



## Exploiting Grouping Capabilities to Optimize Archive Performance

Steve Atkinson, VP Americas, SGL

The progression of video within the broadcast plant from a collection of videotape to a collection of files impacts different operation types and different departments within each operation differently. The traditional thinking of the video archive being a monolithic entity no longer need apply. The fact that the video archive is a repository of files allows differentiation in terms of storage types and storage policies all within the same video archive.

A majority of archive management products for broadcast video support the concept of 'groups'. This concept allows for video of similar attributes to be stored together logically and be allocated common resources. These groups can be contained within a single robotic library, multiple robotic libraries, shelf storage, or a combination of the above. The efficient use of groups not only makes media management easier but can also improve workflow. Rather than have operators lump everything ingested into a single repository, multiple logical repositories can be established. With a little foresight before selecting the archive and a similar amount of operational discipline after installation, the broadcaster can achieve the optimal archive hardware configuration to meet his needs while at the same time smoothing the workflow for everyone involved in the ingest and retrieval process.

Groups within a video archive have three different types of resources allocated to them: storage media (data tapes or DVDs), the drives used to read and write to the media, and the robotic systems to manipulate the media. A single archive management application can manage multiple types. For example, a DVD library may be used for short form material and a high performance tape drive system for long form. If properly implemented, the operator need not have any knowledge about the underlying physical storage. He or she just needs to know that material of type X gets archived to group A and material of type Y gets archived to group B.

While it may seem elementary, many video archives are installed without any pre-analysis. The approach of bigger and faster hardware is used for a 'one size fits all'. Only after a period of use are the lessons of proper media management learned. A number of factors need to be considered when selecting archive hardware and subsequent implementation of group policy. This is best done before installation rather than trying to fix it later.





Exploiting grouping capabilities to optimize archive performance...

### Data storage basics

Magnetic data tape and DVD data storage differ greatly in the way that they write data to their respective media. While certain tape vendors have implemented features to allow segmentation of a single tape into multiple logical segments, all are still basically sequential media. This means that video files are written to tape in an 'append only' manner. When new material is written to tape, it is written to a location on tape succeeding the last previously written video file. Each file written to tape is separated by a single tape mark with the end of the last file written indicated by a double tape mark. When the archive management application looks for a free space to write a new video file, it looks for the double file mark. It then overwrites the double file mark with a single file mark, writes the file to tape, and then writes a double file mark. This process continues until a tape is full and a blank tape is then allocated from the 'free pool' to become the new write tape.

The larger implication of these algorithms comes when material is no longer needed and the associated space on the tape needs to be recovered. Unlike magnetic videotape, where material can be overwritten and inserted in the middle of a tape, data tape does not allow this. The writing of a double file mark indicates that the rest of the material on the tape, from the double file mark to the end of the tape, is no longer valid. This allows the rest of the tape to be overwritten. This means that when writing to tape, new material is only appended to the end of the previously written files. If a file is deleted, no attempt is made to write into that space. Figure 1 shows graphically the append, delete, append cycle.

One other, albeit very desirable but complicating, feature is read after write, rewrite functionality. A majority of tape drives will verify that the data sent to it was actually the data written to tape by reading the data it has just written and comparing it to the data sent in its buffer. If there is an error, the data will be attempted to be written again and again until successful or it reaches a maximum number of retries. This allows the drive to compensate for possibly damaged sections of tape with no impact to the application other than somewhat degraded throughput.

Taken together, this means that if a 30 second spot is deleted, the space that it occupies cannot be overwritten. First, writing into that area will initiate the writing of a double tape mark when finished, in effect deleting the material following it. Even if there is a solution to the double tape mark problem, you still have the non-deterministic space issue. A 30 second spot may take 30MB of space on tape or 70MB of space on tape. The possibility exists of overwriting the material succeeding the open space on tape.





Exploiting grouping capabilities to optimize archive performance...

