



AudioSoft

White Paper

Network and Mass Storage for Evidential Requirements

A three point plan for storing your recordings efficiently

Revision 3.0
September 2008

Author:
Robert Wright (robert.wright@ultra-audiosoft.com), BSc (Hons)

© **Ultra Electronics AudioSoft, October 2008**. Whilst Ultra Electronics AudioSoft makes every attempt to ensure the reliability of the information contained in this document, we cannot assume responsibility for its accuracy. We duly acknowledge all trademarks and names mentioned. Web: www.ultra-audiosoft.com Tel: +44 (0) 1285 883800

Contents

| | |
|---|----|
| 1.0 Executive Summary | 3 |
| 2.0 Step 1: Understand the Requirements on your Organisation..... | 3 |
| 2.1 Courts | 4 |
| 2.2 Air Traffic Control | 4 |
| 2.3 Emergency Services | 4 |
| 3.0 Step 2: Understand the Options available..... | 6 |
| 3.1 RAID technology | 9 |
| 3.2 Archiving | 9 |
| 4.0 Step 3: Choose a complete system | 10 |
| 4.3 Emergency Services | 13 |
| 5.0 Key features to specify | 14 |
| 6.0 Summary..... | 15 |
| Appendix A: Business case for upgrading from DAT | 17 |

1.0 Executive Summary

Recent legislative changes in sectors such as the Courts, Air Traffic Control and Emergency Services mean that audio recordings (and often other data as well) need to be kept for a minimum number of days or years.

In order to maintain a database of recordings, effective solutions for mass storage need to be implemented. This white paper explains the storage requirements for network and mass storage in the following sectors:

- ▶ Courts
- ▶ Air Traffic Control
- ▶ Emergency Services

Many other industries also wish to minimise administration and cost associated with legacy storage solutions whilst others wish to implement new, efficient storage networks. Network solution options for typical situations are evaluated and the best solutions for particular scenarios are then recommended.

The white paper then describes archiving, management and replay features that allow for easy manipulation of your archived recordings to be able to access the information that you require whenever and wherever you need to.

Appendix A includes brief case studies, detailing how two organisations justified an upgrade from their DAT-based recording system by creating a business case based on the savings in media and manpower costs that a digital system would yield.

Audience

This white paper is relevant to:

- ▶ Those responsible for the operation of a recording system
- ▶ Those responsible for the storage of data

2.0 Step 1: Understand the Requirements on your Organisation

Industries such as Air Traffic Control, the Courts and Public Authorities have an increasing requirement to store more data, for longer with better data management facilities as legislation changes and a growing number of distributed recording systems evolve.

For requirements on:

- ▶ Courts, see Section 1a.
- ▶ Air Traffic Control, see Section 1b.
- ▶ Emergency Services, see Section 1c.

2.1 Courts

Legislation on Courts varies from country to country. In some countries recordings are not mandated; in others, audio and video recordings need to be kept for a number of years. Consider a typical scenario in which an organisation manages 100 courtrooms, each recording 8 channels of audio (assuming a compression of 64 KBps) for 8 hours each day. The storage required would be 1.5Gb per day per courtroom, which equates to approximately 400Gb per year for just one of the hundred courtrooms.

If a court also records one channel of video then even greater storage is required. Good quality video can be achieved using 4CIF 704 x 576 resolution at 25 frames per second but this can require as much as 1Gb of storage per hour, which equates to 2Tb per court per year.

2.2 Air Traffic Control

In airports, all audio communications must be recorded and kept for a minimum of 31 days so that any incidents may be replayed and investigated.

Based upon a typical ATC setup with an average 32 audio channels, 25% channel activity per day and assuming a compression of 64 KBps, this would require over 5GB per day of storage. In order to maintain this data for the 31 days as required by ICAO regulations this would require 150GB of storage, a relatively modest amount. In reality, recordings are often kept for longer than this where possible.

