Video Server Architecture

Ralph W. Brown
Time Warner Cable

Abstract

Video-on-demand and other interactive television services are driving the development of video server technology. By their very nature these video servers are multiprocessor systems. This is especially true when the server is required to provide hundreds, thousands, or even hundreds of thousands of video streams. This paper reviews a number of the proposed video server architectures. These architectures differ in some key areas and this paper compares their advantages and disadvantages. Issues regarding their impact on software architecture, reliability, and cost are also discussed. Finally, a summary of areas for future work is presented.

INTRODUCTION

Video-on-demand and other interactive services such as interactive games and interactive shopping are driving the development of video servers [1,2]. Video servers are the computer systems that provide the storage and playback of MPEG compressed video data for interactive services in a broadband network environment. As they are designed to serve a large number of subscribers, video servers are shared resources located at the headend of the cable distribution network [3,4,5].

Figure 1 is a simplified diagram of a hybrid fiber/coax based interactive cable television system. This diagram shows the relationship of the video server to the other components in the end-to-end cable system and identifies the distribution headend and distribution network components. The headend equipment includes one or more video servers and an ATM switch or switching network. For simplicity, the ATM switch is shown as a single switch, but actually may be a network of ATM switches. The number of ATM switches required depends on the service requirements. Alternatively, its function may be subsumed as part of the video server itself. The distribution network is made up of the fiber-optic links between the headend and the neighborhood area network nodes (NAN) and the coax links from the NAN to the subscribers' homes. The home communications terminals (HCT) are located in the subscribers' homes and decode the MPEG data sent to it from the video server at the headend. Not shown in Figure 1 are the modulators, demodulators, amplifiers, and combiner/splitter networks that are also part of the distribution network. The traffic flow in the distribution network is asymmetric as the bandwidth required from the headend to the HCT for MPEG data is much greater than the return path for subscriber requests.

The role of the video server is to provide storage and playback of the video data and to respond to subscriber requests such as "pause", "fast forward", or "view the next product". The ATM switch is responsible for routing data from the video servers to the appropriate neighborhood nodes. The MPEG compressed video data is transmitted using ATM transport [6]. The ATM data is modulated using QAM modulation and transmitted over fiber from the headend to the neighborhood node. At the neighborhood node it is converted from an optical signal to an electrical signal for distribution over coax to the subscriber homes. These signals are then received by the home communications terminal where the video data is decompressed and displayed on the subscriber's television.

The actual number of video servers required in Figure 1 will be determined by the capacity of an individual server and the number of active HCTs being served. In the larger video server architectures the switching function provided by the ATM switch may be subsumed by the interconnection network of the video server.

Video servers are multiprocessor systems. Uniprocessor systems are unable to sustain the high data rates necessary for this type of application. A variety of multiprocessor architectures have been proposed to meet the high bandwidth requirements of a video server. The following sections will compare and contrast these video server architectures.

VIDEO SERVER ARCHITECTURES

All video servers are implemented by some form of multiprocessor architecture. Video server architectures are characterized by two key attributes the type of CPU interconnect that is used and the video data path from disk to distribution network. Based on these two characteristics video server architectures fall into one of three categories:

- Symmetric Multiprocessing (SMP)
- Massively Parallel Processor (MPP)
- Loosely Coupled Computer (LCC)

The following sections discuss these architectural categories in detail, their advantages and disadvantages, and implications for software architecture.

Symmetric Multiprocessing (SMP) Architectures

Shared memory SMP architectures are characterized by a high-speed system bus interconnecting CPUs, memory, and disk and network I/O subsystems. Figure 2 shows a diagram of an SMP architecture. This figure shows only four CPUs, but more are possible. Because the memory and I/O devices are shared equally among all of the CPUs in the system these architectures are called symmetric multiprocessor systems. Typically, these systems scale from one CPU to as many as 36 CPUs. Examples of SMP architecture include the Silicon Graphics, Inc. Challenge server systems and the Sun Microsystems, Inc. SPARCserver systems.

