EXPERIENCES WITH AN ARCHITECTURE FOR A DISTRIBUTED MULTIMEDIA SYSTEM

Nitin S. Naik
NASA Classroom of the Future Program
316 Washington Ave.
Wheeling, WV. 26003

and

Sitharama S. Iyengar
Department of Computer Science
Louisiana State University
Baton Rouge, LA. 70803

November 1996

Introduction

Computer communication demands for higher bandwidth and smaller delays are increasing rapidly as the march into the twenty-first century gains momentum. These demands are generated by visualization applications which model complex real time phenomena in visual form, electronic document imaging and manipulation, concurrent engineering, on-line databases and multimedia applications which integrate audio, video and data. The convergence of the computer and video worlds is leading to the emergence of a distributed multimedia environment.

The initial step in providing multimedia services over computer networks is to ensure bandwidth availability for these services. In this paper a computational framework is developed in the context of the bandwidth needs for traffic generated in a distributed multimedia environment. Thereafter, this model is applied to the real-time problem of designing a backbone for a distributed multimedia environment at the NASA Classroom of the Future Program. The network incorporates legacy LANs and the latest high speed switching technologies.

Performance studies have been conducted with different network topologies for various multimedia application scenarios to establish benchmarks for the operation of the network. In these performance studies it has been observed that network topologies play an important role in ensuring that sufficient bandwidth is available for multimedia traffic.

Multimedia Communication Requirements

Multimedia communication requirements impact various aspects of the integrated network design. The key parameters that must be monitored in this network are bandwidth, transmission delay, capability of multipoint communication, reliability and channel synchronization. Let us consider each of these in further detail.

1. Bandwidth - DVI and MPEG standards for video and audio compression suggest that for good quality the data rate should not be lower than 1.5 Mbps. With regard to wide-area transmission cost, existing H.261 implementations show that 64 Kbps is only acceptable in almost static “talking head” video signals, whereas the 384 Kbps (six bundled ISDN channels) variant yields acceptable results even in a more general environment. This bandwidth is required for simplex, i.e. unidirectional, streams because multimedia traffic is highly asymmetrical. Thus for our design a minimum of 1.5 Mbps must be guaranteed in each direction to each workstation.
2. Transmission Delay - International Communications Union (ITU) standards suggests a maximum total end-to-end delay of up to 150 ms for interactive video applications. Video and audio communication fall into the isochronous class of applications that have an upper bound for the maximum transmission delay and also requires a constant transmission delay for the different packets. Isochrony does not have to be maintained across the entire path from source to sink but only at the information’s final destination. Thus isochrony can be recovered by the use of buffers at the destination. The recovering process will introduce some delay. The end-to-end delay can be broken down into at least four contributing pieces: (i) source compression and packetization delay; (ii) transmission delay; (iii) end-system queuing and synchronization (playout) delay; and (iv) sink decompression, depacketization and output delay. Video streams require the handling of 25 to 30 frames per second. Thus real-time compression or decompression times must not exceed 30 to 40 ms. Using another frame period for queuing and playout delay, leaves 60 ms for the maximum transmission delay.
3. Reliability - Traditional communication aims at providing reliable end-to-end communication between two peers. Existing communication systems always use checksum and sequence numbering for error control and some form of negative or missing positive acknowledgement with packet retransmission handshake for error recovery. If check summing is not performed either in hardware at the media access control or link layer, it can affect the system performance drastically. The acknowledgement with subsequent retransmission handshake adds more than a round-trip delay to the transmission of this data. For time critical data such as video and audio streams, the retransmitted packet may be useless. Therefore, in the case of the time sensitive transmissions, the network must let the error control and recovery schemes be handled by the higher communication layers. They can provide the level of reliability required, taking into account the impact on the delay characteristics. A possible remedy for the conflicting goals of reliability and low delay is to use forward error correction (FEC) techniques. The issue of reliability becomes even more complex for multipoint communication.
4 Channel Synchronization - When audio, video and other data streams are delivered from different sources via different routes, it is necessary to synchronize these different streams at the sink. Synchronization can be achieved using a combination of time-stamping and playout buffers. In our design we assume synchronization to be built into the operating system on the client and the server.
5 Multipoint Communication - Multimedia communication involves audio and video broadcast information. Thus the integrated multimedia communication needs to support multicast communication patterns in addition to normal point-to-

point communication.

