspects and should take into ac-
7 count the recommendations and guidelines for content authoring documented in section 8.3 and
8 HEVC-related issues in section 6.2.

9 If the content is authored such that it also conforms to DASH-264/AVC high as defined in section
10 8.3, then the profile identifier for DASH-264/AVC high shall be added as well. If the profile iden-
11 tifier is missing, the content may be considered as HEVC only content.

12 If content is offered claiming conformance to this IOP, the content author is encouraged to use the
13 HTTP-URL construction as defined in [6], section 5.1.4.

14 **9. Multi-Channel Audio Extensions**

15 **9.1. Scope**

16 The Scope of the Multichannel Audio Extension is the support of audio with additional channels
17 and codecs beyond the basic audio support as specified in the DASH-AVC/264 base, which is
18 limited to Stereo HE-AAC. Multichannel audio is widely supported in all distribution channels
19 today, including broadcast, optical disc, and digital delivery of audio, including wide support in
20 adaptive streaming delivery.

21 It is expected that clients may choose which formats (codecs) they support.

22 **9.2. Technologies**

23 **9.2.1. Dolby Multichannel Technologies**

24 **9.2.1.1. Overview**

25 The considered technologies from Dolby for advanced audio support are:

- 26 • Enhanced AC-3 (Dolby Digital Plus) [35]
- 27 • Dolby TrueHD [36]
- 28 • AC-4 [63]

29 **9.2.1.2. DASH-specific issues**

30 In the context of DASH, the following applies:

- 31 • The signaling of the different audio codecs for the codecs parameters is documented in
32 [35], [36] and [63] which also provides information on ISO BMFF encapsulation.
- 33 • For E-AC-3 and AC-4 the Audio Channel Configuration shall use the "
34 tag:dolby.com,2014:dash:audio_channel_configuration:2011" as
35 defined at <http://dashif.org/identifiers/audio-source-data/>.

36 **Table 25 Dolby Technologies: Codec Parameters and ISO BMFF encapsulation**

Codec	Codec Parameter	ISO BMFF Encapsulation	SAP type
-------	-----------------	------------------------	----------

Enhanced AC-3 [35]	ec-3	ETSI TS 102 366 Annex F [35]	1
Dolby TrueHD	mlpa	Dolby [36]	1
AC-4	ac-4	ETSI TS 103 190-1 Annex E [63]	1

1 **9.2.2. DTS-HD**

2 **9.2.2.1. Overview**

3 DTS-HD [37] comprises a number of profiles optimized for specific applications. More information about DTS-HD and the DTS-HD profiles can be found at www.dts.com.

5 **9.2.2.2. DASH-specific issues**

6 For all DTS formats SAP is always 1.

7 The signaling of the various DTS-HD profiles is documented in DTS 9302J81100 [34]. DTS 9302J81100 [34] also provides information on ISO BMFF encapsulation.

9 Additional information on constraints for seamless switching and signaling DTS audio tracks in the MPD is described in DTS specification 9302K62400 [39].

11 **Table 26: DTS Codec Parameters and ISO BMFF encapsulation**

Codec	Codec Parameter	Pa-	ISO BMFF Encapsu-	SAP type
DTS Digital Surround	dtsc		DTs 9302J81100 [34]	1
DTS-HD High Resolution and DTS-HD Master Audio	dtsh			
DTS Express	dtse			
DTS-HD Lossless (no core)	dtsl			

12 **9.2.3. MPEG Surround**

13 **9.2.3.1. Overview**

15 MPEG Surround, as defined in ISO/IEC 23003-1:2007 [38], is a scheme for coding multichannel signals based on a down-mixed signal of the original multichannel signal, and associated spatial parameters. The down-mix shall be coded with MPEG-4 High Efficiency AAC v2 according to section 5.3.3.

19 MPEG Surround shall comply with level 4 of the Baseline MPEG Surround profile.

20 **9.2.3.2. DASH-specific issues**

21 In the context of DASH, the following applies for audio codecs

- The signaling of the different audio codecs for the codecs parameters is according to RFC6381 [10] is documented in Table 27. Table 27 also provides information on ISO BMFF encapsulation.

- 1 • The content is expected to be prepared according to the MPEG-DASH Implementation
 2 Guidelines [6] to make sure each (sub-)segment starts with a SAP of type 1.

3 **Table 27 Codecs parameter according to RFC6381 [10] and ISO BMFF encapsulation for MPEG Surround
 4 codec**

Codec	Codec Parameter	ISO BMFF Encapsulation	SAP type
MPEG Surround [38]	mp4a.40.30	ISO/IEC 14496-14 [8]	1

5 **Note:** Since MPEG Surround is based on a down-mix coded with AAC-LC and HE-AAC, for the
 6 above mentioned “Codec Parameters” the following is implied:

7 $\text{mp4a.40.30} = \text{AOT 2} + \text{AOT 5} + \text{AOT 30}$

8 **9.2.4. MPEG-4 High Efficiency AAC Profile v2, level 6**

9 **9.2.4.1. Overview**

10 Support for multichannel content is available in the HE-AACv2 Profile, starting with level 4 for
 11 5.1 and level 6 for 7.1. All MPEG-4 HE-AAC multichannel profiles are fully compatible with the
 12 DASH-AVC/264 baseline interoperability point for stereo audio, i.e. all multichannel decoders
 13 can decode DASH-IF IOPS stereo content.

14 **9.2.4.2. DASH-specific issues**

15 In the context of DASH, the following applies for the High Efficiency AAC v2 Profile

- 16 • The content shall be prepared according to the MPEG-DASH Implementation Guidelines
 17 [6] to make sure each (sub-)segment starts with a SAP of type 1.
- 18 • Signaling of profile levels is not supported in RFC 6381 but the channel configuration shall
 19 be signaled by means of the **ChannelConfiguration** element in the MPD.
- 20 • The signaling of MPEG-4 High Efficiency AAC v2 for the codecs parameters is according
 21 to RFC6381 [10] and is documented in Table 28. Table 28 also provides information on
 22 the ISO BMFF encapsulation.
- 23 • For all HE-AAC bitstreams, explicit backward-compatible signaling of SBR shall be used.
- 24 • The content should be prepared incorporating loudness and dynamic range information into
 25 the bitstream also considering DRC Presentation Mode in ISO/IEC 14496-3 [11], Amd. 4.
- 26 • Decoders shall support decoding of loudness and dynamic range related information, i.e.
 27 dynamic_range_info() and MPEG4_ancillary_data() in the bitstream.

28 **Table 28 Codecs parameter according to RFC6381 [10] and ISO BMFF encapsulation**

Codec	Codec Parameter	ISO BMFF Encapsulation	SAP type
MPEG-4 AAC Profile [11]	mp4a.40.2	ISO/IEC 14496-14 [12]	1

MPEG- mp4a.40.5 ISO/IEC 1
4 HE- 14496-
AAC 14 [12]
Profile
[11]

MPEG- mp4a.40.29 ISO/IEC 1
4 HE- 14496-
AAC 14 [12]
v2 Pro-
file [11]

1 Note: Since both, HE-AAC and HE-AACv2 are based on AAC-LC, for the above mentioned “Co-
2 dec Parameters” the following is implied:

3 $\text{mp4a.40.5} = \text{AOT 2} + \text{AOT 5}$

4 **9.2.5. MPEG-H 3D Audio**

5 **9.2.5.1. Overview**

6 MPEG-H 3D Audio is defined in ISO/IEC 23008-3 [64] and is a Next Generation Audio (NGA)
7 codec. MPEG-H 3D Audio encoded content shall comply with Level 1, 2 or 3 of the MPEG-H
8 Low Complexity (LC) Profile, as defined in ISO/IEC 23008-3, clause 4.8 [64]. The sections to
9 follow clarify DASH specific issues for MPEG-H 3D Audio [64].

10 **9.2.5.2. DASH-specific issues**

11 Storage of raw MPEG-H audio frames in the ISO BMFF shall be according to ISO/IEC 23008-3
12 [64], clause 20.5 with the following constraints:

- 13 • One audio ISO BMFF sample shall consist of a single `mpegh3daFrame()` structure, as
14 defined in ISO/IEC 23008-3 [64], clause 20.5.
- 15 • The parameters carried in the `MHADecoderConfigurationRecord()` shall be con-
16 sistent with the configuration of the audio bitstream. In particular, the `mpegh3daPro-`
17 `fileLevelIndication` shall be set to “0x0B”, “0x0C”, or “0x0D” for MPEG-H Audio LC
18 Profile Level 1, Level 2, or Level 3, respectively.
- 19 • The `referenceChannelLayout` field carried in the `MHADecoderConfigurationRe-`
20 `cord()` shall be equivalent to what is signaled by **ChannelConfiguration** according
21 to ISO/IEC 23001-8 [49].
- 22 • The content is expected to be prepared according to the MPEG-DASH Implementation
23 Guidelines [6] to make sure each (sub-)segment starts with a SAP of type 1 (i.e. a sync
24 sample). MPEG-H Audio sync samples contain Immediate Playout Frames (IPFs), as spec-
25 ified in ISO/IEC 23008-3, clause 20.2 [64]. For such frames, the raw MPEG-H audio
26 frames shall contain the `AudioPreRoll()` syntax element, as defined in sub-clause
27 5.5.6 of ISO/IEC 23008-3 [64], and shall follow the requirements for stream access points
28 as defined in clause 5.7 of ISO/IEC 23008-3 [64]. The `AudioPreRoll()` syntax element

1 carried in the IPFs shall contain a valid configuration structure (`AudioPreRoll.Config()`) and should contain one pre-roll frame (`AudioPreRoll.numPreRollFrames = 1`).

4 **Note:** the `mpegh3daConfig()` structure is expected to be different for each representation in
5 an adaptation set.

7 **Table 29 Codecs parameter and ISO BMFF encapsulation**

Code	Codec Parameter	ISO BMFF Encapsulation	SAP Type
MPEG-H 3D Audio LC Profile Level 1	<code>mhm1.0x0B</code>	ISO/IEC 23008-3	1
MPEG-H 3D Audio LC Profile Level 2	<code>mhm1.0x0C</code>	ISO/IEC 23008-3	1
MPEG-H 3D Audio LC Profile Level 3	<code>mhm1.0x0D</code>	ISO/IEC 23008-3	1

8 **9.2.6. MPEG-D Unified Speech and Audio Coding**

9 **9.2.6.1. Overview**

10 MPEG-D Unified Speech and Audio Coding (USAC) has been designed to provide consistently
11 high audio quality with a variety of content that comprises a mixture of audio and speech signals.
12 Using such a codec in a DASH streaming environment enables adaptive switching capability from
13 12 kbps stereo up to transparency.

14 The Common Media Application Format (CMAF) [91] defines a media profile for MPEG-D
15 USAC that is suitable for streaming applications and can be, therefore, referenced here.

16 **9.2.6.2. DASH-specific issues**

17 In the context of DASH-IF IOPs, the following applies to the xHE-AAC profile:

- 18 • Content representations encoded with MPEG-D USAC shall comply with the Extended High
19 Efficiency AAC (xHE-AAC) CMAF media profile '`cxha`', as defined in [91], providing sup-
20 port up to 5.1 multichannel coding.
- 21 • All representations of an adaptation set shall conform to the CMAF switching set constraints.
- 22 • The codec signaling is according to RFC6381 [10] and documented in Table 30.
- 23 • The profiles mime sub- parameter of the `@mimetype` attribute should include '`cxha`'.
- 24 • If the `ChannelConfiguration` parameter is present in the Movie header, then the identi-
25 cal channel configuration shall be signaled by means of the **AudioChannelConfigura-**
26 **tion** element in the MPD, according to the values specified in Table K.2 [91].
- 27 • The CMAF xHE-AAC media profile [91] requires each CMAF Fragment to start with an SAP
28 of type 1.

1 **Table 30** Codecs parameters according to RFC 6381 [10] and ISO BMFF encapsulation for MPEG-D USAC

Codec	Codec Parameter	ISO BMFF Encapsulation	SAP type
MPEG-D USAC	mp4a.40.42	ISO/IEC 23000-19 [91]	1

2

3 **9.3. Client Implementation Guidelines**

4 Independent of the codec, a client that supports one or more codecs of multichannel sound play-
5 back should exhibit the following characteristics:

- 6 • Playback multichannel sound correctly given the client operating environment. As an ex-
7 ample, if the audio track delivers 5.1 multichannel sound, the client might perform one or
8 more of the following: decode the multichannel signal on the device and output either 6ch
9 PCM over HDMI, or pass that multichannel audio with no changes to external AVR, or if
10 the device is rendering to stereo outputs such as headphones, either correctly downmix that
11 multi-channel audio to 2-channel sound, or select an alternate stereo adaptation set, or other
12 appropriate choices.
- 13 • Adaptively and seamless switch between different bitrates as specified in the adaptation
14 sets according to the playback clients logic. Seamless switching is defined as no perceptible
15 interruption in the audio, and no loss of A/V sync. There is no expectation that a client
16 can seamlessly switch between formats.

17 **9.4. Extensions**

18 **9.4.1. General**

19 **9.4.1.1. Definitions**

20 A *multichannel audio client* at least supports the following features:

- 21 • All DASH-related features as defined in section 3 of this document.
- 22 • content protection based on common encryption and key rotation as defined in section 7.
23 And specifically, the client supports MPD-based parsing and movie box based parsing of
24 DRM related parameters for common encryption.
- 25 • The client implementation guidelines in section 9.3.

26 **9.4.1.2. Recommendations**

27 If content is offered claiming conformance to any extension in this section, the content author is
28 encouraged to use the HTTP-URL construction as defined in [6], section 5.1.4.

29 **9.4.2. Dolby Extensions**

30 **9.4.2.1. Introduction**

31 For the support of Dolby advanced audio support, three additional extensions are defined.

1 Conformance to *DASH-IF multichannel audio extension with Enhanced AC-3* (Dolby Digital Plus)
2 [35] may be signaled by an @profile attribute with the value
3 "http://dashif.org/guidelines/dashif#ec-3".

4 Conformance to *DASH-IF multichannel extension with Dolby TrueHD* may be signaled by an
5 @profile attribute with the value "http://dashif.org/guide-
6 lines/dashif#mlpa".

7 Conformance to *DASH-IF multichannel extension with AC-4* may be signaled by an @profile
8 attribute with the value "http://dashif.org/guidelines/dashif#ac-4".

9 **9.4.2.2. Supporters**

10 These extensions are supported by the following DASH IF members: Dolby, DTS, Fraunhofer,
11 BuyDRM, Sony.

12 **9.4.2.3. Definition**

13 Content may be authored claiming conformance to *DASH-IF multichannel audio extension with*
14 *Enhanced AC-3*

- 15 • if the content is multichannel audio content as defined in section 9.4.1, and
- 16 • if a client can properly play the content by supporting at least the following features
 - 17 • all multichannel audio client features as defined in section 9.4.1
 - 18 • Enhanced AC-3 (Dolby Digital Plus) [35] and the DASH-specific features defined in
 - 19 section 9.2.1.2

20 Content may be authored claiming conformance to *DASH-IF multichannel extension with Dolby*
21 *TrueHD*

- 22 • if the content is multichannel audio content as defined in section 9.4.1, and
- 23 • if a client can be properly play the content by supporting at least the following features
 - 24 • all multichannel audio client features as defined in section 9.4.1
 - 25 • Dolby TrueHD and the DASH-specific features defined in section 9.2.1.2

26 Content may be authored claiming conformance to *DASH-IF multichannel extension with AC-*

- 27 • if the content is multichannel audio content as defined in section 9.4.1, and
- 28 • if a client can be properly play the content by supporting at least the following features
 - 29 • all multichannel audio client features as defined in section 9.4.1
 - 30 • AC-4 and the DASH-specific features defined in section 9.2.1.2

31

32 **9.4.3. DTS-HD Interoperability Points**

33 **9.4.3.1. Introduction**

34 For the support of DTS advanced audio support, four additional extensions are defined.

35 Conformance to *DASH-IF multichannel audio extension with DTS Digital Surround* may be sig-
36 naled by a @profile attribute with value "http://dashif.org/guide-
37 lines/dashif#dtsc".

