INTERNATIONAL TELECOMMUNICATION UNION

ITU-T J.127

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(06/2004)

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

Interactive systems for digital television distribution

Transmission protocol for multimedia webcasting over TCP/IP networks

ITU-T Recommendation J.127

ITU-T Recommendation J.127

Transmission protocol for multimedia webcasting over TCP/IP networks

Summary

This Recommendation defines a transmission protocol for multimedia webcasting over TCP/IP networks. The protocol consists of a presentation description and session control. The presentation description is based on the XHTML format, and it specifies properties of the media such as URI, content type, content size, accounting information, and so on. First of all, a terminal receives the presentation description, and sets up in accordance with the description. Then the terminal starts transmission of the media in accordance with the session control, which is based on HTTP. The session control supports starting the session, retrieving the media data and ending the session. In addition, a remote camera control scheme is also defined in the session control. Since the protocol defined in this Recommendation is similar to the WWW browsing system, problems related to the UDP/IP streaming such as the Firewall issue are resolved by this Recommendation, and the implementation will have less affect.

Source

ITU-T Recommendation J.127 was approved on 29 June 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.

CONTENTS

ITU-T Recommendation J.127

Transmission protocol for multimedia webcasting over TCP/IP networks

1 Scope

This Recommendation defines a transmission protocol for multimedia webcasting of various media format over TCP/IP networks. Such formats include J.123 and J.124 as well as other media formats. However, the protocol defined in this Recommendation is independent from the media format, which is outside the scope of this Recommendation.

The protocol consists of a presentation description and session control. The presentation description is based on the XHTML format, and the session control is based on HTTP. Issues not especially specified here shall comply with XHTML and HTTP.

2 References

2.1 Normative 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.

[1] W3C Recommendation, XHTML 1.0: *The Extensible HyperText Markup Language*.

[2] IETF RFC 2068 (1997), *Hypertext Transfer Protocol – HTTP/1.1*.

2.2 Informative references

[3] ITU-T Recommendation J.123 (2002), *Multiplexing format for webcasting on TCP/IP network*.

[4] ITU-T Recommendation J.124 (2004), *Multiplexing format for multimedia webcasting over TCP/IP networks*.

3 Terms and definitions

This Recommendation defines the following terms:

3.1 file downloading: Program transmission method whereby the program starts playing after the entire data has been downloaded.

3.2 live transmission: Program transmission method whereby the program starts playing after a certain amount of data has been buffered while receiving subsequent data in the background, where the program is fed in real time by the content provider.

3.3 VoD transmission: Program transmission method whereby the program starts playing after a certain amount of data has been buffered while receiving subsequent data in the background, where the program is completely created by the content provider.

4 Abbreviations

This Recommendation uses the following abbreviations:

HTTP Hypertext Transfer Protocol

MIME Multipurpose Internet Mail Extensions

VoD Video-on-Demand

XHTML Extensible HyperText Markup Language

5 Presentation Description

5.1 XHTML Presentation Description Format

The overall program and the properties of the media shall be defined by a presentation description in XHTML [1]. The presentation description may be obtained by the receiver using HTTP or other means such as e-mail and may not necessarily be stored on the server.

The presentation description contains a description of the media streams making up the program, including their location, title, encoding types, data size, and other parameters that enable the receiver to start retrieving the most appropriate media.

The presentation description is written by the <object> element with the <param> elements of XHTML.

An example is shown below. Elements defined by this Recommendation are written with bold letters.

```
<?xml version="1.0" ?>
<!DOCTYPE html
        PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en">
<head>
<title>Webcasting Test Page</title>
</head>
<body>
        <object data="http://www.webcasting.org/media.mp4" type="video/MP2T"
                copyright="no" standby="Click Here">
        <param name="disposition" value="devmpzz" valuetype="data" />
        <param name="duration" value="30000" valuetype="data" />
        <param name="size" value="240000" valuetype="data" />
        <param name="title" value="Preview of the movie" valuetype="data" />
        <param name="ac" value="Jc5gUxzTqJ9ebM3U18GEWdKgtiTWR6Fe" valuetype="data" />
        </object>
</body>
</html>
```

Elements used in the presentation description are summarized in Table 1. In Table 1, M/O stands for "Mandatory" or "Optional", respectively.