Multimedia Network Specification

An infrastructure constructed of the legacy LAN technologies such as Ethernet or Token Ring may be sufficient for supporting multimedia applications in a small scale. A small group of users who retrieve video from a server or participate in a video conference may be serviced by a single ethernet segment. However, as the number of multimedia-capable users grows, the aggregate bandwidth of the network must also grow to support their communication requirements. Further, videoconferencing applications are not likely to stay restricted to the individual workgroups or subnetworks but We present a computational model that will enable us to identify the type of network required to be installed in a distributed multimedia system which is an extension of Tobagi's model []. Let us consider the following variables:

- N^{TOT} : Total number of users on the network
- S : Number of subnetworks in the network
- N_i^{SUB} : Number of users in subnetwork i, i= 1.....S
- T^{BACK} : Traffic in the backbone
- B_{MULT} : Data rate of the multimedia stream in bits/second
- γ : Fraction of the traffic generated in the segment destined to users in the same segment
- α : Fraction of the traffic generated in the subnetwork destined to users in the same subnetwork
- δ : Fraction of the users in the network which are multimedia capable

We assume that each multimedia capable user generates a stream of data rate B_{MULT} and consider the subnetwork consisting of switching ethernet hubs then each port is considered as a segment. Further, in a distributed multimedia system due to the bandwidth requirements of transmitting video and audio streams, one may attach only one workstation per port of the hub. In this case, each segment contains only one user and hence $\gamma = 1$. If we consider a completely distributed multimedia system then every station in the subnetwork has the capability of either generating the multimedia stream or at least viewing a multimedia stream. This leads to the variable δ=1. The traffic in the backbone can also be represented by the following equation

$$T^{BACK} = \sum_{i=1}^{S} (1 - \alpha)\, \beta\, N_i^{SUB}\, B_{MULT} \qquad (1)$$

But

$$N^{TOT} = \sum_{i=1}^{S} N_i^{SUB} \qquad (2)$$

we have

$$T^{BACK} = (1 - \alpha)\, \beta\, N^{TOT}\, B_{MULT} \qquad (3)$$

The backbone traffic is found to be proportional to the following: (i) the fraction of traffic generated in the subnetworks that is destined to other subnetworks; (ii) the number of streams that each workstation can handle simultaneously when attached to the network; (iii) total number of users on the network; and (iv) bandwidth of the multimedia streams.

Consider a video server distributing a live broadcast and say half the number of users were participating, in small groups of 4, in video conferences with each other regarding the broadcast then B_{1MULT}=1.33 Mbps, β_1=1 and B_{2MULT}=384 Kbps, β_2=4. The backbone capacity for different number of clients is shown below:

If N^{TOT} = 50 and (1-α) = 0.2 then T^{BACK} Mbps = 28.66

If N^{TOT} = 200 and (1-α) = 0.4 then T^{BACK} Mbps = 229.28

If N^{TOT} = 600 and (1-α) = 0.6 then T^{BACK} Mbps = 1031.76

If N^{TOT} = 1500 and (1-α) = 0.8 then T^{BACK} Mbps = 3439.20

It can be seen that with a small number of workstations on the network we can use ethernet, fast ethernet or FDDI network architectures but as the numbers increase ATM and switching technology need to be considered. The high-speed and high-capacity backplanes of the switching hubs can sustain the bandwidth demands of the multimedia based activity on the network.

NASA Classroom of the Future Program Testbed

An integrated computer-video local area network (C/V LAN) is designed and implemented as part of this research at the Center for Educational Technologies for the NASA Classroom of the Future Program. The LAN is installed using NASA support to facilitate the COTF Programs' mission of enhancing the learning and teaching process for mathematics, science and technology education using advanced computer and telecommunications technologies. The layout of the computer-video LAN implemented at the NASA Classroom of the Future Program has been designed taking into consideration all the issues raised in the previous sections. The network is one of the few networks installed in a production environment that integrates the computer and the video network. A diagram showing the different components is shown in Figure 1.

