Design Issues for Interactive Television Systems

Borko Furht
Florida Atlantic University

Deven Kalra and
Frederick L. Kitson
Hewlett-Packard

Arturo A. Rodriguez and
William E. Wall
Scientific Atlanta

Can TV sets ever be made interactive? It may not be a question of "if" but "when." Learn how different designs are being investigated toward that goal.

Computing, communication, and relevant standards are on the brink of enabling thousands of people to enjoy the services offered by large, distributed multimedia systems in their own homes. Collectively, these services will include

- TV (basic, subscription, and pay-per-view),
- service navigator,
- interactive entertainment,
- digital audio,
- video-on-demand,
- home shopping,
- financial transactions,
- interactive single- and multiuser games,
- digital multimedia libraries, and
- electronic versions of newspapers, magazines, TV program guides, and yellow pages.

It is obvious that current TV systems and architectures must be redesigned to support such services. In this article, we propose potential solutions to modify existing systems to support these new functions, and we discuss related matters.

Cable TV and telephone companies, as infrastructure and content providers, are realizing that potential network technologies, coupled with improved computing and compression techniques, will soon profitably deliver interactive services. Recently, entertainment, cable, phone, and computer companies have formed alliances to design a variety of wide-area multimedia infrastructures. Consequently, universities and industrial laboratories have been working intensively to define a suitable architecture, and a number of groups have developed applicable standards. From these activities, we can see the emerging information infrastructure.

A distributed multimedia system architecture that can support on-demand, interactive TV applications is a hierarchical configuration of multimedia servers and network switches, as we show in this article. In such a system, multimedia data must be compressed, stored, retrieved, transmitted over the network to its destination, then decompressed and synchronized for playback at the receiving site.

After describing a general architecture, we briefly describe several network topologies, now in use, in terms of their applicability to interactive TV (ITV) systems. We then address the main issues in designing a terminal device for a multimedia network and examine potential hardware and software architectures. This device, called a TV *set-top box* (STB), performs two functions. It decodes the information (for example, video and audio) at the subscriber (customer) premises and provides subscribers with interactive capabilities.

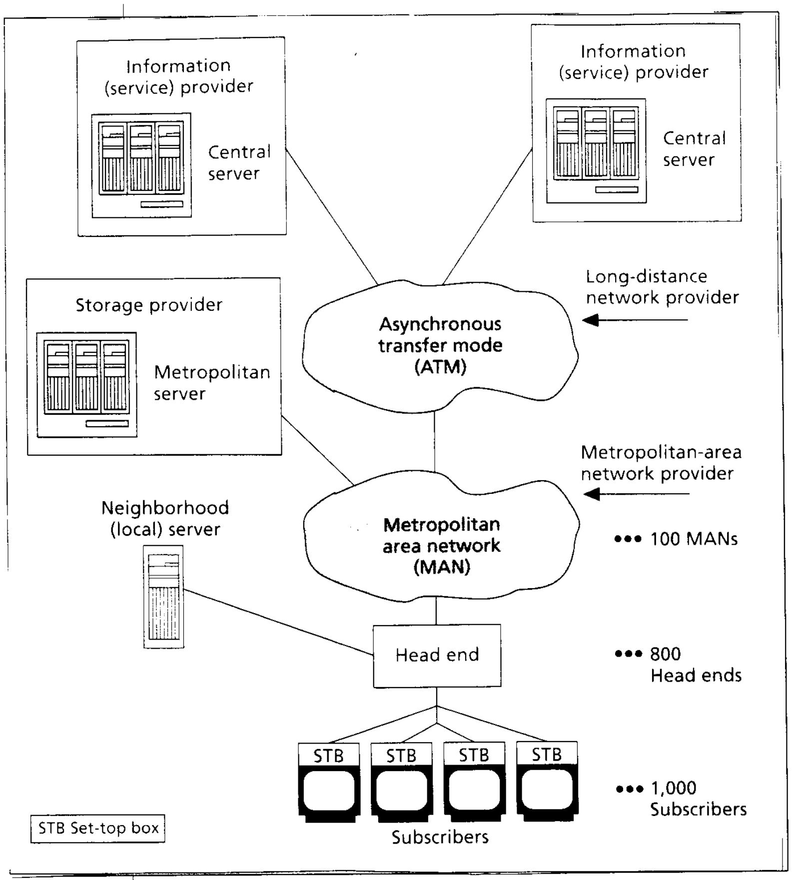

Figure 1. Hierarchical configuration of multimedia servers and network switches for interactive on-demand TV services. A typical system will be able to support about 80 million subscribers in the continental US (one ATM × 100 metropolitan area networks × 800 head ends × 1,000 households).

AN ARCHITECTURE FOR INTERACTIVE TELEVISION

The three main components of an ITV architecture are content (information) servers, a network, and STBs. Content servers are connected to the STBs at the subscriber premises through a network consisting of switches and transmission medium. The transmission medium can be a coaxial cable or a fiber-optic channel. Wireless technology is also being investigated for delivering data to subscribers.

Cable companies already have a large network of coaxial cable, albeit for one-way delivery of video. Although telephone companies are upgrading their long-distance network with fiber optics, the telephone cable connecting to each home is predominantly twisted-pair copper—fiber is still too expensive. The most feasible near-term solution is a *hybrid fiber-coax* (HFC) network, where fiber connects to a small neighborhood node that has a coaxial cable connection to each home.

In a general architecture for deploying interactive multimedia services, the servers are connected to the *head ends* of cable network CATV (community antenna television) trees via a wide area network, most likely of the asynchronous transfer mode (ATM) variety. A head end has slightly different meanings depending on context. In cable TV, a head end is where incoming programming is received by the cable company's satellite dish and TV antenna. In telephone networks, the head end might be a switching office; for on-line services, it's the service provider's databases.

To analyze various system aspects, Nussbaumer, Patel, and Schaffa[1] developed a model that examines (for example) the number of subscribers supported by a CATV-tree, the overall system utilization, the requirements of the broadband WAN, and the required server capacity.

A proposed hierarchical system

The basic architecture can be expanded into a hierarchical configuration of multimedia servers and network switches, as proposed by Ramanathan and Rangan.[2] The system, shown in Figure 1, consists of information

Acronyms

ADSL—Asymmetric digital subscriber line
ATM—Asynchronous transfer mode
BERLU—Broadband-enhanced remote line unit
CATV—Community antenna television
CCITT—Consultative Committee for International Telephone and Telegraphy
FDDI—Fiber Distributed Data Interface
FDM—Frequency-division multiplexing
HAL—Hardware abstraction layer
HDSL—High-speed digital subscriber line
HDTV—High-definition television
HFC—Hybrid fiber coax
ISDN—Integrated Services Digital Network
MIDI—Musical Instruments Digital Interface
MPEG—Moving Pictures Experts Group
NTSC—National Television System Committee
POTS—Plain old telephone service
QAM—Quadrature amplitude modulation
QPSK—Quadrature phase-shift keying
RAID—Redundant arrays of inexpensive disks
Sonet—Synchronous Optical Network
STB—Set-top box
VIP—Video information provider
VTE—Video transfer engine

Glossary

Asynchronous transfer mode—A packet-oriented transfer protocol, ATM was proposed independently by Bellcore and several European telecommunication companies. Information is organized into fixed-length, 53-byte packets called cells and transmitted according to each user's instantaneous need. Each cell comprises a 5-byte header and a 48-byte payload. The header, which contains routing information, identifies cells belonging to the same "virtual channel." Bandwidth in ATM is allocated on demand (as in packet switching), and transmission occurs in fixed-length cells (as in time-division multiplexing).

Asynchronous transfer network—In 1988, the CCITT decided to base the development of B-ISDN on ATM, and therefore the term ATM is sometimes used synonymously with B-ISDN. However, B-ISDN is just one of the services that can use ATM transport technology. An ATM network comprises a set of end systems (terminals) and a set of intermediate nodes (switches), joined by a set of point-to-point ATM links. The two major interface types in ATM networks are the user-to-network interface and the network-to-network interface.

Broadband networks—These networks are characterized by data transfer speeds greater than 45 Mbps.

Fiber Distributed Data Interface—This is a multimodal fiber version of IBM's token ring LAN (IEEE 802.5). It has a bandwidth of 100 Mbps covering a distance to 100 km. FDDI II includes an isochronous, circuit-switched capability with 16 wideband channels, which can be suballocated as $n \times 8$ Kbps.