Table 1/J.127 – Elements defined in this Recommendation

Element	Attibute	M/O	Value	Description
object	data	M	URI String	Actual location of the media file.
object	type	M	MIME Type	MIME type of the media.
object	copyright	O	"yes" | "no"	Copyright control.
object	standby	M	String	The displayed text of the link.
param	name="ac" value="…." valuetype="data"	O	String	Access Ticket.
param	name="bitrate" value="…." valuetype="data"	O	Numeric String	Bit rate of the content in bit/s.
param	name="camctl" value="…." valuetype="data"	O	Numeric String	Camera control capability.
param	name="disposition" value="…." valuetype="data"	M	String	Types of the content distribution, which stands for downloading, VoD transmission, or live transmission.
param	name="duration" value="…." valuetype="data"	O	Numeric String	Duration of the content in milliseconds.
param	name="size" value="…." valuetype="data"	O	Numeric String	File size of the content in bytes. This field is effective for downloading and VoD streaming.
param	name="title" value="…." valuetype="data"	M	String	Title text of the content.

5.2 <object> Element

The following attributes for the <object> element are defined in this Recommendation.

5.2.1 data

This is a mandatory attribute that specifies the URI of the media to be transmitted. In this Recommendation, since the media is transmitted by HTTP, the scheme of the URI shall be http, or the URI shall start with "http://".

5.2.2 type

This is a mandatory attribute that specifies the MIME type of the media to be transmitted. For example, "video/MP2T" is specified for MPEG-2 Transport Stream.

5.2.3 copyright

This attribute takes "yes" or "no", and this is optional. The default value is "no". The copyright attribute takes effect as follows.

yes: The content is protected from storing. The media data cannot be stored in the device after playing.

no: The media data can be stored in the device after playing.

If this attribute is not specified, the terminal shall handle the file as storing allowed.

5.2.4 standby

This is a mandatory attribute that specifies the displayed text of the link to the media. It will typically be "Click Here" or the name of the content.

5.3 <param> Element

Parameters of the media are specified with the <param> element in the HTML description. The following parameters are defined in this Recommendation. Each parameter is identified by the name attribute and the value is specified by the value attribute. For the all parameters, valuetype="data" shall be included in each <param> element.

The terminal shall ignore unknown parameters.

5.3.1 ac

This is an optional parameter and the value attribute specifies the access ticket. The maximum length of the value is 512 bytes. The terminal that obtained the access ticket from the ac parameter in the presentation description shall use this ticket when the terminal carries out the session control as "ac=" parameter in the HTTP request. See also clause 6.

This is used for identification of fee collection.

5.3.2 bitrate

This is an optional parameter. It specifies the total bit rate of the media in bits per second. If the media has video and audio track, the bitrate value will be the sum of the bit rate of each track.

If the media has multiple bit rates for adaptive bit rate changing, all the values are specified with ':' separator. For example,

<param name="bitrate" value="64000:128000:256000" valuetype="data" />

5.3.3 camctl

This is an optional parameter that shows the camera control capability of the content. The value consists of 8-numeric letters and each digit shows the capability of each camera control. The following table explains the camera control value (="XYZABCDE"), where the left-most digit is firstly described.

Table 2/J.127 – Camera control parameters

No.	Value	Description
1	0 or 1	If the value is 1, panning control is supported.
2	0 or 1	If the value is 1, tilting control is supported.
3	0 or 1	If the value is 1, zooming control is supported.
4	0	Reserved
5	0	Reserved
6	0	Reserved
7	0	Reserved
8	0	Reserved

For example, the following shows the case where the content supports the pan and the zoom control.

<param name="camctl" value="10100000" valuetype="data" />

– The "camctl" parameter is effective only for the camera control enabled content, which is specified by the "disposition" parameter described below.

– The terminal shall recognize the content has camera control capability only when the "camctl" parameter exists and the value has an appropriate setting.

– The terminal shall recognize the content has no camera control capability if the value is not specified or the value is "00000000".

– The terminal shall recognize the content has no camera control capability if the value has more than eight digits or less than eight digits.

– The terminal shall recognize the content has no camera control capability if the value contains any digits other than 0 or 1.

– The terminal shall ignore the "camctl" parameter if any reserved field contains a non-zero.

Actual actions for the camera control are described in 6.5.

5.3.4 disposition

The disposition parameter defines the content type, its application, distribution scheme, and so on. The existence of the disposition parameter is mandatory. In this Recommendation, the disposition parameter itself is not defined, but what the parameter specifies is defined as follows.

– Category of the content: Video (including Video and Audio), Audio, Voice, MIDI, Still Image, Animation, Application Program (e.g., JAVA), etc.

– Transmission scheme of the content: File downloading, VoD transmission, Live transmission.

– Purpose of the content: Just viewing, Storing, Particular use (Wallpaper, Screensaver, Alarm, etc.).

5.3.5 duration

This is an optional parameter. It specifies the duration of the media in milliseconds. If the media has different duration of video and audio track, the value is the longest duration in the tracks.