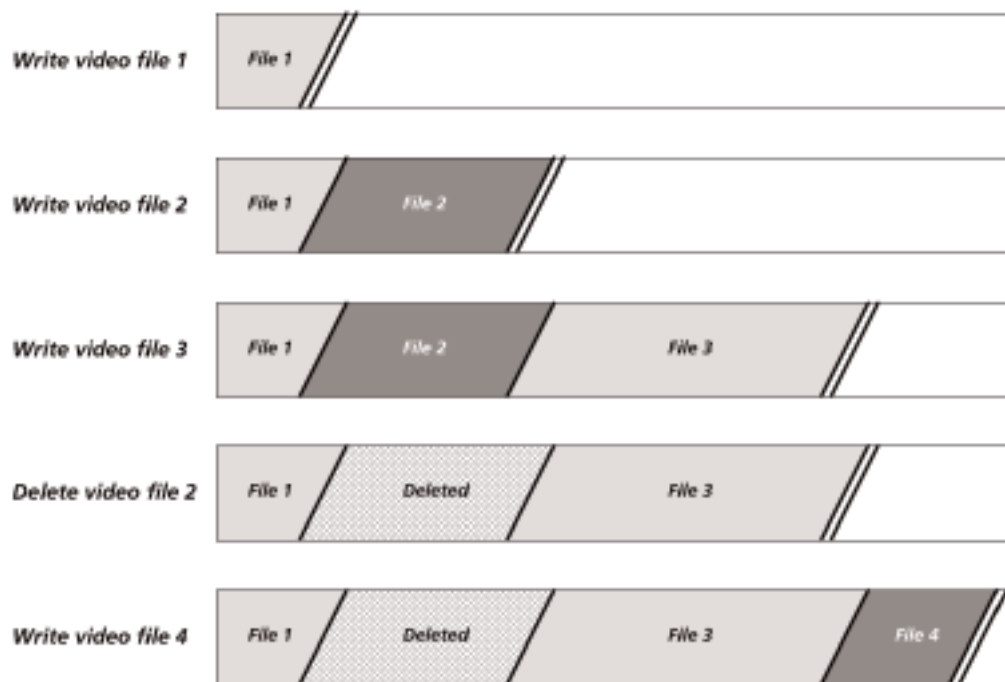


Figure 1

DVD drives are much more like disk drives in a lot of respects. Deleted video files create individual spaces on the DVD disk that can be written to again. This allows the media to be used much more efficiently. The fact that there is less fragmentation does not imply that grouping is unnecessary for DVD archives. As discussed later, efficient media use is just one of the benefits of an effective archive grouping strategy.

### Defragmentation - cleaning up after the fact

Defragmentation by the archive management software mitigates the 'swiss cheese' effect that can impact an archive library over time but it does not come without a cost. Defragmentation works by monitoring the fill level on tape over time as material is deleted from the archive. Deleting in the context of archive management means that references to the material and its location are deleted in the database rather than actual deletion of the material on the archive medium. This means that the material still exists but is inaccessible. The swiss cheese effect comes into play when material which has been 'deleted' is intermixed with material still needed. A broadcaster may have several spots still needed followed by a deleted spot followed by a long form program, etc. Defragmentation is needed to recover the unused storage in the library.





Exploiting grouping capabilities to optimize archive performance...

There are two implementation approaches to defragmentation but both operate basically the same. The most common approach is to cache the material on a fragmented tape on the disk system of the computer server on which the archive management software resides. The second approach is to cache the material on the disk subsystem of one of the video servers networked to the archive server. Before proceeding it is useful to be familiar with the individual components of the archive, video server, edit suites and information technology.

While there are different operations within the broadcast plant, in terms of storage, the components needed for the migration of material are basically the same. For the on-air sid, material which is not needed immediately for playback or which has been played to air and is needed for subsequent playback at a later date is sent to the archive. The migration is controlled more often than not by an automation system rather than by a human operator.

Edit suites are more interactive as there is less predictive information about what material will be needed from the archive and when it will be needed. The migration varies on factors like the time of year and the needs of the individual editor. The migration and retrieval policies tend to be more ad hoc.