Integrated Services Digital Network—ISDN expands the access and signaling of the public switched telephone networks' basic technology, which increases the applications supported by the PSTN, particularly in nonvoice communications. In ISDN, the local loop connection between a subscriber and a switch is digital, with multiple multiplexed information channels supported per access. These channels are logically divided into Bearier (or B-) channels, used to convey user information, and signaling (or D-) channels. The ISDN basic rate interface comprises two B-channels and one D-channel, and the ISDN primary rate interface comprises 23 B-channels and one D-channel.

A *B-channel* is a 64-Kbps (called DS-0) ISDN user-to-network channel that carries a voice, a data, or an image call, but not the signaling for the call.

N-ISDN (Narrowband ISDN) includes a basic interface (2B + D) and primary rate interface (23B + D). It is copper based with rates at or below 1.5 Mbps.

B-ISDN (Broadband ISDN) is a network standard approved by CCITT and the ANSI T-1 committee. It extends N-ISDN with new services and provides voice, data, and video in the same network. It is fiber based with rates at 150 and 600 Mbps.

Local area network—A LAN is a private computer network connecting computers in the same building or campus using coaxial cable, twisted pair, or multimodal fiber. Examples include IEEE 802.3 Ethernet, IEEE 802.4 Token Bus, and IEEE 802.5 Token Ring Arcnet.

Metropolitan area network—This network connects computers in the same city. For example, an ANSI X3T9 FDDI network can be a private MAN, while an IEEE 802.6 network can be a public MAN.

Network characteristics and application demands:

Peak bandwidth—This is the maximum information capacity the application will require per connection.

Bandwidth variability—This indicates whether or not the application requires flexible bandwidth up to the peak level.

Propagation delay—This is the delay that information will experience in traveling from source to destination. Sensitivity to delays varies considerably from application to application.

Connection type—This specifies whether the application has a long-term relationship (permanent connection), a bounded short-term relationship (switched connection), or a boundary-less connection (connectionless).

Sonet (Synchronous Optical Network)—Sonet is a new family of optical transmission channels with speeds currently ranging from 45 Mbps to 2.4 Gbps.

STS-1, STS-3, STS-3c, STS-N (synchronous transport signal levels)— STS is the electrical building block for a Sonet network. The rate of STS-1 is 51.840 Mbps, while STS-3 is 155.520 Mbps. STS-3c combines three STS-1 envelopes into one 150.336-Mbps payload containing an overhead of only one path. It is suitable for three or more NTSC-quality TV signals or one HDTV (compressed) signal. The rate of STS-N is $n \times 51.840$ Mbps.

T-1 carrier—This is a system using time-division multiplexing to carry 24 digital voice or data channels, each operating at 64 Kbps over copper wire. Total speed is 1.544 Mbps, which is called DS-1 (digital signal level 1). One copper wire pair carries the channels in one direction, while a second pair carries the signals in the other.

Wide area network—This network services a geographical area larger than a city or metropolitan area. A wideband service (or system) consists of transmission channels that can support rates between 1.5 and 45 Mbps.

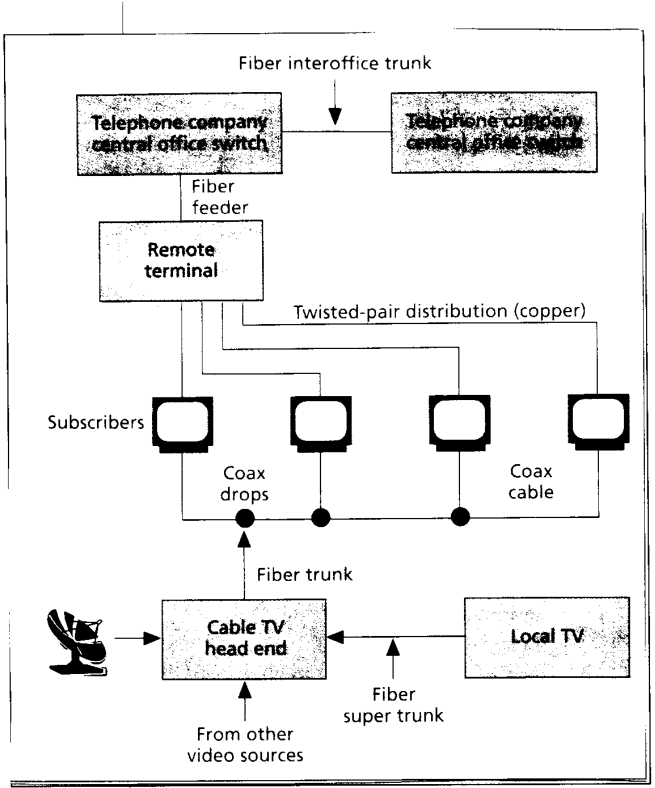

Figure 2. Present telephone and cable network architectures.

providers—such as entertainment houses and television stations—that offer various services, network providers that transport media over integrated networks, and several levels of storage providers that manage data storage in multimedia servers and that contain network switches.

Supporting millions of subscribers

As we show hierarchically in Figure 1, a future interactive ATM-based video-on-demand (VOD) system will likely support up to 100 metropolitan area networks. Each MAN will be linked to an average of 800 head ends, and each head end will support about 1,000 households, which will support a total of 80 million subscribers.[1] The literature has addressed numerous issues concerning such distributed multimedia systems. Examples include media synchronization, multimedia data storage and retrieval, video compression, and network requirements for transmitting continuous media.[3]

Multimedia data—audio, video, text—is stored in multimedia servers at information providers and typically compressed in accordance with the MPEG (Moving Pictures Experts Group) standard. Initially, information is transmitted from large content servers to storage servers (see Figure 1) via a high-speed network such as ATM or Sonet. The data is then transmitted to local servers at neighborhood hubs via MANs. Finally, it is delivered to subscribers over the access line, such as an HFC or asymmetrical digital subscriber line (ADSL), and through the TV STB.

Gateways connect STBs to servers

An STB at the customer's premises is typically connected to the servers through gateway levels. A *level-1 gateway*, similar to a directory service, gives an STB the first level of navigation into the full-service network by offering the subscriber a menu of services. The subscriber selects a service to connect to one or more *level-2 gateways*, which provide access to the multimedia servers containing the needed information.

Several scenarios exist whereby an STB communicates with a level-2 gateway. In one, when a service has been selected, the level-2 gateway communicates directly with the STB. In another scenario, the level-2 gateway always communicates through the level-1 gateway. Hence, the level-1 gateway provides the video dial tone (VDT), which is the signaling required to make a direct request or call. In a direct-call scenario, the STB can bypass the level-1 gateway in calling the level-2 gateway, assuming it has the level-2 gateway's address.

NETWORK TOPOLOGIES FOR INTERACTIVE MULTIMEDIA SERVICES

At present, telephone and cable companies use different topologies and technologies to deliver their services. The phone system is switched, symmetrical, and interactive. The typically fiber backbone ("trunk" link) carries digital signals, and twisted-pair copper wires carry analog sig-

Table 1. Desirable data superhighway features, ranked in low to high applicability according to each network and the importance for ITV.

	Network candidates			
Features	Telephone	Cable	Internet	Importance for ITV
Bandwidth	Very low	High	Very low	Very high
Affordability	High	Medium	Low	Very high
Ease of use	High	High	Low	Very high
Transaction billing	High	Medium	Low	Very high
Availability	High	Medium to low	Medium	High
Information content	Very low	Low to medium	High	High
Security	High	Very low	Low	High
Openness	Medium	Low	High	Low

nals to deliver service into homes and businesses.[4] The cable system, on the other hand, is unswitched and distributive, built on a backbone of analog fiber and satellites, using coaxial cables to connect to subscriber sites.[4] Both systems are shown in Figure 2.

A third network infrastructure possibility is the Internet, which began as a government-subsidized electronic communication network. The Internet remains inconvenient to use because it does not support billing or transmission of real-time data and is both expensive and difficult for many subscribers to access.[4,5] However, due to growing numbers of users, the Internet is still an infrastructure candidate, although a potentially unified cable and phone system could reduce the Internet's importance. What's more likely is that the Internet could become a service of the data superhighway.

In the future, local architectures of both cable and phone systems might be nearly identical. Both systems will perhaps become switched and symmetrical. A hybrid fiber-coax network will transmit two-way voice, data, and cable TV services. Backbone networks, such as the public switched telephone network or a private computer network, will be connected at the central phone office or cable TV head end.

Table 1 compares desirable features of current phone, cable, and Internet networks. These features are critical for the deployment of interactive VOD systems.[4]

Table 2 lists the media, topologies, and protocols used today by telephone systems, cable systems, and the Internet.[4] Additionally, the table includes expected media, topologies, and protocols of future unified cable and phone systems.

The different backgrounds of the main players—cable and telephone companies—are reflected in how they define the future network infrastructure for interactive TV systems.

