**Towards Network-Based Distributed Digital Video Editing**

**Algirdas Pakštas**

School of Computing, Engineering and Technology, University of Sunderland, Sunderland, England

Institute of Mathematics and Informatics, Vilnius, Lithuania

a.pakstas@ieee.org

**Abstract**

For the compilation of the edited video stream, normally at least two others are manipulated (A/B roll). With conventional “linear editing”, video/audio sequences are stored on the magnetic media with compilation via replication in real-time. This causes some limitations in flexibility, since final compilation must be in the correct chronological sequence and is constrained by the serial linear tape access time. “Non-linear editing” systems require all clip/sequence data to be stored on the random access medium, which affords play-out and manipulation/compilation abilities.

Digital formats afford the opportunity for non-linear editing but certain constraints must be considered. Volume of data associated with audio and video streams is very large and requires high capacity storage. While adequate storage is often available, the communications between the workstations and access to the storage becomes a bottleneck.

Paper is devoted to investigation of the existing experience of building Distributed Digital Video Editing Systems and use in them widely available computing platforms such as PCs in order to apply this for the project DIVIDEND (Digital VIDeo Editing based on the Network testbed). Project DIVIDEND addresses the demand of the media industry and video producers in particular, for non-linear editing facilities which are currently unavailable. It intends to build an experimental testbed based on high speed network technology called Dynamic synchronous Transfer Mode (DTM). System architecture and specifications characteristics are described.

**Keywords**: Digital video editing, distributed system, Dynamic synchronous Transfer Mode (DTM) network.

1. **Introduction: From Linear to Non-Linear Digital Video Editing**

   For the compilation of the edited video stream, normally at least two others are manipulated (so called A/B roll) [1]. With conventional “linear editing”, video/audio sequences are stored on the magnetic media. Access to the video/audio sequences (clips) involves physically moving of the media across the magnetic head for reading/writing. Thus, compilation of the edited video stream is achieved via replication from the sources in real time.

   The University of Sunderland currently uses two linear editing systems, namely “Video-Machine” of FAST Multimedia AG, Germany [2], configured as A/B roll “hybrid” systems. This causes some limitations in flexibility, since final compilation must be done in the correct chronological sequence and is constrained by the serial linear tape access time.

   The other approach is called “non-linear” editing. The University of Sunderland currently uses two non-linear editing systems ES-7 of SONY [3]. This technology requires all clip/sequence data to be stored on some medium which allows random access, preferably in real-time, play-out ability, as well as manipulation and compilation the programme sequences from the stored clips according to the editor’s needs [4]. The arrival of digital video equipment in the broadcast industry brought high expectations of increased efficiency and convenience in addition to improved picture quality [5]. Many believed that video and audio material would be treated similarly to computer files for managing, processing, and transferring via computer networks and servers.

   Recent moves towards digitising of the video production as well as start of the digital TV broadcasting (in UK since 1999) shows that the future is going to be digital. However, while digital formats afford the opportunity for non-linear editing, certain constraints must be considered.
This paper aims to investigate existing experience of building Distributed Digital Video Editing Systems and use in them widely available computing platforms such as PCs. The rest of the paper is organized as following. Section 2 is discussing typical features of the video stream and related problems of its storing and retrieving in digital format. Review of the related works aiming use of the inexpensive PC platforms as an editing stations or storages as well as approaches to build video distribution architectures and video delivery protocols are presented in the Section 3. Section 4 introduces DTM (Dynamic synchronous Transfer Mode) network technology and benefits of its use for carrying video traffic. Project DIVIDEND and more details about DTM network technology are presented in Section 5. Section 6 presents conclusions.

2. Video Stream and Problems Related to Digital Video Editing

The analogy to computer data systems, as used for example in office automation, is too simplistic and cannot be applied directly to the demanding video and audio system architecture of a broadcast facility. Stringent requirements must be placed on broadcast A/V systems to deliver high volumes of data, in continuous streams, to a multitude of users. Real-time throughput and predictable delays are all essential for editing and on-air applications [5].

Television program compilation, storage and transmission in the digital domain demands high speed data handling. One of the factors that makes implementation of such facilities especially difficult is the volume of data associated with audio and video streams. For example, the data rate required for 16-bit colour video with 720x570 resolution at 25 frames per second is 27 MBytes per second (as in CCIR601 4:2:2 standard - now ITU-R601 [6]) per stream.

