



INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.124

(03/2004)

SERIES J: CABLE NETWORKS AND TRANSMISSION
OF TELEVISION, SOUND PROGRAMME AND OTHER
MULTIMEDIA SIGNALS

Interactive systems for digital television distribution

Multiplexing format for multimedia webcasting over TCP/IP networks

ITU-T Recommendation J.124

Copyright International Telecommunication Union/ITU Telecommunication Sector

ITU-T Recommendation J.124

Multiplexing format for multimedia webcasting over TCP/IP networks

Summary

This Recommendation provides an extended multiplexing format based on ITU-T Rec. J.123, which is appropriate for audio and video transmission by download-based protocol over TCP/IP without any session control protocols between server and client, a.k.a. "Progressive Download". Fragmented structure is newly introduced into this Recommendation. In the fragment structure, media data is divided into media fragments, and a movie header is also divided into movie fragment headers according to the fragmented media data. Each movie fragment header corresponds to each media fragment, and these elements constitute a movie fragment. By adapting the fragment structure to long duration content, a huge header, which causes initial delay of the progressive streaming, can be avoided. In addition, formatted text information is stored in the media data so that these can be interleaved with one another in a file. This format also carries metadata, digital rights management (DRM) information as well as audio, video and text bitstreams.

Examples of use of this Recommendation are given in Appendices I, II and III.

Source

ITU-T Recommendation J.124 was approved on 15 March 2004 by ITU-T Study Group 9 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

© ITU 2004

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

CONTENTS

Page

1	Scope	1
2	References.....	1
2.1	Normative references.....	1
2.2	Informative references.....	1
2.3	Bibliography	1
3	Terms and definitions	2
4	Abbreviations.....	2
5	Reference architecture	2
6	Multiplexing format.....	3
6.1	Basic structure	3
6.2	Object structure	3
6.3	Box order	5
6.4	Track structure.....	6
6.5	Media data structure	6
6.6	Other descriptions.....	6
7	Box definitions	6
7.1	File type box	6
7.2	Other boxes.....	7
8	Digital Rights Management (DRM) box	7
8.1	Syntax	8
8.2	Semantics.....	8
9	Timed text format	8
9.1	Unicode support.....	8
9.2	Bytes, characters, and glyphs	9
9.3	Character set support	9
9.4	Font support.....	9
9.5	Fonts and metrics.....	10
9.6	Colour support	10
9.7	Text rendering position and composition	10
9.8	Marquee scrolling.....	12
9.9	Language	13
9.10	Writing direction	13
9.11	Text wrap.....	14
9.12	Highlighting, Closed Caption, and Karaoke.....	14
9.13	Media handler	14
9.14	Media handler header	14
9.15	Style record.....	14
9.16	Sample description format.....	15

	Page
9.17 Sample format	16
9.18 Combinations of features.....	20
Appendix I – Application example: Typical VOD transmission.....	21
Appendix II – Application example: Random access transmission.....	22
Appendix III – Application example: Live video transmission.....	23

ITU-T Recommendation J.124

Multiplexing format for multimedia webcasting over TCP/IP networks

1 Scope

This Recommendation defines a multiplexing format appropriate for progressive download, audio and video transmission by download-based protocol over TCP/IP. In contrast to ITU-T Rec. J.123, this Recommendation supports fragmented structure for long duration content. In addition, formatted text information is stored in the media data so that these can be interleaved with one another in a file. Using this format, webcasting of long duration content and live programs is realized.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

2.1 Normative references

- [1] ITU-T Recommendation J.123 (2002), *Multiplexing format for webcasting on TCP/IP network*.
- [2] ISO/IEC 14496-12:2004, *Information technology – Coding of audio-visual objects – Part 12: ISO base media file format*.
- [3] ISO/IEC 14496-14:2003, *Information technology – Coding of audio-visual objects – Part 14: MP4 file format*.

2.2 Informative references

- [4] ITU-T Recommendation J.120 (2000), *Distribution of sound and television programs over the IP network*.
- [5] ISO/IEC 14496-2:2001, *Information technology – Coding of audio-visual objects – Part 2: Visual*.
- [6] ISO/IEC 14496-3:2001, *Information technology – Coding of audio-visual objects – Part 3: Audio*.
- [7] IETF RFC 2068 (1997), *Hypertext Transfer Protocol – HTTP/1.1*.

2.3 Bibliography

- [8] 3GPP TS 26.245:2003, *Transparent end-to-end streaming service; Timed text format*.

