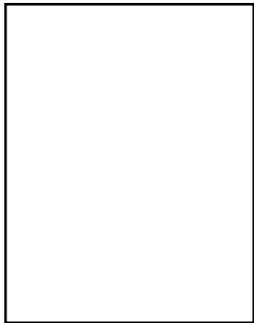


The Evolution of the Interactive Broadband Server

Joan Van Tassel, Ph.D. and Steve W. Rose



Dr. Joan Van Tassel is the author of *Advanced Television Systems: Brave New TV*, published by Focal Press (1996). She writes for *The Hollywood Reporter*, *WiReD*, and other publications on the emerging communications infrastructure. Prior to entering graduate school, she was a documentary filmmaker, known for her work on social issues. Dr. Van Tassel has written for a variety of nonprofit public information campaigns, including the “Friends Don’t Let Friends Drive Drunk”

campaign. She is the recipient of the prestigious Kenny Rogers Media Award, an Emmy nomination, and many public service advertising awards. Dr. Van Tassel received her Ph.D. and M.A. from the Annenberg School of Communications at the University of Southern California.



Mr. Steve W. Rose is an interactive broadband server consultant with Pangrac & Associates. He started his career in the 1970s in video, became the chief engineer of a small Maui cable company, and ended the 1970s designing cable television automation equipment for CRC Electronics. CRC provided the tape automation equipment for the first Warner Qube system. CRC was later purchased by Texscan, so during the 1980s, Steve was a computer guy, working with multiuser and networked computers. In 1991, ATC (later Time Warner Cable) asked him to figure out how to provide video on demand for a system with a million subscribers, which launched him on a new career. In 1994, he worked with CableLabs on their Media Server RFI, visiting various server manufacturers and giving a tutorial at the resulting CableLabs meeting.

This is Part I of a two-part article discussing video enabled servers for metropolitan areas. The division is based on two architectural approaches for building interactive broadband servers. In this first part, we discuss developments leading up to the creation of these devices, and then turn to servers that are constructed by aggregating conventional single bus computers with other necessary components.

In Part II, we cover servers based on massively parallel architectures and describe what we feel is the most appropriate model for future architectures. While we have tried to be objective, we do have a point of view that we will develop in the second article.

This article is predominantly hardware oriented. Although common software platforms are important, the nature and scale of the task at hand make the selection of appropriate hardware more immediate to the successful deployment of an interactive broadband server.

What is an interactive broadband server (IBS)? It is a device that delivers many different kinds of data and provides many simultaneous services. These include but are not limited to:

- Video streams from prerecorded video (ranging from clips to movies).
- Video streams from near-real-time events.
- Two-way voice data.
- Data transport.
- Downloads.
- Database interactions.
- Support for credit card transactions.
- Interactive simulations and games.
- Multimedia.
- Other services—known and to be determined.

In addition, the IBS must accurately track, account for, store, and bill for all services while providing for

network management! Other terms used to describe a fundamentally similar device have been Video Server, Media Server, and Metropolitan Media Server.

It is clear that IBS is at the heart of the new interactive broadband businesses that companies in the cable television, telephone, computer, and wireless industries plan to launch. Attractive programs, services, and applications such as high-speed Internet access over cable, video on demand, interactive shopping and games, and many others depend on the availability of a reliable, cost-effective IBS.

The primary difference between an interactive broadband server and a large conventional computer server, used in many organizations for client/server purposes, is the ability of IBS to provide thousands of simultaneous isochronous data streams. Isochronous ("same time") refers to data streams that are time-sensitive and must be delivered continuously without interruption or they are incoherent. An example of isochronous data streams would be real-time video and audio which are retransmitted as soon as they are received, such as a live television signal.

Stakes and Stakeholders

Many different interests are watching the evolution of IBS. Cable, wireless cable, telephone, and computer companies all want to provide new programs, services, and applications to their customers. Equipment suppliers and content providers hope to provide products, if dependable standards for them are available. Finally, regulators and consumer groups want to clarify such issues as universal access, rate structure, privacy, and security.

