Payoff

Despite the ready availability of multimedia applications, most organizations have been unable to meet the requirements for effectively and efficiently distributing multimedia information. By reviewing the business issues involved in the design of multimedia networks and implementation of multimedia applications, this article prepares IS managers to deliver multimedia applications over the corporate network while providing for guaranteed quality of service.

Introduction

As a highly effective method of communication that simultaneously provides several forms of information to the user—audio, graphics and animation, full-motion video, still images, and text—multimedia offers a departure from the communication confines present in other singular-media applications. Advances in multimedia technology and widespread acceptance of the technology in the business community are driving the need to effectively and efficiently distribute multimedia applications.

Users today have easy and inexpensive access to multimedia-capable equipment. The explosive growth of the Compact Disk Read-Only Memory applications such as multimedia data bases, and the World Wide Web indicate the ease with which multimedia applications proliferate in an organization. Despite this ready availability, however, most multimedia applications are currently not distributed and work on single desktop computers. Most organizations have been unable both to keep up with the requirements to distribute multimedia information and to realize the goal of network computing—to build an infrastructure that supports a cooperative enterprise environment. This article identifies common multimedia applications and business considerations in effective distribution of multimedia information. Article 5-04-15.1 provides an overview of multimedia networking technologies.

Business Drivers of Multimedia Applications

The primary function of multimedia in most applications is as an interface; it allows an unhindered and manageable flow of information to the user that is consistent yet flexible in design and enables the user to accommodate varied work flows. Multimedia must therefore not be a barrier to information transfer; it must be conducive to it.

The following several technology advances have improved multimedia's ability to effectively transfer information and are driving the development of multimedia applications:

- Availability of multimedia-authoring applications.
- Continuously decreasing cost/memory of computer systems.
- Sustained improvements in microprocessor designs that enable multimedia-compression algorithms to be committed to silicon.
- Availability of network and communications equipment that facilitates the storage, management, and transfer of multimedia objects.

- Availability of improved input/output devices such as scanners, voice-and handwriting-recognition systems, and virtual reality.

The most important business driver of multimedia applications is teleconferencing, which saves organizations travel costs by enabling geographically dispersed individuals and groups to communicate in real-time. Multimedia applications development is also driven by the need to access information on the World Wide Web, demand for information and training systems, improved customer services such as advertising and public advice systems, and improved enterprise effectiveness such as correspondence management.

**Applications of Networked Multimedia**

For the purposes of this article, the applications of networked multimedia are divided into two categories:

- People-to-people applications, which assist in interpersonal communication.
- People-to-information server applications, in which the user generally interacts with a remote system to receive information or interact with a server. For example, in a WWW application, clients interact with a web server to provide information to the user.

In addition, networked applications are also classified on the basis of time; they are either immediate or deferred. Immediate applications, in which a user interacts with another person or computer in real-time, must meet latency or delay requirements. Deferred applications imply that the user is interacting with the other user or server in a manner that does not have latency or delay requirements. Messaging applications such as E-mail and voice mail are people-to-people applications in the deferred category.

A useful test for determining whether an application is immediate or deferred is whether the user is only working on the one application (which would make the application immediate) or can move to other applications during its use (making the application deferred). Exhibit 1 depicts the categories of networked multimedia applications.

### Categories of Networked Multimedia Applications

<table>
<thead>
<tr>
<th></th>
<th>Immediate</th>
<th>Deferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>People-to-People</td>
<td>Telephony, Multimedia Conferencing</td>
<td>E-Mail, Voice Mail</td>
</tr>
<tr>
<td>People-to-Information Server</td>
<td>WWW Browsing, Video-on-Demand</td>
<td>File Transfer</td>
</tr>
</tbody>
</table>

---

The following sections examine multimedia applications and their networking requirements. The focus is on immediate applications, because they place greater demands on the network system than do deferred applications.

**People-to-Server Applications**

**Video-on-Demand**

Video-on-demand applications let users request a video from a remote server. In a business setting, this approach is applicable for downloading training videos and viewing them on the desktop. If the videos are stored on a server, video-on-demand can be used for just-in-time training.

Because the disk requirements to download and store a video (before presentation) are expensive, video-on-demand has strict real-time requirements on transmissions but tolerates an initial delay in the playback of the data. Transmission guarantees require that the bandwidth be available for the entire duration of the video. The delay requirements are not stringent, because some data can be buffered at the receiving end.