In the SMP architecture, the processing load is uniformly distributed across

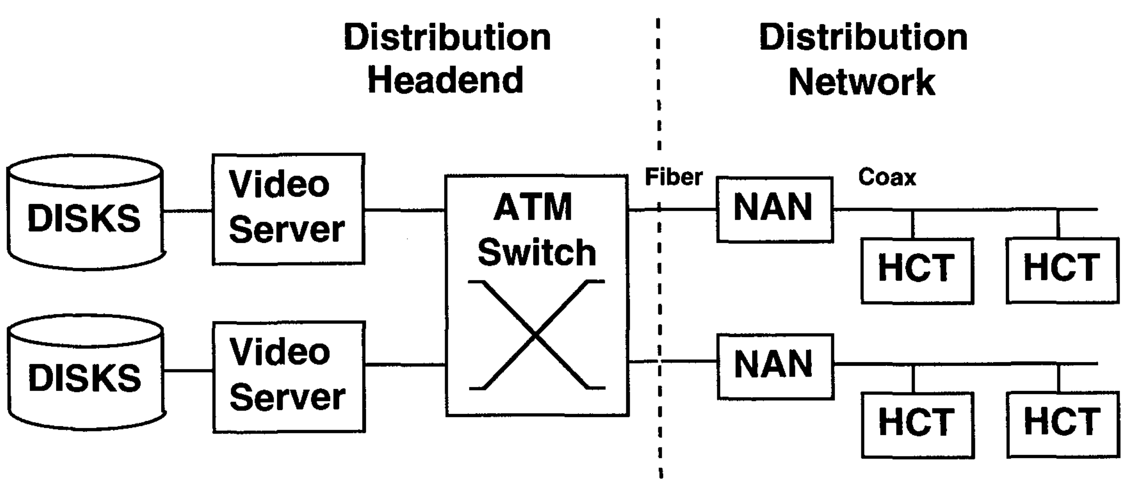

Figure 1 - Interactive Cable Television System

all of the CPUs in the system. Scheduling algorithms in the kernel allocate processing load to unoccupied or lightly loaded CPUs. All CPUs have equal access to the physical memory and see a uniform memory image. Data written to the shared memory by one CPU is immediately available to all the other processors. Software processes, both kernel and user, are able to run concurrently, and can be switched arbitrarily among any of the processors in the computer. Since all of the CPUs are operating out of shared memory, access to critical data and code is protected by locking mechanisms that prevent concurrent access.

Data flow for video data in an SMP system is straightforward. Video data is read from the disk into a buffer area in the shared memory of the system. The data is then transferred from this buffer area to the ATM network interface for transmission to the ATM switch. In this way, the video data crosses the system bus twice in the process of moving from storage on the hard disk to network transmission. As a result the system bus must have relatively high bandwidth to support the video traffic for all streams active in the system as well as all other general processing requirements. An example is the Silicon Graphics, Inc. Challenge server systems, which have a system bus bandwidth of 1.2 GBytes/second.

Massively Parallel Processor (MPP) Architectures

Massively Parallel Processor architectures, as the name implies, use large numbers of processing units. The processors are independent but are interconnected through high-speed networks. Unlike the SMP architecture, each processor in the MPP architecture has its own memory and I/O devices. This memory and I/O devices are local to each CPU and cannot be seen directly by other CPUs in the system.

In large scale MPP systems the issue of inter-processor communication becomes critical. To efficiently share information between processors it is important to have a high-speed interconnect between them. It is impractical to build a high-speed interconnection network that connects every processor to every other processor through a point-to-point network. Rather an approach that minimizes the latencies between processors and scales well is the hypercube architecture. Typically these systems scale from as few as 16 CPUs to as many as 8,192 CPUs. An example of an MPP architecture