The high level of interest is a consequence of the anticipated size of the video on demand market. Existing markets are substantial. The advertising revenues for broadcast and cable television were about \$30 billion in 1995, according to estimates by the Television Advertising Bureau. The National Cable Television Association reports that cable revenues were \$26 billion in 1995. Conservative estimates from *Satellite Business News* magazine indicate there will be 12 million to 13 million direct broadcast satellite (DBS) subscribers by 2000; assuming an average bill of \$40 a month, satellite delivery revenues would be more than \$6 billion by the end of the decade. Wireless cable revenue is expected to be \$600 million in 1996. Finally, Paul Kagan & Associates report that the videocassette rental market earned \$10 billion in 1995, and sell-through of videos to consumers was about \$6 billion.

Packaged programming for other stand-alone devices also brings in significant revenue. Consumers spent \$6 billion for interactive games in 1995 says research firm DataQuest and about \$8 billion to \$9 billion for dedicated game players. The computer is an increasingly profitable venue for video material. There are now 25 million CD-ROM-equipped multimedia machines in consumer hands worldwide and more than 5,000 titles. DataQuest estimates that the market for CD-ROM games is now about \$660 million and growing rapidly.

Based on these figures plus a large grain of salt, the video market was more than \$60 billion in 1995. While that is less than the telephony market (about \$100 billion annually) and even less than the income of power utility companies (about \$200 billion), it is still a substantial enough figure for the various stakeholders to concern themselves with the design, deployment, and implementation of an interactive broadband platform, of which the server is a central element.

Requisite Early Developments

In order to understand the evolution of interactive broadband servers, it is helpful to understand the environment which made it possible, and the elements designers had to work with in building them.

BROADBAND NETWORK ENVIRONMENTS

Broadband networks have evolved over the last 10 years, making it possible to deliver a separate video data stream to each connected subscriber. All the new interactive broadband network technologies offer similar deliverable bandwidth expansion and the ability to carry data upstream and downstream. Most of the innovations have come from the cable and telephone industries, each of which has tried to preserve as much as possible of their existing infrastructure. For the cable industry, it is coaxial cable; for the telephone industry, it is twisted pairs of copper wires. In both cases, the final link into subscriber's homes represents the greatest investment.

The Cable Industry—Hybrid Fiber/Coax. Time Warner Cable pioneered a technology in the late 1980s which has become known as hybrid fiber/coax (HFC). It divides existing coax-based systems into neighborhoods of about 500 to 2,000 subscribers each and sends an individual optical fiber from the headend to each neighborhood. No one in the neighborhood is more than four amplifiers away from the optical fiber.

By minimizing the number of cascaded amplifiers and upgrading them, it is possible to more than double the bandwidth delivered to each subscriber—from about 500 Megahertz currently to as much as 1.2 gigahertz, while reducing system noise.

Due to the reduced noise and direct connection to each neighborhood, getting information back from subscribers becomes practical and fundamentally one-way systems become two-way systems. The HFC architecture is being widely adopted by cable companies and some telephone companies, and won the group at Time Warner Cable an Emmy in 1994.

HFC is significant because it allowed companies to consider delivering custom material to each household. The greater bandwidth provided by the HFC architecture led the press to refer to “the five hundred channel cable universe.” However, this characterization missed the real innovation, which was that operators could deliver 500 different programs simultaneously to each neighborhood node of 500 subscribers. Put another way, the new technologies make it possible to advance from delivering 50 of the same channels to 50,000 viewers to offering an individual channel to each of 50,000 interactive viewers.

Further, because of the way that HFC expands the available bandwidth, cable operators can provide hybrid service: New digital programs can be delivered to individual households over the new bandwidth from a server and long-term storage, while leaving intact the old analog services on the existing bandwidth. This means no changes for subscribers unless they choose to take advantage of new services.

Telcos—ADSL and Fiber-to-the-Curb (FTTC). The telephone industry had a different asset to protect: a network of twisted pairs of copper wire that took more than 100 years and \$1,500 per household to construct. This infrastructure required a different approach, as the bandwidth a twisted pair can support and the distance it can transport a high-bandwidth signal are greatly restricted, as compared with coaxial cable.

As a result, telco designs focused on sending one video signal at a time from the central office to the subscriber over the twisted pair. An example is Asymmetrical Digital Subscriber Line (ADSL) technology, which uses a technique which trades reduced upstream bandwidth for much greater downstream bandwidth. As the limitations of ADSL became apparent, phone companies focused on carrying fiber optic cable deep into the neighborhood, an architecture called Fiber-to-the-Curb. Each FTTC node serves

about 20 subscribers over existing twisted pairs from the curb to the home.