Non-video IT archive functions tend to be very predictive. Rather than archive functions, they are in most cases backup functions. The difference being that archive functions remove the original copy of the material from the source while backup functions make a secondary copy and leave the material on the source. IT backups are commonly a daily function performed at a specific time of day (usually the middle of the night when the machines are not in use) and are either incremental or full backups. The differences between the two being that incremental backups are backups of only files which have been changed and full backups being the entire disk subsystem for each station. Figure 2 shows the different systems in a shared archive environment.

During normal operation (i.e. excluding defragmentation operations), material which is sent to and retrieved from the archive is not cached to the disk on the archive server. A network connection is established between the video server and the archive server. The physical connection is commonly fibre channel or gigabit ethernet. When the media to be written to or read from is mounted in a drive, the material is streamed through the archive server to the destination. This is not only the most time-efficient method to send material to and from the archive but also minimizes the amount of disk storage required for the archive server. Using this methodology dictates that the disk space on the archive server only need be big enough to store the archive database or indices.



# FlashNet

Lock up your data™



Exploiting grouping capabilities to optimize archive performance...

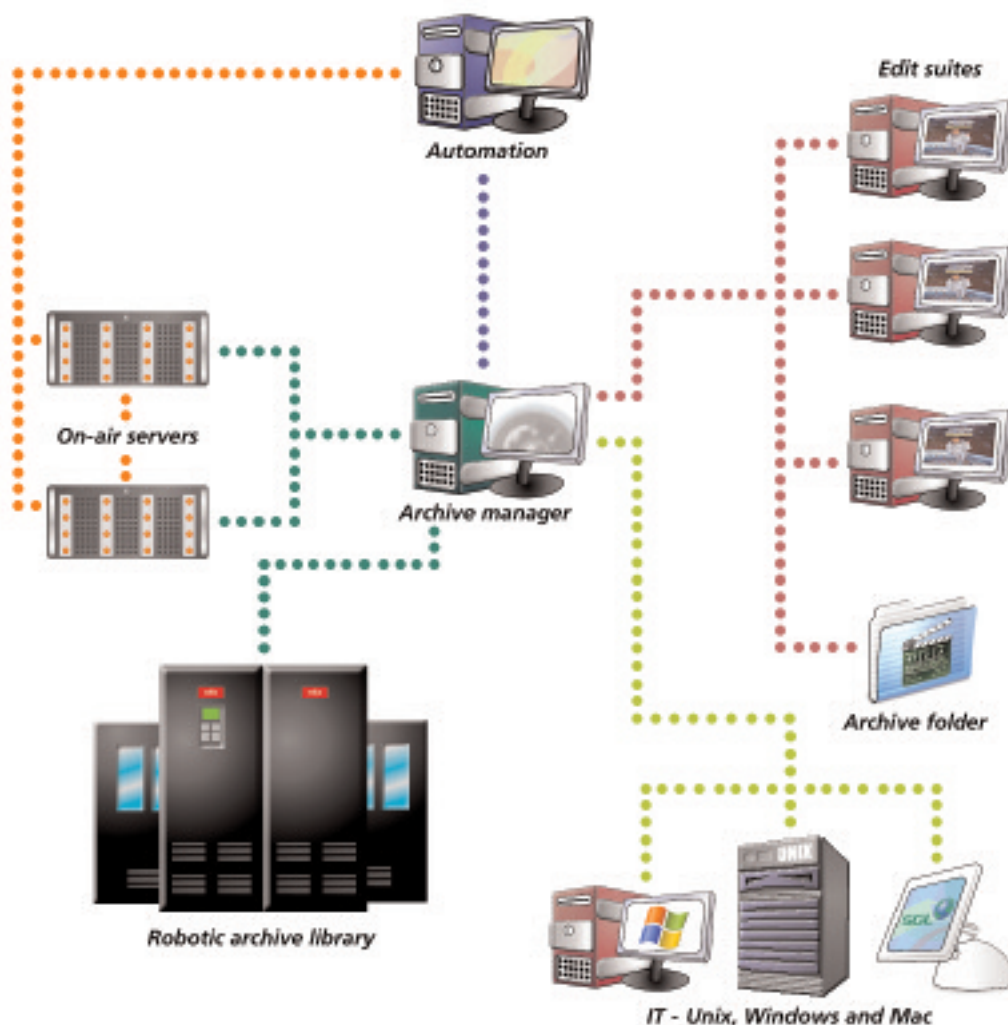


Figure 2

Caching defragmentation material to the archive server obviously requires more disk resource for the archive server. The disk size must be, at a minimum, big enough to hold the largest video file in the archive. When an individual tape or DVD is defragmented, the individual video files on that medium are read out onto the archive server cache and then written out to an unfragmented medium. This can be done singly or in batch mode if there is enough disk space.

The same approach applies to using a video server for cache. Using the video server for cache eliminates the need for additional disk on the archive server but may impact the resources needed for playback. In both cases, the greater impact is on the drive resource in the archive. The most time efficient way to process defragmentation is to mount the medium to be defragmented in one





Exploiting grouping capabilities to optimize archive performance...

drive and the medium to which the material is to be written on another drive. Material is then written to cache and then re-written to a new medium immediately. This ties up two drives while the defragmentation is taking place. Alternatively, a single drive can be used with the material to be defragmented written to disk cache, a new medium mounted and the cached material written back out. This process is slower but minimizes the impact to the drive resources. In both cases, it is good practice to assign the defragmentation process a low priority with read and writes a higher priority so as not to impact operations.

