John R. Smith
IBM

The 3G-324M Protocol for Conversational Video Telephony

Marwan A. Jabri
Dilithium Networks

As mobile operators worldwide migrate to third-generation (3G) networks, conversational video-telephony services are becoming a key differentiator between new 3G offerings and existing 2G/2.5G services. Although it's possible to have limited video-based services—such as a multimedia messaging service—that deliver pictures and video clips over 2.5G services, these are delay-insensitive applications that could run over a packet-based wireless network like General Packet Radio Service (GPRS) or Code Division Multiple Access (CDMA)'s 1XRTT. For delay-sensitive applications such as conversational video telephony, present 3G packet bearers are inadequate, and the Third Generation Partnership Project (3GPP; http://www.3gpp.org) mandates using the 3G bandwidth-guaranteed circuit-switched bearer and the 3G-324M system.

The 3G-324M system is a derivative of the International Telecommunication Union (ITU) H.324 protocol standard for low-bitrate multimedia communication, which ITU-T developed for the public switched telephone network (PSTN). This article describes the 3G-324M system, which has been adopted by both 3GPP and 3GPP2 (htpp://www.3gpp2.org), as well as its H.324 roots.

ITU-T H.324 system

The ITU-T H.324 standard is an umbrella recommendation developed by the ITU-T between 1994 and 1998. It includes several components:

- a multiplexer defined by the ITU-T recommendation H.223;
- a V.34 modem;
- a session command, control, and indication system defined by the ITU-T recommendation H.245;
- audio and video codecs;
- optional data protocol for applications, such as data sharing (for example, the ITU-T T.120 suite); and
- optional encryption support.

An H.324 application isn't required to have each of these functional elements, except for the V.34 modem, H.223 multiplex, and H.245 system control protocol, which are mandatory to all H.324 terminals. However, H.324 mandates that terminals offering audio communication support the G.723.1 audio codec. Various H.324 terminals offering video communication are required to support the H.263 and H.261 video codecs. Additionally, H.324 terminals offering real-time audio-graphic conferencing are required to support the T.120 protocol suite. (As I'll discuss later, 3G-324M mandates specific audio and video codecs irrespective of H.324.)

H.324M system

The H.324 protocol (and in particular the baseline H.223 multiplexer) were designed for PSTNs. Through a series of annexes to H.324 and its components, the protocol has been extended to deal with error-prone mobile networks. Annex C of H.324 and annexes A, B, C, and D of H.223 deal with increasing levels of error resilience. H.324 together with these annexes is usually referred to as H.324M. The most prominent extension features of H.324M are H.223's annexes A through D. For example, annex A of H.223 defines a mobile level 1 (mobile level 0 is the baseline of H.223) and provides 16-bit flags for framing the multiplexer payloads instead of the 8-bit high level data link control (HDLC) flags used in baseline H.223. Annex B defines mobile level 2 and provides 16-bit flags and an extended

1070-986X/04/$20.00 © 2004 IEEE

Published by the IEEE Computer Society

Further Reading

This list includes references for further readings related to 3G-324M video communication. All the 3GPP technical specification references are available from the 3GPP Web site (http://www.3gpp.org).

- D.J. Myers, *Mobile Video Telephony for 3G Wireless Networks*, McGraw-Hill, 2004.
- 3GPP TS 26.110, Codecs for Circuit-Switched Multimedia Telephony Service (General Description).
- 3GPP TS 26.111, Codecs for Circuit-Switched Multimedia Telephony Service and Modifications to H.324.
- 3GPP TS 26.112, Codec(s) for Circuit-Switched Multimedia Telephony Service and Call Set-Up Requirements. (TS 24.008 replaced this document.)
- 3GPP TS 24.008, Mobile Radio Interface Layer 3 Specification and Core Network Protocols (Stage 3).
- 3GPP TS 27.007, AT Command Set for 3G User Equipment.

header for the multiplexer payloads. Annexes C and D define mobile level 3 and provide additional adaptation layers and redundancy coding for highly error-prone channels.

Other prominent differences between H.234 and H.324M reside in the modem requirements. The ITU modem related standard requirements (such as V.34/V.8/V.250) don't apply, although V.250 is commonly available in mobile modems.