Wireless Systems. There are two wireless infrastructures that deliver television. Wireless cable systems (sometimes called MMDS for multichannel, multipoint distribution service) cover about a 35-mile radius and are not likely to become interactive. One reason is that MMDS is typically promoted as a low-cost alternative to wired cable service. In addition, MMDS would need to invest in a cellular-based return path to make two-way communication feasible.

By contrast, LMDS (local multipoint distribution service) systems use a cellular approach, where each transmitter reaches a defined area, as small as two or three miles. Nonadjacent cells can carry different content just as with cellular telephony, and it greatly increases the effective deliverable bandwidth to the overall service area. Wireless data return from subscribers to the cell site is also possible, because the return path from the viewer back to the cell is so short that it needs only a four- to six-inch antenna and a little power, using the cell structure already in place.

ATM SWITCHING

Asynchronous Transfer Mode (ATM) is the first protocol for data transport that allows the mixing of voice, video, audio, and data signals on the same circuit. Switches that incorporate ATM technology have the ability to switch any input circuit to any output circuit at lower costs, higher bandwidth, and greater speed than any preceding technology.

DIGITAL VIDEO COMPRESSION (DVC)

DVC allows the transmission of four to 10 programs in the same bandwidth required by one analog channel. Once an analog video signal is digitized, several techniques can be applied to reduce the amount of bandwidth necessary to transmit it. Compression techniques remove information that cannot be perceived by the human eye, encode redundant information so that it is transmitted only once, and then reconstitute the original image at the receiving end. However, the greater the amount of information that is discarded, the more likely it is that compression artifacts will become noticeable.

Several techniques of compression have been developed, including discrete cosine transform (DCT), wavelet, vector quantization, and fractal schemes. However, the dominant family of standards was developed by the Moving Picture Experts Group (MPEG) based on DCT.

OTHER IMPORTANT DEVELOPMENTS

Robotic Storage Libraries. To compete with the video rental business, an equivalent number of titles must be offered. Since many will be infrequently requested, they must be stored off-line in an automated library. These libraries have been developed for the computer industry, ranging in size from jukeboxes that hold a few hundred optical disks to room-size robots that hold thousands of tape cartridges and optical disks.

Powerful Single Bus Computers. In the computer industry, a battle has waged for years between harnessing the power of tens to thousands of processors running independently in parallel linked by a communication mesh, versus connecting a limited number of processors running on a single bus. Parallel processors have gained a reputation for being difficult to program. However, the rapid development of ever more powerful single bus processors has allowed them to match the computing capacity of any existing parallel processor during its lifetime. Most successful computer servers are single bus, single processor designs. Some use a single bus, but multiple identical tightly-coupled processors, and are referred to as Symmetric Multiprocessors (SMP).

Real-Time Encryption. Protecting intellectual property, when it is represented as a digital data stream, requires that the stream be encrypted in real time so that each instance of the stream can be separately protected. Isochronous digital video streams pose a particular problem due to their high data rate.

Error Correction Codes. There are two points during isochronous stream delivery that errors must be anticipated and corrected in advance. The first high-potential point occurs when the data is played from the hard disk arrays. Hard disk drives are electromechanical, so they can be expected to fail more often than electronic equipment. Their failures are critical due to loss of data and/or interruption of service. As a result, a group of standards for increasing storage reliability, collectively known as RAID (Redundant Arrays of Independent Disks), has evolved which allows disks to be grouped so that the failure of a single drive does not affect the output of the array. The penalty for this reliability is typically an increase of 25% in the number of drives required.

The second error-prone point occurs during network transport. Anticipating these errors involves sending enough redundant information that almost any errors caused by transient noise or interference can be

fixed on the receiving end. This anticipatory error correction is critical for isochronous data, where there isn't enough time to detect an error and ask for a retransmission of the corrupted data. The procedure is called Forward Error Correction (FEC).

Digital Modulation. In order to transmit digital information over analog channels (important for HFC and wireless networks), the digital information must be changed in format. Digital information has only two levels representing 0 and 1. Modern techniques increase the number of levels through phase and amplitude modulation of each cycle to achieve bit efficiencies of up to six to eight bits per Hertz. Typical of these is 64 QAM (Quadrature Amplitude Modulation), which uses 64 unique combinations to transmit six bits per Hertz.

Network Management. As computer networks have grown more complex, the difficulties of monitoring and controlling the equipment that constitute the network have grown. Software and standards now allow central management of networks that span the globe. Chief among the current standards is the Simple Network Management Protocol (SNMP).