1 Conformance to *DASH-IF multichannel audio extension with DTS-HD High Resolution and DTS-*
2 *HD Master Audio* may be signaled by a @profile attribute with value
3 "http://dashif.org/guidelines/dashif#dtsh"

4 Conformance to *DASH-IF multichannel audio extension with DTS Express* may be signaled by a
5 @profile attribute with value "http://dashif.org/guidelines/dashif#dtse"

6 Conformance to *DASH-IF multichannel extension with DTS-HD Lossless (no core)* may be sig-
7 naled by a @profile attribute with value "http://dashif.org/guide-
8 lines/dashif#dtsl"

9 **9.4.3.2. Supporters**

10 These extensions are supported by the following DASH IF members: Dolby, DTS, Fraunhofer,
11 BuyDRM, Sony.

12 **9.4.3.3. Definition**

13 Content may be authored claiming conformance to *DASH-IF multichannel audio extension with*
14 *DTS Digital Surround*

- 15 • if the content is multichannel audio content as defined in section 9.4.1, and
- 16 • if a client can be properly play the content by supporting at least the following features
 - 17 • all multichannel audio client features as defined in section 9.4.1
 - 18 • DTS and the DASH-specific features defined in section 9.2.2.2

19 Content may be authored claiming conformance to *DASH-IF multichannel audio extension with*
20 *DTS-HD High Resolution and DTS-HD Master Audio*

- 21 • if the content is multichannel audio content as defined in section 9.4.1, and
- 22 • if a client can be properly play the content by supporting at least the following features
 - 23 • all multichannel audio client features as defined in section 9.4.1
 - 24 • DTS-HD High Resolution and DTS-HD Master Audio and the DASH-specific features
 - 25 defined in section 9.2.2.2

26 Content may be authored claiming conformance to *DASH-IF multichannel audio extension with*
27 *DTS Express*

- 28 • if the content is multichannel audio content as defined in section 9.4.1, and
- 29 • if a client can be properly play the content by supporting at least the following features
 - 30 • all multichannel audio client features as defined in section 9.4.1
 - 31 • DTS-HD Express and the DASH-specific features defined in section 9.2.2.2

32 Content may be authored claiming conformance to *DASH-IF multichannel extension with DTS-*
33 *HD Lossless (no core)*

- 34 • if the content is multichannel audio content as defined in section 9.4.1, and
- 35 • if a client can be properly play the content by supporting at least the following features
 - 36 • all multichannel audio client features as defined in section 9.4.1
 - 37 • DTS-HD Lossless (no core) and the DASH-specific features defined in section 9.2.2.2

9.4.4. MPEG Surround Interoperability Points

9.4.4.1. Introduction

For the support of MPEG Surround advanced audio support the following extension is defined.

Conformance to *DASH-IF multichannel audio extension with MPEG Surround* according to ISO/IEC 23003-1:2007 [38] may be signaled by an @profile attribute with the value "http://dashif.org/guidelines/dashif#mps".

9.4.4.2. Supporters

These extensions are supported by the following DASH IF members: Dolby, DTS, Fraunhofer, BuyDRM, Sony.

9.4.4.3. Definition

Content may be authored claiming conformance to *DASH-IF multichannel audio extension with MPEG Surround*

- if the content is multichannel audio content as defined in section 9.4.1, and
- if a client can be properly play the content by supporting at least the following features
 - all multichannel audio client features as defined in section 9.4.1
 - ISO/IEC 23003-1:2007 and the DASH-specific features defined in section 9.2.3.2

9.4.5. MPEG HE-AAC Multichannel Interoperability Points

9.4.5.1. Introduction

Conformance to *DASH-IF multichannel audio extension with HE-AACv2 level 4* [11] may be signaled by an @profile attribute with the value "http://dashif.org/guidelines/dashif#heaac-mc51".

Conformance to *DASH-IF multichannel audio extension with HE-AACv2 level 6* [11] may be signaled by an @profile attribute with the value "http://dashif.org/guidelines/dashif#heaac-mc71".

9.4.5.2. Supporters

These extensions are supported by the following DASH IF members: Dolby, DTS, Fraunhofer, BuyDRM, Sony.

9.4.5.3. Definition

Content may be authored claiming conformance to *DASH-IF multichannel audio extension with HE-AACv2 level 4*

- if the content is multichannel audio content as defined in section 9.4.1, and
- if a client can be properly play the content by supporting at least the following features
 - all multichannel audio client features as defined in section 9.4.1
 - HE-AACv2 level 4 [11] and the DASH-specific features defined in section 9.2.4.2

Content may be authored claiming conformance to *DASH-IF multichannel audio extension with HE-AACv2 level 6*

- if the content is multichannel audio content as defined in section 9.4.1, and
- if a client can be properly play the content by supporting at least the following features

-
- 1 • all multichannel audio client features as defined in section 9.4.1
2 • HE-AACv2 level 6 [11] and the DASH-specific features defined in section 9.2.4.2

3 **9.4.6. MPEG-H 3D Audio Interoperability Points**

4 **9.4.6.1. Introduction**

5 Compliance to DASH-IF multichannel audio extension with MPEG-H 3D Audio [64] may be sig-
6 naled by a @profile attribute with the value <http://dashif.org/guidelines/dashif#mpeg->
7 [h-3da](#).

8 **9.4.6.3. Definition**

9 Content may be authored claiming conformance to DASH-IF multichannel audio extension with
10 MPEG-H 3D Audio

- 11 • if the content is multichannel audio content as defined in section 9.4.1, and
- 12 • if a client can properly play the content by supporting at least the following features:
 - 13 ○ all multichannel audio client features as defined in section 9.4.1,
 - 14 ○ MHA and the DASH-specific features defined in section 9.2.5.

15 **9.4.7. MPEG-D USAC Interoperability Points**

16 **9.4.7.1. Introduction**

17 Conformance to DASH-IF audio extension with USAC according to ISO/IEC 23000-19:2018 [91]
18 may be signaled by an @profile attribute with the value "http://dashif.org/guide-
19 lines/dashif#cxha".

20 **9.4.7.2. Supporters**

21 Fraunhofer IIS, Dolby, castLabs, Bitmovin, Sony,

22 **9.4.7.3. Definition**

23 Content may be authored claiming conformance to DASH-IF audio extension with USAC

- 24 • if the content is multichannel audio content as defined in clause 9.4.1, and
- 25 • if the content conforms to what has been defined in clause 9.2.6.2
- 26 • if a client can properly play the content by supporting at least the following features
 - 27 • all multichannel audio client features as defined in clause 9.4.1
 - 28 • DASH-specific features defined in clause 9.2.6.2.

29 **10. DASH-IF UHD Extensions**

30 **10.1. Introduction**

31 This version of the document defines UHD Extensions in this section.

1 **10.2. DASH-IF UHD HEVC 4k**

2 **10.2.1. Introduction**

3 For the support of broad set of use cases the DASH-IF IOP HEVC 4k Extension is defined. UHD
4 HEVC 4k video encoded with H.265/HEVC is an advanced distribution format for TV services
5 that enables higher resolution experiences in an efficient manner.

6 In addition, the features of DASH-IF IOP Main as defined in section 8.5 and DASH-265/HEVC as
7 defined in section 6.2.3, this extension adds the Main interoperability point to include 4k resolu-
8 tions up to 60fps, and restricts the codec support to HEVC Main 10 Level 5.1.

9 The conformance to DASH-IF IOP HEVC 4k may be signaled by a @profile attribute with the
10 value <http://dashif.org/guidelines/dash-if-uhd#hevc-4k>.

11 **10.2.2. Elementary Stream Requirements**

12 **10.2.2.1. Constraints on Picture Formats**

13 NAL Structured Video streams conforming to this Media Profile SHALL NOT exceed the follow-
14 ing coded picture format constraints:

- 15 • Maximum encoded horizontal sample count of 3840 samples
- 16 • Maximum encoded vertical sample count of 2160 samples
- 17 • Maximum Frame Rate of 60000 / 1000.

18 Additional coded picture format constraints:

- 19 • The source video format shall be progressive.
- 20 • Representations in one Adaptation Set shall only differ on the following parameters: Bi-
21 trate, spatial resolution, frame rate
- 22 • The condition of the following SHALL NOT change throughout one HEVC video track:
 - 23 ○ aspect_ratio_idc
 - 24 ○ cpb_cnt_minus1
 - 25 ○ bit_rate_scale
 - 26 ○ bit_rate_value_minus1
 - 27 ○ cpb_size_scale
 - 28 ○ cpb_size_value_minus1
- 29 • The following fields should not change throughout an HEVC elementary stream:
 - 30 ○ pic_width_in_luma_samples
 - 31 ○ pic_height_in_luma_samples

32 Note: A content provider should not change the parameters unless it is aware that the
33 decoder and receiver can handle dynamic resolution switching, in particular switching
34 from lower values to higher values. Clients should implement dynamic resolution
35 switching based on DASH-IF IOP test vectors.

- 36 • YCbCr shall be used as the Chroma Format and 4:2:0 for color sub-sampling. The bit
37 depth of the content shall be either 8 bit or 10 bit. The content shall be restricted to the
38 HEVC video codec. See Section 10.2.2 for details about HEVC encoding.

-
- 1 • The color primaries shall be ITU-R BT.709 [73].

2 **10.2.2.2. Bitstream Requirements and Recommendations**

3 A bitstream conforming to the H.265/HEVC 4k media profile shall comply with the Main10 Tier
4 Main Profile Level 5.1 restrictions, as specified in Recommendation ITU-T H.265 / ISO/IEC
5 23008-2 [19].

6 UHD HEVC 4k bitstreams shall set `vui_parameters_present_flag` to 1 in the active Sequence Parameter Set, i.e. HEVC bitstreams shall contain a Video Usability Information syntax structure.

9 The sample aspect ratio information shall be signaled in the bitstream using the `aspect_ratio_idc` value in the Video Usability Information (see values of `aspect_ratio_idc` in Recommendation ITU-T H.265 / ISO/IEC 23008-2:2015 [19], table E-1). UHD HEVC 4k bitstreams shall represent square pixels indicated by `aspect_ratio_idc` shall be set to 1.

13 In addition to the provisions set forth in Recommendation ITU-T H.265 / ISO/IEC 23008-2:2015 [19], the following restrictions shall apply for the fields in the sequence parameter set:

- 15 - `vui_parameters_present_flag` = 1
16 - `sps_extension_flag` = 0
17 - `fixed_pic_rate_general_flag` = 1
18 - `general_interlaced_source_flag` = 0

19 In addition to the provisions set forth in Recommendation ITU-T H.265 / ISO/IEC 23008-2:2015[19], the following restrictions shall apply for the fields in the `profile_tier_level` syntax structure in the sequence parameter set:

- 22 - `general_tier_flag` = 0
23 - `general_profile_idc` = 2

24 UHD HEVC 4k bitstreams shall obey the limits in Recommendation ITU-T H.265 / ISO/IEC 23008-2:2015 [19], table A.1 and table A.2 associated to Level 5.1. `general_level_idc` shall be less than or equal to 153 (level 5.1).

27 It is recommended that bitstreams which are compliant with the Main or Main10 profile set `general_profile_compatibility_flag`[1] to 1.

29 The chromaticity co-ordinates of the ideal display, opto-electronic transfer characteristic of the source picture and matrix coefficients used in deriving luminance and chrominance signals from the red, green and blue primaries shall be explicitly signaled in the encoded HEVC Bitstream by setting the appropriate values for each of the following 3 parameters in the VUI: `colour_primaries`, `transfer_characteristics`, and `matrix_coeffs`.

34 ITU-R BT.709 [73] colorimetry usage is signalled by setting `colour_primaries` to the value 1, `transfer_characteristics` to the value 1 and `matrix_coeffs` to the value 1.

36 The bitstream may contain SEI messages as permitted by the Recommendation ITU-T H.265 / ISO/IEC 23008-2:2015 [19]. Details on these SEI messages are specified in Recommendation ITU-T H.265 / ISO/IEC 23008-2 / Annex D.

-
- 1 **10.2.2.3. Receiver Requirements**
- 2 Receivers conforming to the HEVC 4k media profile shall support decoding and displaying
3 H.265/HEVC 4k bitstreams as defined in clause 10.2.2.2.
- 4 No additional processing requirements are defined, for example processing of SEI messages is out
5 of scope.
- 6 **10.2.2.4. HDR Dynamic Metadata**
- 7 **10.2.2.4.1. General**
- 8 Bitstreams conforming to the UHD HEVC 4K media profile may contain one or more sets of
9 optional dynamic metadata. Details of the various metadata schemes are detailed below.
- 10 The presence of dynamic metadata is signalled by a Supplemental Descriptor with @schemeId-
11 dUri set to "<http://dashif.org/metadata/hdr>", the @value set to one of the val-
12 ues in the following table:

Dynamic Metadata Scheme	@value
TS 103.433 SEI Messages	TS103433

13 Details of the metadata schemes are provided in the subsections below:

14 **10.2.2.4.2. TS 103.433 HDR dynamic metadata**

15 When the Adaptation Set contains a Supplemental Descriptor with @schemeIdUri set to
16 "<http://dashif.org/metadata/hdr>" and @value set to "TS103433", then the bit-
17 stream shall contain one or more SL-HDR Information SEI messages, as defined in clause A.2.2
18 of ETSI TS 103433-1 [88], and may contain one or more Mastering Display Colour Volume SEI
19 messages, as defined in HEVC specification [19]. The following constraints apply, in addition to
20 those specified above in clause 10.2.2.2:

- 21 – The SL-HDR Information SEI message shall be present at least with every SAP type 1
22 or type 2.
- 23 – When carried, the Mastering Display Colour Volume SEI message shall be present at
24 least with every SAP type 1 or type 2 and shall be used as specified in clause A.3 of
25 ETSI TS 103433-1 [88].

27 **10.2.3. Mapping to DASH**

28 **10.2.3.1. Media Profile Identifier**

29 If all Representations in an Adaptation Set conforms to the elementary stream constraints for the
30 Media Profile as defined in clause 10.2.2.3 and the Adaptation Set conforms to the MPD signaling
31 according to clause 10.2.3.2 and 10.2.3.4, and the Representations conform to the file format con-

1 straints in clause 10.2.3.3, then the @profiles parameter in the Adaptation Set may signal con-
2 formance to this operation point by using "[http://dashif.org/guidelines/dash-
3 if-uhd#hevc-4k](http://dashif.org/guidelines/dash-if-uhd#hevc-4k)".