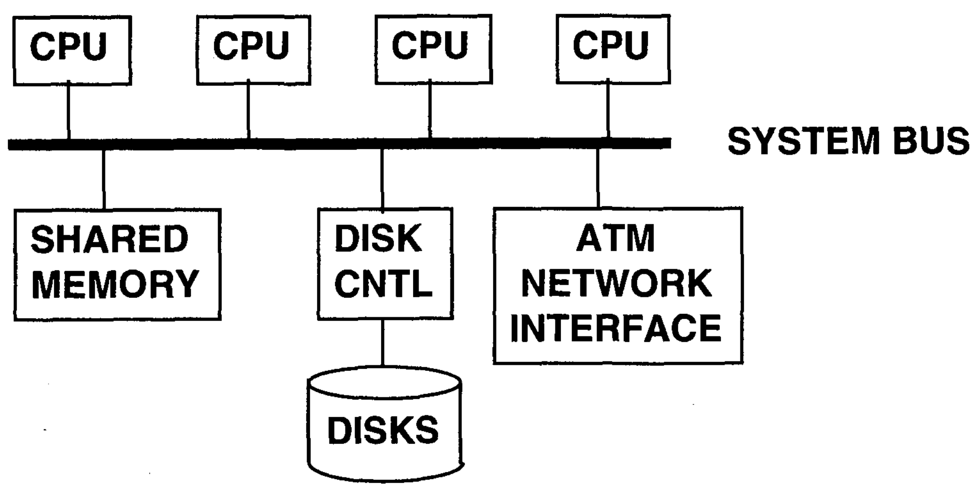

Figure 2 - Symmetric Multiprocessing Architecture

for video server applications is the nCube, Inc. MPP Media Server [7].

In the hypercube architecture each processor is connected to N of its neighbor processing units, where N is the order of the hypercube. The order N determines the number of processors in the system, *Number of Processors* $= 2^N$. The order N also determines the maximum number of links and the maximum delay between any two processors, *Maximum number of links* $= N$. Figure 3 shows the hypercube MPP architecture for orders one, two, and three. Each CPU has its own memory and I/O interfaces. The I/O interfaces will connect either to the disk controller, ATM network interface, or hypercube interconnect. The video data flows from the disk connect to a CPU and is routed either through the hypercube interconnect to the appropriate CPU or directly out through the ATM network interface. When the video data is routed through the hypercube interconnect, it may not be necessary to have an ATM switch in addition to the video server. Since the video data managed by an individual CPU is significantly less than for the video server as a whole, the bus bandwidth for an individual processing unit is correspondingly lower. However, the hypercube interconnect must be of relatively high bandwidth as it must carry a fairly high percentage of the video traffic. For example, the nCube system has a bandwidth of 2.5 MBytes/second for each link in the hypercube interconnect.

Loosely Coupled Computer (LCC) Architectures

Similar to MPP architectures, LCC architectures are characterized by many independent processing modules containing CPU, memory, and I/O subsystem, linked together through a high-speed interconnect network. Unlike MPP architectures, the LCC architectures do not scale to as large a number of CPUs and use low-cost, off-the-shelf components such as Intel 486 or Pentium-based computers. LCC architectures scale to hundreds of CPUs rather than the thousands found in MPP architectures. LCC architectures typically use a mesh based interconnect rather than a hypercube. Examples of LCC architectures include the Intel Corporation Scalable MultiServer [8] and the Digital Equipment Corporation Interactive Information Server [9].