A recent key change in the policy of the International Civil Aviation Organization (ICAO) specifies that Radar recording of primary and secondary radar is mandated and that recordings should be kept for a minimum of 31 days. Radar recording can require significant amounts of storage; a typical channel of 2k x 2k resolution screen capture at one frame per second uses approximately 13Gb per day or 400Gb per month.

Larger airports and control centres may have to deal with significantly larger numbers of channels than the above. For example, Belgocontrol at Brussels currently records 2760 channels of audio data (through redundancy), 31 days of data requiring 13Tb of storage for the audio alone.

2.3 Emergency Services

The requirements on the Emergency Services can vary significantly from one application to another but in many cases there is a need to keep data for seven years. There are a variety of industry standard platforms from which to record data, including analogue and TETRA digital radio, mobile and PSTN telephony. A typical site may require storage of 60 channels with up to 25% channel activity per day. Assuming a compression of 64 KBps, this would require over 10GB per day. In order to maintain this data for seven years, one would need 25Tb of storage.

2.4 Evidentially Admissible Data

In recording environments one issue that must be addressed is that of the integrity of recorded data.

In industries where recordings must be legally admissible, some kind of check to ensure that data is a true copy of the original and has not been tampered with must be in place.

Leading archive software can perform a Cyclic Redundancy Check (CRC) to ensure that archived data is an exact replica of the original recording, including all header information provided by the recording module.

Check with your voice recording supplier what measures are in place to ensure the integrity of your recordings.

3.0 Step 2: Understand the Options available

With typical storage requirements growing by 60% per year (much faster than developments in removable media), networked solutions to this problem must be found, with Network Attached Storage (Option 2) & Storage Area Networks (Option 3) being at the forefront of this race. This section outlines the main options for storage, beginning with the most commonly used Direct Attached Storage.

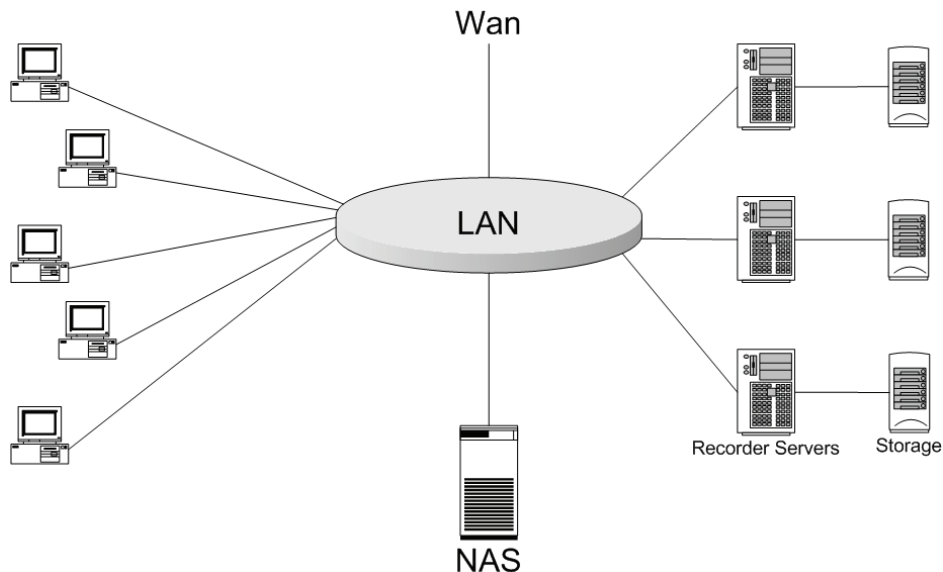
| |
|---|
| Option 1 – Direct attached storage |
| Background |
| Direct Attached Storage (DAS) is a storage device that connects directly to a single server |
| Network Topology |
| ▶ Storage is attached directly to PC |
| Advantages |
| ▶ In addition to handling the storage and retrieval of data files the server will also handle business applications such as e-mail or databases. ▶ Storage is simple to administer since you only have to deal with one server connected to the storage device ▶ Allows the user to add and assign storage where it is most needed ▶ Low incremental cost of purchase ▶ Optimised for local storage |
| Disadvantages |
| ▶ The result of handling multiple applications is that performance of direct storage is not as high as that of networked storage. ▶ Lacks the flexibility of shared networked storage that is required for most businesses |
| Cost Index |
| Cost index = 1 (where 1 = low, 5 = high) |

