An Introduction to Multimedia Services

White Paper

Table of Contents

Multimedia Opportunities Today

The great strides made recently in delivering multimedia, which is the coordination of video, text, and audio media capabilities, is one of the most significant developments in communications today. Multimedia can provide important opportunities for service providers to attract new customers and increase the loyalty of current ones while increasing Average Revenue Per User (ARPU). Businesses can use multimedia to improve employee communication and significantly upgrade the effectiveness of critical functions such as the contact center.

Cellular providers face particularly difficult market conditions because their markets are saturated in many parts of the world. *The Economist* cites very high rates of penetration in many countries. See Table 1.

Country	Mobile Subscribers per 100 Population
Australia	82.8
Czech Republic	105.6
France	73.7
Italy	108.2
Japan	71.6
Taiwan	100.3
United Kingdom	102.2
United States	62.1

Table 1. Mobile Subscribers by Population

Other agencies cite even higher figures. For example, the CTIA cites an 80.5% penetration rate in the United States at the end of June 2007 [CTIA].

Market saturation generally brings an ARPU decrease because more resources are spent on marketing, prices may be lowered for competitive reasons, or additional features may be added at the same or reduced rates to retain customers. To counteract this trend, cellular providers are offering multimedia services as a market differentiator.

At the same time, businesses are looking to multimedia to improve customer service, increase worker productivity, create more effective training, and reduce operating costs. Many facets of a business can be improved, from investor relations to sales, with the implementation of applications such as video portals and video email.

As a long-time technology pioneer, Dialogic is aware of the issues that can arise when emerging technologies converge with existing ones. Because of its experience, Dialogic is committed to helping its customers make their move to multimedia a great success through working with standards organizations, strengthening its ecosystem, and developing innovative new products.

The Move to Multimedia

The adoption of multimedia requires a robust infrastructure that can support powerful multimedia applications, regardless of whether the endpoint is a traditional PC or a mobile device. Bandwidth is an important infrastructure component because of the "real time" nature of many multimedia services.

For PCs, multimedia requires a broadband or equivalent network connection, and broadband subscribers are growing at a healthy rate. According to a recent industry report, the worldwide broadband subscriber base had reached 285 million (May 2007) and was expected to "almost double" to 567 million by 2011 [In-Stat].

Because common broadband implementations (e.g., DSL and cable) have a download rate of 1.5 to 6 megabits per second (Mbps) and upload at a rate of 386 kilobits per second (kbps) to 2 Mbps, bandwidth is adequate to support a number of multimedia services (with the exception of high definition IPTV). The widespread purchase of broadband connections should accelerate the move to multimedia services.

For 3G cellular services, the latest CDMA2000 implementations deliver a peak data rate of 3.01 Mbps with an average throughput of 600 kbps. Bandwidth availability in these ranges should enable service providers to offer a wide range of multimedia services with the possible exception of broadcast-quality video streaming.

As mobile video services proliferate though the network, however, considerable infrastructure enhancement must take place. Dr. Mark Mortensen, senior vice president of marketing at VPIsystems, a company that specializes in network planning, describes three trends: a move from wireline to wireless, from voice to "voice-plus-email" and Internet surfing, and from cell phone as communicator to cell phone as entertainer. All this, says Mortensen, "is helping to drive what we see now, which is a huge 'mobile video build-out'. We find that the carriers are preparing now or are already engaged in a massive build-out on top of the existing infrastructure." Such a build-out is great news for the growth of multimedia services [Mortensen].

Multimedia Services

Development and deployment of new video services is moving forward quickly. A video implementation already available in 2.5G networks is Multimedia Messaging Service (MMS). With new 3G network rollouts, mobile video telephony applications such as video mail, video color ring, video caller ID, video portals, contact center enhancements, video share, and Mobile TV should become widely available. Several of these services are described briefly in this section.

Multimedia Messaging Service

MMS adds images and audio clips to the SMS text messaging model. Some wireless service providers see MMS as a natural evolution from SMS for their large user bases. With appropriate gateway infrastructure, MMS messages can be sent and received between mobile handsets and email, increasing the appeal and user base.

One popular implementation of MMS worldwide is sending pictures taken with a mobile device's camera. Other MMS applications currently deployed include:

- Weather reports sent with map images
- Stock quotes with trending information in graph form
- Key moments from the day's sporting events sent as a "slide show"
- Text messages with animations

Because both the development environments for creating MMS content and its billing mechanisms are well defined, the adoption and popularity of MMS is high.

Video Mail

Most mobile subscribers use voice mail services today. As technology, networks, and devices advance to support real-time video mail, mobile subscribers are likely to use enhanced applications as well. Many callers worldwide can already send and receive video messages, and use playback capabilities similar to those available with voice mail today. Recording a video message for someone who is temporarily unavailable is expected to become as commonplace as leaving them a voice mail message today.

As with other multimedia applications, video mail provides additional revenue for cellular providers because they can charge either per-message or a fixed monthly fee for the enhanced service.

Video email technology, in which an MPEG or other type of video file is attached to an email message, first appeared in the mid-1990s on personal computers. At that time, PCs had to be enhanced with high-end graphics cards, special cameras, and extra memory to use video email systems. Today, video email is far more practical and affordable because of cheaper component costs, improved compression algorithms, and enhanced delivery mechanisms that store messages on a server and stream them to the client.