3G-324M system

Figure 1 shows a block diagram of the 3G-324M system. The 3GPP 3G-324M technical specification defines a video-telephony service based on H.324M as follows:

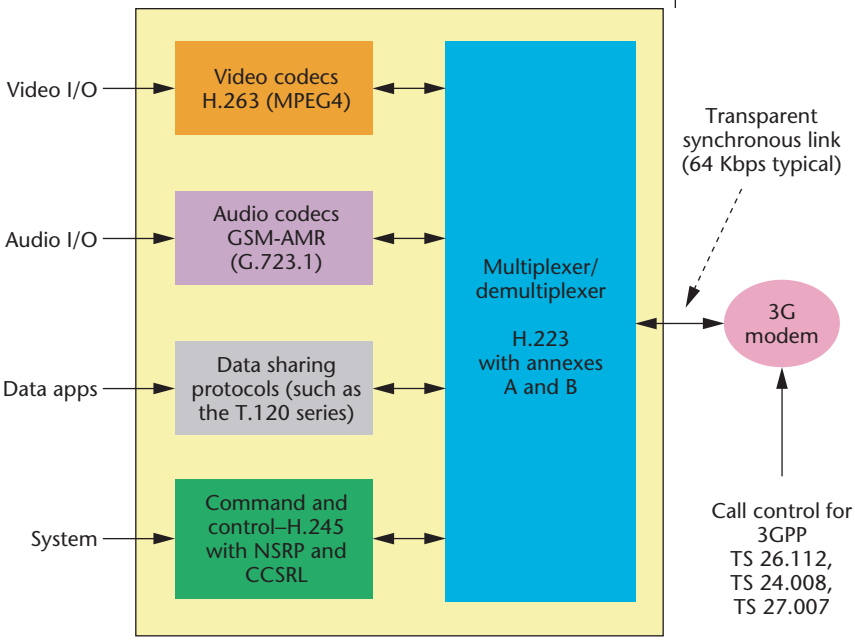

Figure 1. Block diagram of the 3G-324M system. GSM-AMR is the Global System for Mobile Communication; Adaptive Multirate. NSRP is the numbered simple retransmission protocol. CCSRL is the control channel segmentation and reassembly layer. 3GPP is the Third Generation Partnership Project.

- Using the ITU-T H.324 umbrella recommendation and its annex C. This defines the overall video-telephony service, including H.223 and H.245.
- Using annexes A and B of H.223 ITU-T to enhance the framing facilities of the multiplexer in error-prone conditions.
- Using the mobile command and control facilities of H.245.
- Using specific audio and video codecs. For example, it mandates the Global System for Mobile Communication; Adaptive Multirate (GSM-AMR) audio codec and the H.263 video codec. Other audio and video codecs are proposed as options (such as MPEG-4 video and G.723.1 audio codecs).
- Using a transparent circuit-switched channel (bearer) for carrying the overall multiplexed bitstream.
- Interfacing the video-telephony system to a modem and defining modem commands to configure for multimedia operations. The 3GPP defines a set of AT modem commands to configure and manage modems used in the context of 3G-324M.

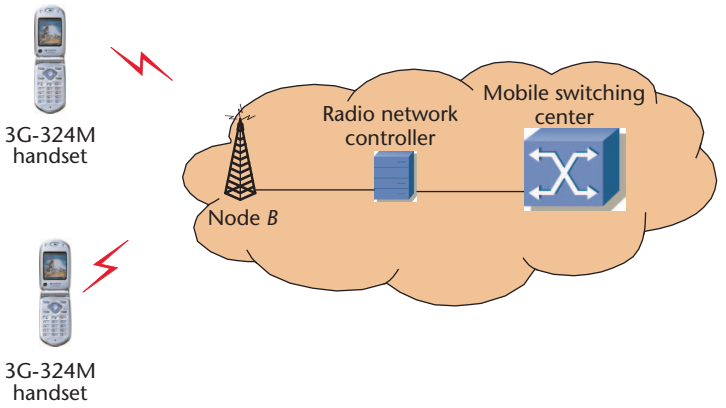

Figure 2. Typical connectivity in a 3G-324M mobile-to-mobile service.

Operations in a 3G-324M session

A typical session between 3G-324M terminals involves the following procedures:

- call signaling (dialing a number and establishing a bearer following a call's acceptance);
- mobile-level negotiation (terminals agree on an H.223 mobile level); and
- using H.245 messages to exchange terminal capabilities, determine mastership of session, open audio and video logical channels, define multiplexing definitions, and process other messages that could be application dependent—for example, transmitting keypad presses for videomail navigation.

Figure 2 shows an overall network diagram.

Call signaling and bearer establishment

Low-level connectivity is about establishing a digital bearer (link) between two handsets. Handsets use 3G call signaling, which involves using code points in the call-signaling records to indicate 3G-324M capabilities. These are the H223 and H245 code points. The code point is communicated when handsets initiate their call signaling through the mobile switching center (MSC), which uses the code point to detect the capabilities of both end points and decides how to handle the call.

Following successful call signaling, a transparent data bearer (at 64 Kbytes per second) is established over which the 3G-324M bitstream (with multiplexer payloads separated by framing flags) will flow.

Mobile-level detection

This is the first step in the protocol-establishment process. Each terminal starts transmitting bytes (flags) corresponding to the highest mobile level it supports (typically mobile level 2, corresponding to annex B of H.223). As soon as both end points detect that they're transmitting and receiving flags corresponding to the same mobile level, they both settle on that as their starting level.