Further, with the volume of data comes the need of high capacity storage [7]. For example, storing of 5 seconds (i.e. size of the typical video clip) of the above mentioned video stream will need 131.8 MB and storing of 5 minutes (= 300 seconds) will need approximately 6GB of memory. For uncompressed studio quality video or High Definition TV (HDTV) this would require correspondingly even greater storage.

For editing workstations where audio/video streams are compiled, access and rendering time is often crucial in real-time operations (such as re-play of situations during the broadcasting of live sports). Both real-time and post-production (non-real-time) environments may have additional need for shared used of materials. Demands for video storage can include re-usable components such as production house logos, credits and graphics. It is beneficial to store these on the shared access storage rather than retain many copies directly on the workstations.

3. Related Works

3.1 Use of the PC Platforms for Digital Video Editing

Interest to the development of the digital video editing system using relatively powerful and inexpensive hardware and software available on personal computers and workstations was expressed relatively early [8]. At that time the video editing application category has been dominated by visual, easy-to-use, direct manipulation interfaces. These systems were able to bring high-bandwidth human-computer interaction to a task formerly characterized by slow, inflexible, indirectly-operated machines. However, the direct manipulation computer interfaces are limited by their manual nature, and cannot easily accommodate algorithmically-defined operations. The paper proposed a melding of the common direct manipulation interfaces with a programming language enhanced to manipulate digital audio and video. The result was a system which can automate routine tasks as well as perform tasks based on sophisticated media recognition algorithms.

Video server system for continuous media (CM) data is described in [9]. Server consists of a low-priced personal computer (PC) and a disk array. The bit rate of the data stream varies
depending on which applications are being used, which include video on demand (VOD) and digital video editing systems. Moreover, due to large data volumes, arranging and controlling data with different bit rates on the disk array presents a number of problems. A new file system which utilizes two-layer hierarchical media management rate has been proposed. It is using the concept of packages based on the management of contents and disk space, and CM-records based on the management of bit-rate and disk access. The size of the packages and the sizes of CM-records are flexible so that a range of bit-rate streams are supported. The site of disk access is decided by taking into account access efficiency, which in turn depends on the continuous media data bit rate. Using these methods it is possible to reduce dramatically the memory space for administration, even for large volumes of data. The file system and the experimental video server have been implemented on a standard PC platform. This prototype system achieved 230.8 Mbps of I/O capability for continuous media.

The necessary high-speed transfer rates and fast access dictates the use of storage devices controlled by the SCSI or even faster interface. Fibre Channel (FC) is usually thought of as a system-to-system or system-to-subsystem interconnection standard. FC uses optical cable in a point-to-point or switch configuration [10, 11, 12]. One of the goals in the development of the FC interface was to improve or eliminate SCSI shortcomings, particularly in the areas of connectivity, performance, and physical robustness. FC is a standard which is especially actively promoted for video storage applications. In theory, FC operates at up to 2 Gbps, however, most installations don't actually achieve this speed [13].

Thus, adequate storage is often available but communications between the different types of the workstations and access time becomes a bottleneck [14, 15, 16]. Specialized RAID systems [17] such as EIDE-based high-speed, low-cost disk arrays VideoRaid pci of Medea Corporation [18] are capable of sustainable 27-40MB/sec datarates.

One of the approaches to speed-up communications with the storage is to design the parallel video server architecture [19]. In order to do so three central architectural issues should be addressed: server striping policies, video distribution architectures, and video delivery protocols.

### 3.2 Video Distribution Architectures and Video Delivery Protocols

Video distribution in a local area network (LAN) can impede or even paralyze normal data transmission activities. The problem can be solved, at least for a while, by compression and by increasing the bandwidth, but that solution can become excessively costly or otherwise impractical. Moreover, experience indicates that usage quickly expands to test the limits of bandwidth.

It is a current practice to view Asynchronous Transfer Mode (ATM) networks as the only choice for high-speed networking [14, 20]. ATM is a link layer protocol which is related to ISDN and broadband ISDN. ATM uses T1 services or SDH/SONET as a physical layer. SDH/SONET, however, lacks switching commands and a signaling protocol. ATM is connection-oriented packet switching technology.

Project [21] is an example of the approach of using ATM broadband networks (LAN and WAN) in building DDVES. This project was a part of the European ACTS project and aimed to build a central to the distributed video production. Various distributed pilot applications for professional digital video production has been developed and described in the paper.