Many cellular providers can send video mail messages to PC-based accounts as video email. Unfortunately, the quality and capacity of the video capture combined with the relative complexity of sending video mail are seen as hurdles to widespread consumer acceptance. These hurdles are likely to be overcome as the technology improves.

Video mail is also gaining popularity beyond the cellular market. Corporations and non-profit organizations have also deployed video mail solutions with positive results. Some examples include:

- Corporations can use video mail to provide employees and industry analysts with information about acquisitions, new product introductions, and earnings.
- Market research firms can use video mail and the Internet to conduct virtual focus groups. First, a video email is sent to focus group participants about a product. After watching the video in the email, participants can provide instant feedback about the product via a website.
- Executives can use video emails to relay complex information instead of spending their valuable time typing long email messages. Video emails are also generally easier to understand.
- Charities can solicit emergency donations by sending video emails containing actual scenes of a disaster immediately after it occurs. Effective thank-you emails often inspire additional donations.

Video Color Ring Back Tones

Color ring back tones replace the ringing sound that the network traditionally supplies while a caller waits for the called party to answer the phone. Color ring back tones range from popular music and theme songs to recorded messages and jokes.

Video color ring adds a pre-selected video to the color ring back tones sent through the network, and the video plays while the caller waits for the call to be answered. Unfortunately, the varied standards and capabilities of handsets may slow the adoption of video color ring. As an interim offering, some service providers are sending photos to the caller's handset. Ultimately, standardization and intelligent gateways, which can decouple the video and audio streams and provide more varied support on the network, are likely to allow video color ring to become as popular as color ring back tones.

Video Caller ID

Video caller ID uses the caller ID function to select and play a pre-recorded video on the mobile device being called before the call is answered. The pre-recorded videos for known callers are stored locally in the mobile device's memory and do not require any intervention by the service provider. Applications loaded and running on the mobile device look up the caller ID, load the proper video, and control play according to a model not unlike handset-based ring tones today.

Video Portal

Both businesses and service providers are beginning to use video portals.

Large corporations are providing information to employees through intranet video portals usually accessed via PCs. Portals can provide company news and keynote speeches by senior management or technical leaders. Portals are also a vehicle well suited for training employees on new products or technologies. Most companies do not currently support mobile access because of security issues, but these issues are expected to be addressed soon.

Meanwhile, service providers in the mobile market are offering video portals. Subscribers can download a range of video clips for playback on their mobile devices. One service provider is offering hundreds of video clips per day that run from two to five minutes. Clips can be browsed and selected via Interactive Voice/Video Response (IVVR), and the content is similar to that of a voice portal: entertainment, news headlines, and clips of key plays in sports.

Contact Center Enhancements

When customers call a contact center today, they usually encounter an Interactive Voice Response (IVR) system because many standard queries can be satisfied without agent intervention. Video can further enhance the customer's IVR experience. A caller on a 3G network can receive streaming video about new products and services while waiting for an agent. A traditional IVR that offers downloadable videos, which further assist or instruct customers, can provide a competitive advantage.

Contact centers often use a coach/pupil technique in training agents. A knowledge expert or supervisor provides coaching on a variety of subjects, some of which can be very complex. These audio sessions can now be enhanced with video to make them more effective. In addition, streaming a video on product installation to an agent during a customer call allows an agent to provide far more accurate and detailed information, which can greatly improve customer satisfaction.

IP-based, SIP-powered, IMS-compatible, 3G contact center software-only solutions are now available, offering full multimedia contact management for multi-location contact centers in a distributed model. Inbound callers to a customer service representative can start off with a chat window, and later escalate the call to an audio or even a video call. 3G mobile operators and Network Service Providers can provide such solutions to contact center outsourcers and "partial" outsourcers as a hosted service or to large businesses that want to fully "own" and control the contact center on their premises.

Video Share

Video Share (VS) — also known as Video Sharing, See What I See, Rich Voice Call, and Push-to-Video (P2video or PTV) — is a service that enables a user engaged in a voice call to stream real-time (live) or pre-recorded one-way video from a handset to any another party or multiple parties on the call regardless of whether they are on a mobile or fixed network.

The sender can see what is being streamed and can provide real-time narration over the audio portion of the call, explaining the content of the video to those receiving it. The recipient(s) can choose not to accept the streamed video and can terminate the VS session at any time.

Since Push-to-Talk has been a success, VS should also become popular, although moving from video mail to live video streaming poses technological challenges.

VS was originally defined by the Global System for Mobile communications Association (GSMA), a global trade association for mobile phone operators, which refers to VS as a "combinational service" because it converges a circuit-switched voice call with a packet-switched multimedia (video) session.

Since it was described in 3GPP specification documents, VS requires a 3GPP-compliant IMS core network system.

VS is currently supported only in 3G UMTS with the W-CDMA interface and in EDGE networks with DTM, a method that enables simultaneous GSM voice and EDGE data connections, which, in turn, enables VS and improves both the end-user experience as well as GSM/EDGE-W-CDMA service continuity. VS does not operate on a GPRS or a CDMA network because CDMA2000 EV-DO (for example) does not support simultaneous circuit and packet connections. If a handset moves from a UMTS environment to GSM, the VS client software will drop the VS session although the voice portion of the call will remain connected.

The anticipated popularity of VS has spurred some mobile operators to offer it as their first IMS service. AT&T (formerly Cingular) led the way with a VS service deployed in 160 major U.S. markets. The service runs over AT&T's 3G W-CDMA/HSDPA network instead of its nationwide, 2G GSM/GPRS network.