Cable companies' viewpoint

The cable TV system in the US is presently one-way, based on analog video broadcasting through a wire. A typical 450-MHz plant uses 6-MHz analog channels, yielding a total capacity of about 70 channels. The cable system's network management and system reliability are relatively primitive. The cable systems tend to be proprietary and not interconnected. Cable companies currently emphasize broadcast-type networks that require minimal switching with near-VOD technology. These digital-analog hybrids use 750-MHz fiber-to-the-node systems to offer more analog channels and movies in staggered schedules with limited random-access capabilities for subscribers. Figure 3 on the next page illustrates cable companies' proposed migration to interactive VOD television.[6]

To allow two-way interactive communication, a return path should be incorporated into the system. Then, digital encoding and video compression must be provided. With 64 QAM (quadrature amplitude modulation), it is possible to get 27 Mbps out of a 6-MHz analog channel, while 256 QAM provides a usable bit rate of more than 40 Mbps. Assuming an MPEG-2 movie of 3.35 Mbps (including video, audio, and control data), this will extend the capacity of the current system by allowing more than 10 MPEG-2 compressed movies to be transmitted via one 6-MHz analog channel (40 divided by 3.35).

Upgrading the cable plant to 750 MHz and fiber-in-the-loop technology will serve 200 to 1,000 households. The so-called 500-channel scenario will then consist of approximately

- 70 analog 6-MHz channels (totaling roughly 450 MHz), and
- 430-plus digital and compressed channels (300 MHz/6 MHz = 50 analog channels, with each analog channel transmitting eight to 10 MPEG-2 movies).

Table 2. Media, topologies, protocols, and users of candidate networks.

	Candidate networks			
Features	**Telephone (today)**	**Cable (today)**	**Internet (today)**	**Unified cable/telephone (future)**
Media/Backbone	Digital fiber optic (97 percent)	Satellite, analog fiber optic	NSFnet (T-3), other telephone companies	Digital/analog fiber optic, satellite
Media/Local	Copper wire, wireless	Coaxial cable	Copper wire, switched T-1	Coaxial cable, fiber-optic copper wire, two-way radio
Topology	Circuit-switched, star	Unswitched, trunk and branch	Packet-switched, routed	Switched/ unswitched, star
Protocols	POTS, ISDN, ATM	Proprietary	TCP/IP	Analog, ADSL, ATM
Key users	Everyone	60 percent of US households	Government, academia, business	Everyone

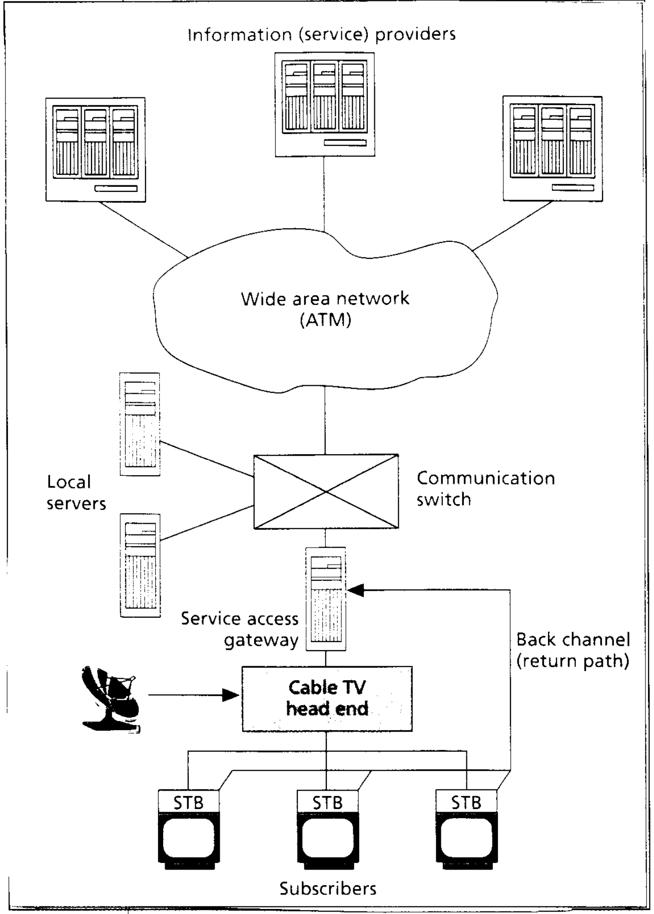

Figure 3. Cable architecture for interactive video-on-demand television.[6]

Cable's video delivery will be a mix of broadcasting and pointcasting (point-to-point, or from a service provider to a subscriber) for VOD. The TV STB is simple for broadcast-only situations, but it is expected to be elaborate and complex for VOD, providing video decompression and communication at a high rate.[6]

A major challenge in terms of expense, time, and feasibility that cable companies are now facing will be to install giant gateways, lease backbone capacity from long-distance carriers, or lay out their own digital fiber trunk lines to construct a nation-wide network.

Telephone companies' viewpoint

Telephone companies have an advantage over cable companies in that they already provide both local and long-distance point-to-point communications. The phone system represents the world's largest switched, distributed network and can handle millions of phone calls simultaneously.[4] However, the phone system's main drawback is low bandwidth. Although the cross-country trunk lines are high-capacity fiber, the local loops to homes and businesses are typically two- or four-wire unshielded copper with limited bandwidth.[4] Figure 4 illustrates the telephone companies' outlook for a future interactive VOD television network infrastructure.[6]

The network in Figure 4 is based on pointcast communication between server and clients (STBs), which tends to favor a server-intensive implementation. Telephone companies could potentially employ client/server STBs with fully switched networks, which would probably be more complex. Although an all-digital STB would be less complex than a hybrid analog-digital STB, it is unlikely that an all-digital STB would be useful because analog support will be required for years to come. The main challenge for telephone companies will be to increase the bandwidth of current local loops.

Access technologies

Let's take a look at several promising technologies now in use by cable and telephone companies for ITV architectures.[7]

HFC. Hybrid fiber coax, an analog access technology for CATV signals, is potentially the emerging standard for both cable and telephone companies. HFC systems typically provide a forward-path frequency range of 50 to 750 MHz and a reverse path of 5 to 30 MHz. Digital transmission is attained by modulating the digital information in a packet format onto analog RF carriers, via QAM. QAM is bit efficient because it generates 4 bits or more out of 1 baud in the forward path and applies quadrature phase-shift keying modulation, which is more robust, in the reverse direction. (See Oetting[8] for a description of modulation techniques.)

HFC's physical architecture typically consists of a gateway device, located at a cable head end or telephone company's central office, that provides RF modulation and termination of the backbone digital network (typically Sonet with ATM). The forward-path composite RF signal, consisting of modulated analog and digital signals, is routed over an analog fiber-optic link to a neighborhood fiber-optic terminating node that typically serves 500 homes. From the fiber-optic node, coax is routed over a star-star architecture to a feeder cable that passes roughly 40 homes. Individual taps connect to those homes.

With multiple access technology, reverse-path RF transmission lets all homes share the available reverse bandwidth. A separate fiber, from the neighborhood node back to the head end or central office, carries the reverse RF path.

Two termination methods at the customer premises are possible. One is a network interface device on the outside of a home. This provides the home with RF modem capabilities to distribute analog video over coax and a LAN to distribute digital data. The second method, now being deployed, directly distributes the coax within the home and has RF modem capability in the STB. Individual home connectivity is accomplished by dynamically assigning both an RF channel and a packet destination to an STB.

CATV technology uses a broadband coaxial cable system for transmitting multiple MPEG-compressed video streams. CATV provides high bandwidth and supports hundreds of simultaneous connections although, as discussed earlier, it requires adaptation to support the two-way interactive communication needed for VOD systems. CATV remains an attractive VOD solution for two reasons:

Roughly 65 percent of US households are already connected to cable, with 90 percent anticipated[7]; and optical fiber is too expensive for wide deployment.

ADSL. Asymmetric digital subscriber line technology[9] enables the phone company to offer affordable VOD and interactive services because existing twisted-pair copper can be used. The term "asymmetric" in ADSL refers to a much higher data rate in the downstream direction (to the subscribers) than in the upstream direction (from the subscribers).

ADSL is a consumer service intended for applications that include transmission of compressed TV-quality video with distribution over almost the entire loop plant. ADSL employs frequency-division multiplexing to transmit a 1.536-Mbps wideband signal downstream, a 16-Kbps subscriber control signal upstream, and a full duplex POTS (plain old telephone service) or 2B+D ISDN signal on a single, copper twisted pair, as illustrated in Figure 5.[9] The downstream wideband signal has an upper limit of 6 Mbps.