### Preventive measure through grouping

Groups can be looked at as separate logical archives within one or multiple physical archives. Each group of material commonly has its own dedicated media. Resources (drives and robots) to write and read to those media can be shared with other groups or dedicated to specific groups. Beyond a common set of physical resources, groups can be assigned different attributes and rules. For example, the metadata stored for one group of video files can be different from a separate group of video files. Production material will need a different set of attributes stored in terms of date, producer, rights etc., than a local ad intended for the on-air server.

Groups can be established and changed at any time but it is best to establish at least a basic framework when the archive is first brought on-line. Establishing groups prior to bringing an archive on-line requires analysis of the expected flow into and out of the archive as well as which departments will be using the archive and when their peak demands will be during the course of a 24-hour day.

### Material length as a group consideration

One of the first parameters to look at in deciding how to establish groups is the length of the material. Given all other factors being equal, material of similar length should be stored as a group. For example, short form material and long form should be allocated separate groups. There are several reasons why this is effective. First, it is an intuitive allocation for station personnel handling the ingest process. Essentially, spots and promos will go to one place, programs will go to another.

Second, keeping like material together will help optimize retrieval times. It is likely that similar material such as ads will be retrieved as a group as part of the look ahead by the automation. If





Exploiting grouping capabilities to optimize archive performance...

no grouping is used, it is also likely that ads, promos and programming will be interspersed throughout the archive library. This can lead to unnecessary mounting and dismounting of tape.

For example, assume Clip1 and Clip2 are 30 second spots which arrive and are ingested and archived early in the day. If groups are not assigned, the archive management software assumes the entire archive is a single group. All incoming material will be written to the current 'write' tape (Tape1, for example) for that single group. The next material in the door happens to be two weeks of a syndicated program. That is ingested and archived. Given the volume of material, it fills the current write tape and part of the next write tape, Tape2. Next, two more 30 second spots, Clip3 and Clip4 arrive and are written to Tape2.

When the playlist arrives, Clip1, Clip2 and Clip4 are spots needed for playback. Automation requests restoration of these clips which initiates retrieval of Clip1 and Clip2 from Tape1 and retrieval of Clip4 from Tape2. Taken singly, this may not seem like much of an impact but given dozens or hundreds of operations a day, it can start to add up. The reason is that with many storage devices, the mount, setup and search times can be significant. In some cases, this can amount to minutes for mount, setup and search to the end of the tape. Again, if the archive retrieval cycle is high, it can have a significant impact on operations. A better approach for the above example is to store Clips1-4 in one group and the syndicated program in its own group. Figure 3 illustrates proper versus improper grouping of short form and long form material.