Mobile operators expect VS to be a popular service in scenarios such as the following:

- A vacationing couple sends live video postcards to their friends
- A proud parent transmits footage of a child in Little League hitting a home run
- A real estate agent streams video of a house to entice a prospective buyer

A variant of this technology (Movial Connect Push-to-Video) seamlessly combines technologies such as Presence, Instant Messaging (IM), and multimedia instant communication. Users who log into the system immediately see a listing that tells them whether or not their previously defined contacts are online and available. A Push-to-Video dialog can be initiated to one or more contacts, or a simple instant messaging dialog can be established. Using VS and IM in tandem creates active dialogues among multiple users, resulting in more revenue-generating content creation along with message and network traffic — and a desire to purchase advanced handsets. Push-to-Video is not only available for SIP/IMS, but also for existing presence standards such as OMA IMPS and XMPP.

Mobile TV (Out of Band Video)

Various test marketing trials and limited deployments indicate that people enjoy watching TV broadcasts on mobile phones, which is made possible by using technology complaint with one of four Mobile TV broadcasting standards: DVB-H, DMB, ISDB-T, and MediaFLO. Trials and limited deployments are under way at many major carriers such as Sprint, Verizon, and AT&T.

Arqiva provides infrastructure for television, radio, and wireless communications in the UK with a presence in Ireland, mainland Europe and the USA. The company has conducted a variety of trials, providing reports on its website. A trial conducted with O2 (UK) in Oxford, UK from October 2005 to February 2006 provided access to a group of well-known UK channels, including BBC1, BBC2, ITV1, Sky, and MTV. A significant number of participants in the trial (72%) expressed interest in subscribing to the service.

The Sprint Nextel Corporation, Comcast Corporation, Time Warner Cable, Cox Communications, and Advance/Newhouse Communications launched a service called "Pivot" in March 2007. Pivot was designed as a "quad-play" offering that combines wireline and wireless networks to bring capabilities such as one-button access to the Internet, home and wireless email, and Mobile TV to Sprint customers.

"Pure" Mobile TV offerings are available from carriers such as Verizon and AT&T. Verizon Wireless offers ESPN MVP sports on V CAST Mobile TV, a service of MediaFLO USA, a subsidiary of Qualcomm. Customers can also view trailers and short films from the Tribeca Film Festival on V CAST-enabled mobile phones.

AT&T is offering full-length episodes of soap operas, situation comedies, sports, and news 24/7 with broadcast-quality picture and sound on AT&T Mobile TV. Programming is available from CBS Mobile, CNN Live Mobile, Comedy Central, ESPN Mobile TV, FOX Mobile, MTV, NBC 2GO, NBC News 2GO, and others.

Although most current deployments in the USA are commercial, other long-term financing models for mobile TV are being explored. Free-to-Air (FTA) Mobile TV, which is unencrypted and available without subscription, is one possibility. FTA could be funded by direct license fees (as in the United Kingdom), by voluntary donations (modeled on the Public Broadcasting Service in the United States), or by advertising and other forms of commercial sponsorship (as in the United States and Japan). In the United States, deployments are only available to mobile phone users by paid subscription, but Mobile TV elsewhere in the world is FTA. As of this writing, FTA Mobile TV services or trials have been launched in Finland, India, Japan, Korea, the Philippines, and Russia.

Mobile TV worldwide has been hampered by lack of spectrum availability, but this situation is expected to change as analog TV services are phased out over the next few years.

Cellular Standards

By continuing to expand and upgrade their services, cellular providers are driving the evolution from simple voice-only mobile devices to sophisticated multimedia devices that are more graphically adept and easier to use. As part of the transition to full multimedia, a wide variety of standards have emerged affecting many aspects of multimedia delivery from network infrastructure to cellular endpoints. Because of competing standards and the gradual nature of the 3G transition, cellular providers today are compelled to support a range of standards.

Key to 3G standards efforts is the work of the 3GPP and 3GPP2. Both organizations began work in 1998 intending to evolve two different cellular networks to 3G capabilities. However, over time, some overlap has developed.

3GPP

Third Generation Partnership Project (3GPP) is a collaboration of several telecommunications standards bodies working to develop a 3G mobile system based on an evolved GSM network. 3GPP is also responsible for maintaining the standards related to the radio access technologies used in GSM. The original scope of this project was later amended to include the maintenance and development of GSM technical specifications. To date, 3GPP has defined two technical specifications for radio access technologies: GPRS and EDGE. Details about these specifications are included below.

3GPP has also done excellent work in developing standards for service delivery mechanisms. One of the most important is the IMS framework, which provides the telecom industry with a modular, standards-based IP/SIP service delivery infrastructure. IMS enables service providers to take advantage of the unique strengths of IP to deliver faster, more flexible, and more cost-effective deployments of new and differentiated multimedia communications services over both mobile and fixed networks.

GPRS

General Packet Radio Service (GPRS) is a non-voice specification that defines how information will be sent and received across a mobile network. The theoretical maximum speed of a GPRS implementation is 171.2 kbps. Although this is a marked improvement over the 9.6 kbps in a GSM network and the 160 character limitation of SMS, bandwidth is still inadequate for a multimedia stream. Because of this, GPRS implementations focus on non-real-time applications such as chat, web browsing, and file transfer.