Option 2 – Network-attached storage

Background

A network-attached storage (NAS) device is a server that is dedicated to nothing more than file sharing.

Network Topology



Advantages

- ▶ Allows more hard disk storage space to be added to a network that already utilizes servers without shutting them down for maintenance and upgrades
- ▶ Does not need to be located within the server
- ▶ Can exist anywhere in a LAN
- ▶ Can be made up of multiple networked NAS devices
- ▶ Optimised for performance and scalability
- ▶ Support for high-speed Fibre Channel media which is optimised for storage traffic, managing multiple disk and tape devices as a shared pool with a single point of control
- ▶ Specialised backup facilities that can reduce server and LAN utilisation

Disadvantages

- ▶ Storage is not an integral part of the server. Instead, in this storage-centric design, the server still handles all of the processing of data but a NAS device delivers the data to the user.
- ▶ NAS does not provide any of the activities that a server in a server-centric system typically provides, such as e-mail, authentication or file management.

Cost Index

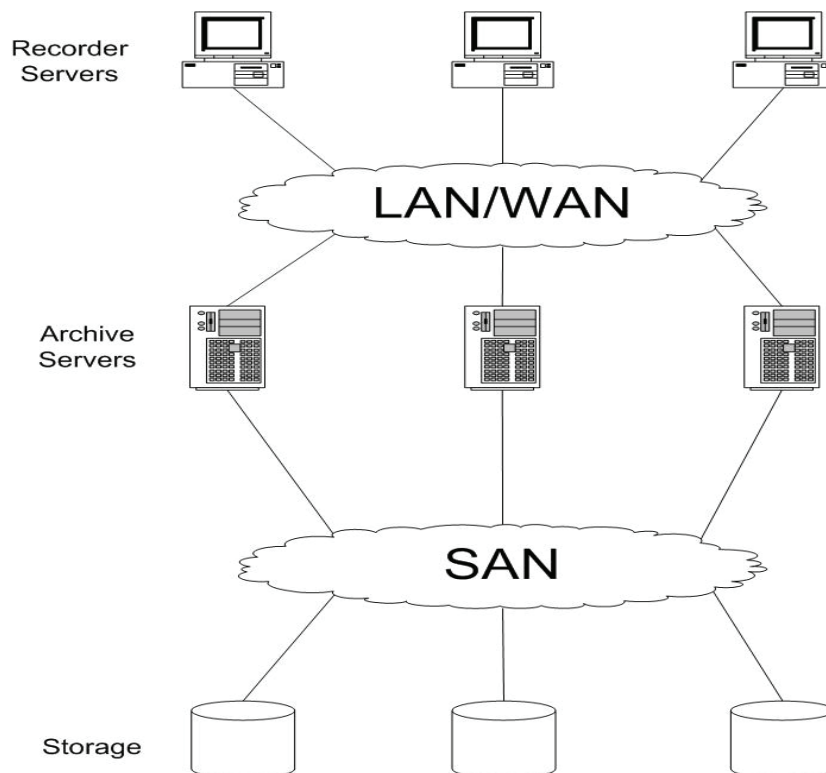
Cost index = 2 (where 1 = low, 5 = high)

Option 3 – Storage Area Network

Background

A Storage Area Network (SAN), is a high-speed sub-network of shared storage devices, each of which contains nothing but disks for storing data. A SAN's architecture works in a way that makes all storage devices available to all servers on a LAN or WAN. The server merely acts as a pathway between the end user and the stored data.