5.3.6 size

This is an optional parameter. It specifies the data size of the media in bytes, which helps the terminal to obtain the content size in advance of transmitting. Regarding the file downloading and the VoD transmission, the file is already created before transmitting. Therefore, the value of the size parameter is the same as the size of the file.

In addition, if the media has multiple bit rates for adaptive bit rate changing, each size corresponding to each bit rate is specified with ':' separator. For example,

<param name="size" value="240000:480000:960000" valuetype="data" />

If this parameter is not specified in the presentation description, the terminal shall require the content size from the server at the beginning of the transmission. This is carried out with the HEAD request of HTTP. Details are described in 6.1.

For the live transmission, the file size cannot be estimated before transmission. In this case, the size value indicates the maximum size of the stream that is continuously transmitted. For example, if the size is 1572864 for the live transmission, the connection will be closed after receiving 1.5 MB of the content.

5.3.7 title

This is a mandatory parameter that describes the title of the content. The maximum length of the value is 40 bytes. The title may be shown on the terminal when the content is playing.

6 Session control

In this Recommendation, all session controls are based on HTTP. Some commands are embedded in the HTTP-URI query part and some are sent to the server in the HTTP request header.

After retrieving the presentation description, the terminal issues the HTTP request in order to start the session. The following defines the session control in this Recommendation.

6.1 Retrieval of the content size

For the file downloading and the VoD streaming, the terminal shall obtain the size of the content unless the size parameter exists in the presentation description. This is done with the HEAD method of HTTP as follows.

Syntax:

HEAD /URI?ac=xxxx&br=128000&ts=1 HTTP/1.1

Item		M/O	Description
Method		M	HEAD
URI		M	URI obtained from the presentation description.
URI Parameters	ac	O	Access ticket obtained from the presentation description.
	br	O	Bit rate selection in accordance with the presentation description.
	ts	M	1 This parameter shall exist if ac is contained.

M/O stands for "Mandatory" or "Optional" respectively.

The following defines the response expected to the request.

HTTP/1.1 200 OK
Content-Length:

Item		M/O	Description
Status Code		M	200 OK
Header	Content-Length	M	The size of the content.

The Content-Length header shall be returned only when the status is successful (200 OK). Other headers might be embedded in the response. The terminal shall ignore unknown headers.

Status codes other than those described above are defined in HTTP [2].

6.2 Data transmission

Request of the data transmission is defined as follows.

Item		M/O	Description
Method		M	GET
URI		M	URI obtained from the presentation description.
URI Parameters	data	O	None: File downloading evdo-2: Live transmission evdo-4: VoD transmission
	ac	O	Access Ticket obtained from the presentation description.
	br	O	Bit rate selection in accordance with the presentation description.
	st	O	Specifies the start position of the transmission in milliseconds. Effective only for the initial transmission request, where ts = 2.
	ts	M	2: The beginning of the transmission 3: Continuous transmission This parameter shall exist if ac is contained.
Header	Range	M	bytes = 0-XXXXX (The beginning of the transmission) bytes = YYYYY-ZZZZZ (Subsequent transmission)
	x-up-devcap-streaming-camctl	O	See 6.5. get_control: Camera control ticket is requested with the HTTP request. pan±X, tilt±X, zoom±X: Camera control commands.

M/O stands for "Mandatory" or "Optional" respectively.

Response to the data transmission request is defined as follows.

Item		M/O	Description
Status Code		M	206 Partial Content
Parameters	Content-Range	M	bytes = xxxx-xxxx/xxxxxx The size transmitted
	x-streaming-camctl	O	See 6.5. XXX (digits, in seconds) Valid period of the camera control ticket. Maximum value is 999 (seconds).
	x-streaming-campos	O	See 6.5. pan±X, tilt±X, zoom±X If the camera control parameter is contained in the request, the current camera position is returned in the response.

Regarding the Range parameter, the start bytes shall correspond to the actual bytes received at the previous request. The following table shows an example. Note that the Range parameter does not specify an actual position of the content in the server but specifies the number of bytes expected through the transmission.

The request number	Specified bytes in the request	Actual bytes in the response
1	0-96767	48000
2	48000-144767	52000
3	100000-196767	50000

Consequently, the request syntax is shown as follows.

GET /URI?data=evdo-4&ac=xxxx&br=128000&ts=2 HTTP/1.1 (The first request)
Range:bytes=0-96767

The following is the response expected to the request.

HTTP/1.1 206 Partial Content
Content-Range: bytes 0-48000/1572864

The subsequent request will be as follows.

GET /URI?data=evdo-4&ac=xxxx&br=128000&ts=3 HTTP/1.1 (The subsequent request)
Range:bytes=48000-144767

The following is the response to the subsequent request.