EDGE

Enhanced Data rates for GSM Evolution (EDGE) enables the delivery of advanced mobile services. These advanced services include enhanced downloading of video and music clips, multimedia messaging, high-speed color Internet browsing, and mobile email access. EDGE provides three times the bandwidth available with a GPRS-based network (typical user data rates are 100 to 120 kbps, and the theoretical maximum is 384 kbps) while freeing up network resources to handle additional voice traffic.

3GPP2

3GPP2 is focusing on Code Division Multiple Access 2000 (CDMA2000). The standard is named after the 3GPP2, a partnership consisting of five telecommunications standards and consortium bodies: CWTS in China, ARIB and TTC in Japan, TTA in Korea, and TIA in North America.

CDMA2000 is a 3G wireless technology that evolved from the existing CDMA specification. Its main features include faster data rates, always-connected data service, and an improved voice network capacity.

CDMA2000 deployment is planned in three general phases:

- **Phase 1** — Data rates up to 144 kbps; doubles the voice capacity over the previous CDMA network
- **Phase 2** — Data rates up to 2.4 Mbps, but must be deployed using a spectrum separate from the voice network; devices for Phase 2 support access to both voice and data spectrums
- **Phase 3** — Circuit and packet data rates of 3 to 5 Mbps; fully integrates the voice network established in Phase 1

Here are the specific CDMA2000 specifications involved:

- **CDMA2000 1X** — First CDMA2000 standard, which enables operators with existing IS-95 systems to double overall system capacity yielding uplink speeds up to 76.8 kbps and downlink speeds up to 153.6 kbps (network dependent).
- **CDMA2000 1XRTT (One-Carrier Radio Transmission Technology)** — A 2.5G network standard adopted by some US wireless carriers, equivalent to GPRS for CDMA. 1XRTT uses a 1.25 MHz radio channel as well as a more sophisticated form of modulation to increase bandwidth for individual users up to about 144 kbps. It requires a different type of phone and a change to some base station equipment, but can double a voice network's capacity and allow data to be packetized and sent without the establishment of a circuit. 1XRTT has three direct descendants: CDMA2000 1xEV, CDMA2000 1xEV-DO and CDNA2000 1xEV-DV.
- **CDMA2000 1xEV** — An evolution (EV) of CDMA2000 1X standard that can support a speed of about 300 kbps on a 1.25 MHz channel.
- **CDMA2000 1xEV-DO (EVolution-Data Optimized)** — A descendant of CDMA2000 deployed in South Korea that is faster than CDMA2000 1X but is not officially a 3G system. 1xEV-DO can support fixed and mobile applications at 700 kbps to 1.2 Mbps on average with a 2.4 Mbps peak on a standard 1.25 MHz CDMA channel. An "always on" wireless technology, 1xEV-DO allows data to be sent over PCS and cellular networks at speeds comparable to DSL or cable modems.
- **CDMA2000 1xEV-DV (EVolution-Data and Voice)** — A real-time, two-way voice communication system that also supports packet data rates up to 3.07 Mbps with an average data throughput of 1.0 Mbps, unmatched by many competing technologies that operate in the 1.25 MHz radio channel. 1xEV-DV has a flexible data air interface that supports a wide range of applications and offers an orderly and cost-effective migration path for existing CDMA2000 1X networks.
- **W-CDMA (Wideband-Code Division Multiple Access)** — A 3G global mobile wireless technology and a fundamental component of UMTS, along with TDMA and CDMA2000. W-CDMA has a 5 MHz air interface and operates in the 1920 to 1980 MHz band for the uplink (Rx) and 2110 to 2170 MHz for the downlink (Tx) situated in the 230 MHz global spectrum identified by the ITU for the worldwide standard called IMT-2000. The air interface occupies more bandwidth than conventional CDMA but offers faster transmission because, unlike CDMA, it optimizes the use of multiple wireless signals, not just one.

3G-324M

Both the 3GPP and 3GPP2 envision supporting both voice and multimedia content over an IP network. However, issues such as the availability of adequate IP addresses and the ability to reliably support time-sensitive applications are likely to delay full deployment of either standard. To resolve the reliability issue, 3GPP and 3GPP2 collaborated to create the 3G-324M standard, which defines real-time streaming of wireless multimedia over existing circuit-switched wireless network links. Multimedia is sent as RTP packets, which provide the timestamp and control mechanisms for synchronizing multimedia content.

3G-324M contains a multiplexing/de-multiplexing protocol that interleaves voice, video, user data, and control signaling for delivery over a transmission channel to the end user. The standard is complex, containing derivatives of three legacy standards revised to provide call control and multimedia delivery.

DVB-H

Digital Video Broadcasting-Handheld (DVB-H) is a technical specification and ETSI standard designed to bring broadcast services to handheld receivers. It is a superset of the Digital Video Broadcasting-Terrestrial (DVB-T) system for digital terrestrial television, with additional features to allow operation in an IP environment and with handheld, battery-powered receivers. DVB-H incorporates advanced audio and video coding schemes, such as H.264, and complements other existing broadcasting technologies.

Of the four competing MobileTV standards (DVB-H, T-DMB, MediaFLO, and ISDB), DVB-H appears to be the most popular. Its closest competitor is T-DMB (and other forms of DMB).

DVB-H is part of the DVB-IP Datacast (DVB-IPDC) set of technologies designed to deliver any type of content to mobile devices over IP when broadcast network output is reformatted for Mobile TV reception. DVB-IPDC specifications provide an Electronic Service Guide that describes the services available to end devices, Content Delivery Protocols to transport content over a broadcast bearer channel, and Service Purchase and Protection to enable over-the-air protection of purchase mechanisms and content.