4 **10.2.3.2. MPD Signaling**

5 The MPD shall conform to DASH-IF HEVC Main IOP with the additional constraints defined in
6 clause 10.2.3.4. The @codecs parameter shall not exceed and should be set to either
7 "hvc1.2.4.L153.B0" or "hev1.2.4.L153.B0".

8 **10.2.3.3. File Format Requirements**

9 Representations used in the context of the specification shall conform to the ISO BMFF Segment
10 format [7], [9] with the following further requirements:

- The value of the duration field in the Movie Header Box ('mvhd') shall be set to a value of '0'
- The Track Header Box ('tkhd') shall obey the following constraints:
 - o The value of the duration field shall be set to '0'.
 - o The width and height fields for a visual track shall specify the track's visual presentation size as fixed-point 16.16 values expressed in on a uniformly sampled grid (commonly called square pixels)
- The Media Header Box ('mdhd') shall obey the following constraints:
 - o The value of the duration field shall be set to '0'.
- The Video Media Header ('vmhd') shall obey the following constraints:
 - o The value of the version field shall be set to '0'.
 - o The value of the graphicsmode field shall be set to '0'.
 - o The value of the opcolor field shall be set to {'0', '0', '0'}.
- The Sample Description Box ('stsd') shall obey the following constraints:
 - o A visual sample entry shall be used.
 - o The box shall include a NAL Structured Video Parameter Set
 - o the maximum width and height values shall correspond to the maximum cropped horizontal and vertical sample counts indicated in any Sequence Parameter Set in the track
 - o It shall contain a Decoder Configuration Record which signals the Profile, Level, and other parameters in the video track.
- The entry_count field of the Sample-to-Chunk Box ('stsc') shall be set to '0'.

-
- 1 - Both the `sample_size` and `sample_count` fields of the Sample Size Box
2 ('`stsz`') box shall be set to zero ('0'). The `sample_count` field of the Sample Size
3 Box ('`stz2`') box shall be set to zero ('0'). The actual sample size information can
4 be found in the Track Fragment Run Box ('`trun`') for the track.
- 5 Note: this is because the Movie Box ('`moov`') contains no media samples.
- 6 - The `entry_count` field of the Chunk Offset Box ('`stco`') shall be set to '0'.
- 7 - Any Segment Index Box ('`sidx`'), if present, shall obey the additional constraints:
- 8 o The `timescale` field shall have the same value as the `timescale` field in the
9 Media Header Box ('`mdhd`') within the same track; and
- 10 o the `reference_ID` field shall be set to the `track_ID` of the ISO Media
11 track as defined in the Track Header Box ('`tkhd`').
- 12 - For HEVCSampleEntry ('`hev1`') NAL Structured Video tracks, the '`first_sample_flags`' shall signal the picture type of the first sample in each movie fragment
13 as specified below.
- 14 o `sample_is_non_sync_sample=0`: If the first sample is a sync sample.
- 15 o `sample_is_non_sync_sample=1`: If the first sample is not a sync sam-
16 ple.
- 17 o `sample_depends_on=2`: If the first sample is an I-frame.
- 18 - The Colour Information Box should be present. If present, it shall signal the transfer
19 characteristics of the elementary stream.
- 20 - The sample timing shall obey the frame rate requirements.

22 **10.2.3.4. Adaptation Set Constraints**

23 For a video Adaptation Set, the following constraints apply, which are identical to the constraints
24 as specified in clause 3.2.10:

- 25 - The `@codecs` parameter shall be present on Adaptation Set level and shall signal the
26 maximum required capability to decode any Representation in the Adaptation Set.
- 27 - The `@profiles` parameter may be present to signal the constraints for the Adapta-
28 tion Set
- 29 - The attributes `@maxWidth` and `@maxHeight` shall be present. They are expected
30 be used to signal the source content format. This means that they may exceed the ac-
31 tual largest size of any coded Representation in one Adaptation Set.
- 32 - The `@width` and `@height` shall be signalled for each Representation (possibly de-
33 faulted on Adaptation Set level) and shall match the values of the maximum width
34 and height in the Sample Description box of the contained Representation.

-
- 1 - The attributes @minWidth and @minHeight should not be present. If present, they
2 may be smaller than the smallest @width or smallest @height in the Adaptation Set.
3 - The maximum frame rate may be signalled on Adaptation Set using the @maxFram-
4 eRate attribute.
5 - The @frameRate should be signalled for each Representation (possibly defaulted
6 on Adaptation Set level).

7 In addition to the above referenced constraints, this profile specifies the following additional con-
8 traints:

- 9 - The Color Space in use may be signalled. If signalled,
10 o an Essential or Supplemental Descriptor shall be used to signal the value by
11 setting the @schemeIdUri attribute to urn:mpeg:mpegB:cicp:Ma-
12 trixCoefficients as defined ISO/IEC 23001-8 [49] and the @value
13 attribute according to Table 4 of ISO/IEC 23001-8 [49]. The values shall
14 match the values set in the VUI.
15 o The signalling shall be on Adaptation Set level, i.e all Representations in one
16 Adaptation Set are required to have the same Chroma Format.
17 - The Color Primaries and Transfer Function may be signalled. If signalled,
18 o Essential or Supplemental Descriptors shall be used to signal the value by set-
19 ting the @schemeIdUri attribute to urn:mpeg:mpegB:cicp: Col-
20 ourPrimaries and urn:mpeg:mpegB:cicp:TransferCharac-
21 teristics, respectively, as defined ISO/IEC 23001-8 [49] and the
22 @value attribute according to the “Colour primaries” Table and the “Trans-
23 fer characteristics” Table of ISO/IEC 23001-8 [49], respectively. The values
24 shall match the values set in the VUI.
25 o The signalling shall be on Adaptation Set level only, i.e all Representations in
26 one Adaptation Set are required to have the same Color Primaries and Trans-
27 fer Function.

28 **10.2.4. Compatibility Aspects**

29 This specification is designed such that content that is authored in conformance to this IOP is
30 expected to conform to the media profile defined by DVB DASH in ETSI TS 103 285 [42] and
31 following the *3GPP H.265/HEVC UHD Operation Point* in section 5.6 of 3GPP TS26.116 [77].
32 However, in contrast to DVB and 3GPP, only BT.709 may be used and not BT.2020.

33 In addition, clients conforming to this extension should be capable to play content authored as
34 conform to the media profile defined by DVB DASH in ETSI TS 103 285 [42] and following the
35 *3GPP H.265/HEVC UHD Operation Point* in section 5.6 of 3GPP TS26.116 [77], if BT.709 colour
36 space is used.

10.3. DASH-IF IOP HEVC HDR PQ10

10.3.1. Introduction

For the support of broad set of use cases addressing higher dynamic range (HDR) and wide colour gamut (WCG), the DASH-IF IOP HEVC HDR Perceptual Quantization (PQ) 10 Extension is defined. This interoperability point allows for additional UHD features including Wide Color Gamut, High Dynamic Range and a new electro-optical transfer curve. These features are in addition to the existing features described in the DASH-IF UHD 4k interoperability point, except that this profile is designed for HDR, and requires the use of SMPTE ST 2084 [71] and Rec. BT-2020 [74] colour space. Note that this is identical to Rec. BT-2100 [80], PQ transfer function, Y'C'BC'R color difference formats, with 10 bit signal representation and narrow range.

Note that this Extension does not require the use of the maximum values, such as 60fps or 4K resolution. The content author may offer lower spatial and temporal resolutions and may use the regular DASH signalling to indicate the actual format of the source and rendering format. Typical cases may be to use HDR together with an HD 1080p signal. Note also that Adaptation Set Switching as defined in section 3.8 may be used to separate different spatial resolutions in different Adaptation Sets to address different capabilities, but still permit the use of lower resolutions for service continuity of higher resolutions.

The compliance to DASH-IF IOP HEVC HDR PQ10 may be signaled by a @profile attribute with the value <http://dashif.org/guidelines/dash-if-uhd#hevc-hdr-pq10>.

10.3.2. Elementary Stream Requirements

10.3.2.1. Introduction

The same requirements as for UHD HEVC 4k as documented in section 10.2 hold, expect for the changes as detailed below.

The changes in the HEVC HDR PQ10 profile that extend it beyond the HEVC 4K profile include:

- NAL Structured Video Streams conforming to this interoperability point SHALL be encoded using the REC-2020 color parameters as defined in [74]. Clients shall be able to correctly decode content that is encoded using that color space.
- NAL Structured Video Streams conforming to this interoperability point SHALL be encoded using the SMPTE ST 2084 electro-optic transfer function as defined in [71]. Clients shall be able to correctly decode content that is encoded using that electro-optic transfer function. Note that one cannot author a single piece of content that is compliant with both this profile and HEVC 4k profile. However, the content may be offered in one MPD in two different Adaptation Sets.

Optional metadata may be present in form SEI messages defined in ITU-T H.265 /ISO/IEC 23008-2:2015 [19].

10.3.2.2. Bitstream Requirements and Recommendations

A bitstream conforming to the HEVC HDR PQ10 media profile shall comply with the Main Tier Main10 Profile Level 5.1 restrictions, as specified in Recommendation ITU-T H.265 / ISO/IEC 23008-2 [19].

-
- 1 In addition the requirements in section 10.2.2.2 apply, except that this profile requires the use of
2 Recommendation ITU-R BT.2020 [74] non-constant luminance colorimetry and SMPTE ST 2084
3 [71].
- 4 SMPTE ST 2084 [71] usage shall be signaled by setting `colour_primaries` to the value 9,
5 `transfer_characteristics` to the value 16 and `matrix_coeffs` to the value 9.
- 6 The bitstream may contain SEI messages as permitted by the Recommendation ITU-T H.265 /
7 ISO/IEC 23008-2:2015 [19]. Details on these SEI messages are specified in Recommendation
8 ITU-T H.265 / ISO/IEC 23008-2 / Annex D. SEI message may for example support adaptation of
9 the decoded video signals to different display capabilities or more detailed content description, in
10 particular those specified in Recommendation ITU-T H.265 / ISO/IEC 23008-2 / Annex D in re-
11 lation to HDR. Other SEI Messages defined in ITU-T H.265 / ISO/IEC 23008-2 / Annex D may
12 be present as well.

13 **10.3.2.3. Receiver Requirements**

- 14 Receivers conforming to the HEVC HDR PQ10 media profile shall support decoding and display-
15 ing HEVC HDR PQ10 bitstreams as defined in section 10.3.2.2.
- 16 No additional processing requirements are defined, for example processing of SEI messages is out
17 of scope.

1 **10.3.2.4. HDR Dynamic Metadata**

2 **10.3.2.4.1. General**

3 Bitstreams conforming to the HEVC HDR PQ10 media profile may contain one or more sets of
4 optional dynamic metadata. Details of the various metadata schemes are detailed below.

5 The presence of dynamic metadata is signalled by a Supplemental Descriptor with @schemeId-
6 dUri set to "<http://dashif.org/metadata/hdr>", the @value set to once of the val-
7 ues in the following table:

Dynamic Metadata Scheme	@value parameter
SMPTE 2094-10 SEI Messages	SMPTE2094-10
SMPTE 2094-40 SEI Messages	SMPTE2094-40
TS 103.433 SEI Messages	TS103433

8 Details of the metadata schemes are provided in the subsections below:

9 **10.3.2.4.2. SMPTE 2094-10 HDR dynamic metadata**

10 When the Adaptation Set contains a Supplemental Descriptor with @schemeIdUri set to
11 "<http://dashif.org/metadata/hdr>" and @value set to "SMPTE2094-10", then
12 the bitstream shall contain SMPTE 2094-10 [83] metadata, provided as a Supplemental Enhance-
13 ment Information (SEI) message containing a DM_data() message (as defined in SMPTE 2094-10
14 [83] Annex C- Display Management Message) in accordance with "User data registered by Recom-
15 mendation ITU-T T.35 SEI message" syntax element.

16 In addition to the Bitstream Requirements defined above in Section 10.3.2.2, when ST2094-40
17 dynamic metadata is carried, exactly one ST 2094-10 SEI message shall be sent for every access unit
18 of the bitstream.

1

2 **10.3.2.4.3. SMPTE 2094-40 HDR dynamic metadata**

3 When the Adaptation Set contains a Supplemental Descriptor with @schemeIdUri set to
4 "http://dashif.org/metadata/hdr" and @value set to "SMPTE2094-40", then
5 the bitstream shall contain SMPTE ST 2094-40 [89] metadata, provided as a Supplemental En-
6 hancement Information (SEI) message (as defined in CTA861-G [90]) in accordance with "User data
7 registered by Recommendation ITU-T T.35 SEI message" syntax element.

8 This SEI message provides information to enable colour volume transformation of the reconstructed
9 colour samples of the output pictures. The input to the indicated colour volume transform process is
10 the linearized RGB colour components of the source content. The semantics and usage of the dynamic
11 metadata shall be in conformance with the specifications in SMPTE ST 2094-40 [89].

12 In addition to the Bitstream Requirements defined above in clause 10.3.2.2, when ST2094-40 dynamic
13 metadata is carried, exactly one ST 2094-40 SEI message shall be present with every SAP of type 1 or
14 type 2.

15 **10.3.2.4.4. TS 103.433 HDR dynamic metadata**

16 When the Adaptation Set contains a Supplemental Descriptor with @schemeIdUri set to
17 "http://dashif.org/metadata/hdr" and @value set to "TS103433", then the bit-
18 stream shall contain one or more SL-HDR Information SEI messages, as defined in clause A.2.2
19 of ETSI TS103433-1 [88], and may contain one or more Mastering Display Colour Volume SEI
20 messages, as defined in HEVC specification[19]. The following constraints apply, in addition to
21 those specified above in clause 10.3.2.2:

- 22 – The SL-HDR Information SEI message shall be present at least with every SAP type 1
23 or type 2.
- 24 – When carried, the Mastering Display Colour Volume SEI message shall be present at
25 least with every SAP type 1 or type 2 and shall be used as specified in clause A.3 of
26 ETSI TS103433-1 [88].

27

28 **10.3.3. Mapping to DASH**

29 **10.3.3.1. Media Profile Identifier**

30 If all Representations in an Adaptation Set conforms to the elementary stream constraints for the
31 Media Profile as defined in clause 10.3.3.2 and the Adaptation Set conforms to the MPD signalling
32 according to clause 10.3.3.2 and 10.3.3.4, and the Representations conform to the file format con-
33 straints in clause 10.3.3.3, then the @profiles parameter in the Adaptation Set may signal con-
34 formance to this operation point by using "http://dashif.org/guidelines/dash-if-
35 uhd#hevc-hdr-pq10".

36 **10.3.3.2. MPD Signaling**

37 The MPD shall conform to DASH-IF HEVC Main IOP as defined with the additional constraints
38 defined in clause 10.3.3.4. The @codecs parameter shall not exceed and should be set to either
39 "hvc1.2.4.L153.B0" or "hev1.2.4.L153.B0".