At the telephone company's central office, an ADSL system performs two functions: demultiplexing the upstream POTS signal and control channel, and multiplexing POTS with the high-bit-rate signal in the downstream direction. At the subscriber level, the incoming signal is demultiplexed into the POTS channel, the high-bit-rate channel, and the low-bit-rate channel.

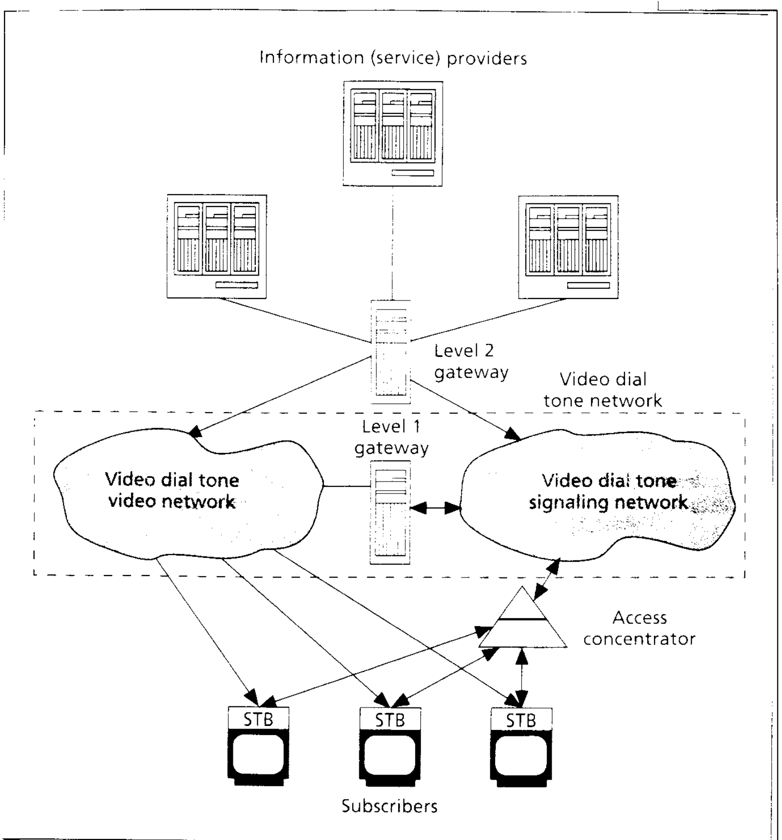

Figure 4. Telephone architecture for interactive video-on-demand television.[6]

There are several possible scenarios for locating ADSL equipment and network interfaces. In one, ADSL and service modules (SMs) are both located in a pedestal in the outside plant. This architecture simplifies upgrading a copper-based network to a fiber-based optical network. In another scenario, the subscriber owns the equipment that provides SM functionality, while the ADSL equipment is provided by the network. The ADSL equipment separates the POTS wiring from the SM interface wiring. In a third scenario, all equipment is plugged into a telephone interface at the subscriber premises. This architecture requires special filters to separate ADSL from telephone services.

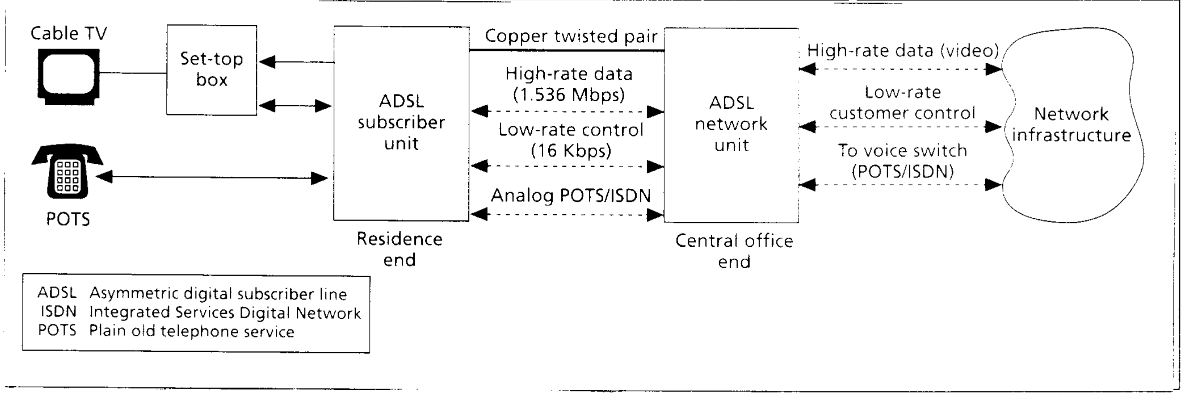

Figure 5. Concept of an asymmetric digital subscriber line. ADSL uses frequency-division multiplexing to transmit a downstream 1.536-Mbps wideband signal, an upstream 16-Kbps subscriber control signal, and a full duplex plain old telephone service or 2B+D ISDN signal on a single, copper twisted wire.

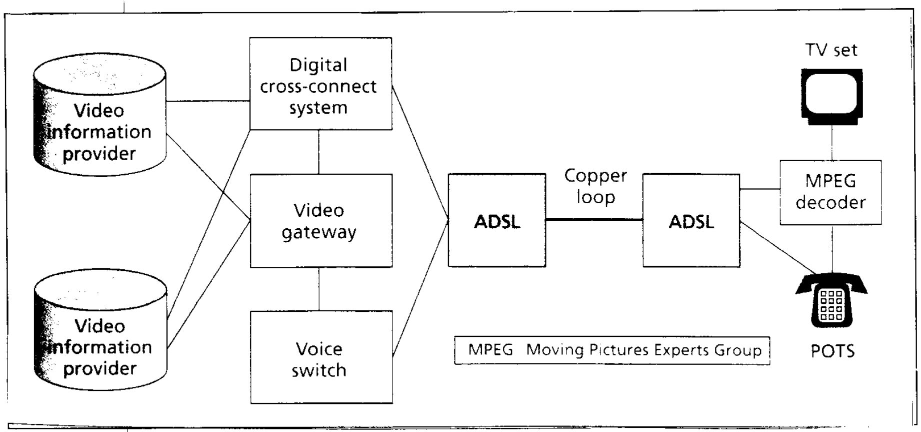

Figure 6. Bell Atlantic's architecture for video-on-demand services.[9] ADSL delivers one channel of MPEG-compressed movies to a single user. A digital cross-connect system switches connections between video information providers and subscriber loops.

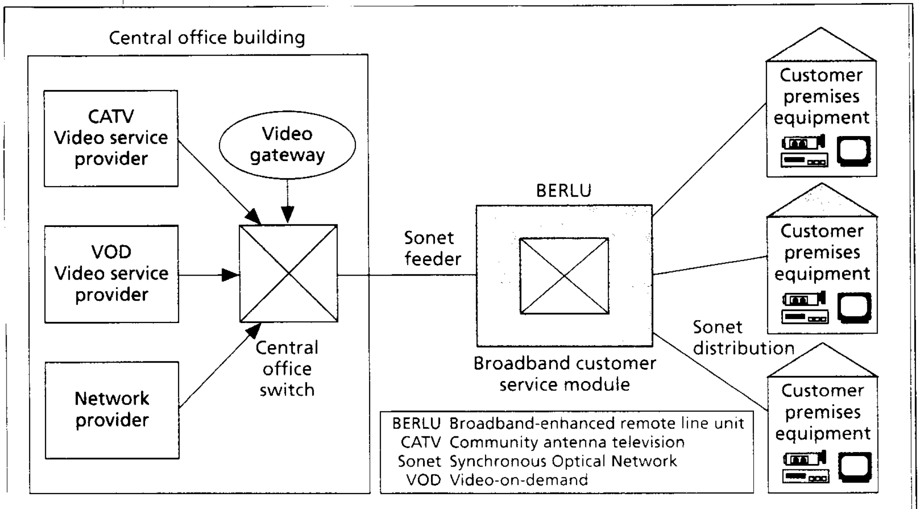

Figure 7. Architecture of GTE's system for video-on-demand services.[11]

One of the first commercial VOD systems, employed by Bell Atlantic and targeted at entertainment, allows residential subscribers to access a video information provider's database to view different movies. The interactive system also offers home shopping and education. (The system architecture appears in Figure 6.[9]) The system, serving about 2,000 subscribers, is being tested in northern Virginia.