DVB-H and its accompanying protocol set are "radio agnostic" and could work with CDMA services and handsets and GSM.

T-DMB

Terrestrial-Digital Multimedia Broadcasting (T-DMB) is a digital radio system for sending multimedia to mobile devices such as mobile phones and is used for Mobile TV. T-DMB uses MPEG-4 Part 10 (H.264) for video and MPEG-4 Part 3 BSAC or HE-AAC V2 for audio. Originally developed in South Korea, DMB can operate via satellite (S-DMB) or T-DMB transmission where it uses radio frequency bands Band III (VHF) and L (UHF). Because the USA and Canada allocate Band III for television VHF broadcast channels 7 to 13, and the USA uses the L band for military applications, North America uses Qualcomm's proprietary MediaFLO system for Mobile TV.

MediaFLO

Media Forward Link Only (MediaFLO) is a proprietary system developed by Qualcomm to broadcast data to mobile devices such as PDAs and cell phones. Multiple real-time audio and video streams, individual pre-recorded video and audio clips, and IP Datacast application data such as stock market quotes, weather reports, and sports scores can be transmitted to mobile devices at approximately 700 MHz (previously allocated to UHF TV Channel 5). MediaFLO is part of Verizon's V CAST offering.

ISDB-T

Integrated Services Digital Broadcasting-Terrestrial (ISDB-T) is the terrestrial version of ISDB, a broadcast format used in Japan to convert radio and TV stations from analog to digital and multiplex channels of data. ISDB-T is used by the Japanese "1seg" mobile terrestrial digital audio/video and data broadcasting service, which is capable of providing Mobile TV to Japanese cell phones, laptop computers, and vehicles. Various versions of the ISDB standard exist in addition to ISDB-T, including ISDB-S for satellite television and ISDB-C for cable.

ISDB is based on MPEG-2 video and audio coding, but both ISDB and DVB allow other video compression methods to be used, such as JPEG and MPEG-4.

Video Standards

Among the various video coding standards available today are MPEG-2, MPEG-4, and VC-1.

MPEG-2

The Moving Picture Experts Group (MPEG) began the development of the MPEG-2 video coding standard (H.262) more than a decade ago in the early 1990s, and it is now used in digital television systems around the world. Although MPEG-2 works well for transmitting over bandwidth capable of supporting large data rates, quality suffers when bandwidth is inadequate.

MPEG-4

In 1998, the Video Coding Experts Group (VCEG) started work on a new standard designed to double the compression efficiency of other video standards available at that time. In 2001, MPEG teamed up with VCEG to finalize the newly renamed H.264/AVC standard. This new standard, also known as MPEG-4, succeeded in doubling the compression of video broadcasts, paving the way for video support in low bandwidth environments such as cellular networks.

Compared to MPEG-2 (the format of traditional digital television and DVDs), H.264/AVC offers 2 to 3 times greater compression, making it attractive for network delivery as well as for high-definition video. MPEG-4 requires only 700 kbps to deliver full-screen, DVD-quality digital video.

VC-1

Microsoft decided to develop its own standard based on its Windows® Media Video 9 coder, which was standardized by the Society of Motion Picture and Television Engineers and became standard 421M, now better known as VC-1. Microsoft claims very high compression rates for VC-1 — three times as many video channels of the same quality as can be delivered with the same amount of bandwidth needed for MPEG-2.

Media Streaming

The Real-time Transport Protocol (RTP) is designed for the transport of real-time data, including audio and video, from endpoint to endpoint over an IP network. RTP supports multimedia by providing timing reconstruction, loss detection, and content identification. Timing reconstruction and loss detection ensure that the multimedia frames are presented in proper sequence and smooth out the multimedia presentation by buffering content. The Real Time Control Protocol (RTCP) is used in conjunction with RTP and provides media session control. RTCP synchronizes different media streams, supports gateways and bridges, and provides QoS feedback to the RTP session. RTP-based streams are generic and flexible enough to allow the definition of a unique profile (stream characteristics) for each target media endpoint.

The Real Time Streaming Protocol (RTSP) is an application layer protocol that controls the real-time data delivery of both live and pre-recorded audio and video. This control is also called "streaming" and an RTSP server can deliver multimedia content to multiple destinations simultaneously. Several popular streaming servers, such as Apple's QuickTime, RealNetworks RealServer, and the Microsoft Windows Media® Player, use RTSP as their control protocol. RTSP can also be used to control media distribution from a media server in a multimedia implementation.