1 **10.3.3.3. File Format Requirements**

2 The file format requirements as defined in clause 10.2.3.3 shall apply.

3 **10.3.3.4. Adaptation Set Constraints**

4 The same requirements as defined in clause 10.2.3.4 shall apply.

5 **10.3.4. Compatibility Aspects**

6 Content authored according to this extensions is expected to be interoperable with the HDR10
7 profile defined in the DECE CFF Content Specification v2.2 [78], although it should be noted that
8 the DECE CFF profile may have additional constraints, such as bitrate restrictions and required
9 metadata.

10 Content authored according to this extensions is expected to be interoperable with the PQ10 pack-
11 age defined in the UHD Forum Guidelines phase A [79].

12

13 **10.4. DASH-IF IOP UHD Dual-Stream (Dolby Vision)¹**

14 **10.4.1. Introduction**

15 For the support of broad set of backward compatible use cases the DASH-IF IOP Dual-Stream
16 (Dolby Vision) Interoperability Point is defined. Backward Compatible refers to a simple method
17 for one delivery format to satisfy both an HDR client and an SDR client. This Interoperability
18 Point allows for two interlocked video streams, as described in the clause 10.4.2 below (restrictions
19 to Enhancement Layers and Annex D 1.1). These two layers are known as the Base and
20 Enhancement layers, where the Base Layer fully conforms to previous non-UHD or UHD DASH-
21 IF Interoperability point. The EL provides additional information, which combined with the BL in
22 a composition process produces a UHD output signal, including Wide Color Gamut and High
23 Dynamic Range signal at the client.

24 The compliance to *DASH-IF IOP Dual-Stream (Dolby Vision)* may be signaled by a @profile
25 attribute on the Enhancement Layer with the value <http://dashif.org/guidelines/dash-if-uhd#dvduallayer>

27 **10.4.2. Definition**

28 **10.4.2.1. General**

29 The dual-stream solution includes two video streams, known as the Base Layer and the
30 Enhancement Layer. The high-level overview of the dual-stream process is shown in Figure 26
31 Overview of Dual-stream System.

32

¹ Note: This extension is designed to be compatible with the “Dolby Vision Media Profile Definition” in DECE “Common File Format & Media Formats Specification” Version 2.2. The name of the DASH-IF extension is inherited from the DECE document in order to indicate the compatibility with this DECE Media Profile.

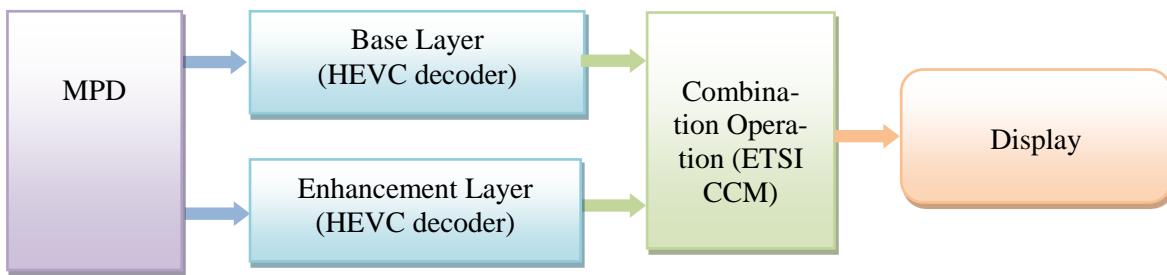


Figure 26 Overview of Dual-stream System

The MPD includes at least two Adaptation Sets as described below, including a Base Layer Adaptation Set and an Enhancement Layer Adaptation Set.

The Base Layer shall conform to the requirements of one of the following Interoperability Points: the DASH-IF IOP Main Interoperability Point, the DASH-IF IOP UHD 4k Interoperability point or the DASH-IF IOP UHD HDR10 Interoperability point. Any client that is able to play DASH-IF IOP Main content, DASH-IF IOP UHD 4k content, or DASH-IF IOP UHD HDR10 content as appropriate will be able to play the content from the Base Layer track as determined by the client capabilities. To be clear, the Base Layer is 100% conforming, with no changes or additional information, to the profile definition. A client that plays content conforming to the Base Layer profile will be able to play the Base Layer content with no modification and no knowledge of the Enhancement Layer or and Dolby Vision specific information. See Annex E, Sample MPD, for an example dual-layer MPD.

In addition, The Enhancement Layer shall conform to H.265/HEVC Main10 Profile Main Tier as defined in Recommendation ITU-T H.265 / ISO/IEC 23008-2, Level 5.1 or lower. The Enhancement Layer shall conform to the following additional requirements:

- The Frame Rate is identical to the Base Layer video track.
- The EL DPB (Decoded Picture Buffer) shall support the same number of maximum frames as the maximum number of frames supported by the BL's DPB.
- If the Base layer sample contains an IDR picture, the Enhancement Layer sample must have an IDR picture at the same presentation time.
- Fragment durations and Presentation times are identical to the Base Layer video track. To clarify, “Presentation times are identical” means that for each picture at one layer, there shall be a picture at the other layer with the same presentation time.
- Each Enhancement Layer track has one and only one associated Base Layer video track (i.e. tracks are paired 1:1).

The client - may either play the Base Layer alone, in which case it complies with the requirements of those interoperability points, or the client plays the Base Layer and Enhancement Layer together, decoding both layers and combining them to produce a 12 bit enhanced HDR signal which conforms to REC.2020 color parameters and SMPTE-2084 electro-optical transfer function. The details of this combination operation are detailed in ETSI Specification “Compound Content Management” [85].

1 Content shall only be authored claiming conformance to this IOP if a client can properly play the
2 content through the method of combining the Base Layer and Enhancement layers to produce an
3 enhanced HDR output. Note that clients who conform to the profile associated with the Base Layer
4 alone may play the Base Layer alone, with no information (and no knowledge) of the Enhancement
5 Layer. In addition, the content shall follow the mandatory aspects and should take into account the
6 recommendations and guidelines for content authoring documented in sections 8 and 10 and
7 HEVC-related issues in this section.

8
9 The dual-stream delivery of Dolby Vision asset uses two tracks; the Base Layer is written into one
10 track according to the profile of the Base Layer, and the Enhancement Layer exists in a second
11 track, per the [TBD Reference on integration, 12] specification and the details in Annex C and
12 Annex D. In particular, details about required mp4 Boxes and sample entries are detailed in Annex
13 C, “Dolby Vision Streams Within the ISO Base Media File Format”

14 The Enhancement Layer is identified by an additional parameter, @dependencyId, which
15 identifies the Base layer which is the match for the Enhancement Layer as described in clause
16 10.4.2.3.

17 **10.4.2.2. Bitstream Requirements for Enhancement Layer**

18 The sample aspect ratio information shall be signaled in the bitstream using the
19 aspect_ratio_idc value in the Video Usability Information (see values of
20 aspect_ratio_idc in Recommendation ITU-T H.265 / ISO/IEC 23008-2:2013 [19], table E-
21 1).

22
23 In addition to the provisions set forth in Recommendation ITU-T H.265 / ISO/IEC 23008-2:2013
24 [19], the following restrictions shall apply for the fields in the sequence parameter set:

- 25 • bit_depth_luma_minus8 shall be set to “2”.
- 26 • aspect_ratio_idc shall be set to “1”.
- 27 • general_interlaced_source_flag shall be set to “0”.

28 **10.4.2.3. Supplemental Enhancement Information for Enhancement Layer**

29 **10.4.2.3.1. General**

30 In addition to the requirements imposed in clause 10.4.2.2, the following additional specifications
31 shall apply to the Enhancement Layer encoding:

32 HEVC Enhancement Layer Bitstreams shall contain the following SEI messages:

-
- 1 • User data registered by Recommendation ITU-T T.35[[IT35](#)] SEI message containing the
2 message CM_data() (named composing metadata SEI message), as described in clause
3 10.4.2.3.3.
- 4 • User data registered by Recommendation ITU-T T.35[[IT35](#)] SEI message containing the
5 message DM_data() (named display management SEI Message), as described in clause
6 10.4.2.3.4.
- 7 • Mastering display colour volume SEI message as specified in Recommendation ITU-T
8 H.265 / ISO/IEC 23008-2 Annex D with the following constraints:
9 o A valid number shall be set for the following syntax elements: display_pri-
10 maries_x[c], display_primaries_y[c], white_point_x,
11 white_point_y, max_display_mastering_luminance and
12 min_display_mastering_luminance.

13 10.4.2.3.2. **SEI User Data Syntax**

14 CM_data() messages and DM_data() messages are carried in the enhancement layer video elemen-
15 tary stream as Supplemental Enhancement Information in HEVC’s “User data registered by Rec-
16 ommendation ITU-T T.35 SEI message” syntactic element. The syntax of the composing metadata
17 SEI message and the display management SEI message is defined in Table 31.

18 **Table 31: Compound Content Management SEI message: HEVC (prefix SEI NAL unit with nal_unit_type =**
19 **39, payloadType=4)**

user_data_registered_itu_t_t35(payloadSize) {	Descriptor
itu_t_t35_country_code	b(8)
itu_t_t35_provider_code	u(16)
user_identifier	u(32)
user_data_type_code	u(8)
user_data_type_structure()	
}	

21 **itu_t_t35_country_code:** This 8-bit field shall have the value 0xB5.

22 **itu_t_t35_provider_code:** This 16-bit field shall have the value 0x0031.

23 **user_identifier:** This 32-bit code shall have the value 0x47413934 (“GA94”).

24 **user_data_type_code:** An 8-bit value that identifies the type of user data to follow in the
25 user_data_type_structure(). The values are defined in Table 32.

26 **Table 32: UserID: user identifier**

user_data_type_code	user_data_type_structure()
0x00 to 0x07	Reserved
0x08	CM_data()

0x09	DM_data()
0x0A to 0xFF	Reserved

1 **user_data_type_structure()**: This is a variable length set of data defined by the value of
 2 user_data_type_code and table C.1 (DM_data()) or table D.1 (CM_data()).

3 **10.4.2.3.3. Composing Metadata SEI Message**

4 The composing metadata SEI message is a “user data registered by Recommendation ITU-T T.35
 5 SEI message” containing a CM_data() message, as specified in Annex F.

6 HEVC Enhancement Layer Bitstreams shall contain composing metadata SEI messages with the
 7 following constraints:

- 8 • It shall be sent for every access unit of the HEVC Enhancement Layer Bitstream.
- 9 • Bitstreams shall conform to ETSI Profile 1 as defined in Annex A of [85] and the value of
 10 the syntax element ccm_profile shall be set to “1”.
- 11 • The value of the syntax element ccm_level shall be set to “0”.
- 12 • The value of BL_bit_depth_minus8 shall be set to “2”.
- 13 • The value of EL_bit_depth_minus8 shall be set to to “2”.
- 14 • The value of the syntax element hdr_bit_depth_minus8 shall be set to “2” or “4”.

15 **10.4.2.3.4. Display Management SEI Message**

16 The display management SEI message is a “user data registered by Recommendation ITU-T T.35
 17 SEI message” containing a DM_data() message, as specified in Annex C.

18 HEVC Enhancement Layer Bitstreams shall contain display management SEI messages with the
 19 following constraints:

- 20 • It shall be sent for every access unit of the HEVC Enhancement Layer Bitstream.
- 21 • app_identifier shall be set equal to “1”.
- 22 • app_version shall be set equal to “1”.
- 23 • The number of extension blocks with ext_block_level equal to “1” shall be con-
 24 strained to be equal to “1”.
- 25 • The number of extension blocks with ext_block_level equal to “2” shall be con-
 26 strained to be less than or equal to “16”.
- 27 • The number of extension blocks with ext_block_level equal to “5” shall be con-
 28 strained to be equal to “0” or “1”.

29 **10.4.3. Mapping to DASH**

30 **10.4.3.1. Media Profile Identifier**

31 If all Representations in an Adaptation Set conforms to the elementary stream constraints for the
 32 Media Profile as defined in clause 10.4.2.1 and the Adaptation Set conforms to the MPD signaling

1 according to clause 10.4.3.2 and 10.4.3.3, and the Representations conform to the file format
2 constraints in clause 10.4.3.4, then

- 3
- 4 - the @profiles parameter in the Adaptation Set may signal conformance to this operation
5 point by using "<http://dashif.org/guidelines/dash-if-uhd#dvdual-layer>" on the Enhancement Layer (the Base Layer uses the normal signaling of the layer
6 as defined in the profile of the Base Layer).
 - 7

8 **10.4.3.2. MPD Signaling**

9 The MPD shall conform to DASH-IF HEVC Main IOP as defined with the additional constraints
10 defined in clause 10.4.2.

11 **10.4.3.3. Codec Parameter Signaling**

12 When the Dual-Stream Dolby Vision asset is delivered as two files, the Enhancement Layer is
13 identified by an additional parameter, @dependencyId, which identifies the Base Layer that is
14 the match for the Enhancement Layer. The Base Layer Representation element must have an @id
15 attribute, and the @dependencyId attribute on the Enhancement Layer Representation shall
16 refer to that @id, to indicate to a client that these two representations are linked. Note that in this
17 case, the @codecs attribute for the Base Layer will have only the Base Layer codec. In this
18 example, the Base Layer @codecs might be:

19

20 codecs="hvc1.1.0.L120.00"

21 And the Enhancement Layer @codecs would be:

22 codecs="dvhe.dtr.uhd30"

23 For both the Base Layer and the Enhancement Layer, HEVC decoders are used in accordance with
24 the @codecs signaling on each layer. The syntax and semantics of the @codecs signaling on
25 the enhancement layer is detailed in Annex D. The output of the decoders are combined by the
26 method detailed in ETSI Specification "Compound Content Management" [85].

27 **10.4.3.4. File Format Requirements**

28 Content shall only be authored claiming conformance to this IOP if a client can properly play the
29 content. In addition, the content shall follow the mandatory aspects and should take into account
30 the recommendations and guidelines for content authoring documented in clause 8 and 10 and
31 HEVC-related issues in clause 6.2.

32 **11. DASH-IF VP9 Extensions**

33 **11.1. Introduction**

34 VP9 [86] is an alternative video codec which may be used for SD, HD, and UHD spatial reso-
35 lutions, as well as HDR10 and HDR12 bit depths (HDR + WCG); and frame rates of 24fps and
36 higher. This codec provides significant bandwidth savings at equivalent qualities with respect to
37 AVC/H.264. While not meant to replace AVC and HEVC, DASH presentations may include ad-
38 ditional VP9 representations for playback on clients which support it.

11.2. DASH-Specific Aspects for VP9 Video

11.2.1. General

For the integration in the context of DASH, the following applies for VP9:

- The encapsulation of VP9 video data in ISO BMFF is defined in the VP Codec ISO-BMFF Binding specification [87]. Clients shall support both sample entries containing 'vp09' and 'vpcC' boxes, i.e. inband storage for VPCodecConfigurationBox + VPCodecConfigurationRecord.
- For delivery to consumer devices, only VP9 profile 0 (4:2:0 chroma subsampling and 8-bit pixel depth), and profile 1 (4.2.0 chroma subsampling and 10- or 12-bit pixel depths) shall be used.
- Stream Access Points shall coincide with the beginning of key frames (uncompressed header field `frame_type = 0`) as defined in the VP9 Bitstream Specification [86] section 7.2. Only type-1 SAPs are supported. Fragmentation and segmentation shall occur only at these points.
- Codec and codec configuration signaling in the MPD shall occur using the codec string defined in the VP Codec Binding Specification [87], DASH Application section.
- Encryption shall be signaled by the same mechanisms as defined in Common Encryption for ISO-BMFF Containers 3rd edition. Subsample encryption is required as per the VP Codec ISO Media File Format Binding spec [87].

