Digital Video on Computer Workstations

Andy Hopper

ORL 92-6

To be published in the Proceedings of Eurographics', September 1992

Digital Video on Computer Workstations

Andy Hopper

University of Cambridge, Computer Laboratory, New Museums Site, Pembroke St., Cambridge, CB2 3QG, United Kingdom, and

Olivetti Research Ltd, 24a Trumpington St., Cambridge, CB2 1QA, United Kingdom.

Abstract

This paper describes the way continuous media can be used on networked digital devices. A number of architectures for incorporating digital video on a workstation are presented. These include systems which control the real-time streams but do not handle the data directly, those that send the streams through normal datapaths, and those that are particularly suited for networking of many real-time streams at once. A summary of experience with the Pandora distributed multimedia system is given, together with an outline of a new system under construction, called Medusa.

1. CONTINUOUS MEDIA

1.1 Communicating devices

Digital devices can be used for handling video and other types of real-time data [IEEE 1990, IEEE 1991*a*]. At present the main interest is in video on PCs and workstations, both of which can be used to control and display video streams. However, it can be expected that portable machines will also be able to display some video material and that cameras can be attached to even smaller computers such as active minipads. At the other end of the spectrum, very large displays such as video walls and 3-D systems will become available and will have their own video stream requirements. At present a threshold is being crossed where a standard PC or workstation is able to handle audio and video in the same way as any other type of data.

For applications which involve viewing only, pre-prepared video material can be made available on CD-ROM or other media and the user is able to navigate using paths of his choice. For applications which involve generation of video, the workstation is equipped with a camera, microphone and loudspeaker. The workstation screen can be used to show the video although for some applications this is not necessary and multiple screen systems have merit. The camera is normally close to one side of the workstation but if required one way mirrors can be used to give the impression of the camera being directly behind the screen.

Fig.1. shows the different devices that can be used for display and sourcing of video streams together with some of the applications that are made possible by use of different speed networks. Such communications applications include video-phones, the use of local or remote filing systems for video-mail, and more general inter-working of live and stored streams.

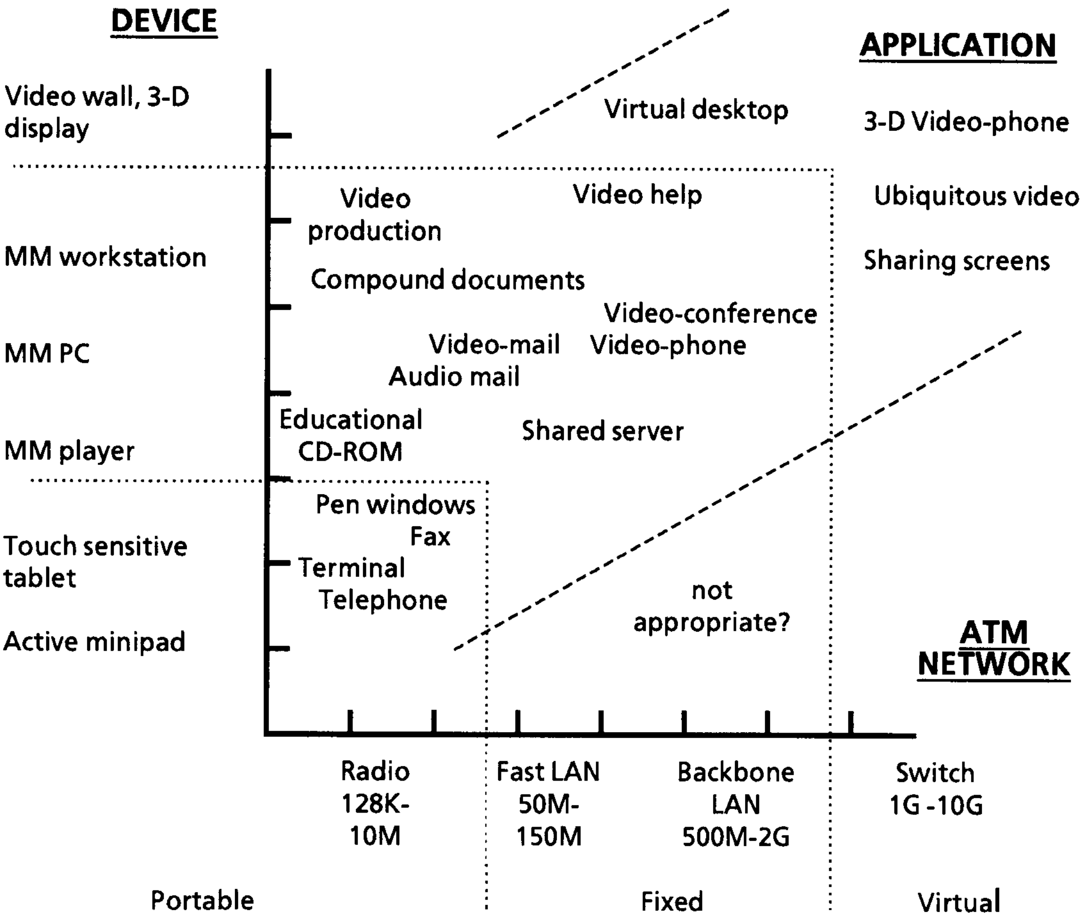

Fig.1. The multimedia space

1.2 Systems issues

Manipulation and Compression

It is important that it should be possible to manipulate video and audio by the computer system in the same way as any other type of data. Thus it should be possible to copy it, send it to file servers, and change the length, size or any other property. Such ease of manipulation will make possible new applications which use the full power of the computer system.

In play-only applications the video content is normally prepared using asymmetric compression, where the amount of processing required to compress is much greater than to decompress. An advantage of asymmetric compression is that the decompression system can be simple and easy to implement in hardware [Luther 1989]. When sourcing, as well as viewing, of video and audio material is required a symmetric compression system is used where compression and decompression take about the same amount of time. This normally means that the compression is done in hardware.