The computer network serves approximately 75 multimedia workstations capable of running both Macintosh based and PC based applications, 75 other workstations used for administration, software development and multimedia authoring. The network is based on switched ethernet and ATM technology. The decision to use these technologies is based on the capacity calculations discussed in the previous section. The calculations were based on the type of activities taking place on the network. The resources on the network include an internet server, a video server with its large storage and a Novell application server. The network was designed on the premise that the servers would have high bandwidth connections to the network and the workstations were serviced by dedicated bandwidth connections. Compression techniques such as MPEG, QuickTime are being employed to distribute video and audio data. The network also requires synchronization features, guarantees for quality of service and congestion control to maintain the bandwidth for the distribution of digital continuous media.

The video network consists of a number of video recording, playback and display devices. It also includes the audio components required in such a system. The facility is equipped with a television studio, a production control area, an offline editing suite, a Ku-band satellite uplink dish, two C/Ku-band downlink dishes and two weather station receiving antennae. Each of the rooms described in the earlier section have equipment and cameras which are connected to the video network for transmission of video signals and for control of the cameras. The equipment are controlled by a card cage and computers which are connected to the computer network. The routing of the video and stereo audio signals is achieved with video routers and a software with a GUI interface running on the control computers.

The video and audio signals are routed to the digitizing workstations for broadcast by the video server or storage in the video server for later use. The digitizing stations compress the video and audio to either QuickTime, Video for Windows or MPEG format for playback on the multimedia workstations in the various rooms and for use by the software developers. The reverse action may also be achieved by performing an analog conversion of the digitized animation from the computer network to the video network for use in video production or for broadcast using the teleport facility. Thus, there is total integration of the two networks.

Benchmarks for the Operation of the Integrated Network

As part of this research, it is important to gather information of the operation of the network described in the previous section. This information is basically a benchmark for the state of the network while users are operating in the distributed multimedia environment. The scenarios considered in the establishment of benchmarks comprised of: (i) a workstation which digitizes the video and audio in real-time and broadcasts it across the network for a video conference situation;
(ii) clients access video files that are already digitized and stored on the multimedia server which has two 10 Mbps ethernet connections; and (iii) a combination of the above two i.e. clients that are watching the video conference broadcast and viewing video files from the server. It is important to consider that the network topologies used in these scenarios were either a shared media hub, or a switching hub connecting two shared media hubs or two switching hubs with an ATM backbone. Thus the clients were either sharing 10 Mbps between each other or had dedicated 10 Mbps links.

The following observations were made during this benchmarking process.

1. As we increase the amount of bandwidth, by reducing the number of stations per segment, the total number of clients viewing a broadcast or interacting with a video clip from the video server increased.
2. It is further observed that as the bandwidth available on the network and the number of users increase, the bottleneck for multimedia traffic migrates from the network to the desktops, server or the broadcaster. This is due to the fact that additional processing power is required to keep up with the data or requests received from the network.
3. The operating system and legacy LAN protocols used in the network implement

the best effort policy. As the number of clients or multimedia capable stations increase on the network, the number of simultaneous requests at the multimedia server increase or the number of clients watching a broadcast or participating in a video conference increase. This results in network bandwidth and system resources being divided to meet the demands of the clients, thereby causing in degradation in service level. In order to overcome this problem a quality of service component must be added to the server and station operating system to ensure that delivery of data simultaneously to some clients and not all clients at the requested or desired levels is maintained at all times.

4. Demand of bandwidth was determined by size of the video stream. The size of video stream is determined by the frames per second at which the video is recorded, the resolution and the number of colors used for depicting depth of picture. Increase in any one of these parameters for any one client results in need for bandwidth increasing quite rapidly.

5. The ideal scenario was 10 where the server had an ATM connection and the individual stations were connected via dedicated 10 Mbps ethernet connections. This is due to the fact that each client talking to the server had a dedicated connection throughout due to the establishment of a virtual circuit from the switching hub to the server. The bandwidth of this circuit was equal to the delivery rate of the stream. It must be noted that the applications and delivery of the video stream is still based on best effort. There needs to be mechanisms designed for guaranteed quality of service for the network and the server operating system.