11.2.2. Bitstream Switching

For VP9 video streams, if the `@bitstreamSwitching` flag is set to true, then the following additional constraints shall apply:

- Edit lists shall not be used to synchronize video to audio and presentation timelines.
- Video Media Segments shall set the first presented sample's composition time equal to the first decoded sample's decode time, which equals the `baseMediaDecodeTime` in the Track Fragment Decode Time Box ('tfdt').
 - o Note: This requires the use of negative composition offsets in a v1 Track Run Box ('trun') for video samples, otherwise video sample reordering will result in a delay of video relative to audio.
- The `@presentationTimeOffset` attribute shall be sufficient to align audio, video, subtitle, and presentation timelines at presentation a Period's presentation start time. Any edit lists present in Initialization Segments shall be ignored. It is strongly recommended that the Presentation Time Offset at the start of each Period coincide with the first frame of a Segment to improve decoding continuity at the start of Periods.
- All representations within the Adaptation set shall have the same picture aspect ratio.
- All VP9 decoders are required to support dynamic video resolutions, however pixel bit-depths may not vary within an adaptation set. Because of this the encoding Profile must remain constant, but the Level may vary.
- All Representations within a video Adaptation Set shall include an Initialization Segment containing an 'vpcC' Box containing a Decoder Configuration Record with the highest, , Level, vertical and horizontal resolutions of any Media Segment in the Representation.

-
- 1 - The **AdaptationSet@codecs** attribute shall be present and contain the maximum
2 level of any Representation contained in the Adaptation Set.
3 - The **Representation@codecs** attribute may be present and in that case shall contain
4 the maximum level of any Segment in the Representation.
5

6 **11.3. DASH-IF VP9 Extension IOPs**

7 **11.3.1. DASH-IF VP9-HD**

8 The scope of the DASH-IF VP9-HD extension interoperability point is basic support of high-
9 quality video distribution over the top based on VP9 up to 1080p with 8-bit pixel depth and
10 up to 30fps. Both, live and on-demand services are supported.

11 The compliance to DASH-VP9 main may be signaled by a @profiles attribute with the
12 value "http://dashif.org/guidelines/dashif#vp9"

13 A DASH client conforms to this extension IOP by supporting at least the following features:

- 14 - All DASH-related features as defined in clause 3 of this document.
15 - The requirements and guidelines in section 4.9.2 for simple live operation.
16 - The requirements and guidelines in section 5.6.1 for server-based ad insertion.
17 - Content protection based on common encryption and key rotation as defined in section
18 7. And specifically, the client supports MPD-based parsing parameters for common en-
19 encryption.
20 - All VP9 DASH IF IOP requirements in clause 11.2.
21 - VP9 Profile 0 up to level 4.1.

24 **11.3.2. DASH-IF VP9-UHD**

25 The scope of the DASH-IF VP9-UHD extension interoperability point is basic support of high-
26 quality video distribution over the top based on VP9 up to 2160p with 8-bit pixel depth and
27 up to 60fps. Both, live and on-demand services are supported.

28 The compliance to DASH-VP9 main may be signaled by a @profiles attribute with the
29 value "http://dashif.org/guidelines/dash-if-uhd#vp9"

31 A DASH client conforms to this extension IOP by supporting at least the following features:

- 32 - All features supported by DASH-IF VP9-HD defined in clause 11.3.1.
33 - VP9 Profile 0 up to level 5.1.

35 **11.3.3. DASH-IF VP9-HDR**

36 The scope of the DASH-IF VP9-HDR extension interoperability point is basic support of high-
37 quality video distribution over the top based on VP9 up to 2160p with 10-bit pixel depth and
38 up to 60fps. Both, live and on-demand services are supported.

1 The compliance to DASH-VP9 main may be signaled by a @profiles attribute with the
2 value <http://dashif.org/guidelines/dashif#vp9-hdr> (up to HD/1080p
3 resolution), or <http://dashif.org/guidelines/dash-if-uhd#vp9-hdr>
4 (up to 4K resolution).

5

6 A DASH client conforms to this extension IOP by supporting at least the following features:

- 7 - All features supported by DASH-IF VP9-UHD defined in clauses 11.3.2.
8 - VP9 profile 2 up to level 5.1.
9 - Pixel depths of 10 bits.

10

Annex A Examples for Profile Signalling

Example 1

In this case DASH-IF IOP content is offered, but in addition a non-conforming Adaptation Set is added.

Here is an example for an MPD:

- **MPD@profiles**="urn:mpeg:dash:profile:isoff-on-demand:2011,
http://dashif.org/guidelines/dash264"
 - **AdaptationSet@profiles**="urn:mpeg:dash:profile:isoff-on-demand:2011, http://dashif.org/guidelines/dash264"
 - **AdaptationSet@profiles**="http://dashif.org/guidelines/dash264"
 - **AdaptationSet@profiles**="urn:mpeg:dash:profile:isoff-on-demand:2011"

Pruning process for IOP <http://dashif.org/guidelines/dash264> results in

- **MPD@profiles**="http://dashif.org/guidelines/dash264"
 - **AdaptationSet@profiles**="http://dashif.org/guidelines/dash264"
 - **AdaptationSet@profiles**="http://dashif.org/guidelines/dash264"

It is now required that the pruned MPD conforms to DASH-IF IOP.

Example 2

In this case DASH-IF IOP content is offered, but in addition a non-conforming Adaptation Set is added and one DASH-IF Example Extension Adaptation Set is added with the virtual IOP signal <http://dashif.org/guidelines/dashif#extension-example>.

Here is an example for an MPD:

- **MPD@profiles**="urn:mpeg:dash:profile:isoff-on-demand:2011,
http://dashif.org/guidelines/dash264, http://dashif.org/guidelines/dashif#extension-example"
 - @id = 1, **AdaptationSet@profiles**="urn:mpeg:dash:profile:isoff-on-demand:2011, http://dashif.org/guidelines/dash264"
 - @id = 2, **AdaptationSet@profiles**="http://dashif.org/guidelines/dash264"
 - @id = 3, **AdaptationSet@profiles**="urn:mpeg:dash:profile:isoff-on-demand:2011, http://dashif.org/guidelines/dashif#extension-example"

Pruning process for profile <http://dashif.org/guidelines/dash264> results in

- **MPD@profiles**="http://dashif.org/guidelines/dash264"
 - @id = 1, **AdaptationSet@profiles**="http://dashif.org/guidelines/dash264"
 - @id = 2, **AdaptationSet@profiles**="http://dashif.org/guidelines/dash264"

It is now required that the pruned MPD conforms to DASH-IF IOP.

Pruning process for profile <http://dashif.org/guidelines/dashif#extension-example> results in

- **MPD@profiles**="http://dashif.org/guidelines/dash264"
 - @id = 3, **AdaptationSet@profiles**="http://dashif.org/guidelines/dashif# extension-example"

It is now required that the pruned MPD conforms to DASH-IF Example Extension Adaptation Set.

1 **Annex B Live Services - Use Cases and Architecture**

2 **B.1 Baseline Use cases**

3 **B.1.1 Use Case 1: Live Content Offered as On-Demand**

4 In this case content that was distributed as live is offered in a separate Media Presentation as On-
5 Demand Content.

6 **B.1.2 Use Case 2: Scheduled Service with known duration and Operating at 7 live edge**

8 In this case a service started a few minutes ago and lasts 30 minutes. The duration is known exactly
9 and also all segment URLs are known. The timeshift buffer is short. This may for example be a
10 live service for which the service provider wants to ensure that only a small window is accessible.
11 The content is typically be pre-canned, but offered in a scheduled manner.

12 **B.1.3 Use Case 3: Scheduled Service with known duration and Operating at 13 live edge and time shift buffer**

14 In this case a service started a few minutes ago and lasts 30 minutes. The duration is known exactly
15 and also all segment URLs are known. The timeshift buffer is long. This may for example be a
16 service for which the service provider wants to ensure that the content is made available in a sched-
17 uled manner, e.g. no client can access the content earlier than scheduled by the content provider.
18 However, after the live edge is completed, the content is available for 24h. The content is typically
19 pre-canned.

20 **B.1.4 Use Case 4: Scheduled Live Service known duration, but unknown Seg- 21 ment URLs**

22 In this case a live service started a few minutes ago and lasts 30 minutes. The duration is known
23 exactly but the segment URLs are unknown, as for example some advertisement may be added on
24 the fly. Otherwise this service is similar to use case 3.

25 **B.1.5 Use Case 5: 24/7 Live Service**

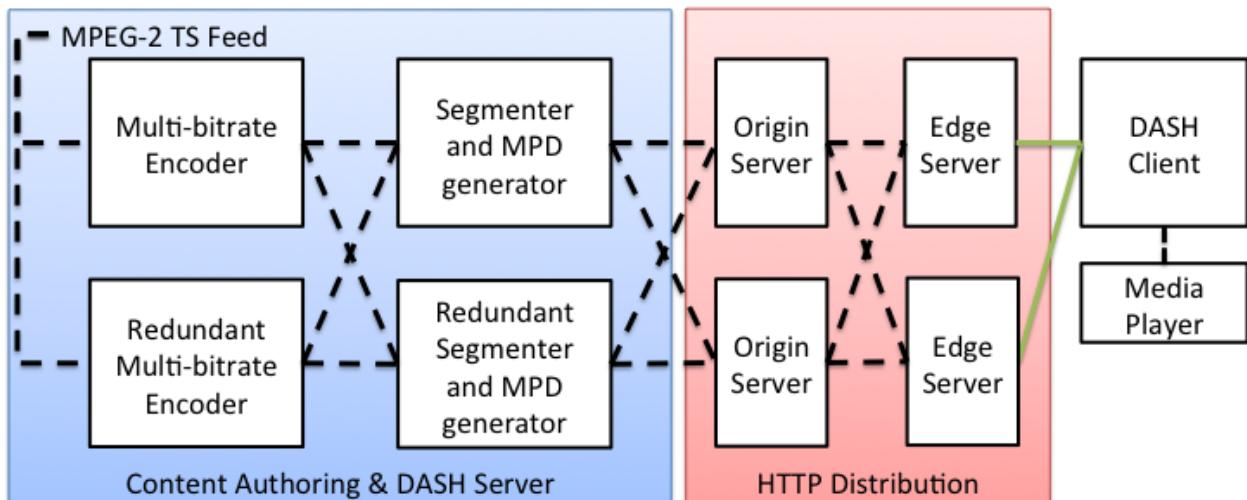
26 In this case a live service started that may have started a long time ago is made available. Ad breaks
27 and operational updates may be done with a 30sec pre-warning. The duration is unknown and also
28 the segment URLs, the exact set of provided media components (different language tracks, subti-
29 tles, etc.) are unknown, as for example some advertisement may be added on the fly. Otherwise
30 this service is similar to use case 3.

31 **B.1.6 Use Case 6: Approximate Media Presentation Duration Known**

32 In this case a live service starts at a specific time. The duration is known approximately and also
33 all segment URLs are known for the approximate duration. Towards the end of the Media Presen-
34 tation, the Media Presentation duration may be extended or may be finally determined by providing
35 an update of the MPD.

36

1 **B.2 Baseline Architecture for DASH-based Live Service**



2 **Figure 27 Typical Deployment Scenario for DASH-based live services**

3 The figure depicts a redundant set-up for Live DASH with unicast. Function redundancy is added
4 to mitigate the impact of function failures. The redundant functions are typically connected to
5 multiple downstream functions to mitigate link failure impacts.

6 An MPEG2-TS stream is used often as feed into the encoder chain. The multi-bitrate encoder
7 produces the required number of Representations for each media component and offers those in
8 one Adaptation Set. In the context of this document is assumed that content is offered in the ISO
9 BMFF live profile with the constraints according to v2 of this document. The encoder typically
10 locks to the system clock from the MPEG2-TS stream. The encoder forwards the content to the
11 segmenter, which produces the actual DASH segments and handles MPD generation and updates.
12 Content Delivery Network (CDN) technologies are typically used to replicate the content to mul-
13 tiple edge servers. Note: the CDN may include additional caching hierarchy layers, which are not
14 depicted here.

15 Clients fetch the content from edge servers using HTTP (green connection) according to the
16 MPEG-DASH and DASH-IF IOP specification. Different protocols and delivery formats may be
17 used within the CDN to carry the DASH segments from the segmenter to the Edge Server. For
18 instance, the edge server may use HTTP to check with its parent server when a segment is not (yet)
19 in the local cache. Or, segments may be pushed using IP Multicast from the origin server to rele-
20 vant edge servers. Other realizations are possible, but are outside of the normative scope of this
21 document.

22 In some deployments, the live service is augmented with ad insertion. In this case, content may
23 not be generated continuously, but may be interrupted by ads. Ads itself may be personalized,
24 targeted or regionalized.

25 **B.3 Distribution over Multicast**

26 This clause describes a baseline architecture for DASH Live Services for broadcast distribution.
27 The intention of the baseline architecture is in particular to identify robustness and failure issue
28 and give guidance on procedures to recover.

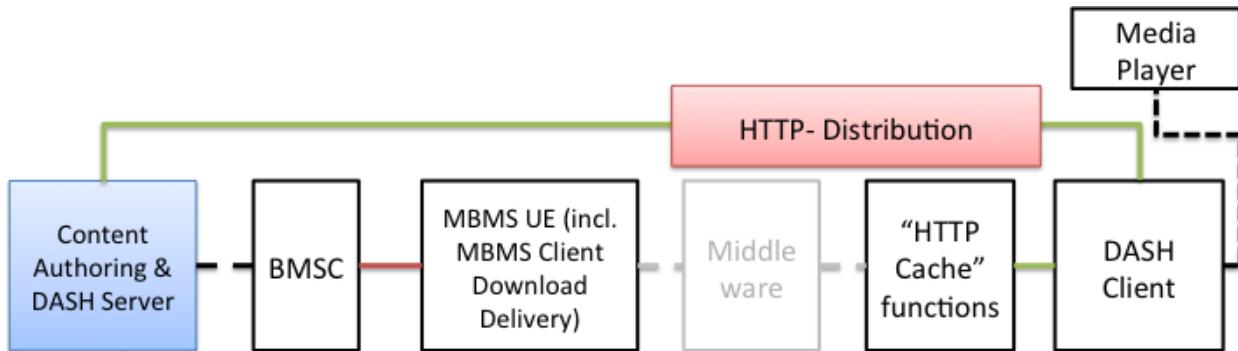


Figure 28 Typical Deployment Scenario for DASH-based live services partially offered through MBMS (uni-directional FLUTE distribution)

The same content authoring and DASH server solution as shown in Figure 1 are considered in this baseline architecture. The DASH Segmente (cf. Fig .1) provides DASH segments of typically one quality representation into the BM-SC, which sends the segments using MBMS Download (as sequence of files using IETF FLUTE protocol) to the MBMS User Equipment (UE). The MBMS UE includes the needed MBMS download delivery client functions to recover the media segments from the FLUTE reception. The MBMS UE makes the segments through a local HTTP Cache function available to the DASH client. The DASH client uses HTTP (green line) to retrieve the segments from the device local cache.