Up until now, the emphasis with compression has been to minimise the number of bits, but it may be that with time, the pressure to do this will be reduced and instead ways of improving the representation of the video will be devised to make manipulation easier [IEEE 1991*b*]. Compression may have to be idempotent and reach some asymptote, beyond which there is no information loss, even if the compression-decompression cycle is performed many times. An attractive approach would be always to transmit the original and in addition to send a list of changes to the recipient. This would mean that, depending on the display device at the receiving end, the video representation could be transformed for display in the appropriate way. If the recipient wished to manipulate and send it elsewhere, or use it in some other way, this would be possible at full resolution. It may be that, for some material, the original version could be replicated in a number of places, after which the changes from it would be the only communication required.

Networking

Conventional networks such as ethernet are able to handle the typical volume of traffic generated by a compressed video stream but their design is such that they do not give firm guarantees on delay and consequent jitter. Thus while it is possible to run a video-phone on a lightly loaded ethernet, once the ethernet becomes heavily used the user is subjected to variable performance. If the video is only being delivered one way, such as viewing television or other sources, buffering can be used to ensure smooth delivery. When the stream is turned off by the user it does not matter if this takes effect after several hundred milliseconds. If many networks of the ethernet type were cascaded, the total amount of jitter generated may require such large buffers as to make switching the stream on and off a problem.

Fixed rate transmission networks are available from telephone carriers for long and medium range transmissions and are suited to fixed rate video traffic. Some forms of ISDN allow bandwidth to be increased or decreased as required and may be well suited to bursty compressed video data.

A method of networking which is becoming popular is Asynchronous Transfer Mode (ATM) [IEEE 1991c, Greaves 1992]. In ATM, packets are made quite short (48 bytes) and each one is prefixed by a header which identifies the packet as belonging to a particular call. All packets for that call take the same path and do not get out of order when moved through the network or from one network to another. In general advantage is taken of the small packet size to design fabrics where the jitter performance is good and well controlled. Under high loads there is a specified upper bound on jitter and delay does not rise significantly. ATM networks of this type are being designed for the local area and will complement, and possibly replace, the networking we see today. Fig.1. shows various ATM network types and the speeds which are likely to become available in the local area.

The ATM approach is also being pursued by the telephone carriers, who internally use ATM to minimise delays in obtaining service on high speed links. It is also planned to make available a public ATM service for long distance traffic. Thus there may be a fortuitous coming together of the telephone and the computer industries around the ATM network approach which will make possible transmission of audio and video in the local and wide areas in the same way as any other type of computer data.

While it is frequently said that video data can be transmitted less reliably than normal data this is not the case if the video data is compressed. Compressed data normally represents changes and lost packets or other faults can cause an error to be propagated. Thus highly compressed video streams have to be transmitted very reliably. Furthermore, if a stream is being recorded, no loss can be tolerated because this would prejudice the use of the video data in some future application.

2. ARCHITECTURAL APPROACHES

2.1 Control of video by the workstation

One way of using video in a workstation is to devise a system where the video does not follow the normal datapaths but rather is part of a peripheral or some other sub-system. In the simplest approach, the display frame store has an extra input to which a camera can be attached

(Fig.2a). The analogue input from the camera is digitised and written to a part of the frame store and subsequently transferred to the monitor. The video data is generated in a way suited for the display in use, the video quality is high, and the total amount of data is large. It is normal to integrate the video with the normal windowing system and thus the video is presented to the user in a natural way. Some simple scaling or manipulation of the video stream is possible, although for bandwidth reasons, typically only one video window is available. A part of the data can be made available to the workstation if there is a path for reading the contents of the display frame store. The device used for sourcing the video can be controlled by the computer and this can give the illusion of an integrated system.

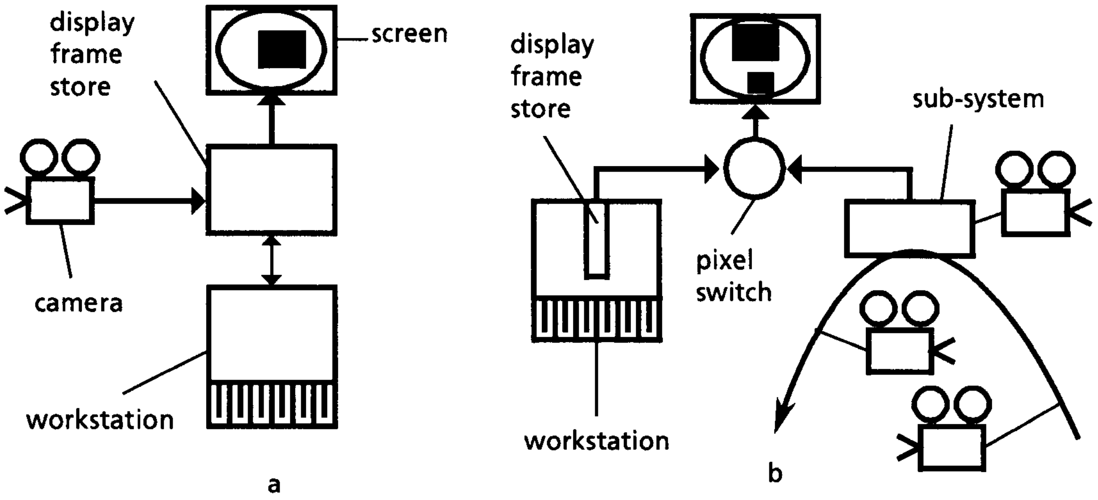

Fig.2. Control of video by workstation

Another approach shown in Fig.2b. is to use two frame stores, the first to hold the picture as generated by the workstation, the second to store the video input. A switch controlled by a pixel map is used to select whether a workstation pixel or a sub-system pixel is shown on the monitor. In general there is no restriction on how the pixel switch is used and any shape or number of video windows can be shown on the screen. Complex video sub-systems can be used, for example a disc or a network can source the video which makes its way to the monitor. The workstation is used to control the video, but does not play any part in the datapath. A difficulty is the handling of the workstation cursor because it is necessary to arrange that, when it is moved to an area which is being sourced from the sub-system, it is displayed correctly. While complex, this approach minimises the impact of the video streams, but it is difficult to make the video data available to the workstation.

