

Video And Streaming Content Creation Guide

Version 1.1; March 19, 2004

Messaging

NOKIA

Contents

1	Introduction	5
2	Understanding the Specification	5
2.1	3GPP Specifications.....	5
2.2	Other Specifications.....	5
3	Streaming Video Creation – Basic Steps	7
3.1	3GPP Compatible Content Authoring Tools	8
3.2	Players for PC Environment.....	9
4	Network and Bit-Rate Options for Streaming	10
5	Content Creation with 3GPP Codecs	11
5.1	Audio	11
5.1.1	Adaptive Multi-Rate Narrowband Speech Codec (AMR-NB).....	11
5.1.2	Adaptive Multi-Rate Wideband Speech Codec (AMR-WB).....	11
5.1.3	MPEG-4 Advanced Audio Codec (AAC).....	12
5.2	Video.....	12
5.2.1	H.263.....	12
5.2.2	MPEG-4 Visual.....	13
5.3	Suggested Video Encoder Configurations	13
6	RealMedia Content Creation with RealProducer and Helix Mobile Producer Authoring Tools	14
6.1	Video encoding	15
6.2	Audio encoding.....	15
7	Tips and Tricks	15
7.1	Selectivity for Authored Sequences.....	15
7.2	Multiple Encodings.....	16
8	Terms and Abbreviations	17

Change History

9 September 2003	V1.0	Initial document release
19 March 2004	V1.1	Minor changes in Chapter 4.

Copyright © 2004 Nokia Corporation. All rights reserved.

Nokia and Nokia Connecting People are registered trademarks of Nokia Corporation. Java and all Java-based marks are trademarks or registered trademarks of Sun Microsystems, Inc. Other product and company names mentioned herein may be trademarks or trade names of their respective owners.

Disclaimer

The information in this document is provided "as is," with no warranties whatsoever, including any warranty of merchantability, fitness for any particular purpose, or any warranty otherwise arising out of any proposal, specification, or sample. Furthermore, information provided in this document is preliminary, and may be changed substantially prior to final release. This document is provided for informational purposes only.

Nokia Corporation disclaims all liability, including liability for infringement of any proprietary rights, relating to implementation of information presented in this document. Nokia Corporation does not warrant or represent that such use will not infringe such rights.

Nokia Corporation retains the right to make changes to this specification at any time, without notice.

License

A license is hereby granted to download and print a copy of this specification for personal use only. No other license to any other intellectual property rights is granted herein.

Video And Streaming Content Creation Guide

Version 1.1; March 19, 2004

1 Introduction

This document is intended for developers and service providers who are planning to produce video and streaming services for mobile devices. It addresses the specifications related to the subject, streaming video creation with standard codecs, and feasible network and bit rate options for streaming.

2 Understanding the Specification

In this section we'll point you to the main 3GPP specifications related to mobile video and streaming and give you a brief synopsis of each one. For a professional study on the topic we recommend additional familiarization with the basic standards set by the other standardization bodies. These specifications are, with some exceptions, available via the Internet, and are widely referenced in the 3GPP service specification.

2.1 3GPP Specifications

3GPP, the standardization body for third-generation mobile networks, has published several specifications related to mobile video and streaming:

Specification	Description
TS 26.233: Transparent End-to-End Packet-Switched Streaming Services (PSS); General Description	A general description of a transparent packet-switched streaming service in 3G networks.
TS 26.234: Transparent End-to-End Packet-Switched Streaming Services (PSS); Protocols and Codecs	Specifies the protocols and codecs for the streaming service, standardized and deployed within the 3GPP compliant systems.

2.2 Other Specifications

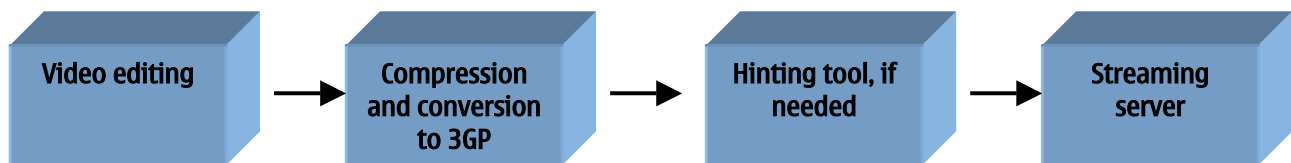
Specification	Description
IETF STD 0006: User Datagram Protocol	These documents specify the means for sending/transporting real-time or streaming data using UDP and streaming control with RTSP and SDP over TCP.
RFC 3550 - RTP: A Transport Protocol for Real-Time Applications	
RFC 3551 -RTP: Profile for Audio and Video Conferences with Minimal Control	