The Bell Atlantic system's front-end is a video gateway that provides subscribers with menus of available video information providers. Digital cross-connect systems facilitate switched connections between the VIPs and subscribers' loops, and a voice switch supports POTS. ADSL technology delivers one channel of MPEG-compressed movies to a single subscriber at a time. The data is sent through the switched phone network to a TV STB that decompresses it and converts it to NTSC analog video for delivery to the TV.

HDSL. A high-speed digital subscriber line supports basic ISDN rates of 1.544 Mbps (full-duplex communication) on existing copper lines; consequently, HDSL supports MPEG video transmission.[9,10] HDSL requires no line repeaters or special circuit design, thereby significantly reducing loop costs for symmetrical 1.544-Mbps service. Ameritech and Bellcore have built a prototype system for high-quality videoconferencing in which HDSL circuits are implemented in Illinois Bell telephone loops.[10]

SONET. The Synchronous Optical Network is a transmission interface standard that specifies multiplexing and one or more 51.84-Mbps channels. GTE Laboratories has developed a Sonet-based VOD system that is currently being tested in Cerritos, California.[11] The system deploys numerous hierarchical switches, similar to those used in telephone networks, to connect many video service providers. The experimental system, shown in Figure 7, connects three groups of equipment with Sonet fiber links.

The first group, central office equipment, comprises three video information providers, a video gateway, and a broadband central office switch. The second equipment group comprises a broadband-enhanced remote line unit (BERLU), which contains a high-capacity broadband switch. This unit is connected to the central office via a Sonet fiber feeder. The third equipment group consists of subscriber premise equipment, which includes ordinary TV sets, VCRs, and video cameras. Each home is connected to the BERLU through a dedicated Sonet distribution fiber. Potential video services, to be delivered via the telephone video dial tone technique, include conventional TV programs, VOD, interactive education, home shopping, video telephone, and video mail.

MULTIMEDIA SERVER ARCHITECTURES

To facilitate shared access to multimedia data, multimedia servers need to control three architectural components: large volumes of continuous data, data storage devices, and the networks and communications that deliver the data to and from the subscribers in real time.

Digital multimedia servers let users incrementally add, delete, or edit stored multimedia content through digital video editing techniques. Digital servers also let multiple clients concurrently access the same media devices, a significant VOD feature.

The basic digital multimedia server architecture—at the information-provider level—consists of a CPU, large capacity storage, and an ATM network adapter, as illustrated in Figure 8. We next discuss storage requirements.

RAID storage for MPEG videos

The disk storage system is typically based on redundant arrays of inexpensive disks. Video programs are stored in MPEG-compressed form, requiring 1- to 2-Mbps transmission for MPEG-1 (and 2- to 20-Mbps transmission for MPEG-2) videos. After disk data is retrieved, it is transmitted through the network at rates exceeding real-time rates, which means a group of video segments can be simultaneously transmitted. For example, a 60-second MPEG-1-compressed video segment transmitting at 1.5 Mbps requires 90 Mbits of storage. The segment can be delivered in a 0.6-second interval at the 150-Mbps transfer rate, which is one hundred times faster than the real-time rate.

However, with MPEG-2 technology that provides HDTV-quality video, the requirements for the multimedia server become stringent. This combination requires large super servers to simultaneously serve thousands of viewers; therefore, the servers must have a large switching capacity to instantly connect data from disks to any output channel. One solution to achieve this might have large super servers based on massively parallel computers that provide high I/O throughput.

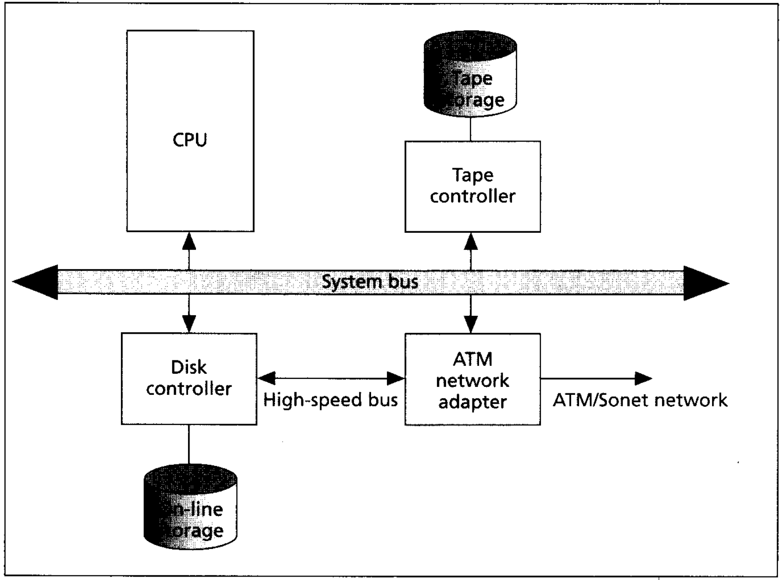

Figure 8. Architecture of a digital multimedia server.

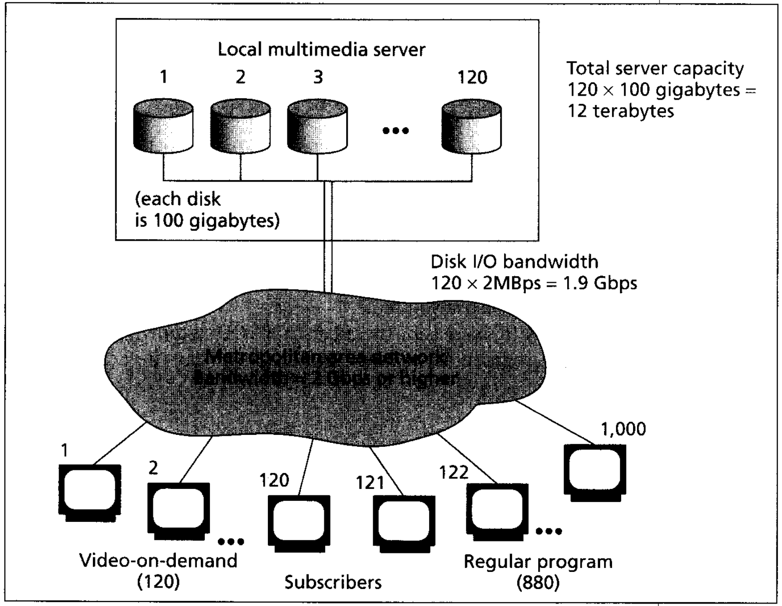

Figure 9. Example of a local multimedia server serving a neighborhood of 1,000 households. Server provides HDTV-type movies-on-demand; only 120 users can receive movies-on-demand simultaneously.

Less-stringent local storage needs

For a multimedia server at local neighborhood storage providers, the throughput requirements are less severe; consequently, a less powerful computer can be deployed. For example, for an HDTV video requiring a 16-Mbps transfer rate, a 100-minute movie will require 12 Gbytes of storage.[12] For a 1,000-household neighborhood, 1,000 different videos will require 12 Tbytes of storage, as illustrated in Figure 9. Assuming that 100-Gbyte disks will soon be available,[12] the server will thus require an array of 120 disks. Regarding transmission capacity, fiber-optic networks are already available that have bandwidth in the tens of gigabits.

In our example, all 1,000 subscribers won't be able to receive video-on-demand simultaneously. Assuming that disk channels can deliver 16 Mbps per channel, 120 disks can deliver 1.92 Gbps (16 Mbps × 120 disks). Therefore, due to the I/O bottleneck, only 120 subscribers will simultaneously receive HDTV video-on-demand (1.92 Gbps/16 Mbps = 120). Other subscribers can, however, receive regular TV broadcasts, or they can play videos on their VCRs.

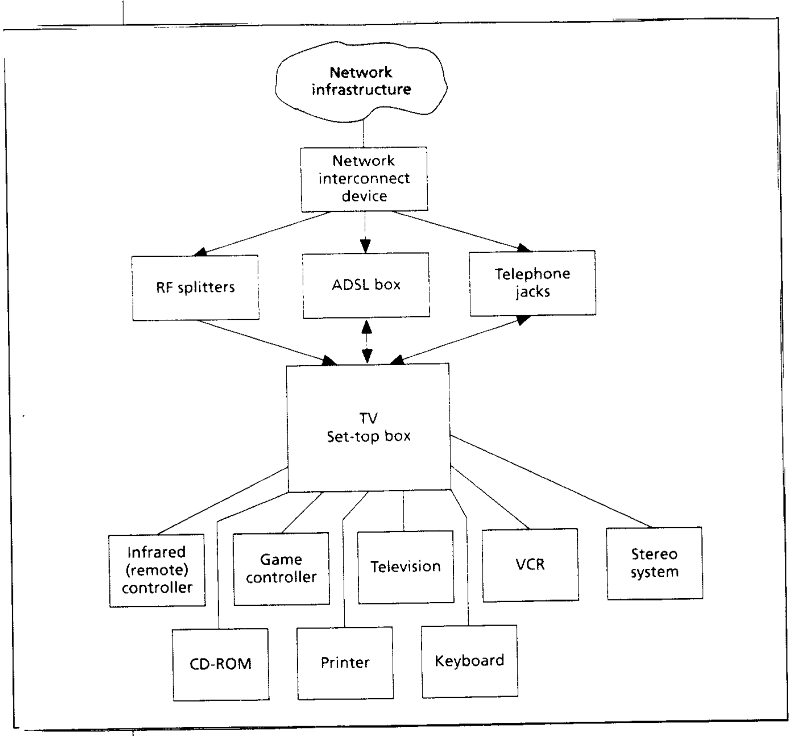

Figure 10. Interactive TV STB interfaces.