2.2 Integrating video within the workstation

By handling digital video using the normal workstation datapaths, more general applications can be developed. One way is to associate expansion and compression hardware with the display frame store as shown in Fig.3a [Parallax 1992]. This means that as well as showing one high quality video stream from the local camera it is possible to make a compressed version available on the workstation bus. This can be stored or networked to another point. Compressed data can be fetched from disc, put through the expansion circuitry and shown on the screen. Normally a switch is provided to permit control of a number of analogue video inputs and outputs to the display frame store. The number of displayed video streams can be increased by reducing the size of each video stream so that the total bandwidth into the display frame store is within limits. An alternative way is to use more than one monitor and more than one expansion/frame store card.

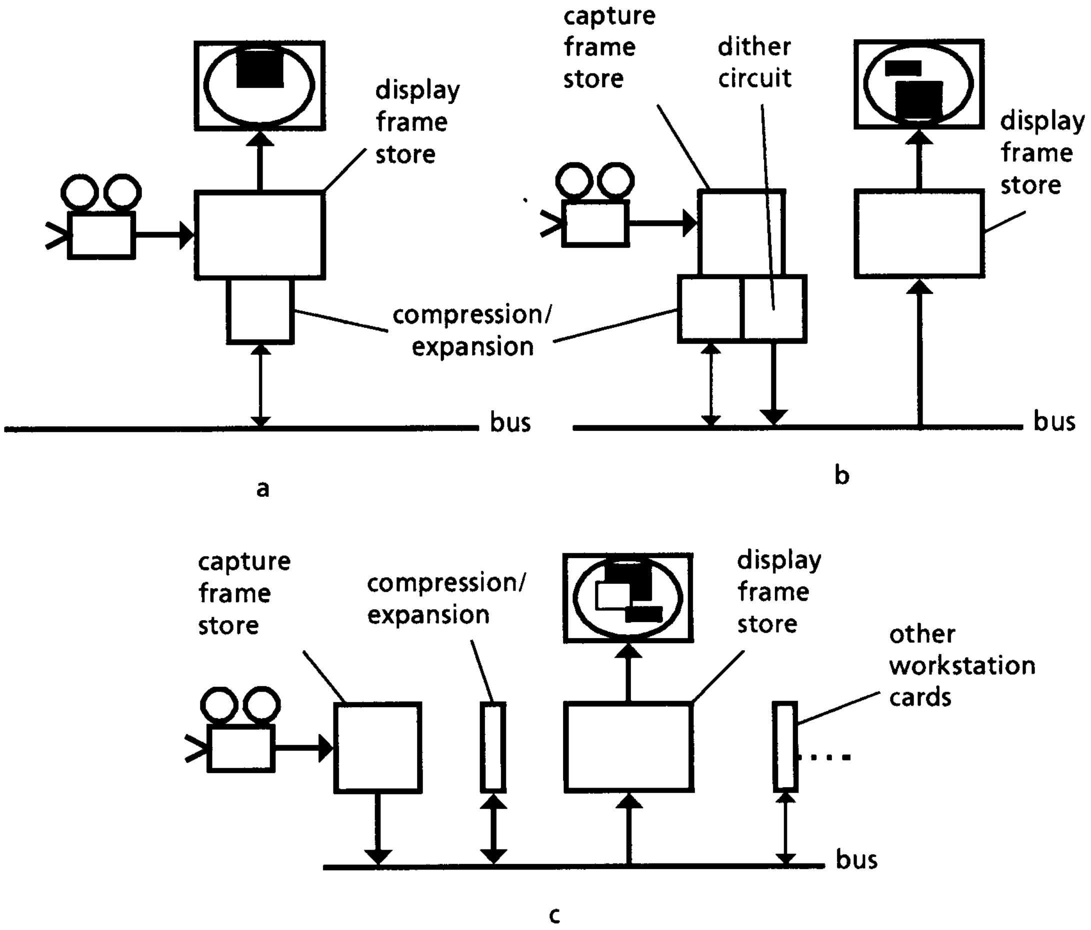

Fig.3. Integrating video within workstation

Another approach shown in Fig.3b separates the video capture system from the display frame store. The expansion/compression circuitry is moved to the capture card which can also generate a form of video suitable for direct writing to the display frame buffer. This is normally an 8 bit representation with dithering and a fixed colour map. Thus a conventional frame store can be used as video data is no longer handled in a special way. While 8 bit video with a fixed colour map produces some artifacts, it is possible that when extended to 16 bits and a larger fixed colour map this approach may be satisfactory, particularly for PCs. The capture card can be used for compression and expansion and for maximum performance this requires a bidirectional high bandwidth connection to the data bus. An 8 bit dithered video representation does not normally overload the bus so it is possible to plug in more than one capture card and send a number of full-sized video streams to the screen. For applications which require multiple streams and minimum latency, the compression system has to operate at the tile, rather than frame, level.

Fig.3c. shows a general (and conventional) approach which requires a bus with enough bandwidth to handle raw video data rates. Although such rates are high, contemporary buses are reaching the point where this approach is becoming feasible. A simple capture card can be used to produce raw video and place it on the bus. This raw video stream can be copied to the display frame store and displayed in the normal way. If bandwidth to the display frame store is a problem, then a dithered version of the video stream may be made available as well. A compression card is also placed on the bus and can absorb raw video and put back on the bus compressed or any other forms that are required. The compressed versions of the video streams would normally be used for networking or storage. While this approach is general and makes possible mixing of graphics and video data easily, it generates large amounts of data and may be expensive.

2.3 Sharing video streams

In future, workstations may become highly video-orientated so that many video connections are maintained, some of which are long term and others switched as required. To make this possible it is attractive to study the design of systems which arrange for the video data to be shared automatically by the participating parties. One way is by providing a switch with a very fast copy operation between inputs and outputs. Alternatively a multi-ported frame store can be designed which can be written and read in parallel. Described below is an approach which pushes control of sharing into hardware and also reduces the traffic on the bus. It is similar to techniques used for snoopy caches on multiprocessors.