In case the MBMS reception is not possible for that Video Session, the DASH client can use unicast HTTP to acquire the stream (according to previous clause).

Note, the objective of the client architecture realization here is on using a generic DASH client for unicast and broadcast. More customized implementations are possible.

B.4 Typical Problems in Live Distribution

B.4.1 Introduction

Based on the deployment architectures in Figure 27 and Figure 28 a few typical problems in DASH-based ABR distribution are explained.

B.4.2 Client Server Synchronization Issues

In order to access the DASH segments at the proper time as announced by the segment availability times in the MPD, client and server need to operate in the same time source, in general a globally accurate wall-clock, for example provided by NTP or GPS. There are different reasons why the DASH client and the media generation source may not have identical time source, such as

- DASH client is off because it does not have any protocol access to accurate timing. This may for example be the case for DASH clients that are running in the browser or on top of a general-purpose HTTP stack.
- DASH client clock drifts against the system clock and the DASH client is not synchronizing frequently enough against the time-source.
- The segmenter synchronized against a different time source than DASH client.
- There may be unknown delay on the ingest to the server/cache whether the segment is accessible. This is specifically relevant if MBMS is used as the contribution link resulting in transport delay.

-
- 1 • It may also be that the MPD provides the availability times at the segmenter, but the actual
2 availability should be the one on the origin server.
3 • There may be a delay from segmenter to the origin server which is known by edge/origin,
4 but there may not be sufficient ways to signal this delay.

5 **B.4.3 Synchronization Loss of Segmenter**

6 The segmenter as depicted in Figure 27 may lose synchronization against the input timeline for
7 reasons such as power-outage, cord cuts, CRC losses in the incoming signals, etc. In this case:

- 8 • Loss of synchronization may result that the amount of lost media data cannot be predicted
9 which makes the generation of continuous segments difficult.
10 • The Segmenter cannot predict and correct the segment timeline based on media presenta-
11 tion timestamps, since the presentation timeline may contain a discontinuity due to the
12 synchronization loss
13 ○ a loss of sync (e.g. CRC failure on the input stream)
14 ○ a power glitch on the source
15 ○ someone pulling a cable
16 • There are cases where no media segments are available, but the MPD author knows this
17 and just wants to communicate this to the receiver.

18 **B.4.4 Encoder Clock Drift**

19 In certain cases, the MBR encoder is slaved to the incoming MPEG-2 TS, i.e. it reuses the media
20 time stamps also for the ISO BMFF.

- 21 • What may occur that the encoder clock drifts between the sender and the receivers (longer
22 term issue) , e.g. due to encoder clock tolerance
23 ○ Example: Encoder produces frame every 39.97ms instead of 40ms
24 ○ Tolerance in MPEG-2 TS: 1 frame every 18 minutes
25 • This may create issues in particular when an existing stream like for satellite is transcoded
26 and segmented into DASH representations.
27 • Annex A.8 of ISO 23009-1 handles drift control of the media timeline, but the impact on
28 the segment availability time (i.e. MPD updates) is not considered or suggested.
29 • In particular when the segment fetching engine of the client is only working with the seg-
30 ment availability timeline (so is not parsing the presentation timeline out of the segments),
31 the segment fetching engine will not fetch the segments with the correct interval, leading
32 to buffer underruns or increased e2e delay.
33 • There is practical evidence that this is a problem in actual deployments, may result in drifts
34 of minutes over hours.

35 **B.4.5 Segment Unavailability**

36 When a server cannot serve a requested segment it gives an HTTP 404 response. If the segment
37 URL is calculated according to the information given in the MPD, the client can often interpret the
38 404 response as a possible synchronization issue, i.e. its time is not synchronized to the time of-
39 fered in the MPD.

40 In the MBMS case, a 404 response is also likely to be caused by non-reparable transport errors.
41 This is even more likely if it has been possible to fetch segments according to the MPD information

1 earlier. Although the client M/W, which is normally located in the same device as the DASH
2 player, knows what segments have been delivered via broadcast and which ones are missing in a
3 sequence, it cannot indicate this to the DASH client using standard HTTP responses to requests
4 for media segments.

5 **B.4.6 Swapping across Redundant Tools**

6 In case of failures, redundant tools kick in. If the state is not fully maintained across redundant
7 tools, the service may not be perceived continuous by DASH client. Problems that may happen at
8 the encoder, that redundant encoders do not share the same timeline or the timeline is interrupted.
9 Depending on the swap strategy ("hot" or "warm"), the interruptions are more or less obvious to
10 the client. Similar issues may happen if segmenters fail, for example the state for segment num-
11 bering is lost.

12 **B.4.7 CDN Issues**

13 Typical CDN operational issues are the following:

- 14 • Cache Poisoning – at times segment generation may be erroneous. The encoder can pro-
15 duce a corrupt segment, or the segment can become corrupted during upload to origin. This
16 can happen for example if encoder connectivity fails in mid segment upload, leading to a
17 malformed segment (with the correct name) being sent to edge and caching servers. The
18 CDN then caches this corrupt segment and continues to deliver it to fulfill future requests,
19 leading to widespread client failures.
- 20 • Cache inconsistency – with a dual origin scheme, identically named segments can be pro-
21 duced with slight differences in media time, due to clock drift or other encoder issues.
22 These segments are then cached by CDNs and used to respond to client requests. If seg-
23 ments from one encoder are mixed with segments of another, it can lead to discontinuous
24 playback experiences on the clients.

25 **B.4.8 High End-to-end Latency**

26 End-to-end latency (also known as hand-waving latency) is defined as the accumulated delay be-
27 tween an action occurring in front of the camera and that action being visible in a buffered player.
28 It is the sum of

- 29 1. Encoder delay in generating a segment.
- 30 2. Segment upload time to origin server from the encoder.
- 31 3. Edge server segment retrieval time from origin
- 32 4. Segment retrieval time by the player from the edge server
- 33 5. The distance back from the live point at which the player chooses to start playback.
- 34 6. Buffering time on the player before playback commences.

35 In steps 1 through 4, assuming non-chunked HTTP transfer, the delay is a linear function of the
36 segment duration. Overly conservative player buffering can also introduce unnecessary delay, as
37 can choosing a starting point behind the live point. Generally the further behind live the player
38 chooses to play, the more stable the delivery system is, which leads to antagonistic demands on
39 any production system of low latency and stability.

1 **B.4.9 Buffer Management & Bandwidth Estimation**

2 The main user experience degradations in video streaming are rebuffering events. At the same
3 time, user experience is influenced by the quality of the video (typically determined by the bitrate)
4 as well as at least for certain cases on the end-to-end latency. In order to request the access bitrate,
5 the client does a bandwidth estimation typically based on the history and based on this and the
6 buffer level in the client it decides to maintain or switch Representations.

7 In order to compensate bandwidth variations, the client buffers some media data prior to play-out.
8 More time buffer results less buffer under runs and less rebuffering, but increases end-to-end latency.
9 In order to maximize the buffer in the client and minimize the end-to-end latency the DASH
10 client would like to request the media segment as close as possible to its actual segment availability
11 start time. However, this may cause issues in the playout as in case of bitrate variations, the
12 buffer may drain quickly and result in playout starvation and rebuffering.

13 **B.4.10 Start-up Delay and Synchronization Audio/Video**

14 At the start-up and joining, it is relevant that the media playout is initiated, but that the delay at
15 start is reasonable and that the presentation is enabled such that audio and video are presented
16 synchronously. As audio and video Representations typically are offered in different sampling
17 rates, and segments of audio and video are not aligned at segment boundaries. Hence, for proper
18 presentation at startup, it is necessary that the DASH client schedules the presentation at the
19 presentation time aligned to the over media presentation timeline.

20 **B.5 Advanced Use Cases**

21 **B.5.1 Introduction**

22 Based on the above issues a few advanced use cases are considered.

23 **B.5.2 Use Case 7: Live Service with undetermined end**

24 In this case a live service started that may have started a long time ago is made available. The MPD
25 update may be done with a 30sec pre-warning. The duration is unknown exactly but the segment
26 URLs are unknown, as for example some advertisement may be added on the fly. Otherwise this
27 service is similar to use case 3.

28 **B.5.3 Use Case 8: 24/7 Live Service with canned advertisement**

29 In this case a live service started that may have started a long time ago is made available. The MPD
30 update may be done with a 30sec pre-warning. The duration is unknown exactly but the segment
31 URLs are unknown, as for example some advertisement may be added on the fly. The advertisement
32 itself is not a dynamic service, but available on a server as a pre-canned advertisement.

33 **B.5.4 Use case 9: 24x7 live broadcast with media time discontinuities**

34 In other use cases, the content provider splices content such as programs and ads with independent
35 media timelines at the content provider.

36 **B.5.5 Use case 10: 24x7 live broadcast with Segment discontinuities**

37 Based on the discussions above, interruptions in encoding, etc., but presentation and media timelines
38 resume after loss of some segments.

1 **Annex C Dolby Vision Streams Within the ISO Base** 2 **Media File Format**

3 **C.1 Introduction**

4 This Annex defines the structures for the storage of Dolby Vision video streams in a file format com-
5 pliant with the ISO base media file format (ISOBMFF). Example file formats derived from the ISO-
6 BMFF include the Digital Entertainment Content Ecosystem (DECE) Common File Format (CFF) and
7 Protected Interoperable File Format (PIFF). Note, that the file format defined here is intended to be
8 potentially compliant with the DECE media specifications as appropriate.
9

10 **C.2 Dolby Vision Configuration Box and Decoder Configuration Rec-** 11 **ord**

12 The Dolby Vision decoder configuration record provides the configuration information that is required
13 to initialize the Dolby Vision decoder.
14

15 **C.2.1 Definition**

16 The Dolby Vision Configuration Box contains the following information:
17

18 Box Type 'dvcC'
19

20 Container
21 DolbyVisionHEVCSampleEntry('dvhe'), or
22 DolbyVisionHVC1SampleEntry('dvh1'), or
23

24 Mandatory Yes
25

26 Quantity Exactly One
27

28 **C.2.2 Syntax**

29 The syntaxes of the Dolby Vision Configuration Box and decoder configuration record are described
30 below.
31

```
32 align(8) class DOVIDecoderConfigurationRecord
33 {
34     unsigned int (8) dv_version_major;
35     unsigned int (8) dv_version_minor;
36     unsigned int (7) dv_profile;
37     unsigned int (6) dv_level;
38     bit (1) dv_metadata_present_flag;
39     bit (1) el_present_flag;
40     bit (1) bl_present_flag;
41     const unsigned int (32)[5] reserved = 0;
42 }
43 class DOVIConfigurationBox
```

```
1   extends Box('dvCC')
2 {
3     DOVIDecoderConfigurationRecord() DOVICConfig;
4 }
5
```

6 C.2.3 Semantics

7 The semantics of the Dolby Vision decoder configuration record is described as follows.

8
9 dv_version_major - specifies the major version number of the Dolby Vision specification that the
10 stream complies with. A stream compliant with this specification shall have the value 1.

11
12 dv_version_minor - specifies the minor version number of the Dolby Vision specification that the
13 stream complies with. A stream compliant with this specification shall have the value 0.

14
15 dv_profile – specifies the Dolby Vision profile. Valid values are Profile IDs as defined in Table B.1 of
16 *Signaling Dolby Vision Profiles and Levels*, Annex B.

17
18 dv_level – specifies the Dolby Vision level. Valid values are Level IDs as defined in Table B.2 of
19 *Signaling Dolby Vision Profiles and Levels*, Annex B.

20
21 dv_metadata_present_flag – if 1 indicates that this track contains the supplemental enhancement in-
22 formation as defined in clause 10.4.2.2.

23
24 el_present_flag – if 1 indicates that this track contains the EL HEVC video substream.

25
26 bl_present_flag – if 1 indicates that this track contains the BL HEVC video substream.

27
28 Note: The settings for these semantic values are specified in Section A.7.1 Constraints on EL
29 Track.

30 C.3 Dolby Vision Sample Entries

31 This section describes the Dolby Vision sample entries. It is used to describe tracks that contain sub-
32 streams that cannot necessarily be decoded by HEVC compliant decoders.

34 C.3.1 Definition

35 The Dolby Vision sample entries contain the following information:

37	Box Type	'dvhe', 'dvh1'
39	Container	Sample Description Box ('stsdb')
41	Mandatory	Yes
43	Quantity	One or more sample entries of the same type may be present

1

2 **C.3.2 Syntax**

3 The syntax for the Dolby Vision sample entries are described below.

4

```
5 class DolbyVisionHEVCSampleEntry() extends  
6     HEVCSampleEntry('dvhe')  
7 {  
8     DOVIConfigurationBox() config;  
9 }  
10  
11 class DolbyVisionHVC1SampleEntry() extends  
12     HEVCSampleEntry('dvh1')  
13 {  
14     DOVIConfigurationBox() config;  
15 }
```

16

C.3.3 Semantics

17 A Dolby Vision HEVC sample entry shall contain a Dolby Vision Configuration Box as defined in
18 C.2.2.

19

20 config - specifies the configuration information required to initialize the Dolby Vision decoder for a
21 Dolby Vision EL track encoded in HEVC.

22

23 Compressorname in the base class VisualSampleEntry indicates the name of the compressor used,
24 with the value “\013DOVI Coding” being recommended (\013 is 11, the length of the string “DOVI
25 coding” in bytes).

26

C.6 Dolby Vision Files

27 The brand ‘dby1’ SHOULD be used in the compatible_brands field to indicate that the file is com-
28 pliant with all Dolby Vision UHD Extension as outlined in this document. The major_brand shall
29 be set to the ISO-defined brand,e.g. ‘iso6’.

30

C.7 Dolby Vision Track In A Single File

31 A Dolby Vision video stream can be encapsulated in a single file as a dual-track file containing separate
32 BL and EL tracks. Each track has different sample descriptions.

33

C.7.1 Constraints on EL Track

34

For the visual sample entry box in an EL track a
35 DolbyVisionHEVCVisualSampleEntry ('dvhe') or
36 DolbyVisionHVC1VisualSampleEntry ('dvh1')
37 shall be used.

38

The visual sample entries shall contain an HEVC Configuration Box ('hvCC') and a Dolby Vision
39 Configuration Box ('dvcC').

40

41 The EL track shall meet the following constraints:

42

- In the handler reference box, the handler_type field shall be set to 'vide'.

- The media information header box shall contain a video media header box.
- The dependency between the Dolby Vision base and enhancement track shall be signaled by the ‘tref’ box in the enhancement track.. The reference_type shall be set to ‘vdep’.
- The dv_profile field in the Dolby Vision Configuration Box (‘dvCC’) shall be set according the encoded Dolby Vision profile.
- The dv_level field in the Dolby Vision Configuration Box (‘dvCC’) shall be set according the encoded Dolby Vision level.
- The dv_metadata_present_flag shall be set to 1.
- The el_present_flag shall be set to 0 or 1.
- The bl_present_flag shall be set to 0.

The following table shows the box hierarchy of the EL track.

Note: This is not an exhaustive list of boxes.