Network Topology



Advantages

- ▶ As more storage devices are added to a SAN, they too will be accessible from any server in the larger network
- ▶ Server power is utilised for business applications and network capacity is released to the end user
- ▶ Provides a common link between multiple servers and storage, allowing administrators to scale only the storage or sever processing power as requirements demand
- ▶ Allows multiple servers to access the same data so that duplication of information can be reduced
- ▶ Facilitates storage back-up directly over the storage channels without adding the bottleneck of the relatively slow Local Area Network (LAN)
- ▶ Optimized for ease-of-management and file sharing using lower-cost Ethernet-based networks

Cost Index

Cost index = 4 (where 1 = low, 5 = high)

3.1 RAID technology

RAID stands for "Redundant Array of Independent Disks". RAID is a method of storing data on multiple hard disks. When disks are arranged in a RAID configuration, the recorder sees them all as one large disk. However, they operate much more efficiently than a single hard drive. Since the data is spread out over multiple disks, the reading and writing operations can take place on multiple disks at once. This can speed up hard drive access time significantly.

The benefits of RAID come from a technique called "striping," which splits up the stored data among the available drives. The "stripes" of data are usually a couple of megabytes large and are interleaved between the drives. The striping system also increases the mean time between failures (MTBF), when reading data. This allows more data to be read accurately in a short period of time.

The benefits of the RAID system are especially noticeable when storing large amounts of data. RAID systems can be deployed locally in a DAS configuration or as part of a NAS or SAN centralised storage solution. They can be scaled up to 24 Tb (1Tb = 1024Gb) and can be supplied with hot swappable, redundant PSU's and Hard Drives.

There are several different 'levels' of RAID as described below:

- ▶ RAID 0 is the fastest and most efficient array type but offers no fault-tolerance. RAID 0 requires a minimum of two drives.
- ▶ RAID 1 is the best choice for performance-critical, fault-tolerant environments. and is the only choice for fault-tolerance if no more than two drives are used.
- ▶ RAID 2, RAID 3 and RAID 4 have been superseded by RAID 1 and RAID 5 for most new operating systems and are not supported by most RAID controllers.
- ▶ RAID 5 combines efficient, fault-tolerant data storage with good performance characteristics. However, write performance and performance during drive failure is slower than with RAID 1. Rebuild operations also require more time than with RAID 1 because parity information is also reconstructed. At least three drives are required for RAID 5 arrays.

3.2 Archiving

Archiving to removable media is one way of achieving security through backing-up important data. Traditionally, recordings were written to a local hard drive and archived to removable media for long-term storage. This used to be to cassette tapes, then DAT tapes but more recently to CDs and DVDs. However, as storage networks have become viable, so organisations have begun to move away from removable media and the costs involved. New technologies such as Blu-Ray and HD-DVD offer 15-25Gb storage per layer (side) but the cost of the media and the time taken to change the media every few hours/days means that some organisations now only archive to fixed storage.

4.0 Step 3: Choose a complete system

Recommended Solution:

It is recommended that Option 1 only be used for small-scale storage purposes where the recording of calls is of low priority. Options 2 and 3 each have advantages that make them appropriate in different medium-large scale applications as per the table below.

| | Option 1: Direct Attached Storage (DAS) | Option 2: Network Attached Storage (NAS) | Option 3: Storage Area Network (SAN) |
|--------------------------------|--|---|---|
| Simplicity of installation | ✓ ✓ | ✓ ✓ | ✓ |
| Flexibility of upgrade options | ✗ | ✓ | ✓ ✓ |
| Reliability | ✓ | ✓ ✓ | ✓ ✓ |
| Speed of access to data | ✓ ✓ | ✓ | ✓ ✓ |
| Impact on network | ✓ ✓ | ✗ | ✓ ✓ |
| Cost Index | ✓ ✓ | ✓ | ✓ |