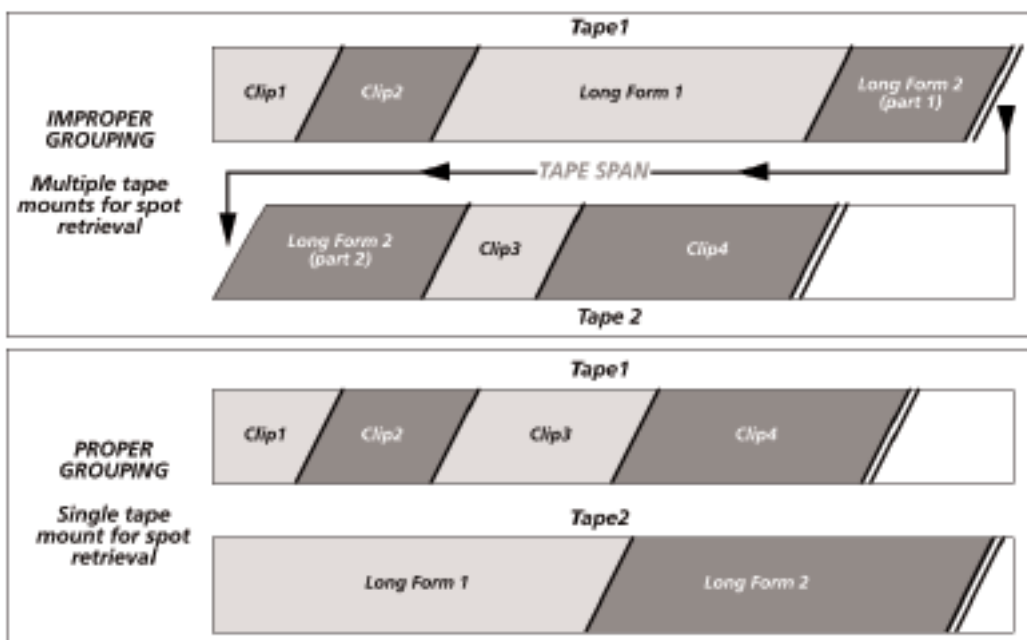


Figure 3





Exploiting grouping capabilities to optimize archive performance...

Third, proper grouping can minimize wear and tear on the archive robotics and drives. As with all devices with moving parts, those moving parts (particularly tape heads) are most likely to wear out. Using grouping in association with another common archive management tool called 'tape efficient retrieval' can minimize the amount of times that a tape and associated tape drive is asked to mount, spin up, search, read or write, and dismount. Most archive management systems will operate in the most 'tape efficient' manner unless told to do otherwise. Tape efficient refers to the fact that given a list of files to retrieve, the archive management software will retrieve all the files on the list from a tape before dismounting the tape and moving on to the next retrieval. Given the previous example, Clips1-4 would be retrieved in the order they are laid down on tape, regardless of the order in which they came in the requested list. The archive management system would then retrieve the next request in the list, process it, and again check to see if there were other items in the list to be retrieved from this tape. This process would be repeated until the request list completes.

Finally, the time of day in which material arrives and is ingested into the archive can affect efficiency without proper grouping. For example, if the first order of business in a normal working day is to archive a network feed, it would go to the current write tape in a monolithic group. If spots tend to come in at the end of the day, they would probably end up on the same write tape and at the end of that tape. What that means is that from a practical standpoint, particularly with single hub tapes, is that the retrieval of a spot at the end of the tape could take minutes to retrieve a short clip. Proper grouping allows for greater determinism in how long it will take to process individual retrievals, particularly short form material.