3 Terms and definitions

This Recommendation defines the following terms:

- 3.1 box:** An object-oriented building block defined by a unique type identifier and length [2].
- 3.2 chunk:** A contiguous set of samples for one track.
- 3.3 container box:** A box whose sole purpose is to contain and group a set of related boxes.
- 3.4 movie box:** A container box whose sub-boxes define the metadata for a presentation ('moov').
- 3.5 media data box:** A container box which can hold the actual media data for a presentation ('mdat').
- 3.6 presentation:** One or more motion sequences, possibly combined with audio.
- 3.7 progressive download:** Streaming by download-based protocol over TCP/IP without any session control protocols. Client can start playing the media before the full file is downloaded.
- 3.8 sample:** An individual frame of video, or a time-contiguous compressed section of audio.
- 3.9 sample description:** A structure which defines and describes the format of some number of samples in a track.
- 3.10 sample table:** A packed directory for the timing and physical layout of the samples in a track.
- 3.11 track:** A collection of related samples, which corresponds to a sequence of images or sampled audio.
- 3.12 webcasting:** Webcasting is defined in ITU-T Rec. J.120, "*Distribution of sound and television programs over the IP network*".

4 Abbreviations

This Recommendation uses the following abbreviations:

DRM	Digital Rights Management
HTTP	HyperText Transport Protocol
IP	Internet Protocol
MP4	MPEG-4 File Format
SMIL	Synchronized Multimedia Integration Language
TCP	Transmission Control Protocol
UTF-8	Unicode Transformation Format (the 8-bit form)
UTF-16	Unicode Transformation Format (the 16-bit form)
UUID	Universal Unique Identifier

5 Reference architecture

This Recommendation assumes that download-based protocol (e.g., HTTP) should be used for multimedia webcasting because it does not require any complex server-client protocols.

The reference architecture for multimedia webcasting on TCP/IP is shown in Figure 5-1.

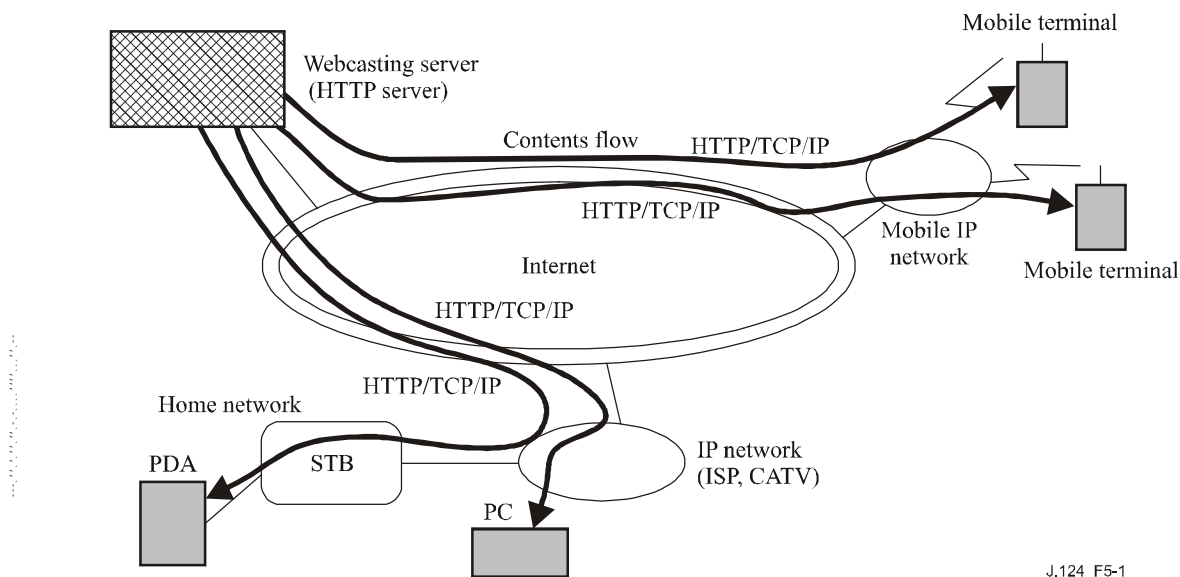


Figure 5-1/J.124 – Architecture of multimedia webcasting on TCP/IP networks

6 Multiplexing format

6.1 Basic structure

The format is structurally based on the ISO base media file format defined in [2]. Basic structure of the format is shown in Figure 6-1, which consists of extension data, contents header and media data.

