INTERNATIONAL STANDARD

ISO/IEC 23009-1

First edition
2012-04-01

Information technology — Dynamic adaptive streaming over HTTP (DASH) —

Part 1: Media presentation description and segment formats

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Technologies de l'information — Diffusion en flux adaptatif dynamique sur HTTP (DASH) —

Partie 1: Description de la présentation et formats de remise des médias

ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

Reference number
ISO/IEC 23009-1:2012(E)

© ISO/IEC 2012

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2012

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

 © ISO/IEC 2012 – All rights reserved

Contents

Page

Foreword v

Introduction vi

1 Scope 1

2 Normative references 1

3 Terms, definitions, symbols and abbreviated terms 2
3.1 Terms and definitions 2
3.2 Symbols and abbreviated terms 5
3.3 Conventions 6

4 Introduction 6
4.1 System description 6
4.2 DASH client model 7
4.3 DASH data model overview 8
4.4 Protocols 11
4.5 Media Stream and Representation properties 11
4.6 Brands 13
4.7 Schemes 13

5 Media Presentation 14
5.1 General 14
5.2 Media Presentation Description 14
5.3 Hierarchical data model 16
5.4 Media Presentation Description updates 54
5.5 MPD assembly 55
5.6 Base URL Processing 57
5.7 Program information 58
5.8 Descriptors 59
5.9 DASH metrics descriptor 65

6 Segment formats 67
6.1 Introduction 67
6.2 Segment types 67
6.3 Segment formats for ISO base media file format 69
6.4 Segment formats for MPEG-2 transport streams 72

7 Combined semantics of MPD and Segment formats 78
7.1 Introduction 78
7.2 General 79
7.3 Media Presentation based on the ISO base media file format 80
7.4 Media Presentation based on MPEG-2 TS 82

8 Profiles 84
8.1 Definition 84
8.2 Full profile 85
8.3 ISO Base media file format On Demand profile 85
8.4 ISO Base media file format live profile 87
8.5 ISO Base media file format main profile 88
8.6 MPEG-2 TS main profile 89
8.7 MPEG-2 TS simple profile 90

Annex A (informative) Example DASH client behaviour 92
A.1 Introduction 92
A.2 Overview 92

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/6e022d63-ac24-4d89-9319-51c97e3d652b/iso-iec-23009-1-2012

© ISO/IEC 2012 – All rights reserved

A.3 Segment list generation 93
A.4 Seeking 96
A.5 Support for trick modes 97
A.6 Switching Representations 97
A.7 Reaction to error codes 98
A.8 Encoder clock drift control 98

Annex B (normative) MPD schema 99

Annex C (normative) MIME type registration for MPD 105
C.1 Introduction 105
C.2 MIME type and subtype 105
C.3 Parameters 106

Annex D (normative) DASH Metrics 107
D.1 Introduction 107
D.2 DASH-Metrics client reference model 107
D.3 Definition of observation points 107
D.4 Semantics of the DASH metrics 108

Annex E (normative) Byte range requests with regular HTTP GET methods 114
E.1 Background 114
E.2 Construction rule 114
E.3 Examples 115

Annex F (informative) Guidelines for extending DASH with other delivery formats 116
F.1 Adding delivery formats to DASH 116
F.2 Media Presentation authoring rules 116

Annex G (informative) MPD Examples and MPD Usage 117
G.1 Example MPD for ISO Base media file format On Demand profile 117
G.2 Example for ISO Base media file format Live profile 118
G.3 Example for MPEG-2 TS Live profile 119
G.4 Example for multiple stereo views 120
G.5 Example for SVC alternative streams 121
G.6 Example for trick play support 122
G.7 Example for content protected by multiple schemes 123
G.8 Example for usage of Role descriptor 124

Bibliography 126

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

 © ISO/IEC 2012 – All rights reserved

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 23009-1 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC 23009 consists of the following parts, under the general title *Information technology — Dynamic adaptive streaming over HTTP (DASH)*:

ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

— *Part 1: Media presentation description and segment formats*

© ISO/IEC 2012 – All rights reserved

Introduction

Dynamic Adaptive Streaming over HTTP (DASH) is intended to support a media-streaming model for delivery of media content in which control lies exclusively with the client. Clients may request data using the HTTP protocol from standard web servers that have no DASH-specific capabilities. Consequently, this part of ISO/IEC 23009 focuses not on client or server procedures but on the data formats used to provide a DASH Media Presentation.