### Material type as a group consideration

Beyond the length of different types of clips, the type of material should also be considered. Specifically, if material is good for only a period of time then it is best to treat that material as a group. The same advantages specified in the length considerations apply here also but there are additional advantages in material type groupings.

Using time basis as a type example is helpful. A common archive could contain local programming, material from the network, syndicated buys, normal ads and seasonal ads. The rights for local programming would be retained by the station and could remain in the archive available for playback indefinitely. The playback of network material and syndicated material will probably be dictated by contract. In some cases of syndicated material, this may call for the return of source



# FlashNet

Lock up your data™



Exploiting grouping capabilities to optimize archive performance...

material and erasure of all copies. Similarly, ads may have usefulness for a specific time like Christmas ads.

Grouping the material allows the media it resides on to be managed as a group. In the case of seasonal adds, they could be removed from the archive in bulk and put on the shelf to be re-introduced into the library when the season arrives again. If the archive is being used for news applications, the same scenarios apply as clips age. It allows the material to remain in the archive database without physically residing in the library. If groups are not used, dated material may reside on the same tape as material needed for playback the next day. This would make it nearly impossible to remove media from the library without causing operational chaos.

The impact is even more pronounced for material which calls for erasure of all copies under contract. Again, when archive material is deleted, it is not erased but is instead made inaccessible. Inaccessibility will probably not fulfil the contract obligations. Even if it did, the mixing of 'erased'

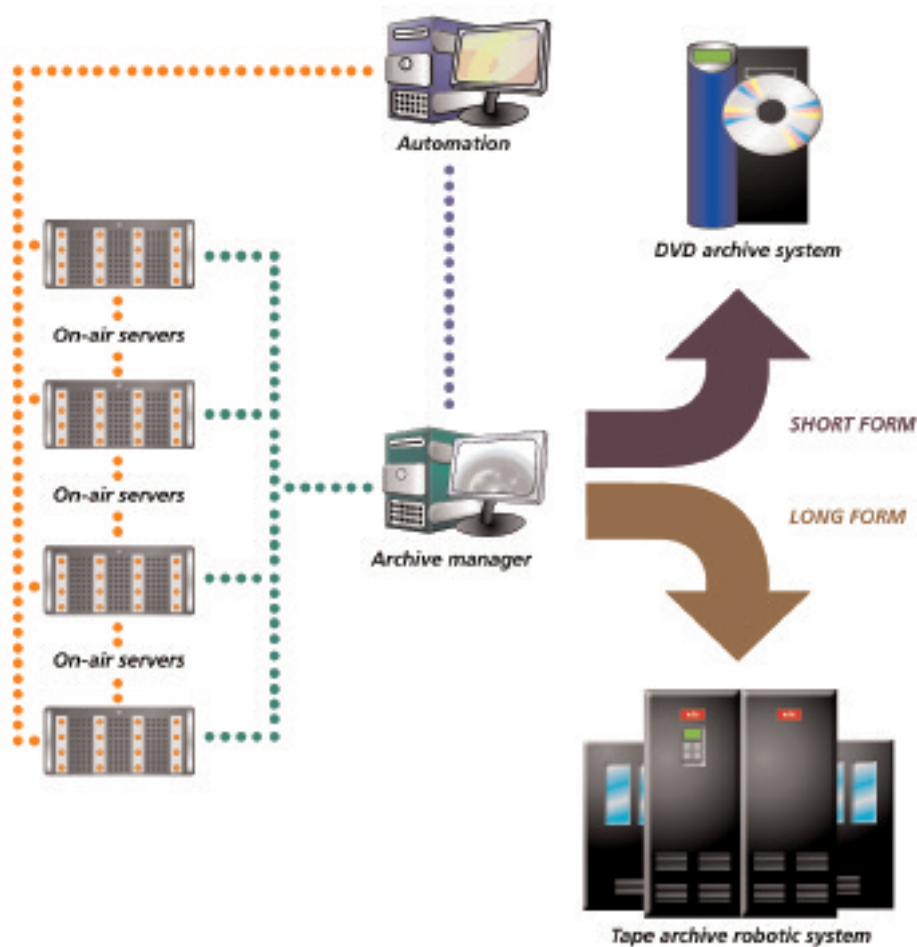


Figure 4





Exploiting grouping capabilities to optimize archive performance...

with valid material will exacerbate the swiss cheese effect discussed earlier. A more effective management scheme is to isolate this type of material on its own group media. When the contractual period expires, the tapes can be group exported from the archive robot, bulk erased and returned to the library for reuse.