Snoopy mechanisms for video on a bus

A design where each incoming video stream is mapped into an incoming frame store and each outgoing stream is mapped to a outgoing frame store as shown in Fig.4. All incoming frame stores are placed in a global address space and the outgoing frame stores contain associative elements similar to caches. Each incoming frame store contains words which represent some part of the incoming picture. Each word is tagged with a small number which indicates how many displays are using it at present. Each outgoing frame store uses an associative field to hold the address of the pixel which is currently being mapped in that location.

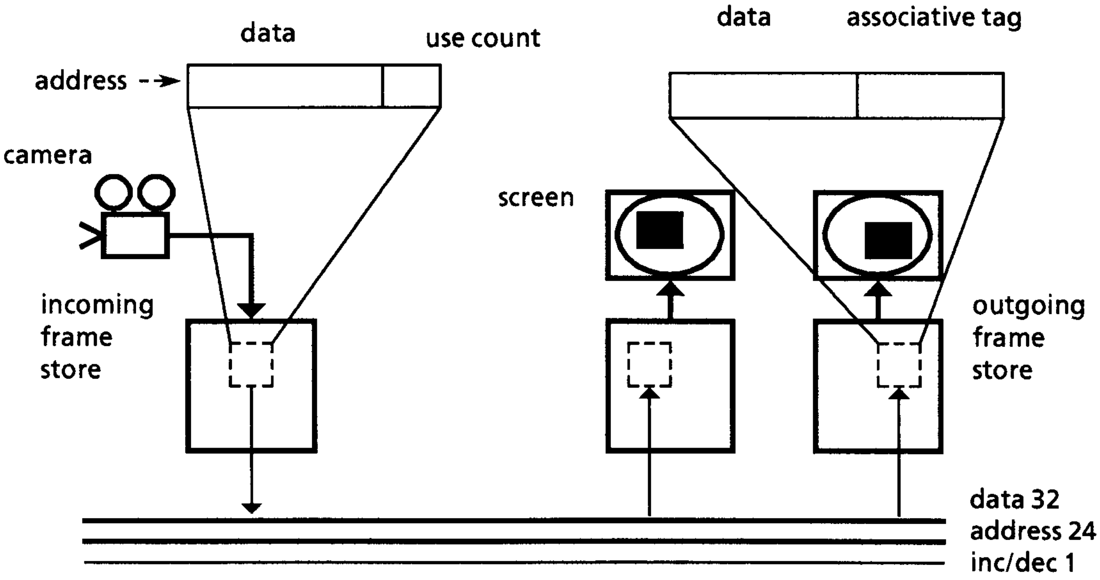

Fig.4. Sharing of video streams by snooping

When a display wishes to start receiving a stream the outgoing frame store reads from the addresses which correspond to the stream in global address space. This is observed by the incoming frame store and the use count is incremented by one. The incoming frame store starts supplying data to the bus every time there is a change to any of the relevant pixels (locations). This can be initiated by any display and continues until the stream is no longer required. Care has to be taken that updates to a location in the source frame store only take place if the value has changed. One way of reducing the use count is by using an additional increment/decrement line and performing a write.

The frame stores can be called snoopy in the sense that on the send side conditional write-through is used and on the receive side automatic copying and mapping into associative memory takes place. For small images a fully associative memory may be feasible but for larger sizes a combination of parallel association and table look up may be more appropriate. An alternative approach similar to write-back, would be for the new pixel value to be supplied across the bus only when required. In summary the hardware mechanism automatically detects what is being shared, and the data is sent in parallel to all places using it.

There are three conditions to consider when looking at the performance of this system. If the camera is sourcing a still image the information will be transmitted only once. If the camera is still but transmitting a moving image only data relating to the movement will be sent. However if the camera moves, all pixels will change and a burst will occur. This may overload the bus and care has to be taken that fair sharing is achieved between different streams. Because the display is scanning sequentially through its frame store, a small amount of time is available to smooth traffic and the transfer will eventually catch up with the burst. If delay is not a constraint, it may be appropriate to hold a number of incoming frames and optimise what is transmitted for each pixel.