Table 33 Sample table box hierarchy for the EL track of a dual-track Dolby Vision file

Nesting Level				Reference
4	5	6	7	
stbl				ISO/IEC 14496-12
	stsd			
		dvhe, or dvh1		
			hvcC	
			dvcC	
	stts			
	stsc			
	stsz			
	stz2			
	stco			
	co64			

C.7.2 Constraints on the ISO base media file format boxes

C.7.2.1 Constraints on Movie Fragments

For a dual-track file, the movie fragments carrying the BL and EL shall meet the following constraints:

- The adjacent movie fragments (‘moof’ and ‘mdat’) for the base and enhancement track shall be interleaved with BL followed by EL. BL and EL samples shall be placed in separate Movie Fragments and that each BL Movie Fragment shall be immediately followed by an EL movie fragment containing the same number of samples with identical composition timestamps.

C.7.2.2 Constraints on Track Fragment Random Access Box

The track fragment random access box (‘tfra’) for the base and enhancement track shall conform to the ISO/IEC 14496-12 (section 8.8.10) and meet the following additional constraint:

-
- 1 • The value of the `time` field in the track fragment random access box indicates the presenta-
2 tion time of a random accessible sample. This `time` value shall be identical for every corre-
3 sponding random accessible sample in the base and enhancement track.
4
5

1 Annex D Signaling Dolby Vision Profiles and Levels

2 This Annex defines the detailed list of Dolby Vision profile/levels and how to represent them in a
3 string format. This string can be used for identifying Dolby Vision device capabilities and identifying
4 the type of the Dolby Vision streams presented to device through various delivery mechanisms such
5 as HTML 5.0 and MPEG-DASH.

6 **D.1 Dolby Vision Profiles and levels**

7 The Dolby Vision codec provides a rich feature set to support various ecosystems such as Over the Top
8 streaming, Broadcast television, Blu-Ray discs, and OTT streaming. The codec also supports many
9 different device implementation types such as GPU accelerated software implementation, full-fledged
10 hardware implementation, and hardware plus software combination. One of the Dolby Vision codec
11 features allows choosing the type of backward compatibility such as non-backward compatible or
12 backward compatible with SDR. A Dolby Vision capable device may not have all the features or op-
13 tions implemented, hence it is critical the device advertises the capabilities and content server provides
14 accurate Dolby vision stream type information.

15 **D.1.1 Dolby Vision Profiles**

16 Following are the currently supported Dolby Vision profiles:
17

18 **Table D.1: Dolby Vision Profiles**

Profile ID	Profile Name	BL Codec	EL Codec	BL:EL	BL Backward Compatibility*	BL/EL Full Alignment**	BL Codec Profile	EL Codec Profile
2	dvhe.der	HEVC8	HEVC8	1:1/4	SDR	No	H.265 Main	H.265 Main
3	dvhe.den	HEVC8	HEVC8	1:1	None	No	H.265 Main	H.265 Main
4	dvhe.dtr	HEVC10	HEVC10	1:1/4	SDR	No	H.265 Main10	H.265 Main10
5	dvhe.stn	HEVC10	N/A	N/A	None	N/A	H.265 Main10	N/A
6	dvhe.dth	HEVC10	HEVC10	1:1/4	HDR10	No	H.265 Main10	H.265 Main10
7	dvhe.dtb	HEVC10	HEVC10	1:1/4 for UHD	Blu-ray HDR	No	H.265 Main10	H.265 Main10
				1:1 for FHD				

19

20 **Legend:**

21 **BL:EL** = ratio of Base Layer resolution to Enhancement Layer resolution (when applicable)

22 **BL/EL Full alignment** = The Enhancement Layer (EL) GOP and Sub-GOP structures are fully
23 aligned with Base Layer (BL), i.e. the BL/EL IDRs are aligned, BL/EL frames are fully aligned in
24 decode order such that skipping or seeking is possible anywhere in the stream not only limited to
25 IDR. BL AU and EL AU belonging to the same picture shall have the same POC (picture order
26 count)

27 **Encoder Recommendations**

28 * Dolby Vision Encoders should only use baseline profile composer for profiles which are non-
29 backward compatible, i.e. the **BL Backward Compatibility** = **None**.

1 ** Encoders producing Dolby Vision dual layer streams should generate BL/EL with full GOP/Sub-
2 GOP structure alignment for all the profiles listed in **Table 4**.

3 **D.1.1.1 Dolby Vision Profile String format**

4 The following is the profile string naming convention:

5 dv[BL codec type].[number of layers][bit depth][backward compatibility]
6 [EL codec type][EL codec bit depth]

Attribute	Syntax
dv	dv = Dolby Vision
BL codec type	he = HEVC
Number of layers	s = single layer d = dual layer without enforcement of BL/EL GOP structure and POC alignment p = dual layer with enforcement of BL/EL GOP structure and POC alignment
Bit depth	e = 8 t = 10
Backward compatibility	n = non-backward compatible r = SDR backward compatible (rec.709, 100 nits) h = HDR10 backward compatible b = Blu-ray backward compatible (Ultra HD Blu-ray™ High Dynamic Range)
EL codec Type (see Note 1 below)	a = AVC h = HEVC
EL codec bit depth (See Note 1 below)	e = 8 t = 10

7 **Notes:**

- 8 1. [EL codec type] and [EL codec bit depth] shall only be present if the EL codec type
9 is different from the BL codec.
10 2. Interlaced: There is no support for interlaced video at this time.
11 3. Codecs other than HEVC or AVC may be supported in future.

12 **D.1.2 Dolby Vision Levels**

13 The Dolby Vision level indicates the maximum frame rate and resolution supported by the device for
14 a given profile. Typically there is a limit on the maximum number of pixels the device can process
15 per second in a given profile; the level indicates the maximum pixels and the maximum bitrate

1 supported in that profile. Since maximum pixels per second is a constant for given level, the
2 resolution can be reduced to get higher frame rate and vice versa. Following are the possible levels:
3

4 **Table B.2: Dolby Vision Levels**

Level ID	Level Name	Example Max Resolution x FPS	Max Bit Rates (BL and EL combined)			
			main tier (Mbps)	high tier (Mbps)	high tier (Mbps)	high tier (Mbps)
1	hd24	1280x720x24	20	50		
2	hd30	1280x720x30	20	50		
3	fhd24	1920x1080x24	20	70		
4	fhd30	1920x1080x30	20	70		
5	fhd60	1920x1080x60	20	70		
6	uhd24	3840x2160x24	25	130		
7	uhd30	3840x2160x30	25	130		
8	uhd48	3840x2160x48	40	130		
9	uhd60	3840x2160x60	40	130		

5

6 **B.1.2.1 Dolby Vision Level String Format**

7 The following is the level string naming convention
8 [resolution] [FPS] [high tier]

Attribute	Syntax
Resolution	hd = 720 fhd = 1080 uhd = 2160
FPS	Frames per second (e.g. 24, 30, 60)
High Tier	Whether or not higher frame rates are supported. If yes, “h” will be appended

9

10 **B.1.3 Dolby Vision Codec Profile and Level String**

11 The profile and level string is recommended to be joined in the following manner:

12 Format:

13 [Profile String].[Level String]

14

15 Examples

-
- ```
1 • dvav.per.fhd30
2 (dual layer avc 8 bit with enforcement of BL/EL GOP Structure and POC alignment, rec709
3 backwards compatible, 1920x1080@30fps)
4 • dvhe.stn.uhd30
5 (single layer hevc 10 bit non-backwards compatible, 3840x2160@30fps)
```

### 6 **B.1.3.1 Device Capabilities**

```
7 The device capabilities can be expressed in many ways depending on the protocol used by the
8 streaming service or VOD service. The device could maintain a list of supported capabilities in an
9 array:
```

```
10 String capabilities [] = {"dvhe.dtr.uhd24", "dvhe.stn.uhd30"}
```

```
11 After receiving the manifest the Player could iterate over the stream types and check whether a
12 stream type is supported by searching the capabilities[].
```

### 14 **User Agent String**

```
16 When using HTTP, the device could send the capabilities via the user agent string in HTTP request in
17 following manner:
```

```
19 Opera/9.80 (Linux armv71) Presto/2.12.407 Version/12.51 Model-
20 UHD+dvhe.dtr.uhd24+dvhe.stn.uhd30/1.0.0 (Manufacturer name, Model)
```

```
22 A server program can search for "+dv" to determine whether Dolby Vision is supported and further
23 identify the profiles and level supported by parsing the characters following the +dv. Multiple
24 profiles/level pairs can be listed with '+' beginning each profile/level pair.
```

```
25
```

```
26
```

---

## 1                   **Annex E   Display Management Message**

### 2                   **E.1 Introduction**

3                   A display management (DM) message contains metadata in order to provide dynamic information  
4                   about the colour volume of the video signal. This metadata can be employed by the display to adapt  
5                   the delivered HDR imagery to the capability of the display device. The information conveyed in this  
6                   message is intended to be adequate for purposes corresponding to the use of Society of Motion Picture  
7                   and Television Engineers ST 2094-1 and ST 2094-10.

8                   The syntax and semantics for DM\_data() are defined in clause C.2.

---

## 1 E.2 Syntax and Semantics

2 **Table C.1: DM\_data()**

| DM_data () {                             | Descriptor |
|------------------------------------------|------------|
| <b>app_identifier</b>                    | ue(v)      |
| <b>app_version</b>                       | ue(v)      |
| <b>metadata_refresh_flag</b>             | u(1)       |
| if( metadata_refresh_flag ) {            |            |
| <b>num_ext_blocks</b>                    | ue(v)      |
| if( num_ext_blocks ) {                   |            |
| while( !byte_aligned() )                 |            |
| <b>dm_alignment_zero_bit</b>             | f(1)       |
| for( i = 0; i < num_ext_blocks; i ++ ) { |            |
| ext_dm_data_block(i)                     |            |
| }                                        |            |
| }                                        |            |
| }                                        |            |
| while( !byte_aligned() )                 |            |
| <b>dm_alignment_zero_bit</b>             | f(1)       |
| }                                        |            |

3

4

5 **Table C.2: ext\_dm\_data\_block()**

| ext_dm_data_block()     | Descriptor |
|-------------------------|------------|
| <b>ext_block_length</b> | ue(v)      |
| <b>ext_block_level</b>  | u(8)       |

---

|                                                                |  |
|----------------------------------------------------------------|--|
| ext_dm_data_block_payload( ext_block_length, ext_block_level ) |  |
| }                                                              |  |

1

2

**Table C.3: ext\_dm\_data\_block\_payload()**

| ext_dm_data_block_payload( ext_block_length, ext_block_level ) { | Descriptor |
|------------------------------------------------------------------|------------|
| ext_block_len_bits = 8 * ext_block_length                        |            |
| ext_block_use_bits = 0                                           |            |
| if( ext_block_level == 1 ) {                                     |            |
| <b>min_PQ</b>                                                    | u(12)      |
| <b>max_PQ</b>                                                    | u(12)      |
| <b>avg_PQ</b>                                                    | u(12)      |
| ext_block_use_bits += 36                                         |            |
| }                                                                |            |
| if( ext_block_level == 2 ) {                                     |            |
| <b>target_max_PQ</b>                                             | u(12)      |
| <b>trim_slope</b>                                                | u(12)      |
| <b>trim_offset</b>                                               | u(12)      |
| <b>trim_power</b>                                                | u(12)      |
| <b>trim_chroma_weight</b>                                        | u(12)      |
| <b>trim_saturation_gain</b>                                      | u(12)      |
| <b>ms_weight</b>                                                 | i(13)      |
| ext_block_use_bits += 85                                         |            |
| }                                                                |            |
| if( ext_block_level == 5 ) {                                     |            |
| <b>active_area_left_offset</b>                                   | u(13)      |
| <b>active_area_right_offset</b>                                  | u(13)      |
| <b>active_area_top_offset</b>                                    | u(13)      |
| <b>active_area_bottom_offset</b>                                 | u(13)      |
| ext_block_use_bits += 52                                         |            |
| }                                                                |            |
| while( ext_block_use_bits++ < ext_block_len_bits )               |            |
| <b>ext_dm_alignment_zero_bit</b>                                 | f(1)       |
| }                                                                |            |

3

4 This clause defines the semantics for DM\_data().

---

1 For the purposes of the present clause, the following mathematical functions apply:

2  $\text{Abs}(x) = \begin{cases} x & ; \quad x \geq 0 \\ -x & ; \quad x < 0 \end{cases}$

3 Floor( x ) is the largest integer less than or equal to x.

4  $\text{Sign}(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x = 0 \\ -1 & ; \quad x < 0 \end{cases}$

5  $\text{Clip3}(x, y, z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad \text{otherwise} \end{cases}$

6 Round(x) = Sign(x)\*Floor(Abs(x)+0.5)

7 / = Integer division with truncation of the result toward zero. For example, 7/4 and -7/-4 are truncated to 1 and  
8 -7/4 and 7/-4 are truncated to -1.

9

10 **app\_identifier** identifies an application in the ST 2094 suite.

11 **app\_version** specifies the application version in the application in the ST 2094 suite.

12 **metadata\_refresh\_flag** when set equal to 1 cancels the persistence of any previous extended display mapping metadata in output order and indicates that extended display mapping metadata follows. The extended display mapping metadata persists from the coded picture to which the SEI message containing DM\_data() is associated (inclusive) to the coded picture to which the next SEI message containing DM\_data() and with metadata\_refresh\_flag set equal to 1 in output order is associated (exclusive) or (otherwise) to the last picture in the coded video sequence (inclusive). When set equal to 0 this flag indicates that the extended display mapping metadata does not follow.

19 **num\_ext\_blocks** specifies the number of extended display mapping metadata blocks. The value shall be in the range of 1 to 254, inclusive.

21 **dm\_alignment\_zero\_bit** shall be equal to 0.

22 **ext\_block\_length[ i ]** is used to derive the size of the i-th extended display mapping metadata block payload in bytes. The value shall be in the range of 0 to 1023, inclusive.

24 **ext\_block\_level[ i ]** specifies the level of payload contained in the i-th extended display mapping metadata block. The value shall be in the range of 0 to 255, inclusive. The corresponding extended display mapping metadata block types are defined in Table E.1.4. Values of ext\_block\_level[ i ] that are ATSC reserved shall not be present in the bitstreams conforming to this version of ATSC specification. Blocks using ATSC reserved values shall be ignored.

29 When the value of ext\_block\_level[ i ] is set equal to 1, the value of ext\_block\_length[ i ] shall be set equal to 5.

31 When the value of ext\_block\_level[ i ] is set equal to 2, the value of ext\_block\_length[ i ] shall be set equal to 11.

33 When the value of ext\_block\_level[ i ] is set equal to 5, the value of ext\_block\_length[ i ] shall be set equal to 7.

1                   **Table C.8: Definition of extended display mapping metadata block type**

| ext_block_level | extended metadata block type     |
|-----------------|----------------------------------|
| 0               | Reserved                         |
| 1               | Level 1 Metadata – Content Range |
| 2               | Level 2 Metadata – Trim Pass     |
| 3               | Reserved                         |
| 4               | Reserved                         |
| 5               | Level 5 Metadata – Active Area   |
| 6...255         | Reserved                         |

2  
3       When an extended display mapping metadata block with ext\_block\_level equal to 5 is present, the  
4       following constraints shall apply:

- 5       • An extended display mapping metadata block with ext\_block\_level equal to 5 shall be preceded by at least  
6       one extended display mapping metadata block with ext\_block\_level equal to 1 or 2.
- 7       • Between any two extended display mapping metadata blocks with ext\_block\_level equal to 5, there shall be  
8       at least one extended display mapping metadata block with ext\_block\_level equal to 1 or 2.
- 9       • No extended display mapping metadata block with ext\_block\_level equal to 1 or 2 shall be present after the  
10      last extended display mapping metadata block with ext\_block\_level equal to 5
- 11      • The metadata of an extended display mapping metadata block with ext\_block\_level equal to 1 or 2 shall be  
12      applied to the active area specified by the first extended display mapping metadata block with  
13      ext\_block\_level equal to 5 following this block.