To increase the number of subscribers simultaneously receiving VOD, system designers have three options:

- The number of disks can be increased, which will increase the I/O throughput.
- Instead of 16-Mbps HDTV movies, 3- to 6-Mbps MPEG-2 movies can be transmitted.
- Techniques for increasing the capacity of VOD systems, such as segmentation and multicasting, can be deployed.

INTERACTIVE TV SET-TOP BOX

To interact with a full-service network offering personalized, on-demand, multimedia services, subscribers need a device for doing so. The most likely digital cable terminal device is an STB. The STB is the bridge between the subscriber's display devices, peripherals, and input devices (such as a hand-held infrared remote controller) and a D/A communication channel. This channel connects the STB to the information infrastructure of service providers (refer to Figures 2, 4, and 5). In addition, the STB gives software developers a means to develop entertaining and compelling application interfaces.

First-generation STBs typically include only decode/play functions and a simple reverse channel. As shown in Figure 10, the ideal STB should be capable of receiving a multimodal data stream from the network consisting of digital video, audio, image, graphic, text, and user-interface components. It should generate such multimodal data streams when sufficient upstream bandwidth becomes available. Until such time as the bandwidth is available, however, STBs will have limited interactivity and be used primarily for control purposes and ordering menus.

Full-service networks are eventually expected to be extensions of current cable TV networks. Current cable systems have a number of channels carrying broadcast programming that are available to subscribers as part of the basic service. In addition, pay-per-view channels are available. The situation will be similar in full-service networks. Some of these channels will be digital, and some will continue to be analog. In addition, interactive channels will offer on-demand services such as video, games, news, and shopping. When tuned to one of these on-demand channels, the user interface will help subscribers navigate and select programs.

The ultimate STB will have two physical interfaces: one to a D/A communication channel, which connects the STB to the information infrastructure, and the other to the TV display, VCR, remote controller, and various computer peripherals. A basic ITV STB minimally has to receive an analog signal from coaxial cable, demodulate and decode it into a digital signal, decompress the digital signal into video and sound, and provide interfaces for remote control (shown in Figure 10). It also offers limited interactivity, so it transmits low-bandwidth upstream data.

Different ITV applications have varied network bandwidth and quality-of-service requirements. Because all these applications are served by the service providers, the network traffic and behavior tends to change rapidly and dynamically. In an ATM network, for example, the QoS parameters, such as cell delay and cell loss rate, are averaged over a period of time and used as criteria for admission control. These parameters do not reflect the current state of the network in a dynamically changing environment. Instantaneous calculation of these parameters will give a more accurate state of the network and should be used by cable and telephone company designers to guarantee reliable QoS.

STB functions

Depending on cost, an STB can offer functionality ranging from minimal to elaborate. As mentioned earlier, some of these functions are the minimal required set for providing standard cable and interactive TV. Other functions augment the STB with the features usually offered by a networked multimedia PC. Table 3 presents a list of STB features.

An STB has higher performance requirements, from a media-handling perspective, than most computers. Today's most powerful processors can barely play SIF (simulator interchange format) resolution format (352 × 240 pixels) at the full-motion rate of 30 frames per second, which is comparable to a VHS-quality tape recording's resolution. Consumers expect visual quality comparable to broadcast quality. NTSC (National Television System

Committee) format requires a resolution of 720 × 486, and HDTV format will require 1,920 × 1,080. In addition, subscribers expect CD-quality audio, which requires a resolution of 16 bits per channel at a data sampling rate of 44.1 KHz or better.

Table 3. Features of an interactive TV STB. (Optional features may be accomplished through STB upgrades.)

Feature	Required	Optional
Ability to download applications and data		✓
Addressability		✓
Analog video and audio outputs	✓	
Capability to control peripherals, such as infrared remote, printer, mass storage, and game controllers	✓	
Capability to store two video frames (MPEG) for continuous playback	✓	
Graphics for user interface and applications	✓	
High-speed data modem		✓
MPEG-2 or similar decoder	✓	
Processing capability	✓	
Security and access control		✓

Security and authentication are key differences between a computer and an STB. Unlike that of a computer, an STB's primary function is to enable subscription services. This vital function ensures that service providers are compensated for their services and customers are fairly charged. Traditionally, cable TV companies have expended considerable effort to develop security systems that prevent unauthorized use of their services. Some of these schemes involve sophisticated video signal scrambling and blocking.

As an output medium for network services, current television technology imposes some limitations. Today's television systems have an interlaced NTSC resolution, whereas computer displays have Super VGA. Even a digital video signal has to be converted to an analog NTSC signal before it is displayed on a conventional TV. The inherent bandwidth limitation of NTSC imposes some quality limitations, especially in the text area. NTSC has been optimized for continuous moving pictures and tends to exhibit visual artifacts on sharp-edged stationary objects like text. Similar limitations exist with color fidelity and bandwidth; consequently, new user interfaces will be required.

To use the STB, a subscriber must have a simple infrared remote controller with perhaps one to 15 buttons. The limited number of buttons and limited functionality of a remote controller impose several challenges to designers of software-based navigation systems. A subscriber will want minimal complexity to select services from an extensive menu. In addition to supporting infrared devices, other interface elements can be provided in the STB at additional cost to connect game controllers and stereo video systems. For example, as an option, a personality module could be added to the STB as either a plug-in card or as nonvolatile memory to store custom subscriber profiles that might include favorite services, game scores, and encrypted billing information.

Hardware architecture of an STB

An example of a proposed STB hardware architecture appears in Figure 11 on the next page. The *network interface* connects the STB to the network infrastructure. The network interface might, as an option, incorporate security services that include permanent and renewable security. (Permanent security would be, for example, a chip mounted on the STB circuit board with hard-wired encryption/decryption, whereas renewable security would be upgradeable devices, such as programmed cards for insertion into STB slots.) Telephone companies in particular have proposed digital delivery mechanisms, such as ATM, to the STB; however, an analog tuner is required because of cost and infrastructure considerations, and because analog channels have to be supported for legacy reasons.

Digital channels are modulated into analog channels, and the digital transport stream is extracted from an analog channel after demodulation. The network interface also incorporates the security services. A back channel transmits data to the head end, gateways and servers, and to other subscribers' STBs. A low data-rate channel is required for basic navigation and service selection. The market for a symmetrical data-rate channel to support bidirectional video communication is rather limited at this point.

A *processing subsystem* runs a small real-time operating system to manage resources and activities in the STB. The system ROM contains boot code and basic OS services, while system DRAM is shared between operating system, applications, and data.

A *peripheral control subsystem* enables subscribers to connect peripherals to the STB. Peripherals can include printers for printing items like coupons and programming guides, CD-ROM and magnetic disks for mass storage, digital video cameras, game controllers, and infrared controllers.

A *video subsystem* decompresses encoded video streams in formats such as MPEG-2. Video streams with MPEG-2 encoding require substantial processing power. Because present processor performance does not handle pure software MPEG decoder implementation, hardware assistance is needed. Several companies have announced MPEG-2 decoding chipsets, including C-Cube Microsystems, Philips, AT&T, LSI Logic, IBM, and SGS Thomson.