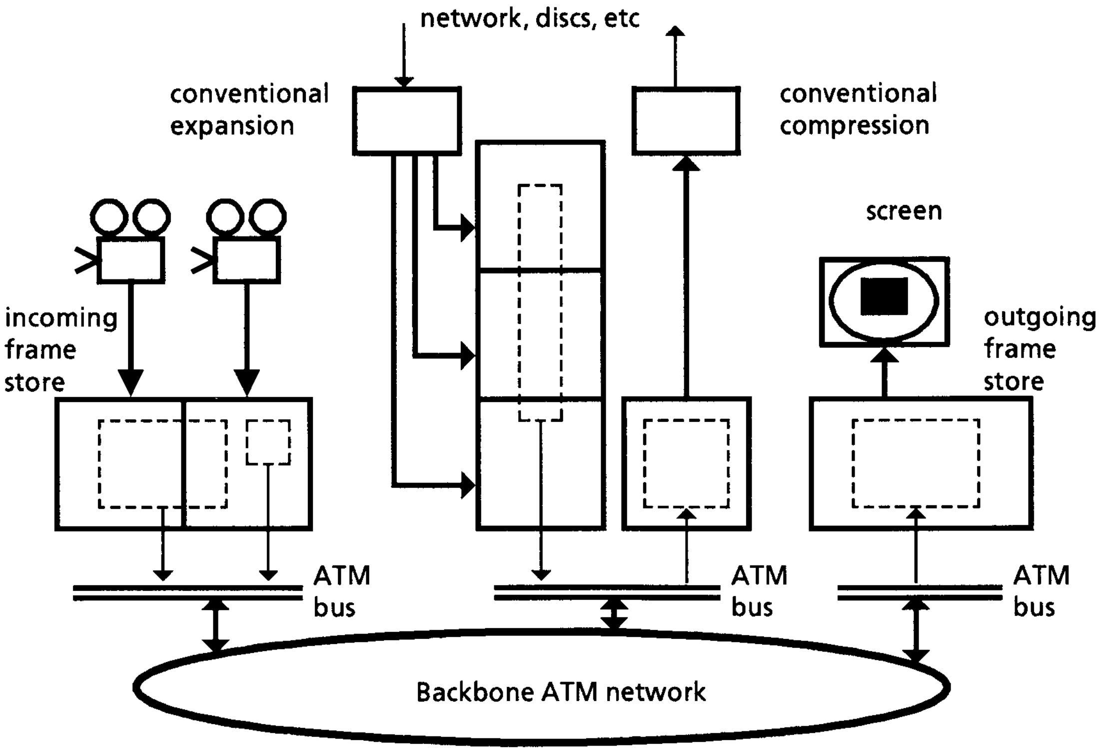

Fig.5. Automatic sharing of video streams in a system

Sharing video in a system

The design can be extended to deal with sources and sinks of video which themselves represent interleaved streams. A system with connections to sources and sinks using conventionally compressed video data and larger frame stores is shown in Fig.5. One of the incoming frame stores is shown interfaced to two cameras with the output to the bus giving a picture which is a combination of both streams. The incoming disc and network streams are shown passing through expansion hardware and are interfaced to a large incoming frame store which can hold three streams. The output to the bus is shown to be a composite across the three streams. For completeness, an outgoing stream is also shown going through a compression system and to the network or disc.

In the snoopy approach it has been assumed that the granularity of transfer across the bus is at the word level. For networked video applications it may be appropriate to use a basic unit of transfer which is the same as that on the ATM network. This would blur the distinction between using a bus and a network for video traffic. The local bus could be used to interconnect the conventional workstation components while the video streams could be sent down the network with ease. Fig.5. shows a system where, instead of a single local ATM bus, several such busses are interconnected using a backbone ATM network. Because ATM networks perform a mapping between the call number and the destination port it may be possible to take advantage of this mechanism to avoid the need for associative memories on the outgoing frame stores.

It may be that the local workstation busses can be dispensed with altogether and all components in a workstation interconnected using ATM networks only [Hayter 1991]. A design of this type would scale well but would shift the emphasis to implementation of suitable caching algorithms for memory to CPU traffic, which would now flow across the network.

3. DESKTOP MULTIMEDIA SYSTEMS

3.1 Pandora.

The Pandora system is a distributed system which makes available networked multimedia workstations [Hopper 1990] to a number of users across two sites and is described in detail elsewhere in this conference proceedings [King 1992]. The system provides a good experimental vehicle for testing applications and it is informative to comment on experience with the design based on its use and evolution over a number of years.

System Design

The Pandora system uses the sub-system approach shown in Fig.2b. and is optimised for handling multiple streams of medium quality. The size of the video streams and frame rates are easily changed so that the total bandwidth required by applications using many streams can be controlled.

The hardware, called Pandora's Box, consists of five cards used for capture, display (mixer), switching (server), audio, and networking. Each card is controlled by a transputer. The main datapaths are implemented by memory mapping the capture, display, network and audio cards in the address space of the switching card. FIFOs are used in these mappings to smooth traffic flows. Serial links between transputers provide an out of band communications path used primarily for signalling and fault reporting. The transputers perform the tasks of controlling special hardware and moving data by block copy from one part of the system to another. The majority of transputer cycles are spent moving data, even though the utilisation of the datapaths is low. This is because video and audio streams consist of a large number of small pieces of data which have to be marshalled through components such as the compression and expansion hardware. The system is relatively well balanced, with the display and network transputers being most heavily loaded, and is designed to operate at maximum efficiency when approaching overload conditions.

The compression system used in Pandora implements fixed rate streams. This has been advantageous, because when many streams are being used, it is easier to monitor and understand what is occurring. However this has meant that datapaths have been optimised to particular speeds and are not easy to change.

Audio is given priority throughout the system by throwing away video packets in preference to audio packets if congestion arises. This has worked well in practice as the audio rarely breaks up, while the users have been surprisingly tolerant of occasional disruptions in the video. The card controlling the audio has to be able to mix several audio streams for video conference applications as well as perform some echo cancellation and feedback removal for hands free operation. This has proved crucial in making the system acceptable to users.

While the use of transputers and many frame stores has allowed the Pandora system to be powerful and flexible, the amount of code which is specific to the system and resides within Pandora's Box is high. The code handles devices and frame stores, and is used to control Pandora's Box. Thus when some change has to be made which requires a different style of use or some new feature, the code has to be recompiled and downloaded again. In a new system it would be attractive to minimise the amount of embedded code and provide an interface which allows for the separation of the devices which are actually used for sourcing and displaying the video data from the way the control system operates. In common with other pixel-switch based systems it is difficult for the data streams to be made available to other computers.

Datapaths

Frames are transferred within Pandora's Box as segments up to 4Kbytes in length. This was chosen to allow interleaving of different streams and corresponds to about 1/4 frame for the highest quality images. This has turned out to be advantageous because for video streams there is no branching implemented, and if a stream is required in a number of places then it is replicated at the source. The main reason for doing this is because users frequently change the size of a video window and thus the characteristics of a stream. If downstream branching of video streams were allowed, then a complex recovery procedure would have to be implemented where the new characteristics of the stream were propagated to the source. With audio, the opposite view can be taken because it exists in one form only and audio streams are branched downstream at the network interface for applications such as video-conferencing.

Within Pandora's Box the segment size is greater than the FIFO size used for intermediate buffering. This has meant that a signalling system for use of the FIFOs had to be implemented. A more efficient design would use larger FIFOs or, if further implementations using transputers were required, higher speed links for both signalling and data traffic.

Networking

The ATM network used in the Pandora system makes available about 7 Mbps to a single user, and 50 Mbps to the whole system. The small (32 byte) packet size means fine sharing and little contribution by the network to the total jitter accumulation in the system. A typical segment transfer within a Pandora's Box maps well onto the blocks implemented at the network level. A maximum size block corresponds to about 150 network packets. However handling small units in the network transputer means that this is a bottleneck. Close to the maximum throughput it is difficult to arrange for the audio to continue without disruption, because of excessive buildup of queue lengths in the network interface.