Article [22] reports about experiments on defining relationship between MPEG video quality and ATM network performance. Recommendations are made on acceptable ranges for cell errors, cell loss, and cell delay variation parameters. Advanced MPEG techniques such as scalability and error concealment are discussed, including their application on ATM networks.

Thus, here has been demonstrated that even with the use of the broadband ATM there is no any talk about loss-less error-free transmission or 100% Quality of Service. Even fast ATM network is not enough for achieving of acceptable quality and some additional efforts are required
anyway. Some other approaches are focusing on the alternative solutions rather than simply applying of the broadband networks.

A telecommunication system for the transmission of video and audio signals over optical fibre is investigated in [23]. A new multiplexing scheme for digitised uncompressed video and audio signals is proposed and is analysed at system level. The special frame developed to support the system allows expandability for the multichannel transmission of audio-video signals multiplexed with other types of data.

In the paper [24] is introduced and analyzed the architecture of a hybrid analog/digital video network (ADViNet) which separates video distribution from standard data-handling functions. The network preserves the features of a standard digital network and, in addition, provides efficient real-time full-screen video transmission through a separate analog communication medium. System employs a specially developed control and management protocol. For all practical purposes, ADViNet may be used when graphical images have to be distributed among many nodes of a LAN. It relieves the burden of video distribution and allows users to combine efficient video data transmission with normal, regular network activities.

While many approaches for implementing of the digital broadcast environments are available, the paper [5] discusses the concept and implementation of a novel peer-to-peer networking system that addresses the specific requirements of the news broadcast environment. The paper emphasizes the merit of a simple solution to manage A/V material between MPEG-2 4:2:2 MP@ML Betacam SX workstations.

Thus, we may conclude that for the isochronous nature of the video stream, some kind of circuit switching rather than connected-oriented packet switching of ATM would be preferred for video transfer across a distributed system.

4. DTM: Alternative Technology for the DDVES Network Backbone

DTM (Dynamic synchronous Transfer Mode) network technology is based on the use of fiber optics [25, 26, 27, 28, 29, 30, 31, 32, 33]. DTM can achieve datarates of 100 Gbits/sec but this is not the most important its feature for DDVES.

4.1 Why DTM is Good for DDVES

The DTM network architecture is based on circuit switching augmented with dynamic and priority reallocation of available bandwidth. The protocol is designed to be used in integrated services networks and has support for point-to-point, multicast and broadcast communication. A DTM network could be used for both distribution and unicast communication. Non-the-less, DTM maintains many of the advantages of circuit-switched networks, namely guaranteed capacity, flow isolation and simple and deterministic Quality of Service (QoS) differentiation.

The isochronous service provided by DTM also ensures a good support of both existing PDH/SDH structures as well as increasing data traffic. Benefits from using a DTM backbone as a DDVES network backbone are:

- DTM enables 270 Mbps ITU-R 601 channels to be switched on demand.
- A high capacity DTM network enables to edit/create films remotely and hence reduce production costs.

Thus, a DTM network also gives a new perspective on feature film production, leading to new services and markets.

4.2 DTM Applications and IP-over-DTM

A DTM network can be used directly for real-time application-to-application communication or can be used as a carrier network for higher layer protocols, such as ATM or IP. The IP switch
using DTM's switching capabilities handles data, voice and video smoothly and solves many of the problems associated with IP traffic in today's networks, e.g., provides bandwidth on demand, allows for preferential resource reservation, and supports multicasting.

4.3 DTM vs. SDH/SONET: Similarity and Interoperability

DTM has the same type of framing structure as existing transport networks, i.e. SDH/SONET, using 125 µs frames that are divided into time slots. Due to the similarity to SDH/SONET data frames, it can interoperate with existing SDH/SONET infrastructures, conversely DTM networks can carry SDH/SONET streams. It belongs to the same class of circuit switching technologies as SDH/SONET but is dynamic not static. DTM, however, operates at OSI layers one to three and thus, in contrast to i.e. SDH/SONET, includes switching commands and a signalling protocol.

5. DIVIDEND: Digital VIDeo Editing over Network testbeD

From the above it could be concluded that currently there are no satisfactory (or for many organizations affordable) solutions for non-linear editing based on the existing distributed networking technologies. There however could be formulated certain principles for building system architecture with defined real-time operating parameters. Project DIVIDEND (Digital VIDeo Editing over Network testbeD) is intended to address some problems of non-linear editing by building an experimental testbed based on the DTM network technology.