Figure 4 shows a diagram of a LCC

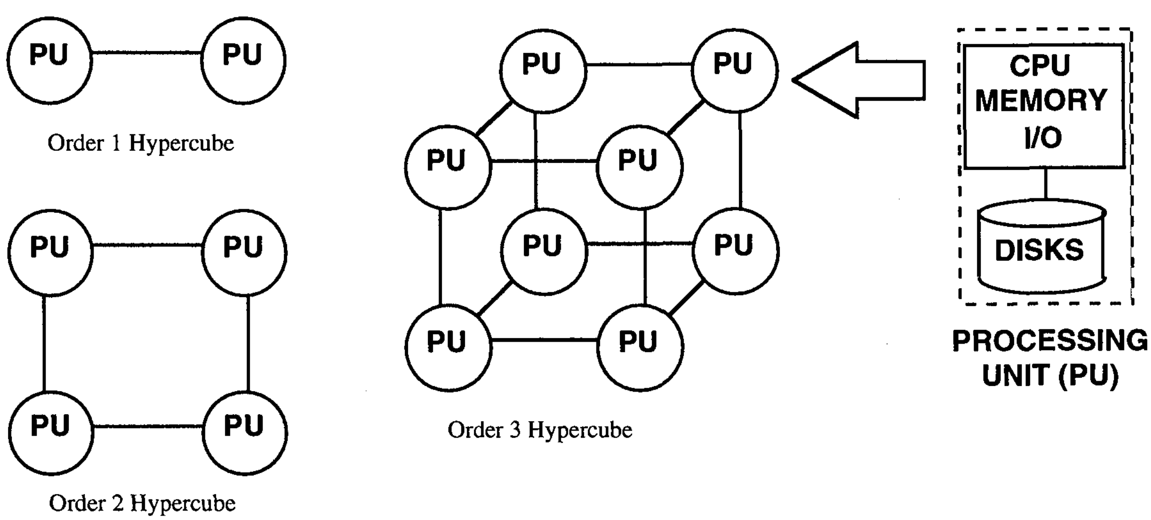

Figure 3 - Massively Parallel Processor Architecture

architecture. In this architecture, each of the processing units represents an Intel 486 or Pentium-based computer. These low-cost computers are connected through a two-dimensional grid, implemented as a cross bar switch or mesh network. This increases the network delay between processing units and decreases the size to which the system can be scaled. Large scale crossbar or mesh networks are impractical to build.

The data flow of video data in the LCC architecture is very similar to the data flow in the MPP architecture. Here, video data is transferred from the disk to the mesh interconnect for routing to the appropriate processing unit for distribution to the ATM network. In some instances, the video data for an individual movie or video clip will be spread across several of the processing units in the system for redundancy and improved concurrent access.

ARCHITECTURE COMPARISONS

The three video server architectures, SMP, MPP, and LCC, also impact four key areas:

- software architecture
- scalability
- reliability
- cost

The following sections compare the three architectures' impact in each of these areas.

Impact on Software Architecture

The primary difference between the SMP, MPP, and LCC architectures that affects the video server software architecture is the memory and CPU interconnect; the shared-memory model of the SMP architecture versus the local-memory model of the MPP and LCC architectures. In the SMP architecture there are multiple processors looking at a unified physical memory space. In other architectures each processor is only looking at its own local memory. This single aspect, more than any other drives the video server software architecture.

The high-speed system bus allows much faster interprocessor communication in the SMP architecture than in either the MPP or LCC architectures making data sharing

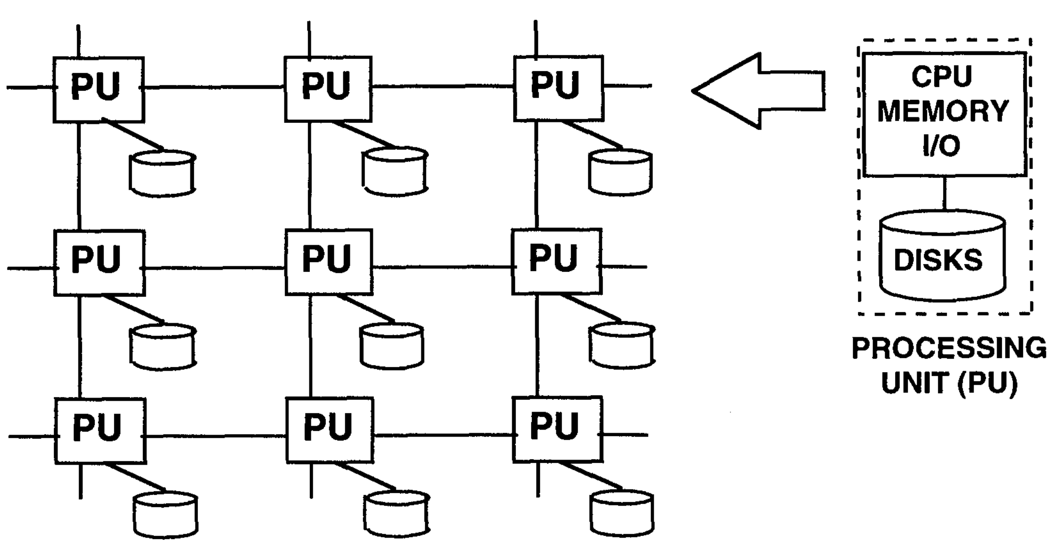

Figure 4 - Loosely Coupled Computer Architecture

between processors is much simpler and faster. The SMP architecture is easier to program and provides faster synchronization between processors and processes. The other architectures have a more difficult time providing a simple multiprocessor programming environment and fast process synchronization.