Components for Delivering Multimedia Services

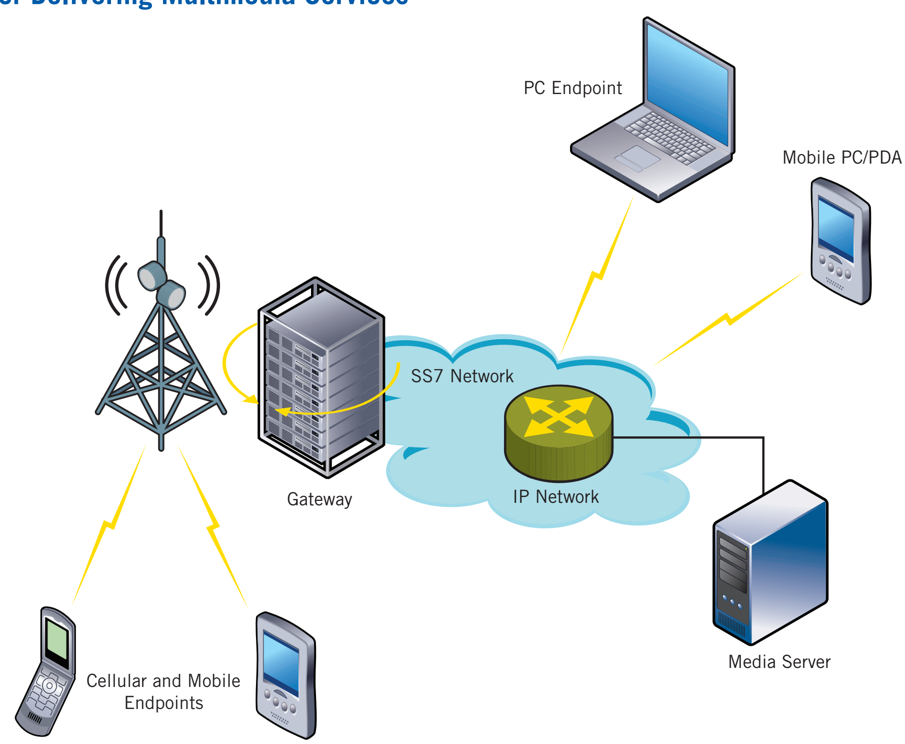

Figure 1. Simplified Topology for Multimedia Service Delivery

Figure 1 illustrates components defined in standards for both 3G networks and the IMS service delivery framework. Several of these components are discussed in more detail below.

The capabilities of SS7 are essential for some multimedia applications if the transactions require using the information from remote database lookups together with call signaling to perform predefined actions related to a call. In multimedia services, the predefined actions involve providing context-specific multimedia presentations to both the caller and the called parties.

MultiProtocol Label Switching (MPLS) is used to manage IP network bandwidth and helps ensure that adequate bandwidth for QoS is available to support a multimedia application. Video and high-quality audio streams require a high level of guaranteed bandwidth, and MPLS provisions (reserves) a fixed amount of bandwidth to meet bandwidth needs. Although MPLS is an established method of ensuring bandwidth availability, other methods are available. For example, ATM and the Resource reSerVation Protocol (RSVP) can be used to pre-provision paths in the network and ensure QoS.

Gateways

A gateway is a node on a network that serves as an entrance to another network. PC-based users may use gateways as proxy servers that direct traffic to a service provider network. Services can include delivering a simple web page or a complex conference involving several participants in a video call. In 2.5G and 3G cellular networks, gateways provide inter-working between cellular phones and VoIP, PSTN, and IP content. Gateway inter-working functions include negotiating picture size and translating or transcoding audio and video to a format supported at the endpoints.

In addition to inter-working functions, a 3G video gateway is responsible for the composition and decomposition of traffic headed to and from a cellular handset. For video, the composition function (multiplexing) is the blending of audio and visual packets from the network into a single stream for delivery to the wireless handheld. Decomposition (de-muxing) is the process of separating the video and audio received from the wireless handheld into separate packet streams for delivery through the network.

Composing and decomposing media streams at the gateway provides two notable benefits. First, the protocol used to perform the composing and decomposing resides solely in the gateway. No device other than the gateway needs to participate in the composition/decomposition process. The multimedia protocols used can be totally transparent to the rest of the network. Second, decoupling the video and audio for a multimedia session at the gateway ensures greater interoperability with other devices. For example, video-based conference calls can be extended to PSTN phone sets that do not have video capabilities.

Because the 3GPP Release 5 definition of a 3G cellular network specifies that all voice and multimedia traffic be transported over IP, gateways will be one of the first devices encountered by packets moving to and from a cellular or mobile device in a carrier network. The expected high volume of call sessions and the functionality needed will require that gateways be carrier-class, high-end servers with high network bandwidth connections to the cellular base equipment.

Media Servers

In IMS networks, media servers are known as Media Resource Function Processors (MRFPs) and Media Resource Function Controllers (MRFCs). Regardless of what they are called, media servers provide access to media processing resources when a client application requests them. The media can be a simple “busy signal” or “extension is no longer in service” message or a complex video stream or video conference. The requesting client can be a video-enabled IP phone, traditional PC, or a 3G phone set. Because 3G and mobile network implementations and IMS standards are still emerging, the exact functionality provided by a media server varies; however, convergence is coming quickly.

Media servers can use any one of a number of standards-based implementations to support the delivery of media to a PC-based endpoint. Apple’s QuickTime and RealNetworks RealPlayer are two popular media server implementations in the PC end-user market today. The relatively new Microsoft® Windows Media® Player 10 Series provides streaming capabilities, utilizing bandwidth management techniques that seem likely to increase its presence as a streaming server. Streaming to 3G endpoints adds complexity to an implementation because media streams must be transcoded into a format that is supported by the mobile network and endpoint. 3GPP Release 4 and beyond define media transcoding as a function of the media server.

Because 2.5G and 3G network implementations differ, media servers are required to support additional capabilities. For example, video mail exchanged between 3G and 2.5G cellular users requires media servers to convert H.263-based messages to MMS messages for delivery to a 2.5G user. Transcoding of multimedia streams and the need to support different video and audio file formats among cellular endpoints has led cellular providers to restrict both the endpoints and network interoperability provided.