This part of ISO/IEC 23009 primarily specifies formats for the Media Presentation Description and Segments. It is applicable to streaming services over the Internet.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

© ISO/IEC 2012 – All rights reserved

Information technology — Dynamic adaptive streaming over HTTP (DASH) —

Part 1:
Media presentation description and segment formats

1 Scope

This part of ISO/IEC 23009 primarily specifies formats for the Media Presentation Description and Segments for dynamic adaptive streaming delivery of MPEG media over HTTP. It is applicable to streaming services over the Internet.

2 Normative references

iTeh STANDARD PREVIEW
(standards.iteh.ai)

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ITU-T Rec. H.222.0 | ISO/IEC 13818-1, *Information technology – Generic coding of moving pictures and associated audio information: Systems*

ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

ISO/IEC 14496-10, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*

ISO/IEC 14496-12, *Information technology – Coding of audio-visual objects – Part 12: ISO base media file format (technically identical to ISO/IEC 15444-12)*

ISO/IEC 23003-3, *Information technology – MPEG audio technologies – Part 3: Unified speech and audio coding*

IETF RFC 1521, *MIME (Multipurpose Internet Mail Extensions) Part One: Mechanisms for Specifying and Describing the Format of Internet Message Bodies*, September 1993

IETF RFC 1738, *Uniform Resource Locators (URL)*, December 1994

IETF RFC 2141, *URN Syntax*, May 1997

IETF RFC 2616, *Hypertext Transfer Protocol – HTTP/1.1*, June 1999

IETF RFC 3023, *XML Media Types*, January 2001

IETF RFC 3406, *Uniform Resource Names (URN) Namespace Definition Mechanisms*, October 2002

IETF RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax*, January 2005

IETF RFC 4122, *A Universally Unique IDentifier (UUID) URN Namespace*, July 2005

IETF RFC 4337, *MIME Type Registration for MPEG-4*, March 2006

© ISO/IEC 2012 – All rights reserved

IETF RFC 5646, *Tags for Identifying Languages*, September 2009

IETF RFC 6381, *The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types*, August 2011

W3C XLINK *XML Linking Language (XLink) Version 1.1*, W3C Recommendation 06, May 2010

3 Terms, definitions, symbols and abbreviated terms

3.1 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1.1
access unit
unit of a media stream with an assigned Media Presentation time

3.1.2
accessibility
degree to which a media content or certain media content components are available to as many people as possible

3.1.3
Adaptation Set
set of interchangeable encoded versions of one or several media content components

iTeh STANDARD PREVIEW
(standards.iteh.ai)

3.1.4
available Segment
Segment that is accessible at its assigned HTTP-URL and a possibly assigned byte range, that is the request with an HTTP GET results in a reply of the Segment and 2xx status code

ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

3.1.5
Bitstream Switching Segment
Segment that if present contains essential data to switch to the Representation it is assigned to

3.1.6
complementary Representation
Representation which complements at least one dependent Representation

3.1.7
continuous media
media with an inherent notion of time, for example, speech, audio, video, timed text or timed metadata

3.1.8
DASH metric
metric identified by a key and defined in this part of ISO/IEC 23009

3.1.9
dependent Representation
Representation for which Segments from its complementary Representations are necessary for presentation and/or decoding of the contained media content components

3.1.10
earliest presentation time
smallest presentation time of any access unit of a Media Segment or Subsegment for a media stream

3.1.11
group
collection of Adaptation Sets that are not expected to be presented simultaneously

 © ISO/IEC 2012 – All rights reserved

3.1.12
HTTP-URL
URL with a fixed scheme of “http” or “https”

3.1.13
Index Segment
Segment that primarily contains indexing information for Media Segments

3.1.14
Initialization Segment
Segment containing metadata that is necessary to present the media streams encapsulated in Media Segments

3.1.15
media content
one media content period or a contiguous sequence of media content periods

3.1.16
media content component
one continuous component of the media content with an assigned media component type that can be encoded individually into a media stream

3.1.17
media content component type
a single type of media content such as audio, video, or text

iTeh STANDARD PREVIEW
(standards.iteh.ai)

3.1.18
media content period
set of media content components that have a common timeline as well as relationships on how they can be presented

ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

3.1.19
Media Presentation
collection of data that establishes a bounded or unbounded presentation of media content

3.1.20
Media Presentation Description
MPD
formalized description for a Media Presentation for the purpose of providing a streaming service

3.1.21
Media Presentation timeline
concatenation of the timeline of all Periods which itself is common to all Representations in the Period