Business Support. The last few years have seen the creation of reliable computer software to accommodate various complex forms of billing. Cable television billing typically includes a flat monthly fee, fixed monthly additions for additional services, plus individual billing for special events. Telephone billing includes flat monthly charges, plus billing on the basis of utilization on a minute-by-minute or second-by-second basis, plus billing on behalf of secondary companies (e.g., long distance providers). Telemarketing services require immediate credit verification and real-time interaction with financial and fulfillment (inventory and shipping) systems. Finally, just as neighborhood shopping centers must monitor the sales of associated businesses to enable billing on a percentage of revenue, so must virtual shopping centers on the network track sales. All of these complex billing models will be required in an interactive broadband network.

Conventional Interactive Broadband Servers

DEFINING THE TASK

Regardless of its architecture, an IBS must carry out a precise set of tasks. In addition, it must perform them reliably and cost-effectively. These functions are:

- Find the correct data: storage and importation of content into stream generator.
- Reassemble content into a stream: generation.
- In many cases, encrypt the stream: encryption;
- Route stream to the correct subscriber: stream sorting and multiplexing.
- Encode stream for transport: monitor quality of service, forward error correction, and modulation.
- System management: operations and business support.

A BRUTE FORCE SERVER MODEL

In order to conceptualize the cost and complexity of a metropolitan IBS, we take a brute force approach (see Figure 1). This model is a heuristic device, not an actual (or even possible) design for a real-world server.

We assume a population of 500,000 subscribers, 200,000 streams, and a library of 300 movies. Each movie requires about three gigabytes of storage, resulting in about one terabyte of memory to store all the movies. Although it would be too expensive to actually store this content with RAM chips, we will store the movies in RAM, using 16 megabit DRAMs for comparison purposes. This storage will require about 500,000 RAM chips.

To sort each resulting four megabits per second (Mb/s) stream to the viewer requesting it, we hypothesize a switch made of 100 input by 100 output crosspoint switch chips. To actually deliver each

stream to the requester, we will need a routing switch with 200,000 inputs (the number of streams) and 500,000 outputs (we have to be able to switch any stream to any subscriber). With our 100 × 100 switch chips, we will need 2,000 times 5,000 chips, or 10,000,000 switch chips—20 times more than the number of RAM chips!

It is clear that the big problem is not storage, but switching. The conclusion that emerges is: The cost of storage increases linearly with capacity, and the cost of switching increases geometrically with capacity.

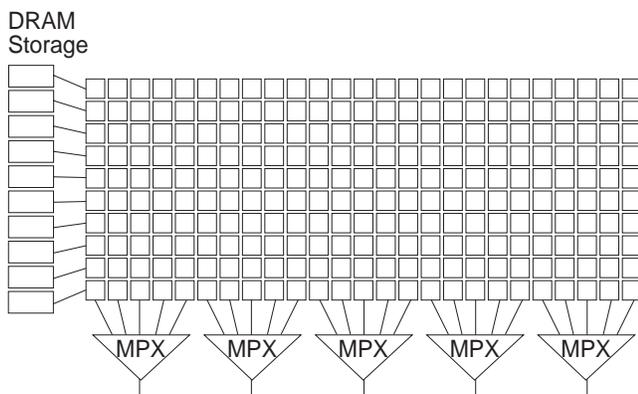
Assumptions

Compressed digital video streams have data rates varying from 1.2 Mb/s to nine Mb/s, with four Mb/s being a typical choice for an excellent quality video plus audio signal. The definition of excellent quality is that the image is equivalent to Super-VHS or Hi8 video and CD-quality multichannel audio. In the remainder of this article, calculations will be based on four Mb/s constant bit rate MPEG-2 encoding.

Assuming that an average movie lasts 100 minutes, it will consume three gigabytes of storage. The estimates of the maximum number of subscribers using interactive digital services simultaneously have ranged from 7% to 40% of total subscribers. We will use a figure of 20% to represent the peak capacity design point for our calculations.

We will also assume that the subscriber population for a single server will be between 20,000 and 100,000 subscribers and the typical size is 50,000 households. This results in a server design that must be capable of back-to-front throughput of 16 gigabits per second (Gb/s) to 80 Gb/s; throughput of 40 Gb/s is needed to serve a 50,000 household system.