The transcoding of media streams is a key function of video gateways, but a media server can also perform this function. Performing transcoding in the media server may be more cost-effective for homogenous networks, and enable lower cost, higher density video gateways to provide connectivity to disparate IP and PSTN networks.

Multimedia Endpoints

A PC is a multimedia endpoint. Embedding video or other forms of multimedia in emails, accessing multimedia content streamed from web servers, and participating in multimedia PC-hosted meetings are all commonplace today. Recent enhancements to PDAs, cell phones, and the cellular infrastructure, which supports these devices, are also bringing multimedia to mobile devices. Because of the adoption of 2.5G and 3G wireless standards, several companies are developing technologies that deliver rich content in a race to create the latest must-have applications and wireless devices.

Unfortunately, interoperability is an inhibiting factor for wireless multimedia endpoints today. The conditions in the mobile marketplace currently resemble those in the early years of the PC for the following reasons:

- Multiple mobile network standards are being used
- Handset and mobile devices offer differing sets of capabilities
- Equipment vendors are interpreting standards differently

An increased level of standardization and widespread adoption of these standards is necessary to ensure that the multimedia services will work across carriers and mobile devices from different vendors.

A brief review of the evolution of the cellular network shows a trend towards mobile devices becoming multimedia endpoints equal to or exceeding a PC's capabilities. See Table 2 for a summary.

Network Generation	Data Rates	Description	Standards
1G	Separate 9.6 kbps external modem	Analog cell phones designed for voice	NMT TACS AMPS
2G	64 kbps	Digital voice plus messaging, voice mail, and caller ID	GSM D-AMPS PDC CDMA
2.5G	115.2 kbps	Data enhancements to 2G	GPRS HSCSD
3G	144 kbps to 2 Mbps	Broadband data and VoIP	EDGE W-CDMA CDMA2000

Table 2. Cellular Network Characteristics by Generation

In the First Generation (1G), cellular phones offered basic analog phone service with data capabilities provided through an external modem. 2G was launched with the definition of the GSM standard in Europe and D-AMPS and PDC elsewhere. The GSM standard was patterned after ISDN, with support for 64 kbps of data.

With the recent definition of GPRS, packet switching replaced circuit switching in cellular calls. The 2.5G network and 2.5G devices are the first to use packet switching similar to that used on the Internet. Because 2.5G cellular devices send bursts of data only as needed and data exchange does not require an entire circuit, the devices can always be ready to receive and send data.

3G cellular service is moving to a converged IP transport network that handles data, voice, and video. This represents a shift from voice-centric TDMA services, the focus of previous cellular generations, to multimedia-centric services. Unlike in circuit-switched networks, cellular bandwidth is only consumed when call participants speak. As a result, one circuit can support multiple conversations and meet the needs of the latest multimedia applications.

A paradigm shift has occurred. Data access is no longer an adjunct function, but part of a feature set that supports a rich multimedia communications experience. 3G cellular phones promise to provide a multimedia user experience that is always on, easily transported, and universally reachable.

Market Outlook

Both technology innovation and marketing ingenuity are clearly moving forward in the mobile marketplace. The stumbling block for equipment vendors and network operators is finding the application that will excite "must-have" customer demand and combine the daily usage model of broadcast television with the ubiquity of mobile phones. A massive adoption of multimedia services would require a major network infrastructure expansion, but the increased monthly ARPU for operators as users consume hour after hour of video content could make the infrastructure investment pay handsome rewards.

Although the promise of individualized, mobile video technology has been enticing speculation and investment for two decades, two essentials have been missing: a winning business model and an enthusiastic subscriber base. The good news is that the availability of multimedia-enabled handsets and a new generation of technically savvy consumers seem to be combining to create a huge pool of potential subscribers — assuming the right price point can be found.

Analysts such as Insight Research foresee a worldwide growth of 3G subscribers at a total Compound Average Growth Rate (CAGR) of 46.2% through 2011 [Insight Research, p. 209]. Insight Research's numbers are conservative since some commentators are predicting a growth rate twice as large, depending on how quickly the number of subscribers grows in the world's largest mobile market, China.

Although 3G is not a requirement for delivering multimedia content and 2G and 2.5G phones have delivered multimedia more than adequately for almost a decade, the higher bandwidth of 3G networks is important for the high-quality video content required for video downloads that last longer than a few minutes or for Mobile TV.

Once the technology is in place, the critical issue is how much subscribers are willing to pay for video service. The current trend seems to be a subscription fee that allows unlimited viewing that is added to the monthly wireless bill. In the USA, rates are averaging $4 to $15 per month. Alternatively, multimedia is being included in an "unlimited" monthly plan for US$99. As was discussed earlier in this paper, some providers have been achieving success with reformatted broadcast TV programming, news, and sports. Since technology is progressing rapidly to allow high-quality content to be delivered in a consistent, portable manner, it is not surprising that Insight Research predicts a worldwide CAGR of 70.1% for wireless residential video telephony service worldwide through 2011 with the greatest growth in North America (285.4% CAGR) where revenue is currently the lowest of all regions [Insight, p. 227].

Insight Research also predicts robust growth through 2011 for audio/video streaming services with a total CAGR of 68.7% worldwide [Insight, p. 250].