Otherwise, after repeating the same sequence several times without success, each handset drops its mobile level and starts transmitting the flags corresponding to the lower mobile level. At the end of the procedure, the handsets can connect at mobile level 2, 1, or 0.

Transporting MUX-PDUs

After the terminals have established a common mobile level, the H.223 multiplexer (MUX) immediately becomes operational on the data link. From this point, we can only transmit H.223 protocol data unit (MUX-PDU) frames separated by framing flags on the link.

A MUX-PDU is a frame that consists of a header and an information field. The header includes a multiplexing table entry index and a cyclic redundancy check (CRC) code associated with that index. At this early stage of the communication, this corresponds to multiplexer table entry zero, which corresponds to the built-in multiplexing mode that lets the MUX-PDU carry H.245 command and control messages encapsulated into special frames.

H.245 Numbered Simple Retransmission Protocol

Assuming MUX-PDUs are well-formed, the payload of the MUX frames at this stage of the link will carry H.245 messages (using the predefined logical channel zero and the multiplex table entry index zero). The H.245 messages are encapsulated using the Numbered Simple Retransmission Protocol (NSRP). An NSRP request frame consists of a header and a payload. The H.245 messages are contained in the payload as an array of bytes in the abstract syntax notation 1 (ASN.1, ITU recommendation) format. That is, H.245 messages are first serialized using an ASN.1 encoder that converts messages into a string of bits, and the resulting octet array—which can hold one or more H.245 messages—makes up the payload field of the NSRP request frame.

According to NSRP specifications in H.324, an H.245 entity can't transmit a new NSRP request frame until it receives an NSRP response (acknowledgment) frame for the last NSRP request sent.

Terminal capability exchange

The first H.245 message to be transmitted (or received) by an end point is the terminal capability request message that declares the capability of a terminal (handset, gateway, and so on). A terminal capability request message advertises a terminal's capability in terms of the audio and video codecs it supports, its multiplexer's capability in terms of carrying audio and video over the three adaptation layers (whether it supports simple or nested multiplexing options), and the extent of the mobile options it supports.

Master-or-slave determination

The master-or-slave determination (MSD) H.245 procedure determines whether an end point is a master or slave. This status is essential for the protocol's operation as far as conflict resolution in the context of the bidirectional logical channels' opening procedures.

Conflict might occur when an end point tries to open a bidirectional logical channel that conflicts with an ongoing logical channel procedure that the peer end point has launched.

Media logical channels

Once the outcome of a MSD procedure is known, handsets are ready to open logical channels to exchange voice, video, and data. To achieve this, two procedures must be completed: opening the logical channels and defining multiplexing schemes. These two procedures can occur in any order.

I assume here that logical channels are opened first. 3G-324M logical channels can be unidirectional or bidirectional. To open a unidirectional logical channel, the requester must define the logical channel number, its media capability, and its multiplexer parameters. The media capability defines what the channel will carry (such as H.263 or GSM-AMR), the bit-rate, and other codec parameters. A bidirectional logical channel is opened using a bidirectional open logical channel (BOLC) H.245 request message similar to the unidirectional case, except that the request contains the requested reverse logical channel parameters. On acknowledgement of the BOLC, the far end returns the logical channel number of the reverse channel.

Multiplex tables

A 3G-324M end point must define the peer end point to the multiplexer table configuration that it will use to communicate the media in MUX-PDU frames. A H.223 multiplex table can have up to 16 entries, 15 of which are user definable (entry 0 is for logical channel 0, which is reserved for H.245 messaging). A multiplex table entry (MTE) is a structure that has an index and a list containing the combination of logical channel data in the multiplex frame. An end point defines its entries to the far-end multiplexer using H.245 procedures.

Media exchange

Once logical channels are open and MTEs are defined, handsets are in a position to exchange voice, video, and data.

User experience and quality-of-service issues

Numerous important factors affect the user's experience with 3G-324M video telephony, the most important of which are the call setup time and the audio–video lip synch and delays. Call-setup time is the duration the user must wait, following dialing or answering, to hear audio and see video on the handset. 3G-324M protocol compliant techniques to hasten the call setup time are available and should be used because they significantly enhance the user's experience.

For audio–video synch and delays, the 3G-324M stack implementation needs to ensure that the H.223 multiplexing tables are adequately designed and used to ensure minimal delays and proper audio and video synchronization, and that media codecs meet their target bit rate.

Conclusions

The H.324/H.324M and 3G-324M protocols enable the deployment of new conversational videotelephony services. Early deployment feedback shows enthusiastic response from users. Conversational videotelephony will revolutionarize personal communication and is likely to have the same impact and success as television had with respect to radio. **MM**