<p>RFC 2429: RTP Payload Format for H.263 Video</p> <p>RFC 3016: RTP Payload Format for MPEG-4 Audio/Visual Streams</p> <p>RFC 3267: RTP Payload Format for AMR and AMR-WB Audio Codecs</p>	<p>These documents specify the real-time transport (RTP) payload formats to be used for the multimedia coded bit streams during streaming or at the storage phase.</p>
<p>IETF STD 007: TCP Protocol</p> <p>RFC 2326: Real-Time Streaming Protocol (RTSP)</p>	<p>Specifies an application-level protocol for control over the delivery of prestored or real-time data. It enables VCR-like control capabilities for the user.</p>
<p>RFC 2327: Session Description Protocol (SDP)</p>	<p>Specifies a protocol that enables the description of the multimedia session with the help of a well-defined and formatted structure. This protocol is mainly utilized by the RTSP protocol.</p>
<p>ITU-T Recommendation H.263 Video</p> <p>ISO/IEC 14496-2:2001: Coding of Audio/Visual objects; Part 3: Audio (MPEG-4 Audio, a.k.a., AAC)</p> <p>ISO/IEC 14496-2:2001: Coding of Audio/Visual Objects; Part 2: Visual (MPEG-4 Visual)</p>	<p>Documents specify the syntax of the encoded audio and video media bit streams.</p>

3 Streaming Video Creation – Basic Steps

The following procedure outlines the basic steps in streaming video creation:

1. Use any of the commonly used video editing tools, e.g., QuickTime, Adobe, etc., to create the video content.
2. Export your content to the compressed format (RealMedia or 3GPP formats) and perform the actual crunching compression.
3. Here the 3GPP file *may* be "hinted" in preparation for the streaming session.
4. The server extracts the media content from the file and sends it to the network, according to the appropriate payload formats.



Depending on the implementations, some of the steps described above can be merged. A server may be capable of doing media transport protocol layer packetization on the fly, i.e., it may be able to understand a non-hinted file and simply stream it. Encoding and hinting can also be merged in the same element. This is the case, for example, in the Helix Mobile Producer (for 3GPP media), the Philips Platform 4 tools, and the Apple QuickTime Pro tool (version 6.3 and above) with 3GPP support. Since there's no common hinting language for streaming server files yet, all hinting tools today address one specific server.

3.1 3GPP Compatible Content Authoring Tools

	Video Conversion	Input	Output	Basic Editing (Cut and Paste)	Hint Track Generation	Server
RealNetworks	Helix Producer	.AVI, MPEG, etc.	*.rm	No	No, not needed for Real Streaming	Yes, for Real Media
RealNetworks and Envivio	Helix Mobile Producer 2.0 Powered by Envivio	.AVI, MOV, MPEG, etc.	*.3gp	No	Yes, for 3GPP-serving Real servers	Yes, Helix Universal Server Mobile (3GPP-serving Real servers)
Nokia	Nokia Multimedia Converter 2.0	.AVI, .MOV, MPEG, etc.	*.3gp	No	No	Non-commercial test server. Packetization on the fly – omits any hint tracks
Philips	Platform 4 Encoder	.AVI, MPEG, etc.	*.3gp	Limited	Yes, for Philips' server	Yes
Apple	QuickTime 6.3 Pro	.AVI, .MOV, .MPEG, etc.	*.3gp	Yes	Yes, for Apple Darwin server	Yes, "Darwin" open source server
PacketVideo	PV author	.AVI .MOV .MPEG .etc	*.3gp	No	???	PV server v4.x and upwards?

3.2 Players for PC Environment

	Player Name	Free Version	Commercial Version	3GPP Support	Web Link
RealNetworks	RealOne Player	Yes, basic version	Yes, RealOne Player Plus	Yes	www.real.com/
Apple	QuickTime 6.3	Yes	QuickTime Pro (allows exporting to 3GP)	Yes	www.apple.com/
PacketVideo	PV player	60-day trial	Yes	Yes	www.packetvideo.com/
Nokia	Nokia Multimedia Player	Yes	No	Yes	www.nokia.com/
Philips	Platform 4 Player	Yes	No	Yes	www.platform4.com/

Note: Nokia does not guarantee the correctness and completeness of the information provided above about third-party software. This list is meant as an aid to developers who desire to probe further, and does not imply any endorsement by Nokia, nor any guarantee about the availability, performance, compliance with 3GPP specifications, and conditions of use of this software.