**Figure 1
Brute Force Diagram**



MPX = Multiplexed Output to Neighborhoods

Source: S. Rose

STORAGE AND IMPORTATION OF CONTENT INTO STREAM GENERATOR

How should library content be stored? Server manufacturers' designs for library storage range from a single tape drive to complete robotic tape libraries. Due to their high cost, robotic libraries are generally proposed as an option rather than a necessity. Typically, successive generations of storage equipment offer ever-increasing capacity per tape or disc. Ironically, when greater capacity allows more than one program to be stored on a single tape or disc, then the so-called improvement becomes a liability. Library storage media allow only one program to be accessed at a time; if a request arrives for another title on the

same medium while the first one is still loading, the second title cannot be read until the first one is finished. Thus, the only reason to store a second title on a tape or disc should be as a physical backup for the primary copy of the title stored elsewhere.

REASSEMBLE DATA INTO A STREAM: GENERATION

It takes Herculean engineering, especially in the software arena, to turn a conventional computer into a server that can produce isochronous video streams. In spite of the difficulties, many of the major computer vendors have done so. Hewlett-Packard, Sun, Silicon Graphics, and DEC all see the problem as one of I/O (input/output) bandwidth and have designed servers based on conventional computers. This type of server typically has a single CPU or several tightly-coupled identical processors, a high-speed bus, an array of associated hard disk drives, and substantial I/O (disk I/O on one side and network I/O on the other). The server will generate from 16 to 125 video streams of video from a single copy of a program.

When more streams are required, these stream generators are grouped. Since each has its own hard disk storage, a title which is popular enough to generate demand for more streams than one unit can provide must be replicated to as many units as necessary to meet the demand. In fact, since many titles can be stored on one stream generator, it must be managed so that a popular title doesn't block access to other titles stored uniquely on that unit by consuming all of its output streams.

In a 50,000 subscriber area, the peak server load will be 10,000 streams. This means that, to satisfy peak load requirements, we will need 100 stream generators, each with a 100 stream capacity. First-run movie demand, as indicated by box office receipts, shows that one title can account for as much as 40% of demand. With a peak server load of 10,000 streams, a single title could account for 4,000 of the streams.

Assuming that we will allow up to 80% of the streams on one unit to come from a single title, that title will have to be replicated on at least 50 of the stream generators (4,000 streams divided by 80 streams per generator). At three gigabytes per copy, at least 150 gigabytes of storage will have to be devoted on the system to that title.

If the operator decides to store it on every server, the storage requirement rises to 300 GB. In a large metropolitan system with 500,000 subscribers, this redundant storage would expand to at least 1.5

terabytes (TB), regardless of whether it was on one large server or 10 smaller ones.

ENCRYPTION

Encryption protects the privacy and security of both downstream delivery to customers and upstream communication coming from them. Surveys of actual and potential users of interactive services consistently find that security and privacy are high on the list of consumers' concerns. In addition, when interactive broadband systems are used for telecommuting so that employees can work at home, these issues are extremely important to the employer companies. For the system operator, security is crucial to protect the system's assets—programming—from signal pirates and to make sure that only the subscriber who pays for the delivery of the signal is able to decode it. The logical place to apply encryption to a stream is at the point of its generation.

STREAM SORTING, ROUTING, AND MULTIPLEXING

As demonstrated by the brute force model, generating streams is the easy part, compared with sorting them so that the correct stream is delivered to the subscriber requesting it. The switch which sorts the streams may be thought of as having as many inputs as there are streams to be generated under peak conditions, and as many outputs as there are subscribers. Any input can be connected to any output, so the size of the switch is proportional to the number of inputs times the number of outputs.

The volume of information in video streams, the high transport speed required, and the ability to switch all types of information has caused most server architects to incorporate ATM switching into their designs to sort and route all the streams and to multiplex streams going to the same neighborhood. In addition, ATM makes interconnection with other services using the ATM protocol straightforward (for example, long distance service providers).

ENCODING FOR TRANSPORT: MONITOR QUALITY OF SERVICE, FEC, AND MODULATION

The final point to monitor the quality of the digital signal is just before forward error correction (FEC). Following quality monitoring, a processor generates redundant information that enables the FEC. The modulation procedure begins with encoding the digital signal to an analog signal, which improves the modulation efficiency (bits per Hertz), then to impress an RF

carrier with the resulting analog signal for distribution over an HFC or wireless system.