3.1.22
Media Segment
Segment that complies with media format in use and enables playback when combined with zero or more preceding segments, and an Initialization Segment (if any)

3.1.23
media stream
encoded version of a media content component

3.1.24
Media Subsegment
Subsegment that only contains media data but no Segment Index

© ISO/IEC 2012 – All rights reserved

3.1.25
MPD start time
approximate presentation start time of a Media Segment signalled in MPD

3.1.26
MPD duration
approximate presentation duration of a Media Segment signalled in MPD

3.1.27
Period
interval of the Media Presentation, where a contiguous sequence of all Periods constitutes the Media Presentation

3.1.28
presentation time
time associated to an access unit that maps it to the Media Presentation timeline

3.1.29
remote element
element that is not fullly contained in the MPD document but is referenced in the MPD with an HTTP-URL

3.1.30
Representation
collection and encapsulation of one or more media streams in a delivery format and associated with descriptive metadata

iTeh STANDARD PREVIEW
(standards.iteh.ai)

3.1.31
Segment
unit of data associated with an HTTP-URL and optionally a byte range that are specified by an MPD

ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

3.1.32
Segment availability start time
time instant in wall-clock time at which a Segment becomes an available Segment

3.1.33
Segment availability end time
time instant in wall-clock time at which a Segment ceases to be an available Segment

3.1.34
Segment Index
compact index of the time range to byte range mapping within a Media Segment separately from the MPD

3.1.35
stream access point
SAP
position in a Representation enabling playback of a media stream to be started using only the information contained in Representation data starting from that position onwards (preceded by initializing data in the Initialization Segment, if any)

3.1.36
Sub-Representation
part of a Representation described in the MPD that is present in the entire Period

3.1.37
Subsegment
unit within Media Segments that is indexed by a Segment Index

© ISO/IEC 2012 – All rights reserved

3.1.38
valid Segment URL
HTTP-URL that is promised to reference a Segment during its Segment availability period

3.1.39
wall-clock time
time as stated by UTC

3.2 Symbols and abbreviated terms

For the purposes of this document, the following symbols and abbreviated terms apply.

AVC	advanced video coding
CAT	conditional access table
DASH	dynamic adaptive streaming over HTTP
DM	DASH Metrics
DRM	digital rights management
ECM	entitlement control message
HTTP	hypertext transfer protocol
IDR	instantaneous decoding refresh
ISOBMFF	ISO base media file format
MPD	Media Presentation Description
MVC	multi-view video coding
PAT	program association table
PCR	program clock reference
PES	packetized elementary stream
PID	packet identifier
PMT	program map table
PSI	program specific information
PTS	presentation time stamp
SAP	stream access point
SEI	supplementary enhancement information
SVC	scalable video coding
TCP	transmission control protocol
TLS	transport layer security

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

© ISO/IEC 2012 – All rights reserved

TS	transport stream
URI	uniform resource identifier
URL	uniform resource locator
URN	uniform resource name
UTC	coordinated universal time
UUID	universally unique identifier
XML	extensible mark-up language

3.3 Conventions

The following naming conventions apply in this document.

— Elements in an XML document are identified by an upper-case first letter and in bold face as **`Element`**. To express that an element **`Element1`** is contained in another element **`Element2`**, we may write **`Element2.Element1`**. If an element's name consists of two or more combined words, camel-casing is typically used, e.g. **`ImportantElement`**. Elements may be present either exactly once, or the minimum and maximum occurence is defined by <minOccurs> ... <maxOccurs>.

— Attributes in an XML document are identified by a lower-case first letter as well as they are preceded by a '`@`'-sign, e.g. `@attribute`. To point to a specific attribute `@attribute` contained in an element **`Element`**, one may write **`Element`**`@attribute`. If an attribute's name consists of two or more combined words, camel-casing is typically used after the first word, e.g. `@veryImportantAttribute`. Attributes may have assigned a status in the XML as mandatory (M), optional (O), optional with default value (OD) and conditionally mandatory (CM).

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

— Namespace qualification of elements and attributes is used as per XML standards, in the form of **`namespace:Element`** or `@namespace:attribute` The fully qualified namespace will be provided in the schema fragment associated with the declaration.

— Variables defined in the context of this document are specifically highlighted with *italics,* e.g. *InternalVariable*.

— Structures that are defined as part of the hierarchical data model are identified by an upper-case first letter, e.g. Period, Adaptation Set, Representation, Segment, etc.

— The term "this clause" refers to the entire clause included within the same first heading number. The term "this subclause" refers to all text contained in the subclause with the lowest hierarchy heading.