The reader is encouraged to retrieve updated information directly from the relevant companies.

4 Network and Bit-Rate Options for Streaming

The following table includes identified network service configuration options for streaming, differentiated by the network technology over which the service is intended to be used.

Table 1 Network and Bit-Rate Options for Streaming

Configuration Number	Network	Configuration	Available Bandwidth, kbit/s	Suggested Bit Rate for Streamable Content, kbit/s
1**	HSCSD	2 TS DL	28.8	20
2**	HSCSD	3 TS DL	43.2	35
3	GPRS	2 TS DL (CS 2)	26.8	20
4	GPRS	3 TS DL (CS 2)	40.2	28
5	GPRS	4 TS DL (CS 2)	53.6	38
6	EGPRS	2 TS (MCS 6)	59.2	42
7	EGPRS	3 TS (MCS 6)	88.8	63
8	EGPRS	4 TS (MCS 6)	118.4	84.2
9	WCDMA	GBR=MBR=64	64	47
10	WCDMA	GBR=MBR=128	128	92
11	CDMA20001xRTT		144	40
12	CDMA20001xEV-DO		2400	64 – 100
13	CDMA2000 1xEV-DV		3100	64 – 100

** Connection method could be ISDN V.110, ISDN V.120, or normal modem type

Disclaimer: Note that other manufacturer's devices may be providing time slot combinations or WCDMA downlink bearers not listed in the table, and network operators may limit their service offering, QoS parameters such as maximum or guaranteed bit rate, or allocation of time slots per (E)GPRS subscriber. Thus, consultation with the operators is recommended. Real-time dataflow, for example video/audio streaming, requires a more stable bit rate than best-effort data, such as browsing or e-mail. Because of this, the suggested bit rate is a bit lower than the bandwidth that is available for the configuration. Note also that the bit rates suggested in the previous table are media bit rates (audio+video) and don't include packetization headers.

5 Content Creation with 3GPP Codecs

This chapter lists the basic technical information on available coding parameters per each codec, and gives parametrization guidance on most important adjustable media parameters. The same media types are usually available for local playback on the device and/or via streaming services from a remote server, thus readers should note that some parametrization applies only for streaming.

Note that optional media types may not be supported depending on the client device capabilities.

5.1 Audio

5.1.1 Adaptive Multi-Rate Narrowband Speech Codec (AMR-NB)

- Mandatory speech codec in 3G PSS
- Mainly used for speech encoding at very low bit rate
- Sampling frequency: 8000 Hz, mono
- One AMR frame = 20 ms = 160 samples
- Can be used for both Constant Bit Rate (CBR) and Variable Bit Rate (VBR) encoding/decoding
- Main CBR encoder bit-rate switches (kbit/s): 4.75 – 5.15 – 5.90 – 6.70 – 7.40 – 7.95 – 10.2 – 12.2
- VBR can be achieved by dynamically changing the bit-rate switch. Switching can be done on a frame-by-frame basis
- Possible relevant encoder switches:
 - DTX (ON/OFF): When ON, uses silence indicators (SID) for silence frames, i.e., consumes fewer bits

5.1.2 Adaptive Multi-Rate Wideband Speech Codec (AMR-WB)

- Mainly used for high-quality speech/low-quality audio encoding at low bit rate
- Sampling frequency: 16000 Hz, mono
- One AMR-WB frame = 20 ms = 320 samples
- Can be used for both Constant Bit Rate (CBR) and Variable Bit Rate (VBR) encoding/decoding
- Main CBR encoder bit-rate switches (kbit/s): 6.60 – 8.85 – 12.65 – 14.25 – 15.85 – 18.25 – 19.85 – 23.05 – 23.85
- VBR can be achieved by dynamically changing the bit-rate switch. Switching can be done on a frame-by-frame basis
- Possible relevant encoder switches:

- DTX (ON/OFF): When ON, uses silence indicators (SID) for silence frames, i.e., consumes fewer bits