A *graphics subsystem* supplies user interface elements

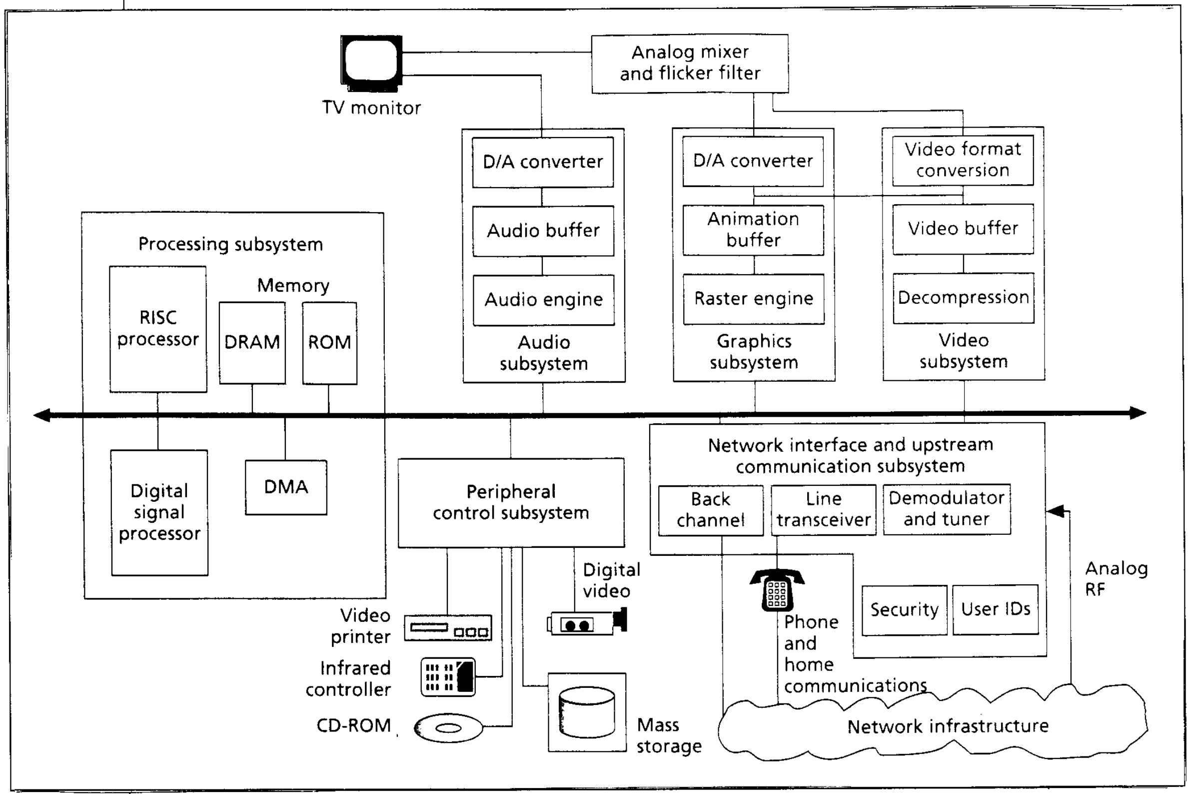

Figure 11. Proposed hardware architecture of an interactive TV STB.

for navigation and presentation. In addition, this subsystem provides acceleration for two- and three-dimensional graphics applications, such as video games. Eventually, it would be capable of rendering several hundreds of thousands of texture-mapped shaded polygons per second.

An *audio subsystem* decodes the audio corresponding to the video stream and generates audio locally in the form of MIDI (Musical Instruments Digital Interface) and wavetable synthesis to accompany games and other applications. The subsystem might optionally also enable high-fidelity digital audio-only services.

To produce a low-cost STB for consumers, the architectural blocks will be integrated differently from the way they appear now in Figure 11. Most likely the major functions will be integrated into four VLSI chips: processor, media processing, media access, and communication. The processor chip will include a powerful RISC processor, clock, and memory caches. The media processing chip will include the video decompression subsystem, a graphics accelerator, a display controller, color space conversion, and D/A output to the RF modulator. The media access chip will integrate the peripheral control functions. The communication chip will integrate all communication and network interface functions, such as demodulation and modulation, encryption and decryption, and security.

Software architecture of an STB

A proposed software architecture of an STB, shown in Figure 12, is layered to show abstract device-dependent features.

The *hardware abstraction layer* provides a low-level programmer's interface to different hardware subsystems such as the video, audio, network interface, and graphics subsystems. This interface is similar to the BIOS in a PC. The HAL provides an interface that ensures software compatibility with new hardware versions because this layer hides hardware specifics such as registers and memory mapping.

In the *microkernel*, a small real-time operating system resides on top of the HAL for services such as process creation and execution, interprocess communications, and resource allocation and management. The resources managed by the OS include memory, communication channels, network bandwidth, and peripheral access.

Drivers and library routines provide common, frequently requested services to applications. These include APIs for network and session management, video control, graphics, and the user interface. Applications reside on top of the software hierarchy.

Considering the price points, network architecture, and expected applications for a set-top, it seems likely that an STB will have limited memory and, eventually, local mass storage. The application code will be downloaded to the STB from media servers over the network. The HAL and user interface for simple channel selection will reside in ROM on the STB. A bootstrap loader will download the OS and on-demand applications from the server.

Partitioning an application between the STB and the multimedia server is an interesting issue. Because of cost

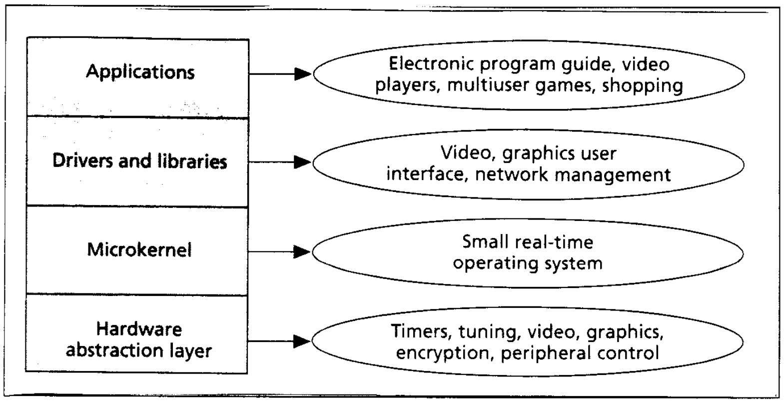

Figure 12. Proposed software architecture of an interactive STB.

constraints, the STB will generally have limited processing capabilities. On the other hand, multimedia servers connected to the STBs will have significant computational and storage resources, although shared by thousands of subscribers. A well-designed application will judiciously use the resources of both STB and server. The main design trade-offs would involve network latency and bandwidth, resources in the STB, and sharing of server resources over multiple applications.

As an example, consider VOD, in which video player applications will be downloaded to the STB. The STB has substantial computational resources (video decoding hardware). The application will perform network management and user-interface functions. Because most full-service networks are being designed for video delivery, and latency is not an issue in video delivery, trade-offs for STB/server partitioning do not exist.

In an application like interactive video games, STB/server partitioning trade-offs are significant. A low latency is required to provide an acceptable response to the game player. Game quality might be enhanced through sophisticated graphics and video. For graphics functions, the application will need to balance the capabilities of the STB's graphics hardware, the STB's computational resources, the server's computational resources, and the network's video bandwidth.

Interoperability issues

Given the potentially large and profitable market for ITV applications, many vendors will likely provide the servers, networks, STBs, services, and applications. Each of these components must support many different devices; for example, each STB must support multiple video information providers. Similarly, each VIP must provide its video information on STBs from multiple vendors. For reasons of cost and protocol compatibility, standards will be necessary (see "Standards groups" sidebar). Although many network operators and equipment vendors are pushing de facto standards, a number of forums have been established whereby industry, university, and government representatives are developing and approving standards.

Interoperability advantages are that subscribers can use the same equipment as they move from one place to another (moving to a different region serviced by a different cable or telephone company shouldn't have any effect on usability). Competition and high volumes of usage arising because of interoperability also bring down equipment costs. Also, definition of standards gives content creators a stable platform toward which they can target their products.

Another issue is STB addressing. In current cable systems, each analog decoder box is addressable so that the cable company can selectively turn services on and off and control descrambling circuitry. However, the domain of the boxes is usually limited over a small area, and assigning unique STB addresses is not difficult. In the world of full-service networks where services are delivered over a long-haul network, unique addressing must be supported over a large area, possibly extending to a world-wide addressing scheme. One approach is to extend Internet protocol addresses to encompass STBs.

WITH A TREMENDOUS INVESTMENT TAKING place in interactive TV, it is only a matter of time before this technology enters the home. The interactive TV STB will one day become just another home appliance and will probably affect the culture and user habits just as ordinary television and VCRs have. Today, it remains a formidable technological and business challenge.