Table 1: Analysis of different options for Radar Recording

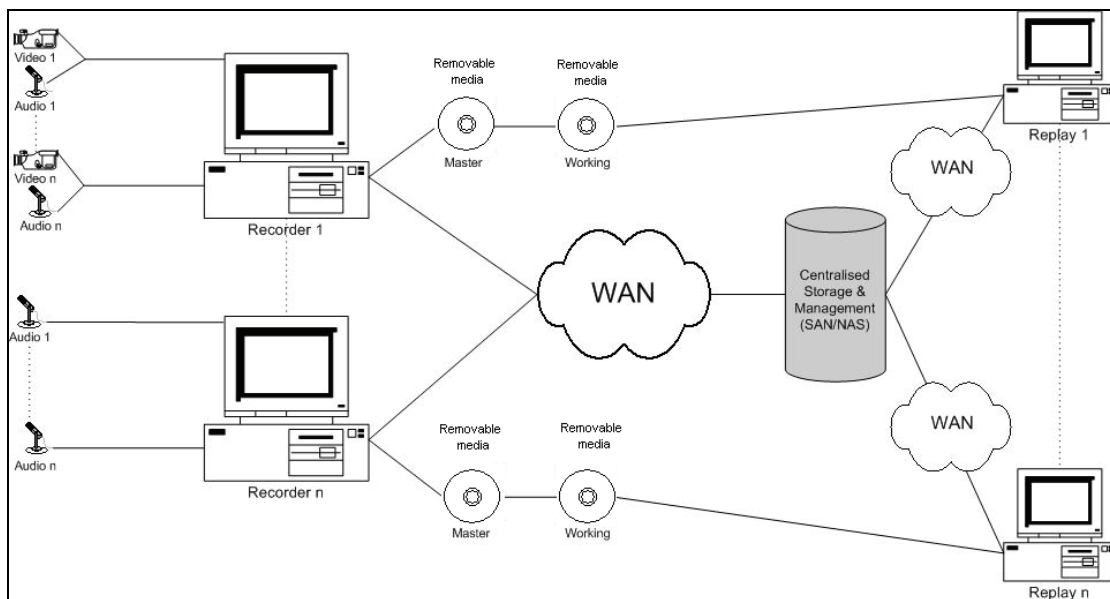
✗ = unsatisfactory, ✓ = satisfactory, ✓ ✓ = excellent

How to archive, store and retrieve data depends on the nature of the data, regulatory requirements applying to the industry and network topology but below are some examples of how the storage could be managed for particular industries.

4.1 Courts

For up to 7 years required archiving:

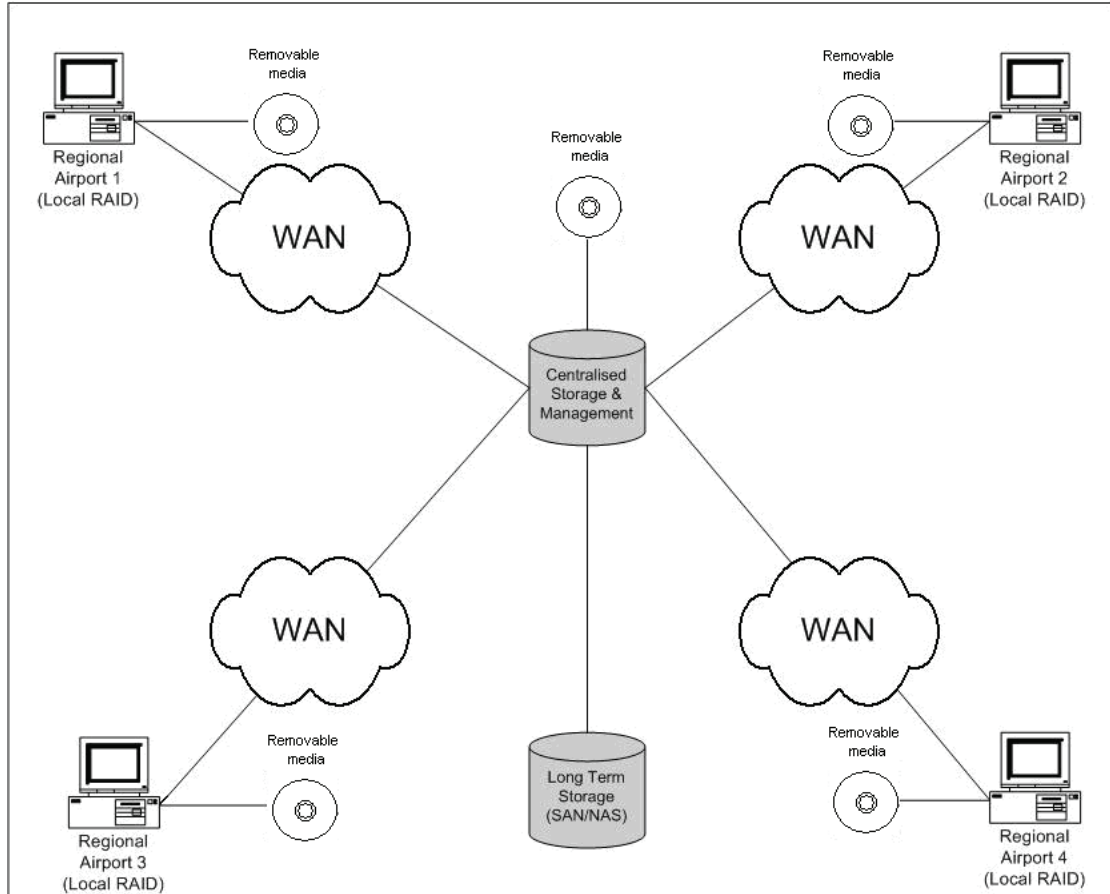
- ▶ Distributed secure access across LAN/WAN
- ▶ Large capacity storage as video complements audio



4.2 ATC

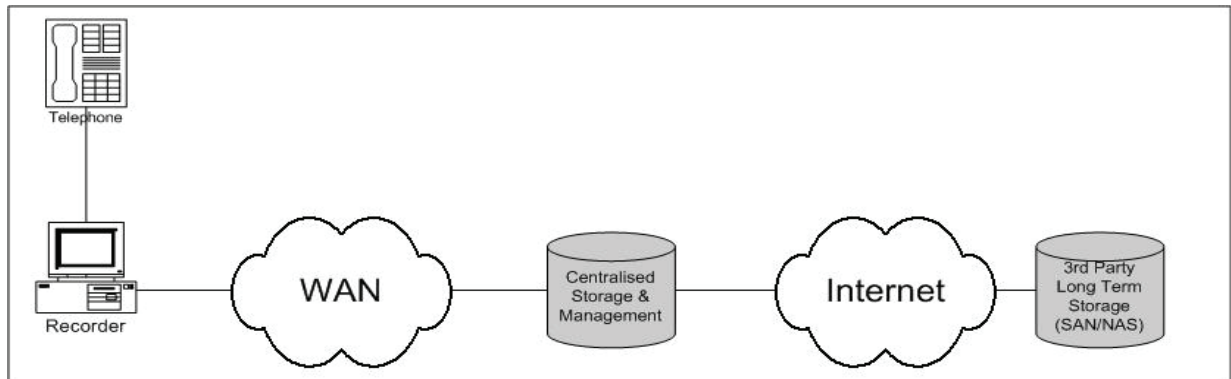
ATC (31 days required archiving)

- ▶ Distributed Local RAID
- ▶ Centrally managed storage and archiving (NAS or SAN) with local DVD archiving