6. If we compare the network utilization in the scenarios with the data obtained from our computational model, it is observed that the equation needs to be modified as:

$$T^{BACK} = (1 - \alpha)\, \beta\, N^{TOT}\, B_{MULT} + \phi \tag{5}$$

where ϕ represents the overhead required for the transport protocol as a fraction of the total bandwidth capacity required in the backbone and is normally about 15%.

Conclusion

The implementation of the network architecture for a distributed multimedia environment described in the previous chapter was found, during the benchmark operations, unable to effectively use the ATM features such as transporting constant bit rate video and audio data. This drawback is mainly due to the fact that the network is a hybrid topology consisting of ATM and ethernet, and uses best effort policy. These types of networks will be implemented in the field for a long time due to the large cost factor that is encountered for a pure ATM network. Thus, to maximize this combination of network technologies we need to modify the operating system and the ethernet technology to handle multimedia traffic. The multimedia operating system (MOS) must also be able to request quality of service in the distributed environment. The QoS mechanism can be easily implemented in the ATM world with the protocol stack definition but in the legacy LAN world it needs considerable thought. Thus we need to consider how these QoS mechanisms can be implemented in a heterogenous network consisting of ethernet and ATM.

Bibliography

[1] J. F. Adam, H. Houh and D. Tennenhouse, *"A Network Architecture for Distributed Multimedia Systems,"* Proceedings of IEEE International Conference on Multimedia Computing and System, IEEE CS Press, Los Alamitos, California 1994, pp. 76-85.

[2] ATM Forum, *"ATM: User-Network Interface Specification, Version 3.0,"* Prentice Hall, Englewood Cliffs, N.J. 1993.

[3] D.D. Clark, S. Shenker and L. Zhang, *"Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanism,"* Proc. SigComm 92, ACM, New York, 1992, pp. 14-26.

[4] R.L. Cruz, *"A Calculus for Network Delay and a Note on Topologies of Interconnection Networks"*, Report No. UILU-ENG_87-2246, University of Illinois, July 1987.

[5] D. Deloddere, W. Verbiest and H. Verhille, *"Interactive Video on Demand"*, IEEE Communications, Vol. 32, No. 5, May 1994, pp. 82-88.

[6] D. Ferrari, *"Client Requirements of Real-Time Communication Services,"* IEEE Communications, Vol. 28, No. 11, Nov. 1990, pp. 65-72.

[7] R. Garg, *"Characterization of Video Traffic,"* TR-95007, International Computer Science Institute, Berkeley, January 1995.

[8] D. LeGall, *"MPEG: A Video Compression Standard for Multimedia Applications"*, Communications of the ACM, April 1991.

[9] T.D.C. Little and A. Ghafoor, *"Network Considerations for Distributed Multimedia Object Composition and Communication"*, IEEE Network Magazine, pp. 32-49, November 1990.

[10] M. de Prycker, *"Asynchronous Transfer Mode: Solution for Broadband ISDN,"* Ellis Norwood, 1993.

[11] V. P. Rangan, H. M. Vin and S. Ramanathan, *"Designing an On-Demand Multimedia Service,"* IEEE Communications, Vol. 3, No. 7, July 1992, pp. 56-64.

[12] R. Steinmetz and K. Nahrstedt, *"Multimedia: Computing, Communications and Applications,"* Prentice Hall, New Jersey, 1995.

[13] H. J. Stuttgen, *"Network Evolution and Multimedia Communication,"* IEEE Multimedia, Fall 1995, pp. 42-59.

[14] A. S. Tanennbaum, *"Computer Networks,"* Second Edition, Englewood Cliffs, New Jersey, Prentice Hall, 1988.

[15] F. A. Tobagi and J. Pang, *"StarWorks-A Video Applications Server,"* Proceedings of IEEE Compcon Spring 93, CS Press, Los Alamitos, Calif., 1993, pp. 4-11.