4 Introduction

4.1 System description

Dynamic Adaptive Streaming over HTTP (DASH) specifies XML and binary formats that enable delivery of media content from standard HTTP servers to HTTP clients and enable caching of content by standard HTTP caches.

This part of ISO/IEC 23009 primarily defines two formats:

— The Media Presentation Description (MPD) describes a *Media Presentation*, i.e. a bounded or unbounded presentation of media content. In particular, it defines formats to announce resource

 © ISO/IEC 2012 – All rights reserved

identifiers for *Segments* and to provide the context for these identified resources within a Media Presentation. These resource identifiers are HTTP-URLs possibly combined with a byte range.

— The Segment formats specify the formats of the entity body of the HTTP response to an HTTP GET request or a partial HTTP GET with the indicated byte range using HTTP/1.1 as defined in RFC 2616 to a resource identified in the MPD. Segments typically contain efficiently coded media data and metadata according to or aligned with common media formats.

The MPD provides sufficient information for a client to provide a streaming service to the user by accessing the Segments through the protocol specified in the scheme of the defined resources. In the context of this part of ISO/IEC 23009 the assumed protocol is HTTP/1.1. Such a client is referred to as a DASH Client in the remainder of 23009-1. However, this part of ISO/IEC 23009 does not provide a normative specification for such a client.

Figure 1 shows a possible deployment architecture in which the formats defined in this part of ISO/IEC 23009 may be used. Boxes with solid lines indicate devices that are mentioned in this specification as they host or process the formats defined in this specification whereas dashed boxes are conceptual or transparent. This part of ISO/IEC 23009 deals with the definition of formats that are accessible on the interface to the DASH Client, indicated by the solid lines. Any other formats or interfaces are not in scope of this Part of ISO/IEC 23009. In the considered deployment scenario, it is assumed that the DASH Client has access to an MPD. The MPD provides sufficient information for the DASH Client to provide a streaming service to the user by requesting Segments from an HTTP server and demultiplexing, decoding and rendering the included media streams.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