HTTP/1.1 206 Partial Content
Content-Range: bytes 48000-96000/1572864

Request and response are repeated until the media data is completely received.

6.3 Normal termination

For the VoD transmission and the live transmission, the terminal shall issue the termination request after receiving data of the requested size.

Syntax:

GET /URI?ac=xxxx&ts=4 HTTP/1.1

Item		M/O	Description
Method		M	GET
URI		M	URI obtained from the presentation description.
	ac	O	Access Ticket obtained from the presentation description.
	ts	M	4 This parameter shall exist if ac is contained.

– For the file downloading, it is not necessary to send the termination request.

– When the terminal receives the response to this request, the VoD/live transmission is finished. The response will be "200 OK" or similar, and the terminal should ignore it.

6.4 Abnormal termination

For the VoD transmission and the live transmission, the terminal shall issue the abnormal termination request when a transmission error occurs. For example, the case is regarded as a transmission error where the expected size of the content cannot be received during VoD transmission or live transmission.

Syntax:

GET /URI?ac=xxxx&ts=5 HTTP/1.1

Item		M/O	Description
Method		M	GET
URI		M	URI obtained from the presentation description.
	ac	O	Access Ticket obtained from the presentation description.
	ts	M	5 This parameter shall exist if ac is contained.

– For the file downloading, it is not necessary to send the abnormal termination request.

– When the terminal receives the response to this request, the VoD/live transmission is finished. The response will be "200 OK" or similar, and the terminal should ignore it.

6.5 Camera control

This clause describes camera control functions. If the "camctl" attribute is specified in the presentation description, the terminal shall recognize the content as camera control enabled. The camera control is carried out in accordance with the following procedure:

– request the camera control ticket;

– retrieve the ticket;

– send the camera control command;

– receive the notification of completion;

Details are defined in the following subclauses.

6.5.1 Request of the camera control ticket

The request of the camera control ticket is issued with the data transmission request. This is done by adding a header to the HTTP request. The following shows the format of the header.

x-up-devcap-streaming-camctl: get_control

If the ticket is successfully retrieved, the following header will be embedded in the HTTP response of the data transmission.

x-streaming-camctl: XXX

where XXX is the period of validity in seconds of the camera control assigned to the terminal. The range of the XXX is 0 to 999.

If the XXX is less than zero or if the header is not contained in the HTTP response, the camera control ticket request has been rejected by the server.

The following shows an example of the request and the response with the camera control ticket retrieval.

(*Request*)

GET /URI?data=evdo-4&ac=xxxx&br=128000&ts=3 HTTP/1.1
Range:bytes=48000-144767
x-up-devcap-streaming-camctl: get_control

(*Response*)

HTTP/1.1 206 Partial Content
Content-Range: bytes 48000-96000/1572864
x-streaming-camctl: 30

Note that if the terminal receives the "x-streaming-camctl" header without sending the request of the camera control ticket, it shall be ignored.

6.5.2 Send the camera control command

The request of the camera control is issued with the data transmission request as well as the ticket request. This is done by adding a header to the HTTP request during the ticket validation period. The following shows the format of the header.

x-up-devcap-streaming-camctl: pan±X, tilt±X, zoom±X

where X is relative camera position, which is an integer value from 0 to 5. Adaptation of the value to the actual camera position is outside the scope of this Recommendation. Only the supported parameter shall be sent. The supported parameter is specified by the "camctl" attribute in the presentation description. See 5.3.3. Parameters that are not changed do not have to be sent. In this case, making X to 0 is also accepted.

If the control is successfully performed, the following header will be embedded in the HTTP response of the data transmission.

x-streaming-campos: pan+3, tilt-1, zoom+2

where each parameter is an actual camera position as a consequence of the camera control.

The following shows an example of the request and the response with the camera control parameters.

(*Request*)

GET /URI?data=evdo-4&ac=xxxx&br=128000&ts=3 HTTP/1.1
Range:bytes=96000-192767
x-up-devcap-streaming-camctl: pan+1,tilt-1,zoom+2

(*Response*)

HTTP/1.1 206 Partial Content
Content-Range: bytes 96000-134000/1572864
x-streaming-campos: pan+3,tilt+0,zoom+1

Note that if the terminal receives the "x-streaming-campos" header without sending the request of the camera control ticket, it shall be ignored. If the server receives the "x-up-devcap-streaming-camctl" parameter without a valid ticket, it shall be ignored. The terminal shall not re-send the "x-up-devcap-streaming-camctl" parameter unless the terminal receives the "x-streaming-campos" response.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	**Cable networks and transmission of television, sound programme and other multimedia signals**
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure, Internet protocol aspects and Next Generation Networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2004