In the SMP architecture the application software is shielded from the details of which CPU is actually running the software at any given time and synchronization with other processes on other CPUs is relatively transparent. The application developer can view the system as a uniprocessor environment rather than a distributed processor environment. In addition, the underlying operating systems for SMP architectures are well understood and many robust implementations can be found in the SMP server market. In MPP architectures the application developer must design the application to make use of the distributed processing environment. The MPP architecture must also have a distributed operating system implementation to provide the infrastructure on which applications are layered. Similarly, LCC architectures require applications developers to design applications to take advantage of the distributed processing environment. LCC architectures generally use a nondistributed operating system and place the distributed communications functions at the application level. This provides less support for interprocessor communication than in MPP architectures.

The following is a list of the software architectures that are most commonly mentioned in the video server arena:

- Microsoft Tiger - This software architecture is targeted to an LCC architecture and is based on the Windows NT operating system. The Windows NT operating system is currently ported to the Intel X86 and MIPS processors.

- Silicon Graphics, Inc. / Interactive Digital Solutions - This software architecture was developed for the Time Warner Cable Full Service Network trial in Orlando and is tightly coupled to the Silicon Graphics SMP server architecture [10].

- Oracle Media Server - This software architecture is based on the Oracle relational database management system (RDBMS) and may be ported to a number of different hardware platforms. The article "Oracle Media Server: Providing Consumer Based Interactive Access to Multimedia Data" [11] discusses the Oracle Media Server in detail. This software has been implemented initially on the nCube MPP architecture.

- Sybase Intermedia Server - This software architecture is based on the Sybase RDBMS, although, it proposes to integrate the RDBMS with an independent video server rather than provide the video server software itself.

Impact on Scalability

Scalability is the measure of how the capacity of video server can be increased to support additional video streams. This can be achieved by adding CPUs, network interfaces, disk interfaces, additional servers, or any combination of the preceding elements.

In general, the SMP architecture scales well up to the maximum number of streams supported within one server. Provided the system bus bandwidth is sufficient, additional CPUs, network interfaces, and disk interfaces can be added to the system without requiring software changes. Beyond the limits of one server, additional servers must be added to increase capacity.

The LCC and MPP architectures have the potential to scale to larger number of streams within one video server than SMP architectures. In the SMP architecture you can only add so many CPUs and interfaces to a server before you have to add more servers. MPP and LCC architectures are designed to scale to larger numbers of CPUs and interfaces within one video server.

Impact on Reliability

Since a video server is a shared resource that may serve thousands of customers, reliability is critical. A single CPU failure or an uncorrectable memory error can bring down the entire video server in an SMP architecture. Also, since the memory and CPUs are tightly-coupled, errors may propagate to other parts of the server. A similar failure in the MPP or LCC architecture will only kill one of the processing units. If the operating system is designed to detect and correct for this problem the video server can continue to function with this loss. While most SMP servers do have the ability to map out failed hardware as the system boots, temporary loss of service still results as the system is rebooted. To accommodate this aspect of SMP architectures, redundant video servers are used. This will dramatically affect reliability and mean-time-between-failures (MTBF) figures.

Impact on Cost

There are two critical costs to consider when evaluating video servers:

- Entry-level costs
- Incremental-scaling costs

The entry-level cost for SMP architectures is moderate due to the cost of the high-bandwidth, system bus infrastructure. The SMP architecture carries lower incremental expansion cost, however, up to the limit of the server capacity, after which a new server must be added.

The entry-level cost for MPP architectures is fairly high due to the cost of the infrastructure to support the large hypercube interconnect. However, the MPP architecture carries lower incremental expansion cost to scale to larger numbers of video streams.