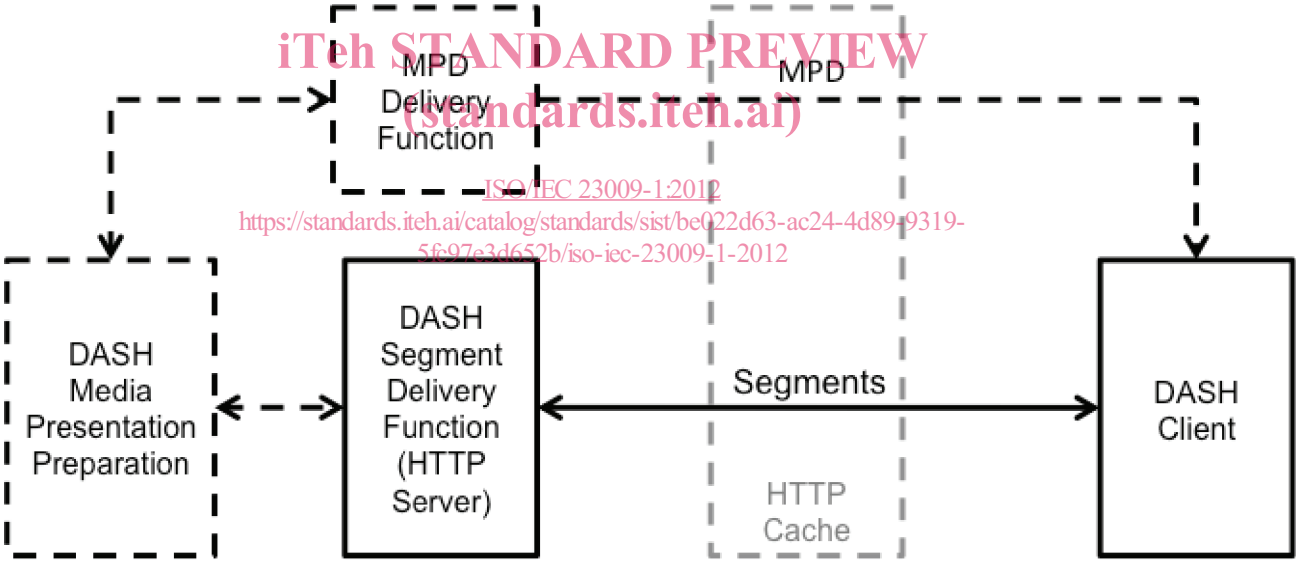


Figure 1 — Example system for DASH formats

4.2 DASH client model

The design of the formats defined in this part of ISO/IEC 23009 is based on the informative client model as shown in Figure 2. The figure illustrates the logical components of a conceptual DASH client model. In this figure the DASH Access Engine receives the Media Presentation Description (MPD), constructs and issues requests and receives Segments or parts of Segments. In the context of this part of ISO/IEC 23009, the output of the DASH Access Engine consists of media in MPEG container formats (ISO/IEC 14496-12 ISO Base Media File Format or ISO/IEC 13818-1 MPEG-2 Transport Stream), or parts thereof, together with timing information that maps the internal timing of the media to the timeline of the Media Presentation. In Annex F of this part of ISO/IEC 23009, guidance on enabling the use of this part of ISO/IEC 23009 with other container formats is provided.

© ISO/IEC 2012 – All rights reserved

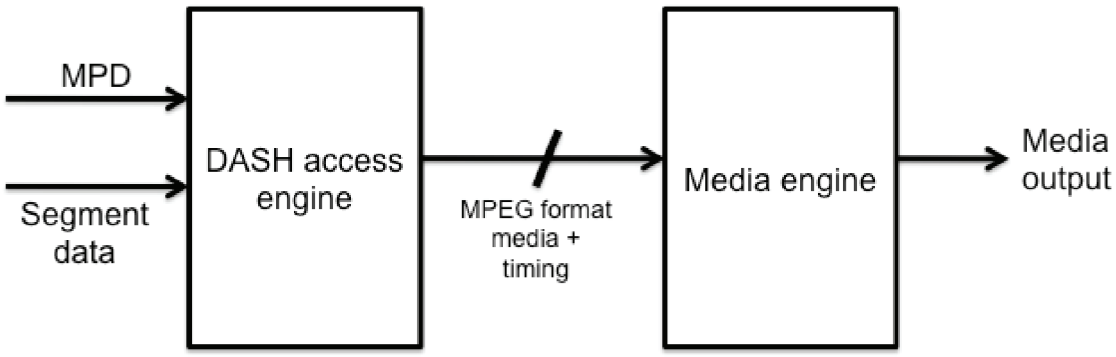


Figure 2 — DASH Client Model

4.3 DASH data model overview

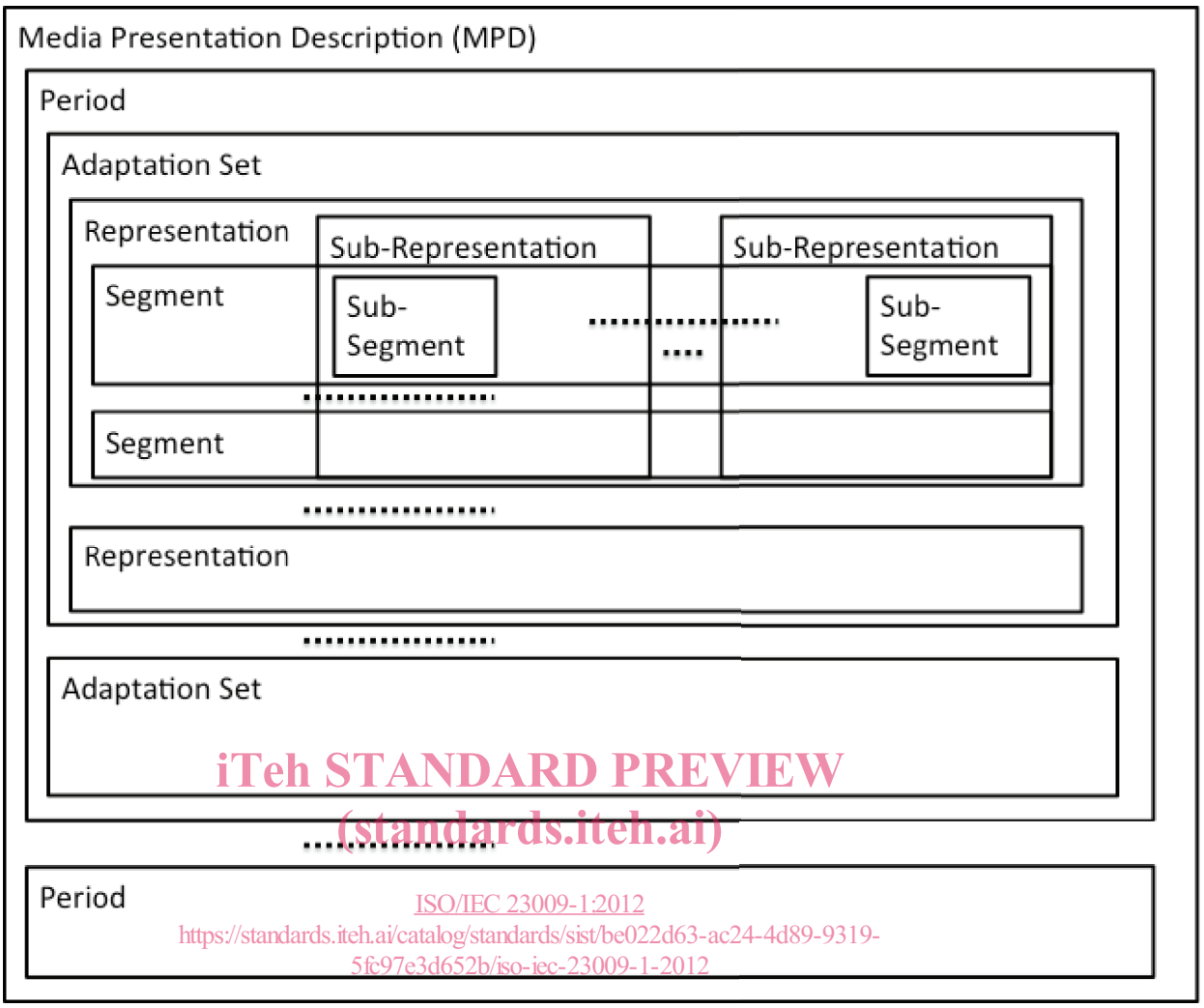
DASH is intended to support a media-streaming model for delivery of media content in which control lies exclusively with the client. Clients may request data using the HTTP protocol from standard web servers that have no DASH-specific capabilities. Consequently, this standard focuses not on client or server procedures but on the data formats used to provide a DASH Media Presentation.