Dialogic in Multimedia

Dialogic has a history of delivering innovative technology for Enhanced and Value-Added Services. Dialogic's media processing and signaling products form the foundation for mobile applications worldwide — from advanced Ring Back Tone solutions to emerging video portals. Dialogic partners use Dialogic® components based on open standards to develop breakthrough mobile applications.

Today, IMS-compliant, and IMS-enabled technologies are driving many multimedia innovations. The software-based Dialogic® IP Media Server supports video as well as the industry-standard interfaces of VoiceXML and MSCML needed for IMS. Advanced video manipulation, including video conferencing and transcoding, can be delivered for high availability carrier deployments with the Dialogic® Multimedia Platform for AdvancedTCA.

Because Dialogic adheres to the goals of open standards in communications technology, its products support advanced coders and IP and TDM call control standards. Mobile multimedia services appear poised for tremendous growth, and forward-looking Dialogic products are ready for use in delivering these exciting services.

References

[CTIA] CTIA figures reported by Joseph Palenchar in “Cellular Penetration Hits Record Despite Net-New Subs Drop” (TWICE, 11/5/2007) at http://www.twice.com/article/CA6497408.html. See also CTIA filing with the FCC at http://files.ctia.org/pdf/080108_US-OECD_10_Comparison_Ex_Parte.pdf.

[Economist] *The Economist Pocket World in Figures 2007 Edition*. London: Profile Books, 2006.

[Insight] *IP-Based Application Services Market 2006-2011*. The Insight Research Corporation (November 2006).

[In-Stat] Figures quoted in In-Stat press release “Global Broadband Subscriber Base to Nearly Double by 2011” at http://www.instat.com/press.asp?Sku=IN0703510MBS&ID=2016.

[Mortensen] Interview with Mark Mortensen in Rich Grigonis, “VPIsystems on ‘WiFi Band-Aids’ for Cellular Networks,” October 29, 2007 at http://www.tmcnet.com/wifirevolution/articles/13410-vpisystems-wifi-band-aids-cellular-networks.htm.

Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project 2
AdvancedTCA	Advanced Telecom Computing Architecture
AMPS	Advanced Mobile Phone Service
ARPU	Average Revenue Per User
ATCA	Advanced Telecom Computing Architecture
ATM	Asynchronous Transfer Mode
AVC	Advanced Video Coding
CAGA	Compound Average Growth Rate
CDMA	Code Division Multiple Access
CDMA2000	Code Division Multiple Access 2000
CDMA2000 1X	First CDMA2000 standard
CDMA2000 1xEV	EVolution (EV) of CDMA2000 1X
CDMA2000 1xEV-DO	1X-EVolution-Data Optimized
CDMA2000 1xEV-DV	EVolution-Data and Voice
CDMA2000 1XRTT	One-Carrier Radio Transmission Technology
CDMA3xRTT	Three-Carrier Radio Transmission Technology
D-AMPS	Digital Advanced Mobile Phone Service
DMB	Digital Multimedia Broadcasting
DSL	Digital Subscriber Loop
DTM	Dual Transfer Mode
DVB-H	Digital Video Broadcasting-Handheld
DVB-IPDC	DVB-IP Datacast
DVD	Digital Versatile Disc or Digital Video Disc

Acronyms (continued)

DVG-T	Digital Video Broadcasting-Terrestrial
EDGE	Enhanced Data rates for GSM Evolution
ETSI	European Telecommunications Standards Institute
FTA	Free-to-Air
GPRS	General Packet Radio Service
GSM	Global System for Mobile communications
GSMA	Global System for Mobile communications Association
HSCSD	High-Speed Circuit-Switched Data
HTTP	HyperText Transfer Protocol
IM	Instant Messaging
IMS	IP Multimedia Subsystem
IP	Internet Protocol
ISDB-T	Integrated Services Digital Broadcasting-Terrestrial
ISDN	Integrated Services Digital Network
IVVR	Interactive Voice/Video Response
IVR	Interactive Voice Response
JPEG	Joint Photographic Experts Group
MediaFLO	Media Forward Link Only
MMS	Multimedia Messaging Service
MPEG	Moving Picture Experts Group
MPLS	MultiProtocol Label Switching
MRFC	Media Resource Function Controller
MRFP	Media Resource Function Processor
NMT	Nordic Mobile Telephone
NSP	Network Service Provider
OMA IMPS	Open Mobile Alliance Instant Messaging and Presence Service
PC	Personal Computer

Acronyms (continued)

PDA	Personal Digital Assistant
PDC	Personal Digital Cellular
PSTN	Public Switched Telephone Network
PTT	Push-to-Talk
PTV	Push-to-Video (also P2video)
QoS	Quality of Service
SS7	Signaling System 7
RSVP	Resource reSerVation Protocol
RTP	Real-time Transport Protocol
RTCP	Real Time Control Protocol
RTSP	Real Time Streaming Protocol
S-DMB	Satellite-Digital Multimedia Broadcasting
SIP	Session Initiation Protocol
SMS	Short Message Service
SMPTE	Society of Motion Picture and Television Engineers
TACS	Total Access Communications System
T-DMB	Terrestrial-Digital Multimedia Broadcasting
TDMA	Time Division Multiple Access
UHF	Ultra High Frequency
UMTS	Universal Mobile Telecommunications System
W-CDMA	Wideband-Code Division Multiple Access
VCEG	Video Coding Experts Group
VHF	Very High Frequency
VoIP	Voice over IP
VS	Video Share
XMPP	eXtensible Messaging and Presence Protocol