**WWW Browsing**

The World Wide Web is a system that allows clients to retrieve and browse through hypertext/hypermedia documents stored on remote servers. Currently most material on the Web is text- or graphic-based; the data storage and transmission requirements associated with audio and video allow only a few users to add these features. Video access is also awkward because the video player that supports the video format (i.e., software plug-in) must exist on the client.

Many businesses, however, use the Web to provide organizational information and to support access to data bases (mostly internal to the company). The Web creates different challenges for business organizations, which must conduct comprehensive capacity planning (i.e., of bandwidth, types of applications, and security) before roll out. It is bandwidth and the number of simultaneous users accessing the service that play a critical role in Web site management. Applications providing multimedia data require greater bandwidth that supports transmission-delay requirements.

**People-to-People Applications**

Two common people-to-people applications are multimedia conferencing and groupware.

---

Multimedia Conferencing

Multimedia conferencing is used by many businesses to support the collaborative efforts of growing numbers of virtual groups and organizational teams. The benefits of multimedia conferencing include:

- Reduction or elimination of costly business trips.
- Facilitation of collaborative work such as computer-aided design/computer-aided manufacturing CAD/CAM
- Increased productivity resulting from the ability to meet deadlines and less chance for miscommunication.

Types of Multimedia Conferencing Systems

There are three main types of multimedia conferencing systems: point-to-point, multipoint, and multicast.

**Point-to-Point Systems.**

Point-to-point systems involve two persons communicating interactively from the desktop or groups of people communicating from a conference room. Point-to-point desktop conferencing systems are becoming popular because of the availability of inexpensive, low-overhead digital cameras such as the Quickcam and associated videoconferencing software. These systems let users share screens of data, text, or images.

**Multipoint Systems.**

Multipoint conferencing involves three or more locations that are linked either through a local area network or a wide area network and can each send and receive video. Such systems have several unique characteristics, including presentation of multiple media, management and transport of multiple media streams, distributed access to multiple media, high bandwidth requirements, and low-latency bulk data transfer and multipoint communications. Because desktop systems quickly run out of screen space, multipoint conferencing is more effectively conducted in conference rooms with video walls.

**Multicasting Systems.**

Multicasting involves the transmission of multimedia traffic by one site and its receipt by other sites. Rather than sending a separate stream to each individual user (i.e., unicasting) or transmitting all packets to everyone (i.e., broadcasting), a multicasting system simultaneously transmits traffic to a designated subset of network users. Many existing systems use broadcasting and let the receivers sort out their messages. This inefficient practice fails to maximize use of network bandwidth and poses potential security problems.
**Groupware**

The notion of conferencing is changing to include such features as shared windows and whiteboards enabled by distributed computing. In addition, the use of computer mediation and integration is increasing.

A shared application or conferencing system permits two or more users at separate workstations to simultaneously view and interact with a common instance of an application and content. With such applications, users working on a report, for example, can collectively edit a shared copy. In general, documents used in groupware are active (i.e., the document displayed on the screen is connected to content in a data base or spreadsheet).

Groupware provides such features as support for group protocols and control, including round-robin or on-demand floor-control policy, and both symmetric and asymmetric views of changes. In symmetric view, a change that is made is immediately shown to other users. In asymmetric view, applicable in applications involving teacher-student or physician-patient interactions, the changes made in one window may not be shown in another window. Groupware systems also support such issues as membership control (i.e., latecomers) and media control (i.e., synchronization of media).

**Technical Requirements for Networked Multimedia Applications**

The immediate multimedia applications discussed (i.e., video-on-demand, multimedia conferencing, groupware, and web browsing) have several technical requirements.

**Latency**

Latency refers to the delay between the time of transmission from the data source to the reception of data at the destination. Associated with delay is the notion of jitter. Jitter is the uncertainty of arrival of data. In the case of multimedia conferencing systems, practical experience has shown that a maximum delay of 150 milliseconds is appropriate. Synchronous communications involve a bounded transmission delay.

**Synchronization**

Existing networks and computing systems treat individual traffic streams (i.e., audio, video, data) as completely independent and unrelated units. When different routes are taken by each of these streams, they must be synchronized at the receiving end through effective and expeditious signaling.

**Bandwidth**

Bandwidth requirements for multimedia are steep, because high data throughput is essential for meeting the stream demands of audio and video traffic. A minimum of 1.5M bps is needed for MPEG2, the emerging standard for broadcast-quality video from the Moving Picture Experts Group. Exhibit 2 depicts the storage and communications requirements for multimedia traffic streams.