Finally, the entry-level cost for LCC architectures is low due to the leveraging of standard off-the-shelf component computers. The incremental expansion costs are also low for the same reason.

ISSUES FOR FUTURE DEVELOPMENT

There are a number issues that require further research and development:

- Use of tertiary storage
- Cost
- Interoperability

The following sections discuss each of these issues in turn.

Use of Tertiary Storage

The cost of disk storage is a major component of the total video server cost. The use of tertiary storage media such as magnetic tape or CD-ROM can significantly reduce the cost-per-megabyte of storage. Unfortunately, these media do not have the access latency and bandwidth capacities to directly support video delivery. In the future it will be necessary to develop sophisticated video data caching schemes to take advantage of these low-cost storage media.

Cost

As discussed earlier, entry-level cost and incremental-scaling costs are the important cost factors. Video server costs will continue to be a large portion of the over all end-to-end cost of providing interactive television services. Thus, it is necessary to make the most effective use of the video server capability as is possible. The article "Making a Cost-Effective Video Server" [12] discusses how placement of video data files in either dynamic random access memory (DRAM), disk storage, or tertiary storage affects the cost-effectiveness of a video server implementation.

Interoperability

There at two types of interoperability that are important. The first is interoperability between video servers and HCTs and the second is interoperability between the servers themselves. In order to make the most cost effective decisions, the network operator must be able to choose the most cost effective settop or server with out regard to interoperability. It is desirable to have the ability to mix and match video servers over time to increase overall capacity and to take advantage of the latest improvements in technology. Standards groups such as DAVIC are working towards defining standards for these levels of interoperability.

ACKNOWLEDGMENTS

The author would like to acknowledge the tremendous efforts of the Time Warner Cable Denver and Orlando teams in building the Time Warner Cable Full Service Network.

REFERENCES

1. Thomas D. C. Little and D. Venkatesh, "Prospects for Interactive Video-on-Demand," *IEEE Multimedia*, Vol. 1, No. 3, Fall 1994, pp. 14-24.

2. Daniel Deloddere, W. Verbiest, and H. Verhille, "Interactive Video On Demand," *IEEE Communications Magazine*, Vol. 32, No. 5, May 1994, pp. 82-88.

3. Winston W. Hodge, *Interactive Television - A Comprehensive Guide for Multimedia Technologies*, McGraw-Hill, Inc., New York, 1994.

4. Yee-Hsiang Chang, D. Coggins, D. Pitt, David Skellern, M. Thapar, and C. Venkatraman, "An Open-Systems Approach to Video on Demand," *IEEE Communications Magazine*, Vol. 32, No. 5, May 1994, pp. 82-88.

5. Louis D. Williamson, "FSN Technology," *Proceedings: Society of Cable Television Engineers 1995 Conference on Emerging Technologies*, Jan. 4-6, 1995, Orlando, FL, pp. 27-35.

6. Michael B. Adams, "Full Service Network Software Architecture," *Proceedings: Society of Cable Television Engineers 1995 Conference on Emerging Technologies*, Jan. 4-6, 1995, Orlando, FL, pp. 13-26.

7. Robin Bloor, "The Coming of Parallel Servers," *DBMS Magazine*, May 1994, Vol. 7, No. 5.

8. Intel Corporation, Supercomputer Systems Division, "Reply to Cable Television Laboratories, Inc. Request for Information - Digital Media Servers," June 30, 1994.

9. Digital Equipment Corporation, " Response to Request for Information for Digital Media Servers," June 30, 1994.

10. Ralph W. Brown, "Full Service Network Software Architecture," *Proceedings: Society of Cable Television Engineers 1995 Conference on Emerging Technologies*, Jan. 4-6, 1995, Orlando, FL, pp. 6-12.

11. Andrew Laursen, J. Olkin, and M. Porter, "Oracle Media Server: Providing Consumer Based Interactive Access to Multimedia Data," *Proc. SIGMOD 94*, Minneapolis, May 1994, pp. 470-477.

12. Yurdaer N. Doganata and Asser N. Tantawi, "Making a Cost-Effective Video Server," *IEEE Multimedia*, Vol. 1, No. 4, Winter 1994, pp. 22-30.