The collection of encoded and deliverable versions of media content and the appropriate description of these form a Media Presentation. Media content is composed of a single or multiple contiguous media content **periods** in time. Each media content period is composed of one or multiple **media content components,** for example audio components in various languages and a video component. Each media content component has an assigned **media content component type**, for example audio or video.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97a3d652b/iso-iec-23009-1-2012

Each media content component may have several encoded versions, referred to as **media streams**. Each media stream inherits the properties of the media content, the media content period, the media content component from which it was encoded and in addition it gets assigned the properties of the encoding process such as sub-sampling, codec parameters, encoding bitrate, etc. This describing metadata is relevant for static and dynamic selection of media content components and media streams.

 © ISO/IEC 2012 – All rights reserved



iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC 23009-1:2012
https://standards.iteh.ai/catalog/standards/sist/be022d63-ac24-4d89-9319-5fc97e3d652b/iso-iec-23009-1-2012

Figure 3 — DASH High-Level Data Model

DASH is based on a hierarchical data model aligned with the presentation in Figure 3. A DASH Media Presentation is described by a **Media Presentation Description** document. This describes the sequence of **Periods** (see 5.3.2) in time that make up the Media Presentation. A Period typically represents a media content period during which a consistent set of encoded versions of the media content is available i.e. the set of available bitrates, languages, captions, subtitles etc. does not change during a Period.

Within a Period, material is arranged into **Adaptation Sets** (see 5.3.3). An Adaptation Set represents a set of interchangeable encoded versions of one or several media content components (see 5.3.4). For example there may be one Adaptation Set for the main video component and a separate one for the main audio component. If there is other material available, for example captions or audio descriptions, then these may each have a separate Adaptation Set. Material may also be provided in multiplexed form, in which case interchangeable versions of the *multiplex* may be described as a single Adaptation Set, for example an Adaptation Set containing both the main audio and main video for a Period. Each of the multiplexed components may be described individually by a media content component description.

An Adaptation Set contains a set of **Representations** (see 5.3.5). A Representation describes a *deliverable encoded version* of one or several media content components. A Representation includes one or more media streams (one for each media content component in the multiplex). Any single Representation within an Adaptation Set is sufficient to render the contained media content components. Typically, clients may switch from Representation to Representation within an Adaptation Set in order to adapt to network conditions or other factors. Clients may also ignore Representations that rely on codecs or other rendering technologies they do not support or that are otherwise unsuitable.

© ISO/IEC 2012 – All rights reserved