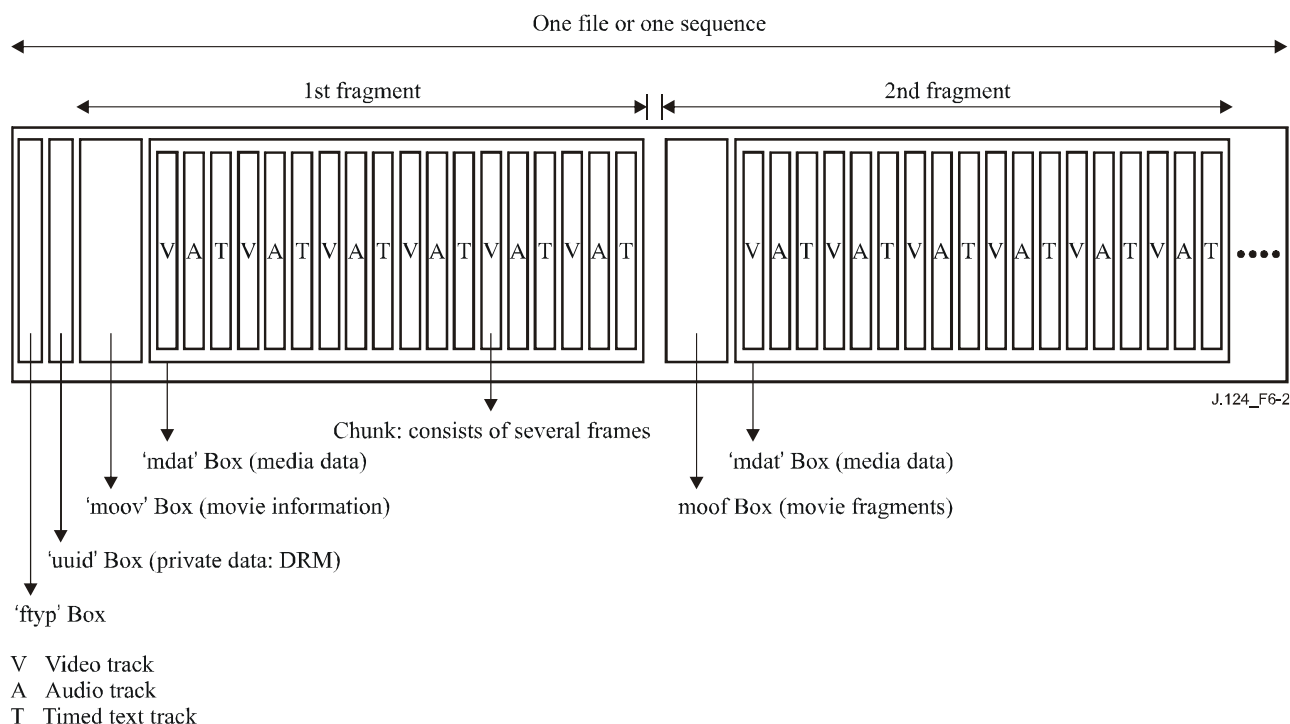


Figure 6-1/J.124 – Basic structure of a file format

6.2 Object structure

The file is structured as a sequence of objects called "Box"; some of these objects may contain other objects. The sequence of objects in the file shall contain exactly one presentation metadata wrapper

(the movie box 'moov'). It should be close to the beginning of the file. The other objects found at this level may be file type box 'ftyp', 'uuid' box, movie fragments 'moof', and media data boxes 'mdat'.

All boxes defined in this Recommendation are listed in Table 6-1, and are indicated by grey shading.

Table 6-1/J.124 – Box types and structure

ftyp				file type and compatibility					
uuid				uuid box for DRM (see clause 8)					
moov				container for all the information					
	mvhd				movie header, overall declarations				
	trak				container for an individual track or stream				
		tkhd				track header, overall information about the track			
		tref				track reference container			
		edts				edit list container			
			elst			an edit list			
		mdia				container for the media information in a track			
			mdhd				media header, overall information about the media		
			hdlr				handler, declares the media (handler) type		
			minf				media information container		
				vmhd			video media header, overall information		
				smhd			sound media header, overall information		
				hmhd			hint media header, overall information		
				nmhd			null media header, overall information		
				dinf			data information box, container		
					dref			data reference box, declares source(s) of media data in track	
				stbl			sample table box, container for the time/space map		
					stsd			sample descriptions (codec types, initialization etc.)	
					stts			(decoding) time-to-sample	
					ctts			(composition) time to sample	
					stsc			sample-to-chunk, partial data-offset information	
					stsz			sample sizes (framing)	
					stz2			compact sample sizes (framing)	
		stco			chunk offset, partial data-offset information				
		co64			64-bit chunk offset				
		stss			sync sample table (random access points)				
	stsh			shadow sync sample table					
	padb			sample padding bits					
	stdp			sample degradation priority					
	mvex				movie extends box				
		mehd				movie extends header box			
		trex				track extends defaults			