It is assumed that the compilation of the video data stream is done on non-linear editing workstations with access to the Video Storage via the Network Backbone. System Architecture is shown in Fig.1. Compilation of the video data stream is done at the workstations (WS1,... WSn) comprising the non-linear editing facilities with shared access to the Video Storage via Network Backbone.

![Figure 1. DIVIDEND System Architecture](image)

Workstation specifications include:
- Limited local disk storage (sufficient for at least 5 minutes of video stream) with high-speed transfer/access rates.
- Local operating memory (to allow manipulation of data equal to at least 5 seconds of the video stream).
- High speed CPU.
- The video editing software package.

From this specification and data rate for CCIR601 4:2:2 video stream [6], it seems reasonable that currently available PCs with 256 MB RAM and 20 GB of hard disk storage would be sufficient for manipulating and editing (including manipulating in RAM and storing on the disk) the video data required. However, even some “normal” IDE disk servers with 9.5msec access time and 10GB capacity could be acceptable in low-end configurations, where real-time access is not critical, as a building component for larger storage.

Typical distributed video editing activity is shown in Fig.2 with Editing Unit, Film Bank and Recording Studio being connected to the switches of the DTM network. This configuration uses the 270 Mbps Video Module (referred as 601) which is an Interface Card for the DTM-switch the
Nimbra One. Based on the BT.601 video standard (derivative of CCIR601), the 270 Mbps Video Module is targeted at the video/TV production and distribution market. A configuration consisting of the Nimbra One and the Video Module supports full uncompressed streaming of studio quality video. Due to its modularity, the switch can be set to function as a pure video switch or as an integrated switch for video, voice, and data. It is capable of switching and transporting multiple streams of video (BT.601/(656), 270 Mbps), data (Ethernet 10/100 Mbps) and voice (E1/T1) simultaneously on the same physical fiber and with 100% guaranteed Quality-of-Service (QoS).

The modules have two Serial Digital Interfaces for incoming traffic, and two for outgoing traffic. Fully configured for video, the Nimbra One switch supports 12 interfaces for incoming video, and 12 for outgoing video.

As a true multi-service platform, the Nimbra One DTM Switch transports IP traffic for purposes such as controlling video equipment, carrying MPEG streams for compressed video, and telephony traffic for inter PBX communication with guaranteed QoS alongside the 270 Mbps video streams.

The minimum channel size is 512 Kbps, which can be scaled in increments of 512 Kbps up to the total capacity of the fiber cable. The channel can be connected between one sender and one receiver, or between one sender and many receivers.

To set up a channel with a certain capacity, DTM allocates the required number of time slots to the channel. For example:

- 4 time slots are necessary to set up a 1.5-Mbps channel for a PSTN switch or other leased line.
- 6 time slots is enough to send 3-Mbps MPEG video with guaranteed service quality.
- 530 time slots are required for transferring studio-quality video at 270 Mbps.
- 1900 time slots are needed to set up a channel for the capacity of the Gigabit Ethernet traffic.

Different traffic types run simultaneously in the network, each in its own channel and at the desired bandwidth. In this way, DTM complies with all the demands of a well functioning multiservice network.

Since the Nimbra One switches are cascadable and switching is distributed, co-located switches can increase the number of possible interfaces significantly. This allows to have several possible video units connected at the same time and choose, via the management system, the interfaces to which the streams should be directed, either locally or remotely.

Using the Nimbra One and its modules in a WAN or MAN environment (see example in Figure 3) allows save money, thanks to centralization and more cost-effective use of expensive editing and storage equipment.

6. Conclusions

The high speed communications capability of the DTM system, makes the concept of high resolution multi-channel uncompressed video transfer between nodes in real-time not only possible, but feasible. Non-linear editing and graphics for instance would be afforded a much greater degree of flexibility when data is stored and manipulated between available server capacity anywhere on a network in real-time.

Project DIVIDEND addresses the demand of the media industry and video producers in particular, for non-linear editing facilities which are currently unavailable. Additionally, the School of Computing, Engineering and Technology at the University of Sunderland plans to use this testbed for the following goals:

- As an educational facility;
- For production of the educational video projects.
7. References