Synchronisation

There are three levels at which synchronisation is maintained in the Pandora system. These are maintenance of the correct relationship of segments to a particular frame, the relationship between different frames (or voice samples) in a single stream, and the relationship between related audio and video streams.

Different segments in a stream are marked with sequence numbers so it is possible to identify which segments belong to a frame. However a difficulty arises when several streams arrive at the display frame store and need to be shown correctly, because they are not synchronised and may have different frame rates. A double buffer arrangement with new arrivals assembled before being displayed could not be used, because there is no time at which it can be guaranteed that all streams will have complete frames available. Synchronisation is done with respect to the display scan line and complete frames are copied from shadow to display memory just after the previous data has been scanned out. In practice this requires careful tuning of the code and if not implemented correctly can lead to loss of quality.

For inter-frame synchronisation, the approach taken throughout the Pandora system (both within a Pandora's Box and on the network) is to provide fine sharing of resources and thus avoid the need for explicit synchronisation. Different streams are launched into the system and it is ensured that the delays introduced by all components enroute to the receiver have a well controlled upper bound on delay. This avoids the need for more sophisticated source synchronisation mechanisms, and while fine sharing does mean some parts of the system are highly loaded, it has been one of the most important ways of simplifying the design of the Pandora system.

The only place in the system where time stamps are used, is when recording pairs of audio and video streams to the file server. Time is available within a Pandora's Box and time stamps are placed in segments by the originating transputers. Because these cycle through a large number space, it is necessary to fix a relationship between the audio and video time stamps. This is done by the video file server just as the recording starts. When the video file server replays the pair of streams the data is launched into the network with the correct time relationship and at the correct rate.

Applications

The Pandora system is deployed at the Olivetti Research Laboratory and at the University of Cambridge Computer Laboratory [Hopper 1992]. Fig.6. shows the different components that make up the full system which is available to a group of about 50 users. Gradually some of the applications which previously required the full Pandora sub-system are shifting to standard platforms such as Multimedia PCs and Unix workstations with synchronisation extensions to their X system [Glauert 1992].

In the Pandora system, the camera is placed slightly to one side of the monitor because it was considered that a one way mirror arrangement would be too bulky for normal desktop use. In practice this has not been a problem and users learn to glance towards the camera from time to time. Microphones are potential bugs and some users have required that a mechanical microphone switch be made available, which they switch off in circumstances where the audio system is not in actual use. In practice, users feel there is no need for a similar arrangement with the video camera and get used to people looking in occasionally. The camera is a relatively obvious object and the user is made aware of it by a red "on air" light.

It is difficult to balance the audio throughout the system as different users record at various volume settings and the microphone and speaker controls are continuously changed. The different types of room in which the systems are deployed mean that a control system would need to be very sophisticated. Hands-free operation is popular and means the round trip delay has to be minimised to make any echo less obtrusive.

Normal techniques of iconification are used, both to reduce the amount of screen space taken up with video, and to minimise traffic by reducing the frame rate to zero for iconified streams. This is also done when one video window is placed on top of another in a way that completely obscures the lower one. An alternative scheme would permit the second window to be seen through the first one in a translucent manner so that the user was made aware of the contents of both streams. It is not clear whether this would be a useful feature and what the priorities of many such overlaid video windows should be. Similarly it is not clear whether audio should be stopped when the stream is iconified. If the reason for iconification is momentarily to reduce the amount of video displayed on the screen, then the associated audio stream should be kept going. If, however, the iconification is a way of parking the application for quick subsequent use, then both the audio and video streams should have their rates reduced to zero. Perhaps the optimum would be for the user to have control and be able to iconify video only or both video