5.1.3 MPEG-4 Advanced Audio Codec (AAC)

- Optional audio codec in 3G PSS
- Mainly used for high-quality audio encoding at reasonable bit rates
- Sampling frequency: Up to 48000 Hz, mono/stereo
- Support for AAC-LC (Low Complexity Mode); may also support AAC-LTP (Long Term Prediction)
- One AAC frame = 1024 samples
- Can be used for both Constant Bit Rate (CBR) and Variable Bit Rate (VBR) encoding/decoding
- Reasonable encoder bit-rates switches (kbit/s): 24 – 48 – 56 – 96
- VBR can be achieved by dynamically changing the bit-rate switch. Switching can be done on a frame-by-frame basis

5.2 Video

5.2.1 H.263

- Mandatory video codec in 3G PSS
- Supports H.263 Profile 0 Level 10; may also support Profile 3 Level 10
- Frame size: QCIF (176 x 144) or Sub-QCIF (128 x 96)
- Frame rate: Max. 15 fps
- Bit rate: Max. 64 kbit/s*
- Content can be encoded using constant frame rate or variable frame rate, not exceeding the selected frame rate
- Content can be encoded using constant bit rate or variable bit rate, not exceeding the selected bit rate
- Reasonable CBR encoder bit-rate switches (kbit/s): 15 – 20 – 25 – 30 – 35 – 40 – 50 – 60 (the actual bit rate takes into account the presence of other medium and the bearer size)

* Bit rate can be exceeded in local playback beyond the conformance level point up to ~128 kbit/s, while frame rate is saturated in 15 fps. Performance may degrade if player settings are set to prefer sharp image to fast frame rate or associated audio content will cause player to consume more computational power than available at the device hardware.

5.2.2 MPEG-4 Visual

- Optional video codec in 3G PSS
- Support Visual Simple Profile Level 0
- Frame size: Max. QCIF (176 x 144). Suggested QCIF or Sub-QCIF
- Frame rate: Max. 15 fps
- Bit rate: Max. 64 kbit/s*
- Content can be encoded using constant frame rate or variable frame rate, not exceeding the selected frame rate
- Content can be encoded using constant bit rate or variable bit rate, not exceeding the selected bit rate
- Reasonable CBR encoder bit rate switches (kbit/s): 15 – 20 – 25 – 30 – 35 – 40 – 50 – 60 (the actual bit rate takes into account the presence of other medium and the bearer size)

* Bit rate can be exceeded in local playback beyond the conformance level point up to ~128 kbit/s, while frame rate is saturated in 15 fps. Performance may degrade if player settings are set to prefer sharp image to fast frame rate or associated audio content will cause player to consume more computational power than available at the device hardware.

5.3 Suggested Video Encoder Configurations

Table 2 The table below applies to both H.263 and MPEG-4 Visual.

Video Bit Rate (only video, kbit/s)	Frame Size (pixels)	Suggested Intra Refresh Period Range (seconds)*	Suggested Frame Rates (constant fps)
15	Sub-QCIF (128 x 96)	5 – 10	5 to 7.5
	QCIF (176 x 144)	5 – 10	2.5 to 3
20	Sub-QCIF (128 x 96)	5 – 10	7.5 to 10
	QCIF (176 x 144)	5 – 10	3 to 5
25	Sub-QCIF (128 x 96)	5 – 10	10 to 15
	QCIF (176 x 144)	5 – 10	5 to 7.5
30	Sub-QCIF (128 x 96)	5 - 10	10 to 15
	QCIF (176 x 144)	5 - 10	7.5 to 10
35	Sub-QCIF (128 x 96)	5 - 10	10 to 15

	QCIF (176 x 144)	5 - 10	7.5 to 10
40	Sub-QCIF (128 x 96)	5 - 10	10 to 15
	QCIF (176 x 144)	5 - 10	10 to 15
50	Sub-QCIF (128 x 96)	5 - 10	15
	QCIF (176 x 144)	5 - 10	15
60	Sub-QCIF (128 x 96)	5 - 10	15
	QCIF (176 x 144)	5 - 10	15

* The intra frames help resynchronization at rebufferings during streaming, as well as qualitative improvement if packet losses occur. The intra period may be decreased/increased to gain from quality. The suggested value for content served on-demand is 10 seconds.