---

### Reliability

The high data-presentation rate associated with uncompressed video means that errors such as a single missed frame are not readily noticeable. Most digital video is compressed, however, and dropped frames are easily noticeable. In addition, the human ear is sensitive to loss of audio data. Hence, error controls (such as check sums) and recovery mechanisms (i.e., retransmission requests) need to be built into the network. Adding such mechanisms raises a new complexity, because retransmitted frames may be too late for real-time processing.

### Guaranteeing Quality of Service

Quality-of-service guarantees aim to conserve resources. In a broad sense, quality of service enables an application to state what peak bandwidth it requires, how much variability can it tolerate in the bandwidth, the propagation delay it is sensitive to, and the connection type it requires (i.e., permanent or connectionless, multipoint). The principle of quality of service states that the network must reliably achieve a level of performance that the user/application finds acceptable, but no better than that. Network systems can either guarantee the quality of service, not respond to it, or negotiate a level of service that they can guarantee.

Quality of service has several components, which are depicted in Exhibit 3 and described in the sections that follow.46

---

Quality-of-Service Components in Networked Multimedia Applications

Application Parameters

Application quality-of-service parameters describe requirements for applications, such as media quality and media relations. Media quality refers to source/sink characteristics (e.g., media data-unit rate) and transmission characteristics (e.g., end-to-end delay). Media relations specifies media conversion and inter-and intrastream synchronization.

System Parameters

System quality-of-service requirements are specified in qualitative and quantitative terms for communication services and the operating system. Qualitative parameters define the following expected level of services:

- Interstream synchronization, which is defined by an acceptable skew relative to another stream or virtual clock.
- Ordered delivery of data.
- Error-recovery and scheduling mechanisms.

Quantitative parameters are more concrete measures that include specifications such as bits per second, number of errors, job processing time, and data size unit.

Network and Device Parameters

Network quality-of-service parameters describe requirements on the network, such as network load (i.e., ongoing traffic requirements such as interarrival time), and performance or guaranteed requirements in terms of latency and bandwidth. In addition, traffic parameters such as peak data rate, burst length or jitter, and a traffic model are specified. Traffic models describe arrival of connection requests or traffic contract based on calculated expected traffic parameters.

Device quality-of-service parameters typically include timing and throughput demands for media data units.

Determining Service Levels

Several different types of service levels can be negotiated, including:

- Guaranteed service, which establishes quality of service within deterministic or statistical bounds.
- Predictive service, which estimates quality of service from past behavior.
- Best-effort service, which is used in the absence of available quality parameters.
System Considerations

Designing networks to support multimedia applications involves more than just networking requirements; attention must also be paid to the entire system. Network configurations, for example, do not treat how the bandwidth is handled once it reaches the desktop. Bus speeds and input/output operation throughput are part of the link between the data source and the users’ screen. There are two possible approaches to handling bus speeds and throughput:

- Faster bus and I/O hardware.
- Desktop LAN that eliminate the workstation bus altogether and replace it with an internal packet switch linking the motherboard and peripheral system.

Bandwidth is also handled through compression techniques for images and video that radically reduce the amount of data transmitted and consequently lower bandwidth requirements. Multimedia information is bursty, meaning that some parts of it require higher bandwidth than others. Dynamic bandwidth allocation is useful to lessen the network burden.

Another consideration involves the accommodation of real-time requirements by the operating system. For example, the jitter and slowdown of a movie player may not be due to availability of resources but to a lack of proper scheduling. A music play program often pick ups speed when contending programs terminate. Principal requirements for multimedia-capable operating systems are as follows:47

- Operating-system resource management must now be based on quality of service and respond to a new class of service that satisfies time constraints and negotiable service levels.48
- Real-time CPU scheduling, memory buffer and file management policies to support real-time processing, and support for real-time synchronization.
- Support for standard applications in addition to real-time multimedia applications.
- Low-overhead task management resulting from the need for frequent switching.

Barriers to Multimedia Networking

The extensive bandwidth and storage required to transmit multimedia streams coupled with the insufficient bandwidth of existing networks pose one of the major barriers to multimedia networking. The tendency of existing networks to treat individual streams as independent and unrelated units underscores the challenge of and need for effective synchronization.

Another major roadblock in networking existing applications is caused by proprietary development environments, data formats, and runtime environments, and by incompatible proprietary client-server models. The tight coupling among existing devices, data formats, and application program interfaces API makes it even more difficult to devise a standard.