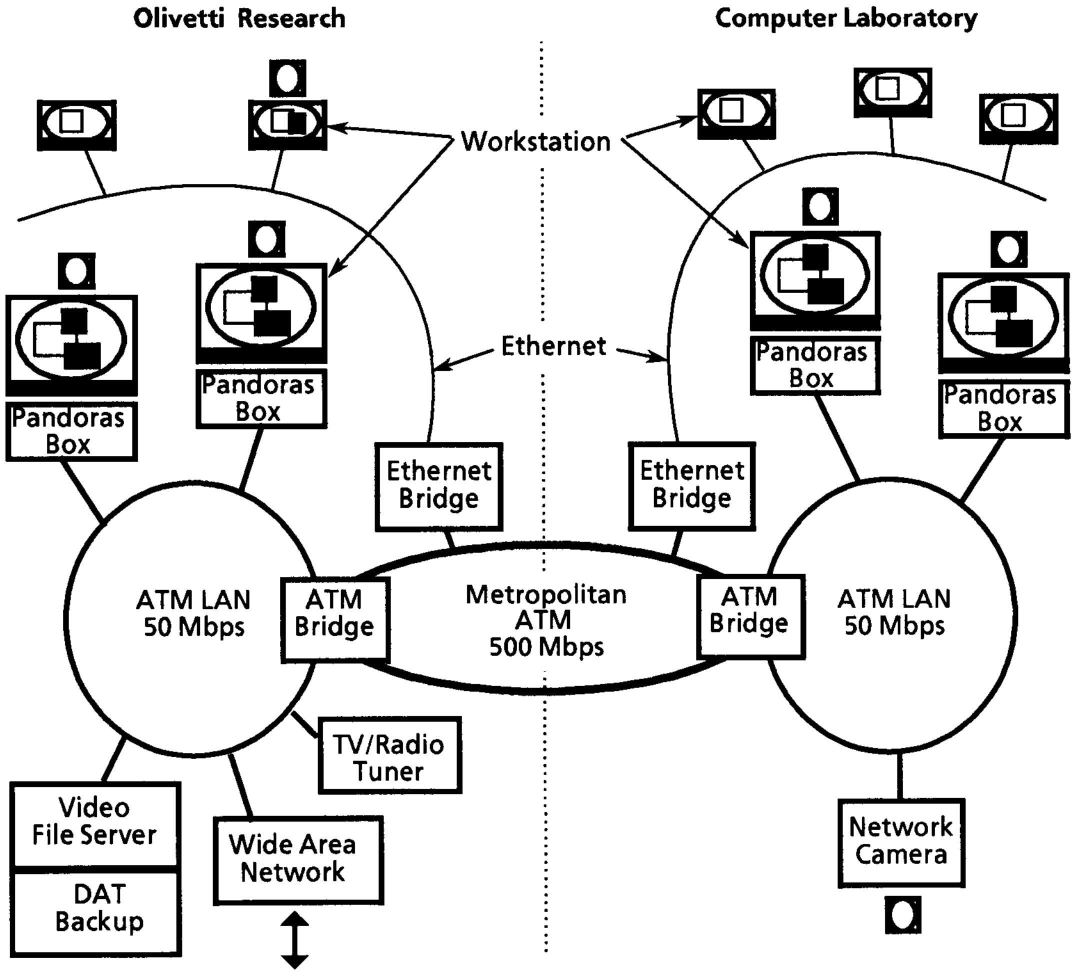

Fig.6. Pandora multimedia system

and audio, with a visual indication of which iconified streams are currently receiving changing data.

The ability to use many streams in Pandora is widely used. For example a video conference between four parties requires twenty four network streams as well as four additional streams to show the local image. If several of these are taking place across the network at once then it is conceivable that over a hundred streams are in existence. Another application which generates many streams is the ability to look at all cameras in the system at once. At present this generates twenty or thirty small video windows on the screen.

The most popular applications in the Pandora system are video-phone and video-mail. Video-phone calls typically last two to three minutes and are normally between two parties only. This may be because the size of the group is not large enough to support three-way or four-way video-phones frequently. The style of the video-phone is relaxed and more of a chat than a formal transaction. Video-mail is typically thirty seconds long and has a substantive part which if written would require careful drafting. Video-mail seems to be popular because it equalises the time to generate and time to read, it can be viewed by the recipient at leisure, and it carries over much of the body language which makes communication easier. Other applications include live services such as television and radio, and stored services such information from the video file server. Surprisingly a system which permits mixing of text and video has not proved popular. This may be because the user makes a choice of either video-mail or text-mail based on what is to be communicated, and is not prepared to put in the effort of making a more sophisticated compound document.

3.2 Medusa System

The Medusa project follows on from the Pandora project and is designed to continue the use of digital audio and video in a networked workstation environment. It is designed to blur further the difference between local and networked applications by making many more audio and video devices available to the user both remotely and locally.

The Pandora system made available many streams to the application but did not emphasise their synchronisation. Rather it took advantage of the underlying parameters of the hardware and software to guarantee performance within bounds. In the Medusa system the same approach is being taken, except that provision is being made for many more streams of different speeds to be available to the user. This will mean extending the use of time stamps and the mechanisms for launching streams at the correct times into the system.

It is envisaged that the number of cameras per workstation will increase from one to eight or sixteen. Each camera will produce a video stream which will be sent to the receiver. The recipient will be presented with all streams and can make a choice which to watch. The cameras may represent different views of the scene or may be used to cover a wider area. The transmitter can send a navigation path for the way the streams should be used, but the ultimate decision is left to the receiver. The receiver can be a piece of hardware, software or an actual user. So, for example, in a video-phone all eight streams will be made available to the receiver, who will choose the one to watch.

Another way of using many streams at once is in making lectures, or other meetings, available to the system. The lecture theatre will be equipped with many (64) cameras all trained in different ways. All users will have streams from all cameras available simultaneously and can choose the ones to watch. Thus the remote audience will be able to participate in the same way as the actual audience by looking around the lecture theatre, or focusing in on the speaker, all at the same time.

Allowing each source to provide many streams, together with the receiver choosing what to watch, increases the total number of bits on the network considerably. However, because networks of large capacities are becoming available, this no longer seems an unusual approach.

Components

There will be two types of components making up the Medusa system: those that attach to the network directly and those that attach to a workstation cluster. Within the workstation cluster raw uncompressed video data will be used to make possible applications which mix graphics, video and allow for general manipulation. For devices which attach directly to the network, JPEG compression will be used so that data volume is reduced, while manipulation at the frame level is still possible.

The components in a workstation cluster use the architecture shown in Fig.3c. The first component is a capture card which is able to digitise and place raw video on the bus. It performs no compression but is able to fulfil various scaling and rotation functions. The second component is a compression card which reads raw video, and can place compressed streams back on the bus. The compression card is able to handle many video streams at once and, by operating at the tile level, is able to switch between them rapidly thus introducing minimum delay in processing each stream. The compressed streams can be networked by being sent to a conventional ATM network card. In order to make it possible for many video streams to coexist, a high speed ATM-oriented interconnect is being designed which makes possible transfer of units from the network interface card to the final destination with minimal manipulation. Audio is handled by using conventional cards plugged into the workstation.

Network devices consist of a camera cluster which can aggregate eight cameras onto a single network connection, and perform compression if required. Similarly microphone and loudspeaker clusters are being designed which attach directly to the network so that array and phase techniques for handling audio and video can be investigated. Video and audio file servers will be able to handle many streams in parallel. This involves the use of arrays of discs (RAID) for bulk storage as well as high performance ATM network interfaces designed to handle synchronised parallel streams to many destinations at once.

A framework for a complete Medusa system is shown in Fig.7. and will consist of a number of clusters attached to the network. Workstation clusters will handle raw data internally for ease of inter-working with graphics and other conventional workstation systems. Networked clusters will handle compressed streams and provide many ways of sourcing video and audio data.