14      When the active area defined by the current extended display mapping metadata block with  
15      ext\_block\_level equal to 5 overlaps with the active area defined by preceding extended display  
16      mapping metadata blocks with ext\_block\_level equal to 5, all metadata of the extended display  
17      mapping metadata blocks with ext\_block\_level equal to 1 or 2 associated with the current ex-  
18      tended display mapping metadata block with ext\_block\_level equal to 5 shall be applied to the  
19      pixel values of the overlapping area.

20      **min\_PQ** specifies the minimum luminance value of the current picture in 12-bit PQ encoding. The  
21      value shall be in the range of 0 to 4095, inclusive. Note that the 12-bit min\_PQ value with full  
22      range is calculated as follows:

$$23 \quad \text{min\_PQ} = \text{Clip3}(0, 4095, \text{Round}(\text{Min} * 4095))$$

24      where Min is MinimumPqencodedMaxrgb as defined in clause 6.1.3 of SMPTE ST 2094-10.

25      **max\_PQ** specifies the maximum luminance value of current picture in 12-bit PQ encoding. The  
26      value shall be in the range of 0 to 4095, inclusive. Note that the 12-bit max\_PQ value with full  
27      range is calculated as follows:

$$28 \quad \text{max\_PQ} = \text{Clip3}(0, 4095, \text{Round}(\text{Max} * 4095))$$

29      where Max is MaximumPqencodedMaxrgb as defined in clause 6.1.5 of SMPTE ST 2094-10.

30      **avg\_PQ** specifies the midpoint luminance value of current picture in 12-bit PQ encoding. The  
31      value shall be in the range of 0 to 4095, inclusive. Note that the 12-bit avg\_PQ value with full  
32      range is calculated as follows:

$$33 \quad \text{avg\_PQ} = \text{Clip3}(0, 4095, \text{Round}(\text{Avg} * 4095))$$

---

1 where Avg is AveragePqencodedMaxrgb as defined in section 6.1.4 of SMPTE ST 2094-10.

2 **target\_max\_PQ** specifies the maximum luminance value of a target display in 12-bit PQ encoding.  
3 The value shall be in the range of 0 to 4095, inclusive. The target\_max\_PQ is the PQ encoded  
4 value of TargetedSystemDisplayMaximumLuminance as defined in clause 10.4 of SMPTE ST  
5 2094-1.

6 If there is more than one extended display mapping metadata block with ext\_block\_level equal to  
7 2, those blocks shall have no duplicated target\_max\_PQ.

8 **trim\_slope** specifies the slope metadata. The value shall be in the range of 0 to 4095, inclusive. If  
9 trim\_slope is not present, it shall be inferred to be 2048. Note that the 12-bit slope value is calcu-  
10 lated as follows:

$$11 \quad trim\_slope = Clip3(0, 4095, Round((S-0.5) * 4096))$$

12 where  $S$  is the ToneMappingGain as defined in clause 6.2.3 of SMPTE ST 2094-10.

13 **trim\_offset** specifies the offset metadata. The value shall be in the range of 0 to 4095, inclusive.  
14 If trim\_offset is not present, it shall be inferred to be 2048. Note that the 12-bit offset value is  
15 calculated as follows:

$$16 \quad trim\_offset = Clip3(0, 4095, Round((O+0.5) * 4096))$$

17 where  $O$  is the ToneMappingOffset as defined in clause 6.2.2 of SMPTE ST 2094-10.

18 **trim\_power** specifies the power metadata. The value shall be in the range of 0 to 4095, inclusive.  
19 If trim\_power is not present, it shall be inferred to be 2048. Note that the 12-bit power value is  
20 calculated as follows:

$$21 \quad trim\_power = Clip3(0, 4095, Round((P-0.5) * 4096))$$

22 where  $P$  is the ToneMappingGamma as defined in clause 6.2.4 of SMPTE ST 2094-10.

23 **trim\_chroma\_weight** specifies the chroma weight metadata. The value shall be in the range of 0  
24 to 4095, inclusive. If trim\_chroma\_weight is not present, it shall be inferred to be 2048. Note that  
25 the 12-bit chroma weight value is calculated as follows:

$$26 \quad trim\_chroma\_weight = Clip3(0, 4095, Round((CW+0.5) * 4096))$$

27 where  $CW$  is the ChromaCompensationWeight as defined in clause 6.3.1 of SMPTE ST 2094-10.

28 **trim\_saturation\_gain** specifies the saturation gain metadata. The value shall be in the range of 0  
29 to 4095, inclusive. If trim\_saturation\_gain is not present, it shall be inferred to be 2048. Note that  
30 the 12-bit saturation gain value is calculated as follows:

$$31 \quad trim\_saturation\_gain = Clip3(0, 4095, Round((SG+0.5) * 4096))$$

32 where  $SG$  is the SaturationGain as defined in clause 6.3.2 of SMPTE ST 2094-10.

33 **ms\_weight** specifies the multiscale weight metadata. The value shall be in the range of -1 to 4095,  
34 inclusive. If ms\_weight is not present, it shall be inferred to be 2048. Where ms\_weight is equal  
35 to -1, the bit stream indicates ms\_weight is unspecified. The 13-bit multiscale weight value is  
36 calculated as follows:

$$37 \quad ms\_weight = -1 \text{ OR } Clip3(0, 4095, Round(MS * 4096))$$

---

1 where *MS* is the ToneDetailFactor as defined in clause 6.4.2 of SMPTE ST 2094-10.  
 2 **active\_area\_left\_offset**, **active\_area\_right\_offset**, **active\_area\_top\_offset**, **active\_area\_bottom\_offset** specify the active area of current picture, in terms of a rectangular region specified in  
 3 picture coordinates for active area. The values shall be in the range of 0 to 8191, inclusive. See  
 4 also UpperLeftCorner and LowerRightCorner definitions in ST 2094-1.  
 5  
 6 If **active\_area\_left\_offset**, **active\_area\_right\_offset**, **active\_area\_top\_offset**, **active\_area\_bottom\_offset** are not present, they shall be inferred to be 0.  
 7  
 8 The coordinates of top left active pixel is derived as follows:  
 9                    $X_{top\_left} = active\_area\_left\_offset$   
 10                   $Y_{top\_left} = active\_area\_top\_offset$   
 11 The coordinates of top left active pixel are defined as the UpperLeftCorner in clause 9.2 of SMPTE  
 12 ST.2094-1.  
 13 With Xsize is the horizontal resolution of the current picture and Ysize is the vertical resolution of  
 14 current picture, the coordinates of bottom right active pixel are derived as follows:  
 15                    $X_{bottom\_right} = XSize - 1 - active\_area\_right\_offset$   
 16                   $Y_{bottom\_right} = YSize - 1 - active\_area\_bottom\_offset$   
 17 where  $X_{bottom\_right}$  greater than  $X_{top\_left}$  and  $Y_{bottom\_right}$  greater than  $Y_{top\_left}$ .  
 18 The coordinates of bottom right active pixel are defined as the LowerRightCorner in clause 9.3 of  
 19 SMPTE ST.2094-1.  
 20 **ext\_dm\_alignment\_zero\_bit** shall be equal to 0.  
 21

## 22 Annex F Composing Metadata Message

### 23 F.1 Introduction

24 A composing metadata (CM) message contains the metadata which is needed to apply the post-  
 25 processing process as described in the ETSI [ETCCM] specification to recreate the HDR UHDTV  
 26 pictures.

### 27 F.2 Syntax and Semantics

28 The syntax for CM\_data() is shown in table D.1. The number of bits “v” used to represent each of the syntax  
 29 elements of CM\_data(), for which the parsing process is specified by the descriptor u(v), is defined in table D.2.  
 30  
 31

**Table D.1: CM\_data()**

| CM_data()                     | Descriptor |
|-------------------------------|------------|
| <b>ccm_profile</b>            | u(4)       |
| <b>ccm_level</b>              | u(4)       |
| <b>coefficient_log2_denom</b> | ue(v)      |

|                                                                                 |       |
|---------------------------------------------------------------------------------|-------|
| <b>BL_bit_depth_minus8</b>                                                      | ue(v) |
| <b>EL_bit_depth_minus8</b>                                                      | ue(v) |
| <b>hdr_bit_depth_minus8</b>                                                     | ue(v) |
| <b>disable_residual_flag</b>                                                    | u(1)  |
| for( cmp = 0; cmp < 3; cmp++ ) {                                                |       |
| <b>num_pivots_minus2[ cmp ]</b>                                                 | ue(v) |
| for ( pivot_idx = 0; pivot_idx < num_pivots_minus2[ cmp ] + 2; pivot_idx ++ ) { |       |
| <b>pred_pivot_value[ cmp ][ pivot_idx ]</b>                                     | u(v)  |
| } // end of pivot points for BL three components                                |       |
| } //cmp                                                                         |       |
| for ( cmp = 0; cmp < 3; cmp++ ) { //mapping parameters                          |       |
| for ( pivot_idx = 0; pivot_idx < num_pivots_minus2[ cmp ] + 1; pivot_idx++ ) {  |       |
| <b>mapping_idc[ cmp ][ pivot_idx ]</b>                                          | ue(v) |
| if( mapping_idc [ cmp ][ pivot_idx ] == MAPPING_POLYNOMIAL ) {                  |       |
| <b>poly_order_minus1[ cmp ][ pivot_idx ]</b>                                    | ue(v) |
| for( i = 0 ; i <= poly_order_minus1[ cmp ][ pivot_idx ] + 1; i ++ ) {           |       |
| <b>poly_coef_int[ cmp ][ pivot_idx ][ i ]</b>                                   | se(v) |
| <b>poly_coef[ cmp ][ pivot_idx ][ i ]</b>                                       | u(v)  |
| }                                                                               |       |
| else if( mapping_idc [ cmp ][ pivot_idx ] == MAPPING_MMR ) {                    |       |
| <b>mmr_order_minus1[ cmp ][ pivot_idx ]</b>                                     | u(2)  |
| <b>mmr_constant_int[ cmp ][ pivot_idx ]</b>                                     | se(v) |
| <b>mmr_constant[ cmp ][ pivot_idx ]</b>                                         | u(v)  |
| for( i = 1; i <= mmr_order_minus1 + 1; i ++ ) {                                 |       |
| for (j = 0; j < 7; j++) {                                                       |       |
| <b>mmr_coef_int[ cmp ][ pivot_idx ][ i ][ j ]</b>                               | se(v) |
| <b>mmr_coef[ cmp ][ pivot_idx ][ i ][ j ]</b>                                   | u(v)  |
| } // the j-th coefficients                                                      |       |
| } // the i-th order                                                             |       |
| }                                                                               |       |
| } // pivot_idx                                                                  |       |
| } // cmp                                                                        |       |
| if ( !disable_residual_flag ) {                                                 |       |
| for ( cmp = 0; cmp < 3; cmp++ ) { //quantization parameters                     |       |
| <b>nlq_offset[ cmp ]</b>                                                        | u(v)  |
| <b>hdr_in_max_int[ cmp ]</b>                                                    | ue(v) |
| <b>hdr_in_max[ cmp ]</b>                                                        | u(v)  |
| <b>linear_deadzone_slope_int[ cmp ]</b>                                         | ue(v) |
| <b>linear_deadzone_slope[ cmp ]</b>                                             | u(v)  |
| <b>linear_deadzone_threshold_int[ cmp ]</b>                                     | ue(v) |
| <b>linear_deadzone_threshold[ cmp ]</b>                                         | u(v)  |

---

|                              |                         |
|------------------------------|-------------------------|
| }                            | // cmp                  |
| }                            | // disable_residue_flag |
| while( !byte_aligned() )     |                         |
| <b>cm_alignment_zero_bit</b> | f(1)                    |
| }                            |                         |
|                              |                         |

1  
2

**Table D.2: Specification of number of bits “v” for CM\_data() syntax elements with descriptor u(v)**

| Syntax element            | Number of bits “v”      |
|---------------------------|-------------------------|
| pred_pivot_value          | EL_bit_depth_minus8 + 8 |
| poly_coef                 | coefficient_log2_denom  |
| mmr_constant              | coefficient_log2_denom  |
| mmr_coef                  | coefficient_log2_denom  |
| nlq_offset                | EL_bit_depth_minus8 + 8 |
| hdr_in_max                | coefficient_log2_denom  |
| linear_deadzone_slope     | coefficient_log2_denom  |
| linear_deadzone_threshold | coefficient_log2_denom  |

3

4     The definitions of the header parameter values are contained in [ETCCM], Section 5.3.2, “CM  
5     Header Parameter Definitions”.

6     The definitions of the mapping parameter values are contained in [ETCCM], Section 5.3.3, “CM  
7     Mapping Parameter Definitions”.

8     Parameter **cm\_alignment\_zero\_bit** shall be equal to 0.

9

## 10           Annex G Sample Dual-layer MPD

11     Below is an example dual-layer MPD, with dual adaptation sets – both a Base layer and an  
12     Enhancement Layer. Items of note are highlighted:

```
14 <Period>
15 <!-- Video -->
16 <AdaptationSet subsegmentAlignment="true"
17 subsegmentStartsWithSAP="1" frameRate="24000/1001">
18 <Representation mimeType="video/mp4" codecs="
19 hvc1.2.100000000.L150.B0" id="base-layer"
20 bandwidth="14156144" width="3840" height="2160">
21 <BaseURL>BL_dual_track_BC.mp4</BaseURL>
22 <SegmentBase indexRange="795-1210">
23 <Initialization range="0-794"/>
24 </SegmentBase>
25 </Representation>
```

---

```
1 <Representation mimeType="video/mp4" codecs="dvhe.dtr"
2 id="enhancement-layer"
3 dependencyId="base-layer" bandwidth="3466528"
4 width="1920" height="1080">
5 <BaseURL>EL_dual_track_BC.mp4</BaseURL>
6 <SegmentBase indexRange="704-1119">
7 <Initialization range="0-703"/>
8 </SegmentBase>
9 </Representation>
10 </AdaptationSet>
11 <!-- Audio -->
12 <AdaptationSet mimeType="audio/mp4" codecs="ec-3" lang="und"
13 subsegmentAlignment="true" subsegmentStartsWithSAP="1">
14 <Representation id="2" bandwidth="192000">
15 <AudioChannelConfiguration
16
17 schemeIdUri="tag:dolby.com,2014:dash:audio_channel_configuration:20
18 11" value="F801"/>
19 <BaseURL>audio.mp4</BaseURL>
20 <SegmentBase indexRange="652-875">
21 <Initialization range="0-651"/>
22 </SegmentBase>
23 </Representation>
24 </AdaptationSet>
25 </Period>
26 </MPD>
27
```