4.3 Emergency Services

- ▶ Selective long term archiving
- ▶ 3rd party companies involved for long term storage



5.0 Key features to specify

When choosing your recording and archiving system, it is important to be able to easily manipulate the data, using automatic processes to comprehensively archive data and a user-friendly interface to manage and replay your data. Listed below are a number of features available in a next generation digital recording system that mean no more searching through all of your old archive media to access the information that you require:

1. Multiple archive processes

- ▶ Automatically archive data to storage locations as required
- ▶ Multiple locations, multiple devices: Local, central and/or emergency backup
- ▶ Media of all types: Separate processes available for audio, video, other media and/or associated documents
- ▶ Schedule archiving to take account of low network activity (e.g. overnight)
- ▶ Support for industry standards
- ▶ Error handling (e.g. for losing network drives)

2. Web-based streaming (“thin-client”) for replay of recordings; ensure that your recording system allows you to replay on any networked PC (with the appropriate user access) using web-streaming through a browser such as Internet Explorer v7. Not only does this make replay easier, it means no unnecessary copying of files, thus preventing careless loss of sensitive data.

3. Local database of archived recordings: Allows you to easily find the recording that you require without extensive searching.

4. Web-based interface for unmanned operation: Manage recorders and recordings whenever and wherever required

5. Seamless integration with audio mining tools: Whilst huge quantities of data from a variety of sources can now be recorded and stored, the fact is that most organisations do not have sufficient manual resource to analyse the recorded data and may be missing vital information. Best in class audio mining tools allow rapid identification of important information that may be buried within thousands of hours of these recordings. Whether you are searching for a word, a phrase, a person or just trying to understand complex relationships audio analytics can greatly enhance your operational capability and save significant costs.

It may be that audio mining is a consideration for the future or alternatively, you may need to archive recordings to a central location for immediate key word spotting by a server farm. In either case, it is important to choose a recording and archiving system that has been designed to work with analytical tools, optimising the recording process to allow useful interrogation of the data through tools such as key word spotting, text to speech, speaker identification and language identification.

For more information, please ask for our white paper on Audio Mining – How to automatically locate vital information: A review of technologies that will allow you to optimise your analysis and transcription resource.

6.0 Summary

Courts, Air Traffic Control and Emergency Services need to maintain a database of recordings for set amounts of time in order to meet legislative requirements.

This white paper has set out options for this mass storage including a RAID array of hard drives, Network Attached Storage and Storage Area Networks.

When choosing your recording and archiving system, it is important to specify features such as multiple archiving processes for comprehensive management of your data and a web interface for easy access to and replay of your archived recordings.

For an effective return-on-investment, solutions for mass storage need carefully planned so that they can be implemented in stages or scaled as storage requirements grow. An effective solution can significantly reduce both media and manpower costs.

7.0 Bibliography

4CIF: 704 x 576 pixels screen resolution, roughly four times the area of CIF (Common Intermediate Format; 352 x 288 pixels)

ATC: Air Traffic Control

Blu-Ray: Disc format that uses blue laser rather than the red laser used in standard DVDs. Storage capacity of 25 GB (single-layer) or 50 GB (dual-layer).

CIF: Common Intermediate Format screen resolution; 352 x 288 pixels

CRC: Cyclic Redundancy Check. A CRC is a type of check value designed to catch most transmission errors. A decoder calculates the CRC for the received data and compares it to the CRC that the encoder calculated which is appended to the data. A mismatch indicates that the data was corrupted in transit.

DAS: Direct Attached Storage

DVD: Digital Versatile Disk

Gb: Gigabyte = 1024 Megabytes

HD-DVD: High Definition/High Density-DVD: Disc format that can store 15GB (single-layer) or 30GB (dual layer)

ICAO: International Civil Aviation Organisation

LAN: Local Area Network

MTBF: Mean time between failures

NAS: Network-Attached Storage

Primary Radar: A Radar in which the return signals are the echoes obtained by reflection from the target.