Conventional designs usually place FEC and quality monitoring functions within the modulator. As a result, an “intelligent modulator” that accepts standard ATM input and provides sufficient processing power to accomplish the additional information-modifying tasks is needed for each output channel.

Typically, the modulator puts 27 Mb/s of data on a six MHz chunk of spectrum (the space allotted for a conventional analog channel). This allocation allows for six 4-Mb/s compressed digital streams plus overhead (four Mb/s, divided into 27 MHz equals 24 MHz, with three MHz capacity left over for overhead).

OPERATIONS SUPPORT AND BILLING

Each element of the server must provide standard management information. This information, carried on a separate network, is processed by an independent computer running network monitoring and management software, and is referred to as the Operations Support System (OSS). It is crucial for operators to be able to flag error conditions immediately and to access remote operation of the management system. When an error occurs, it is inconvenient and time-consuming to have to go to the physical premises to begin to diagnose and solve the problem.

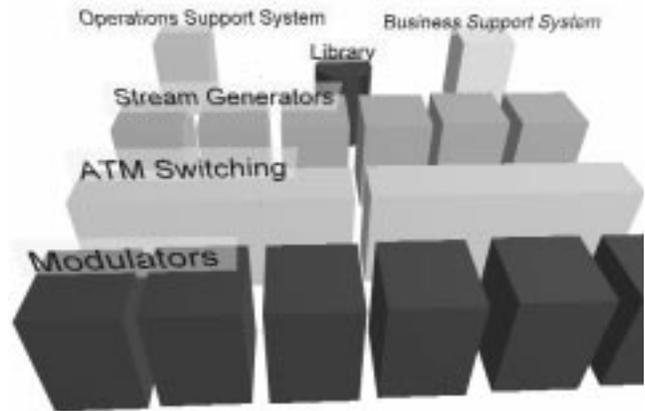
Another separate control computer, or sometimes multiple computers arranged in a hierarchy, direct the operation of the stream servers and switch. This system manages content and directs normal operation of the system. Billing information is sent from the control computer to a separate billing computer system. For the sake of reliability, an isolated machine is used to bill for services. This physical separation of the billing data from the content also ensures that access to the content server cannot be used to hack billing data. The billing system is referred to as the Business Support System (BSS).

Additional Considerations for Interactive Broadband Servers of Conventional Architecture

LIBRARY STORAGE VERSUS PRIMARY STORAGE

The speed of importing content from storage into the stream generator is a key variable in server design because it determines how quickly the viewer can see requested material. The delay between request and fulfillment is called “latency.”

Figure 2
The Conventional Trial Size Server



Source: S. Rose

The critical speed for importation is real time (e.g., four Mb/s). If the content can be downloaded at the same or a faster rate than the rate of delivery to the subscriber, then it is possible to begin to deliver the content shortly after beginning the download. However, if the import rate is much slower than real time, then the entire program must be loaded before delivery to the subscriber can begin.

For example, a dominant vendor’s product takes 3× real-time to download from a single tape drive to the stream generator. This speed means that if a customer requests a 100-minute movie that is not already stored in the stream generator, it will be five hours after loading begins from tape before the movie can even begin to be delivered to the subscriber. A further consequence of the 3× real-time import speed and a single drive is that when the system launches, if it has a capacity of only 200 hours of content (120 movies), it will take 600 hours to load them into the stream generators, or 15 work-weeks of eight-hour work days, or five weeks of 24-hour days!

A rapid import rate confers important advantages that consumers like, such as “VCR functionality.” If the material is imported from the library at a real-time rate, the subscriber can be allowed VCR functionality (except fast forward). If it is imported at or above the fast forward rate, the subscriber can receive the requested program within seconds of the time that the download begins—with full VCR functionality.

It is not possible to implement fast forward functionality directly from the library by jumping ahead in the material because the read element of storage devices (CD-ROM, DVD, tape drive, or hard disk) addresses only one stream at a time. If the library read element jumped ahead to support a viewer's request for fast forward, there would be a gap in the material stored in the hard disk array. If a second customer were to request the same program, it would not be possible to deliver it until the first viewer was done, as the copy stored on hard disk (which can support multiple viewers) would be incomplete.

The nature of the transfer from the library to hard disk storage makes an enormous difference in the nature of the server. If the transfer occurs at a rate equal to or faster than real time and the server supports immediate delivery, then the library becomes the primary storage of the system, and consumers have immediate access to the full content of the library.