A major design issue is how to achieve desired STB functionality at the right cost. Although it might be difficult to meet that price with existing components and technology,

Standards groups

The following groups are developing standards related to a full-service network:

- ATM Forum
- Corporation for Open Standards
- Digital Audio Visual Council
- Interactive Multimedia Association
- Moving Pictures Experts Group
- Multimedia Communications Community of Interest
- Multimedia Communications Forum
- National Cable Television Association
- Society of Motion Picture and Television Engineers
- Video Electronics Standards Association

cost-reduction strategies, such as integration of major functions into a small number of VLSI chips, will very likely satisfy the price goal. On the other hand, market pressures acutely affect the ITV infrastructure. Technology and implementation are being driven by large communication and computer companies. The entire strategy depends on consumers having to order and use the services offered by the full-service network to make the network and service providers commercially viable. Because of unexpected delays in developing and deploying the prototype systems, companies are moving cautiously in implementing these networks and related on-demand systems. ❚

Acknowledgments

The authors would like to thank all reviewers, including coguest editor Larry Rowe, for their constructive comments and suggestions.

References

1. J.P. Nussbaumer, B.V. Patel, and F. Schaffa, "Capacity Analysis of CATV for On-Demand Multimedia Distribution," *Proc. First ISMM Conf. on Distributed Multimedia Systems and Applications*, Acta Press, Anaheim, Calif., Aug. 1994, pp. 97-100.
2. S. Ramanathan and P. Venkat Rangan, "System Architectures for Personalized Multimedia Services," *IEEE MultiMedia*, Vol. 1, No. 1, Spring 1994. pp. 37-46.
3. B. Furht, "Multimedia Systems: An Overview," *IEEE MultiMedia*, Vol. 1, No. 1, Spring 1994, pp. 37-50.
4. A. Reinhardt, "Building the Data Highway," *Byte*, Mar. 1994, p. 46.
5. L. Press, "The Internet and Interactive Television," *Comm. ACM*, Vol. 36, No. 12, Dec. 1993, pp. 19-23.
6. M. Milenkovic, "Multimedia Workstation and Interactive TV," in M. Milenkovic and B. Furht, "MultiMedia Systems and Networks: A Guided Tour," *Tutorial Notes, 2nd ACM Conf. on Multimedia 94*, Oct. 1994, pp. 1-62.
7. T.D.C. Little and D. Venkatesh, "Prospects for Interactive Video-on-Demand," *IEEE MultiMedia*, Vol. 1, No. 3, Fall 1994, pp. 14-24.
8. J.D. Oetting, "A Comparison of Modulation Techniques for Digital Radio," *IEEE Trans. Comm.*, Vol. COM-27, No. 12, Dec. 1979, pp. 1,752-1,762.
9. J. Sutherland and L. Litteral, "Residential Video Services," *IEEE Comm.*, Vol. 30, No. 7, July 1992, pp. 36-41.
10. W. Walkoe and T.J. Starr, "High Bit Rate Digital Subscriber Line: A Copper Bridge to the Network of the Future," *IEEE J. Selected Areas Comm.*, Vol. 9, No. 6, Aug. 1991, pp. 765-768.
11. J. Woodward et al., "A Broadband Network Architecture for Residential Video Servers Featuring Distributed Control of Broadband Switches and Video Dial Tone Functionality," *Proc. IEEE Int'l Conf. Comm. (ICC 93)*, Product No. CH32615, IEEE Service Center, 445 Hoes Ln., Piscataway, N.J., May 1993, pp. 853-857.
12. H.M. Vin and R.V. Rangan, "Designing a Multiuser HDTV Storage Server," *IEEE J. Selected Areas Comm.*, Vol. 11, No. 1, Jan. 1993, pp. 153-164.

Borko Furht *is a professor of computer science and engineering at Florida Atlantic University in Boca Raton, where he is also the founder and director of the multimedia laboratory. Previously, he held senior positions at Modcomp, a computer division of Daimler Benz in Germany; at the University of Miami in Coral Gables, Florida; and at the Institute Boris Kidric-Vinca in Belgrade, Yugoslavia. His research is in multimedia systems and their applications.*

Furht has published over 120 papers and eight books, and holds two patents. He is a co-author of A Guided Tour of Multimedia Systems and Applications *(IEEE Press, 1995), an editor-in-chief of the* J. Multimedia Tools and Applications, *and associate editor of* Real-Time Systems J. *and* Real-Time Imaging J. *He received BS (Dipl.Eng.), MS, and PhD degrees in electrical and computer engineering from the University of Belgrade, Yugoslavia.*

Deven Kalra *is project manager for computer graphics at Hewlett-Packard Laboratories in Palo Alto, California. His research interests include algorithms and architectures in computer graphics for workstations, PCs, and for consumer platforms such as set-top boxes. He has worked and published in constrained- and physically-based modeling, simulation, animation, and rendering.*

Kalra received MS and PhD degrees in computer science from the California Institute of Technology, and a BTech degree, with honors, in electrical engineering from the Indian Institute of Technology in Delhi.

Arturo A. Rodriguez *is a coguest editor of this theme issue. His biography appears following the guest editors' introduction on p. 22.*

Further reading

1. J.Y.L. Boudec, "The ATM: A Tutorial," *Computer Networks and ISDN Systems*, Vol. 24, 1992, pp. 279-309.
2. E.D. Sykas, K.M. Vlakos, and M.J. Hillyard, "Overview of ATM Network: Functions and Procedures," *Computer Comm.*, Vol. 14, No. 10, Dec. 1991, pp. 615-626.
3. R. Rooholamini, V. Cherkassky, and M. Garver, "Finding the Right ATM Switch for the Market," *Computer*, Vol. 27, No. 4, April 1994, pp. 16-28.
4. W. Stallings, "ISDN and B-ISDN," Macmillan, New York, 1992.
5. C.S. Thachenkary, "Integrated Services Digital Networks (ISDN): Six Case Study Assessment of a Commercial Implementation," *Computer Networks and ISDN Systems*, Vol. 25, No. 8, Mar. 1993, pp. 921-932.
6. *Data Communication Handbook*, Siemens Stromberg-Carlson, Boca Raton, Fla., Feb. 1994.
7. R. Jain, "FDDI: Current Issues and Future Plans," *IEEE Comm.*, Vol. 31, No. 9, Sept. 1993, pp. 98-105.

William E. Wall *is chief scientist for the Digital Video Group at Scientific Atlanta. His principal research interests are in communications and network design for satellite and terrestrial systems. Wall's current work focuses on interactive television systems.*

Wall received his PhD in physics from the Georgia Institute of Technology in 1978.

Frederick L. Kitson *is the manager of the Visual Computing Department at Hewlett-Packard Laboratories in Palo Alto, California, where he works with graphics algorithms and architectures, image processing, video/audio compression and scientific visualization. He is also an adjunct professor of the Georgia Institute of Technology and a part-time lecturer at the University of California at Berkeley.*

Kitson received a BS degree, with honors, in electrical engineering from the University of Delaware in 1974; an EE degree from the Georgia Institute of Technology in 1975, and a PhD degree in digital signal processing from the University of Colorado in 1981. He was the 1991/1992 Clyde Chair visiting associate professor at the University of Utah Computer Science Department.

Engineering Opportunities

Rocketdyne, a division of Rockwell Aerospace, needs engineers to help in the development of complex electro-optical systems.

- Engineering skills needed are electronics, controls, opto-mechanics, software, computer, and electro-optics.
- The work location is in an outlying part of the Antelope Valley near Palmdale, California.
- A minimum of a B.S. degree required.
- U.S. citizenship required. Applicants selected may be subject to security investigation. If so, applicant must meet eligibility requirements for access to classified information.

In return for your contributions, we offer an outstanding salary and benefits package. For immediate consideration, please forward your resume to: **Rockwell International, Rocketdyne Division, Dept. IEEE, 6633 Canoga Avenue, P.O. Box 7922, Canoga Park, CA 91309-7922, Attn: Staffing Services. Fax (818) 586-9380.** We are an equal opportunity employer.

Rockwell Aerospace

Rocketdyne

Does Your Software Have Bugs?

Have you spent countless hours trying to track down that elusive memory leak or nasty pointer error? *Insure++* detects bugs for you automatically, in a fraction of the time.

AdvancedSystems BEST PRODUCT 1994

"Parasoft has produced the product to beat in the memory debugging field."
Advanced Systems
Best Product Award 1994

Software Tested with INSURE

Insure++ includes runtime error detection, test coverage analysis, malloc monitor, algorithm validation, data visualization, and more!

Mention this ad for 1 year FREE maintenance.
For a limited time only.

Insure++™ 2.1 *(formerly* ***Insight++****)*

The most thorough runtime error detection available. Period.

E-mail: Insure@ParaSoft.com
Web: http://www.ParaSoft.com

ParaSoft Corporation

Phone: (818) 305-0041
FAX: (818) 305-9048

Reader Service Number 3

See Us at SunWorld Booth #514