PSU: Power Supply Unit

RAID: Redundant Array of Independent Disks

SAN: Storage Area Network. A high-speed sub-network of shared storage devices, each of which contains nothing but disks for storing data

Secondary Radar: A Radar in which the return signals are obtained from a beacon, transponder, or repeater carried by the target

Tb: Terabyte = 1024 Gigabytes

WAN: Wide Area Network

Appendix A: Business case for upgrading from DAT

To find the best storage configuration for your organisation, first assess how archiving is currently performed and the costs involved. For example:

Organisation A

Organisation A had a relatively large number of channels but only needed to keep data for a short period of time.

Organisation A were:

- ▶ Recording ~200 channels of audio at any one time, 24 hours a day, 7 days a week, 365 days a year
- ▶ Archiving to DAT tapes, 2 hours per £4.35 tape
- ▶ Reusing each DAT tape 100 times (despite manufacturer's recommendation of only reusing DAT tapes 15 times)
- ▶ This equated to a cost of **£38,106 per annum** in replacement DAT tapes

Intangible costs also included:

- ▶ Manpower involved in replacing and labelling DAT tapes every two hours
- ▶ Manpower involved in searching through a vast quantity of DAT tapes for old recordings
- ▶ Couriering of data on DAT tapes from one site to another as data could not be sent electronically
- ▶ Supporting legacy DAT equipment

Risks involved:

- ▶ Loss of data through overuse of DAT tapes
- ▶ Obsolescence of equipment meant it would soon no longer be supported by manufacturer

Organisation A chose to move to a DVD-RAM archiving solution:

- ▶ Each DVD-RAM could store 60 hours of data, costing £10 each
- ▶ Each DVD-RAM can be reused 10,000 times
- ▶ Replacement DVD-RAMs would therefore only cost **£30 per annum**
- ▶ No couriering of data – all files could be sent electronically using existing infrastructure
- ▶ Archiving software performed a Cyclic Redundancy Check (CRC) to ensure the integrity of the archived recordings
- ▶ A local database of all recordings meant it was much quicker to find old recordings
- ▶ **Annual media costs and manpower costs (and maintenance of recording system costs) had therefore been vastly reduced**

Organisation B

Organisation B had 25 sites and needed to keep a small number of audio (and in some cases video channels) from each site for 7 years. These recordings needed to be accessible from anywhere.

Organisation B were:

- ▶ Recording 4 channels of audio at each of 25 sites for 8 hours a day, 5 days a week, 250 days a year
- ▶ Recording 2 channels of video at 4 sites for 8 hours a day, 5 days a week, 250 days a year
- ▶ Archiving audio to DAT tapes, 2 hours per £4.35 tape
- ▶ Archiving video to VCR tapes, 4 hours per £0.99 tape
- ▶ This equated to a cost of **£435,000** per annum in DAT tapes plus £3,960 per annum in VCR tapes

Intangible costs also included:

- ▶ Manpower involved in replacing and labelling DAT tapes every two hours
- ▶ Manpower involved in searching through a vast quantity of DAT tapes for old recordings
- ▶ Couriering of data on DAT tapes from one site to one central location, every day

Organisation B chose to move to a SAN storage solution:

- ▶ All recordings could be transferred across the WAN when required, both by an automated archiving process for storage and by operators (with the appropriate access) using a web-based interface from any of the 25 sites for replay
- ▶ **Reduced storage costs** – all archiving is to fixed storage, which is significantly cheaper than removable storage
- ▶ No couriering of data
- ▶ Archiving software performed a Cyclic Redundancy Check (CRC) to ensure the integrity of the archived recordings
- ▶ All the recordings were stored centrally so search times were reduced to a few seconds
- ▶ Audio and video could be synchronised and exported to a common format (such as .avi or .mpeg) along with associated notes