If the transfer is slower than real time or immediate delivery is not supported, then the only titles which may be offered to subscribers for immediate consumption are those already loaded in hard disk storage. This usually represents the difference between 100 or 200 and thousands of titles.

STREAM GENERATOR SIZE

Although 100 is a large number of isochronous streams from the perspective of conventional computer technology, it is a drop in the bucket relative to the needs of a metropolitan interactive broadband server. In a 50,000 subscriber area, 10,000 peak streams means 100 conventional stream generators will be needed to fulfill subscriber requests for video. These stream generators each occupy from one-fourth of a rack to three full racks (about the size of a phone booth), so it will require 25 to 300 racks of equipment to generate video streams.

ATM SWITCH LIMITATIONS

The ATM switches used to sort the video streams bring their own set of problems. The ATM protocol is quite rigid, which means that stream generators must produce fully standards-compliant output. This compliance adds considerably to the cost of the switch. Another difficulty is that even a large ATM switch is small when used for compressed digital video. For example, a large 16 Gb/s ATM switch, even if it could be fully utilized, would provide about one-third of the 40 gigabit bandwidth needed by a 10,000 stream server. The sheer volume of video data makes

it necessary to partition each server complex into multiple independent servers and switches. This partitioning is expensive, inefficient, and difficult to manage.

Another limitation is the nature of the traffic, which is largely unidirectional because so much of the information is the downstream delivery of high bandwidth video on demand to subscribers. The design of ATM switches assumes that there will be approximately the same amount of traffic in both directions, and each downstream channel is paired with an upstream channel.

It would seem logical to reverse some of the upstream channels and use them to support the downstream traffic. Unfortunately, when an ATM switch is wired this way, it triggers SNMP (Simple Network Management Protocol) error messages. Disabling the error messages removes the ability to manage and monitor the switch, which is then operating in the "crossed-fingers" management mode.

As a result, upstream and downstream channels must remain paired, and almost half of the I/O capacity of the switch (and in some cases, some of its throughput bandwidth) goes to waste. This results in up to twice as many switches being required.

INTELLIGENT MODULATOR SIZE

If each modulator outputs six streams, it means that 17 modulators will be required for each neighborhood to meet the 100 stream anticipated peak demand ($6 \times 17 = 102$). In a headend with 100 neighborhoods, 1,700 modulators of this type are required. One brand of modulators fits four on a standard equipment rack; the smallest modulators fits about 22. Thus, the system operator needs 77 to 425 equipment racks to provide for just the downstream traffic of a 10,000 stream server. In fact, control and signalling channels for each household call for additional modulators and demodulators, accounting for at least 10 more racks of equipment.

SCALABILITY

The fundamental problem with conventional designs is that they don't scale. They can be made larger by full replication of small servers, but this creates substantial redundant (and expensive) storage. In addition, it creates new problems of headend design and cost of operation.

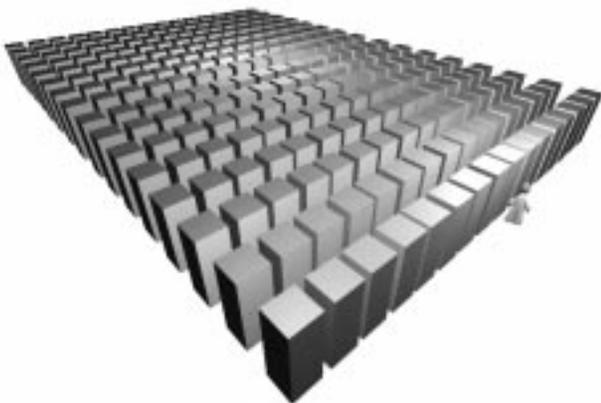
For example, our considerations so far have led us to conclude that the conventional interactive broadband server designed for a 50,000 subscriber area calls

for between 25 and 300 racks of equipment for stream generation, and 77 to 425 racks for downstream modulation and associated processing. Switching equipment is relatively small, requiring about 10 to 20 racks of gear. The sum of these requirements amounts to between 112 and 745 racks of equipment.

If we allow 12 square feet per rack including service access isles, and 15% for office space, air and power conditioning equipment, then a facility for 50,000 subscribers will occupy from 1,600 to 10,280 square feet. While these areas are less than huge, they are substantial when compared with the customary 600 square feet used by the headend of an average 50,000 subscriber cable system. For a 500,000 subscriber system, the total area would be about an acre, full of electronic equipment. Zoning approval delays and ongoing real estate costs must be factored into the planning for a conventional IBS.