- Streaming use case: Selected video bit rate = suggested total bit rate (network) – selected audio bit rate
- A selection of two-pass encoder could increase the visual quality with the same bit-rate budget
- Usage of variable fps could increase image quality, but could also introduce time-wise non-constant frame updates on the display screen
- Intra refresh period can be changed to tune the visual quality. For 3G PSS streaming, intra frames are generally anchor points for stream synchronization for a server
 - A maximum Intra period of 10 seconds is suggested for on-demand and live streaming configurations
 - For local video clips, sent, for example, via MMS, forced periodic Intra frames are unnecessary, given that the used encoder process always provides Intra frame after a scene change event in the source sequence

6 RealMedia Content Creation with RealProducer and Helix Mobile Producer Authoring Tools

This chapter presents parametrization guidance for generation of mobile optimized content with RealMedia content types via the dedicated tools available from RealNetworks. In general, official content authoring guidelines are also available directly from RealNetworks, from <http://www.realnetworks.com/resources/howto/mobile/index.html>.

6.1 Video encoding

- Use RealVideo G2 or RealVideo 8 encoding. Decoding of RealVideo 9, the latest video coding technology from RealNetworks shipped on the basic version of RealProducer, is not currently supported by the RealOne Player for Nokia mobile terminals.
- Use two-pass encoding
- Enable Variable Bit Rate encoding (VBR)
- Use VBR startup latency of 10 seconds
- Use key-frame difference of 10,000 milliseconds
- Enable loss protection
- Suggested motion-related parameter: “Normal Motion Video”
- Frame size (suggested): QCIF (176 x 144) or Sub-QCIF (128 x 96)
- Maximum frame rate suggested: 15 fps (10 fps is the safest from RealOne Player Mobile performance point of view)
- For local playback use case, the maximum bit rate suggested is 80 kbit/s
- For streaming use case, see the Table 2 in Section 5.3, which is also applicable to RealVideo

6.2 Audio encoding

- Use RealAudio G2, RealAudio 8, or RealAudio Voice (for speech)
- Use frequency response from 8 to 16 kHz for speech and from 20 to 44 kHz for music
- Streaming use case: Selected audio bit rate = selected total bit rate (network) – selected video bit rate

7 Tips and Tricks

There are many small details and bits of information that will make the end-user experience of mobile video more appealing. Below are several very basic authoring guidelines for selecting and editing content.

7.1 Selectivity for Authored Sequences

Authors should avoid source video sequences with:

- Blank or black first frame (resulting from a video edit)
- High scene-change frequency (one per 1- 2 seconds or fractions of a second)
- Fast horizontal/vertical or diagonal panning

- Fast zooming
- Encoding at 20 – 50 kbit/s and 3 –5 fps achievable via streaming over GPRS cannot reproduce TV-like end-user experience, irrespective of technology used.
- A badly chosen video sequence can ruin the result of an otherwise optimal content creation process.
- Avoid re-encoding from an already encoded content. Always use the source audio and video sequences.

7.2 Multiple Encodings

It may be advisable to generate multiple bit-rate versions of the content at the same time, once it is aimed for streaming and for several terminal types that can utilize various radio technologies (GPRS, EGPRS, W-CDMA, CDMA in U.S. and Korea, etc.). The Table 1 in Section 4 gives a good estimate of the suitable total session encoding rates, including the packetization overheads.

As long as the 3GPP service is unable to signal the alternative bit-rate sessions for the mobile client (Rel-4, Rel-5 versions of the PSS service), the session offering towards the mobile client must be based on the client's indicated capabilities. Nokia streaming-enabled terminals will provide the necessary profiling and capability information via the CC/PP- and UAPProf-based mechanism defined in the 3GPP technical specification 26.234.

The aforementioned Rel-4 and Rel-5 limitation also applies to the file management — lack of a generally defined server file format gives the server manufacturers freedom to select their preferred solution for content management on and preparation for the specific streaming server. For content authors, the guideline is to generate a 3GP file with or without hint track, and include only the specific media track on the file. Tool and server manufacturers may have their own guidance on the file preparation requirements in order to ensure that the files are streamable from the manufacturer's server platform.

8 Terms and Abbreviations

Term or Abbreviation	Description
3GPP	Third-Generation Partnership Project
CBR	Constant Bit Rate
CDMA	Code Division Multiple Access
CS	Coding Scheme
DL	Downlink
FPS	Frames Per Second
GBR	Guaranteed Bit Rate
GPRS	General Packet Radio System
HSCSD	High Speed Circuit Switched Data
MBR	Maximum Bit Rate
MCS	Modulation and Coding Scheme
PSS	Packet Switched Streaming
QCIF	Quarter Common Intermediate Format
RTP	Real-Time Transport Protocol
TS	Time Slot
UAProf	User Agent Profile
UL	Uplink
VBR	Variable Bit Rate
WCDMA	Wideband CDMA