Heterogeneous delivery platforms pose networking challenges even when multimedia applications are not involved.

Other related concerns that aggravate existing problems associated with networked multimedia applications result from the lack of uniform standards in the following areas:

- **Data capture and recording.** Uniform data formats for graphics, sound, music, text, video, animation, and still images are needed.

- **Data compression and decompression.** Although there is no predominant standard, several organizations have proposed data compression and decompression standards. Among them are MPEG2 and Joint Photographic Experts Group (Joint Photographic Experts Group).

- **Media storage and retrieval.** In the case of CD-ROMs, digital video discs are needed for portability across platforms.

- **Edit and assembly.** Content description and container standards are either specified as universal object types or through the conventional method of surrounding dissimilar information types with wrappers. Scripted, structural tagging, identification tagging, and other language constructs have emerged as the conventional tools for cross-platform applications development and parameter passing.

- **Presentation.** Uniform customizable presentation standards are needed for maintaining a common look-and-feel across platforms.

- **Transfer or networking.** Asynchronous transfer mode Asynchronous Transfer Mode and TCP/IP-NG are emerging standards in this area.

- **Multimedia signaling.** Setting up a multimedia conference automatically, without being routed through a common conference bridge, requires a sequence of signaling events, such as call setup/connect and messages. New approaches are needed to support the capability of originating, modifying, and terminating sessions in a network.

### Issues in Multimedia Systems Planning

The challenge in networking applications is to develop a strategy that works with existing technology and enables management to provide gradual enhancements to the existing infrastructure. Both scalability and integration must be considered during the planning process. In terms of network management support, the technology chosen should support the entire infrastructure.

#### Scalability

The network must be capable of scaling smoothly when new nodes, or applications are added; the goal is to simply add resources to the system rather than change the technology. A desirable form of scalability is that resource costs be linear in some measure of performance or usage.
Integration

Networked applications are designed along the principles of either vertical or horizontal integration. In vertical integration, a dedicated infrastructure is designed for each application so that, for example, the telephone network is kept separate from the computer network. In contrast, horizontal integration is based on open standards and provides for complexity, management, and portability. It has the following characteristics:

- Use of integrated networks that handle data, audio, and video and are configured to meet application requirements.

- Use of middleware software to provide, for example, directory and authentication services. Although the underlying network may be heterogeneous, middleware services provide a set of distributed services with standard programming interfaces and communication protocols.

- The user is provided with a diverse set of applications. 49

Application/Content Awareness

In general, vertically integrated networks are application-aware (e.g., videoconferencing networks in general know the media type), and horizontal networks are application-blind (e.g., the Internet). Networks are also sometimes content-aware (e.g., a video-on-demand network knows what video is being downloaded).

Planning Strategies

Several approaches to multimedia systems planning are available. 50 Assuming that network resources will grow to meet the demand of the most-stringent combined user-applications, best-effort schemes are sufficient as currently used. Examples include low-profile, low-cost videoconferences, multimedia E-mail, and downloadable files. This approach, however, does not take the entire multimedia system into account.

Another alternative is to overengineer the networks, making bandwidth shortage a rare problem and access to services almost always available. This approach entails providing the user with the latest technologies and a high cost. In addition, leapfrogging computer power and applications resource requirements makes this approach highly susceptible to problems.

A third and generally more effective approach is to use either quality-of-service parameters or resource reservation, or both. For example, high-profile users (e.g., users of investment banking and medical applications) require that some form of resource reservation occur before the execution of multimedia applications.

Conclusion

As networking evolves and voice, video, and data are handled together, networking systems are expected to handle data streams with equal efficiency and reliability from the temporal, synchronization, and functional perspectives. Technology advancements have

49 Messerschmidt, pp. 1167-1186.
made users less tolerant of response delays, unreliable service, and lossy data occurring either at the network or at the desktop system.

For these reasons, it is essential that multimedia networks be designed with all computing resources in mind and provide some resource reservation or quality-of-service guarantee. The task before organizations is to find the most flexible, cost-effective method for delivering multimedia applications over networks while also providing for guaranteed quality of service.\textsuperscript{51}

**Author Biographies**

T.M. Rajkumar  
T.M. Rajkumar is associate professor of MIS at Miami University in Oxford OH.

Amitava Haldar  
Amitava Haldar is a senior systems analyst for AIC Consulting in Chapel Hill NC currently working for the network software division of IBM.

---