Now, let us turn to the problem of power that the conventional design requires. In the design plans for an actual trial-size facility of an interactive broadband server, estimates of power requirements for the stream generators, switches, and modulators added up to 200 watts per video stream. This figure means that a 50,000 subscriber server putting out 10,000 digital streams would require two megawatts. Even assuming there might be some economies that could be employed to reduce the usage to 50 to 100 watts per stream, it would still result in a monthly bill of \$70,000 to \$140,000 including air conditioning.¹

Figure 3
The 256-Rack Headend



Source: S. Rose

SYSTEM RELIABILITY

The need to compete with videocassette rental outlets results in a strange paradox: The digital service that brings in the least revenue (per megahertz of bandwidth) demands the greatest system reliability. Video on demand service requires that three gigabytes of data be delivered over a 100-minute period for a total revenue of as little as \$0.99.

In order for the viewer to receive a coherent picture, the delivery must be nearly perfect. What is transported is not digital video, where a single error results in a bad pixel that disrupts a tiny portion of the picture for 1/60th of a second. Rather, the data is a compressed digital data stream, where a modest error rate can affect a large part of the screen for up to several seconds or even disrupt the session altogether. If the disruption results in a complaint from the subscriber, it costs more than \$1.00 to handle the phone call.

Even without the loss of the revenue from the movie, there is an extremely narrow profit margin because the net revenue from the \$0.99 to \$4.00 charge is so low. Any disruption of service will offset any possible profit. The server itself contributes little to overall system reliability; however, it affects the new digital services from which operators hope to derive additional revenue. The server must be designed to deliver signals without disrupting service, in spite of the failure of any one of its components.

Conventional servers are sometimes designed with RAID technology to accommodate the failure of single hard disks. However, the remainder of the server is susceptible to single point failures. The sheer amount of equipment and the complex interfacing of equipment from multiple suppliers called for by the conventional approach multiplies the probability of failures and the difficulty of correcting them when they occur.

Summary

The evolution of the conventional IBS has been essentially linear. Each function has been addressed by adding on a new layer of hardware and software. The design begins logically with storage and stream generation. Then, the switch is added for sorting, routing, and multiplexing. Appended downstream modulators encode the output for transport. Then, the need for end-to-end management results in the overlay of an operational support system. Finally, business requirements demand an additional system to monitor usage, store the data, and charge customers for service.

Table 1 recaps traditional approaches to the design of interactive broadband servers:

Table 1
Conventional Server Design Solutions

Function	Conventional Solution
Importing content	Slower than real-time from tape
Generating isochronous streams	Ganged single bus computers
Sorting and routing streams	ATM switch
Multiplexing output streams	ATM switch
Quality monitoring, FEC	Intelligent modulators
Facility design (50K subs)	112 to 745 racks of equipment
Network management (OSS)	Separate computer
Billing, other business	Separate computer support (BSS)
Reliability	Problematic: vulnerable to single failures

Source: Van Tassel & Rose

Preview of Part II

In the second part of this article that will appear in the next issue, we present our point of view about how to design an integrated IBS that might perform these functions in more efficient, cost-effective, and reliable ways. We will look at the application of massively parallel processing technology to an IBS and describe a design that looks at the task as a throughput problem rather than as an input/output problem. We will also present innovative ideas for configuring storage. Finally, our suggestions will address the problems of headend design and power consumption. In the process, we will present a metaphor to provide common ground for future discussions. **NTQ**

Authors' Note—Questions or comments can be sent to the authors via e-mail to Joan Van Tassel at jvantass@pepperdine.edu and Steve Rose at roses@maui.com.

¹ Here is how the numbers play out. Electrical usage takes place whether or not the operator is actually generating streams; that is, the equipment is always on drawing electricity. Each of the 10,000 streams consumes 100 watts of power. Thus, at any instant, the system is using a megawatt of electricity. Over the course of an hour, this consumption becomes a megawatt hour, or 1,000 kilowatt

hours for all the streams. At \$0.10 per kWh (1,000 ¥ .10 = \$100), the systems consumes \$100 an hour. Given 24 hours in a day, electrical power alone will cost \$2,400 per day, or \$72,000 per month. Air conditioning the equipment stored on this many racks will double the electricity needed by the headend.