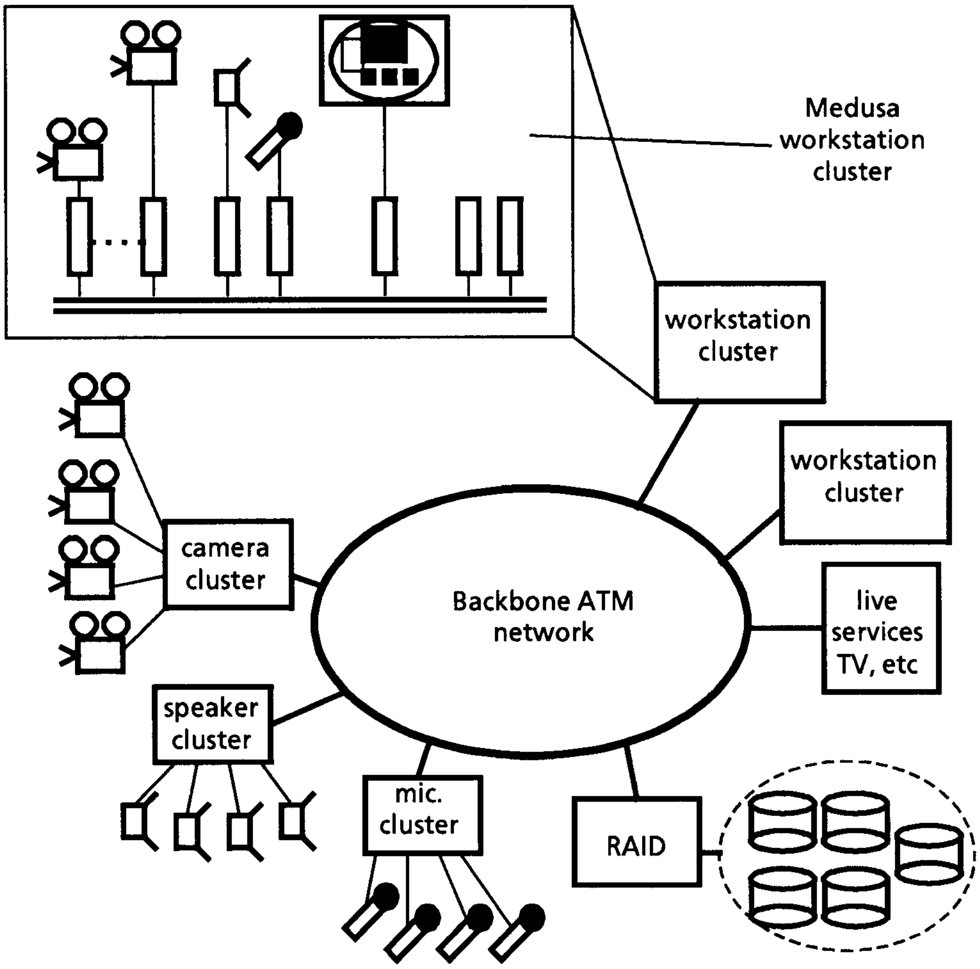

Fig.7. Framework for Medusa multimedia system

Applications

The Medusa architecture enables many different types of applications to be considered. The use of multiple streams allows the design of a video space within which the user can consider himself as always in the field of view of a camera. During a video-phone the user can move around the space and the recipient can automatically follow or use the video and audio data as required. Not all streams in a Medusa application need to be of the same type. For example infra-red cameras or devices which give depth information can be incorporated in a coherent and synchronised way if applications so require. Slower speed streams can also be sent; data from the active badge system which gives information about the positions and movement of individuals and equipment can be another stream type [Want 1992]. If sourcing of the video streams is arranged appropriately it may be possible to use the Medusa system to drive a 3-D display. In this case, rather than providing ubiquity of video and audio, the bandwidth will be used to supply 8 parallel synchronised streams, each of which forms one particular view for the 3-D display [Travis 1991]. Finally because the total capacities available in the system are high it may be possible to aggregate packets back into a single stream, and use it for HDTV or other very high quality video transmissions.

4. ACKNOWLEDGEMENT

My thanks go to the many colleagues at both Olivetti Research and the University of Cambridge Computer Laboratory, whose joint efforts have made much of the work described in this paper possible.

5. REFERENCES

[Glauert 1992] Glauert,T.H. X Synchronisation Extension Protocol Version 2.0. *MIT X Consortium Public Review Doc.*, 1992, (To be published).

[Greaves 1992] Greaves,D.J.,McAuley,D. Private ATM Networks. *Proc. 3rd Int. IFIP WG6.1/6.4 Workshop on Protocols for High-Speed Networks,* Stockholm, May 1992.

[Hayter 1991] Hayter,M.,McAuley,D. The Desk Area Network. *ACM Oper. Sys. Rev.,* Oct.1991.

[Hopper 1990] Hopper,A. Pandora - An Experimental System for Multi-Media Applications. *ACM Oper. Sys. Rev.,* Apr.1990.

[Hopper 1992] Hopper,A. Improving Communications at the Desktop. *Rep. of Royal Soc. Discussion Meeting on Communications After 2000 AD,* Chapman and Hall, 1992.

[IEEE 1990] Multi-Media Communications. *IEEE J. on Sel. Areas in Comm.,* Apr.1990.

[IEEE 1991*a*] The Promise of the Next Decade. *Computer, IEEE Comp. Soc.,* Sep.1991.

[IEEE 1991*b*] Digital Video. *IEEE Spectrum*, Oct.1991.

[IEEE 1991c] Architectures and Protocols for Integrated Broadband Switching. *IEEE J. on Sel. Areas in Comm.,* Dec.1991.

[King 1992] King,T. Pandora: An Experiment in Distributed Multimedia, *Proc. Eurographics'92,* Cambridge, Sep. 1992.

[Luther 1989] Luther,A.C. Digital Video in the PC Environment, Intertext Publications, McGraw Hill, 1989.

[Parallax 1992] XVideo, Parallax Graphics Inc., 1992.

[Travis 1991] Travis,A.R.L.,Lang,S.R. The Design and Evaluation of a CRT-based Autostereoscopic 3-D Display, *Proc. Soc. for Inf. Disp.,* Oct.1991.

[Want 1992] Want,R.,Hopper,A.,Falcao,V.,Gibbons,J. The Active Badge Location System. *ACM Trans. on Inf. Sys.,* Jan.1992.