Grouping material types can yield greater efficiencies in a heterogeneous archive library environment also. Most archive management systems support multiple library types that can be controlled from a single software package. This might include both a DVD library and a tape library. Given that tape libraries generally have higher throughput but slower access and vice versa for DVD libraries, material can be directed to the appropriate storage subsystem by using groups.

Short form material may be best served being sent to a DVD library while long form goes to tape library. Specifically, the 20 times real time transfers achievable with a number of tape devices make little sense for ten second spots where the mount and search times are a minute or two while the time to transfer the material is milliseconds. Conversely, retrieving long form at one half real time from a DVD device is not an option for most operations. Directing material to the appropriate device allows material to be stored on the medium best suited for storage and retrieval.

The material type varies greatly from station to station, PBS versus network affiliate, and network versus local operations. Even within a single broadcast operation, news and on-air may share a single archive. The analysis of material type should ideally be done before the archive hardware selection is made.

### Retrieval pattern as a group consideration

Retrieval patterns are most well defined in news operations but should also be considered for on-air. It is well documented that material for news loses its relevance quickly. Immediate access is prudent for material less than a week old, reasonable access to material for a month, and stored in a deep archive thereafter. Allocating groups on a date and time basis can help facilitate the migration of material from one level of the archive to another.

Material less than an operationally set age should remain on disk. This allows immediate access for the journalists in the case of news and playback for on-air operations. When material reaches a certain age, it is migrated to the next level of the hierarchy. Traditionally, this has been the videotape vault. The advent of archive management systems incorporating near-line archive robots have been used to reduce retrieval times to minutes. In lieu of other group criteria, material of a





Exploiting grouping capabilities to optimize archive performance...

certain age can then be exported from the library. As mentioned earlier, most archive management systems allow for the tracking of off-line media for subsequent re-introduction into the library. If grouping is not used, exportation for the library can be problematic. Material from today may be on the same media storing a piece of material from two months ago. Heterogeneous archive storage may also apply to these scenarios also as a DVD library may be one level of the hierarchy in front of the tape library and off-line storage.

This is not to imply that using date-based grouping solely is wise. Certain news stories in particular may have the life of a month or more. Date-based grouping should be used in conjunction with other criteria to create sub-categories of the archive which can be customized to the specific type of operation and department within that operation.

### Allocation of archive resources to specific groups

The allocation of archive resources applies to both heterogeneous and homogeneous archive hardware environments. As discussed earlier, a specific library type may be dedicated to a group (DVD for spots) and another library type to another group (tape for long form). The resources within single homogeneous archive systems can be allocated to groups also. This issue is particularly pronounced in operations where the archive is being shared by multiple departments.

Let's examine a case where on-air, news, and the IT department are sharing an archive library. Each may have a time period during the day when they absolutely have to have access to the archive library. Some operations may take obvious preference over others. Retrieval for on-air is more important than daily backup of the IT department computers. For that reason, it often makes sense for specific drives to be allocated to a specific group.

For this particular case, a single physical archive library is being shared by multiple departments. Again, most archive management products support the view of a single logical archive which is really several physical archives and/or offline storage. Figure 5 show a single archive allocated to different departments. without grouping, all resources (drives and robots) are available to process all requests. A request to backup a personal computer is equally important as retrieving a spot to the on-air server in lieu of specific priority assignments. Given a shared environment, it is worthwhile considering assigning specific resources to a group. Assigning a specific set of drives to



# FlashNet

Lock up your data™



Exploiting grouping capabilities to optimize archive performance...

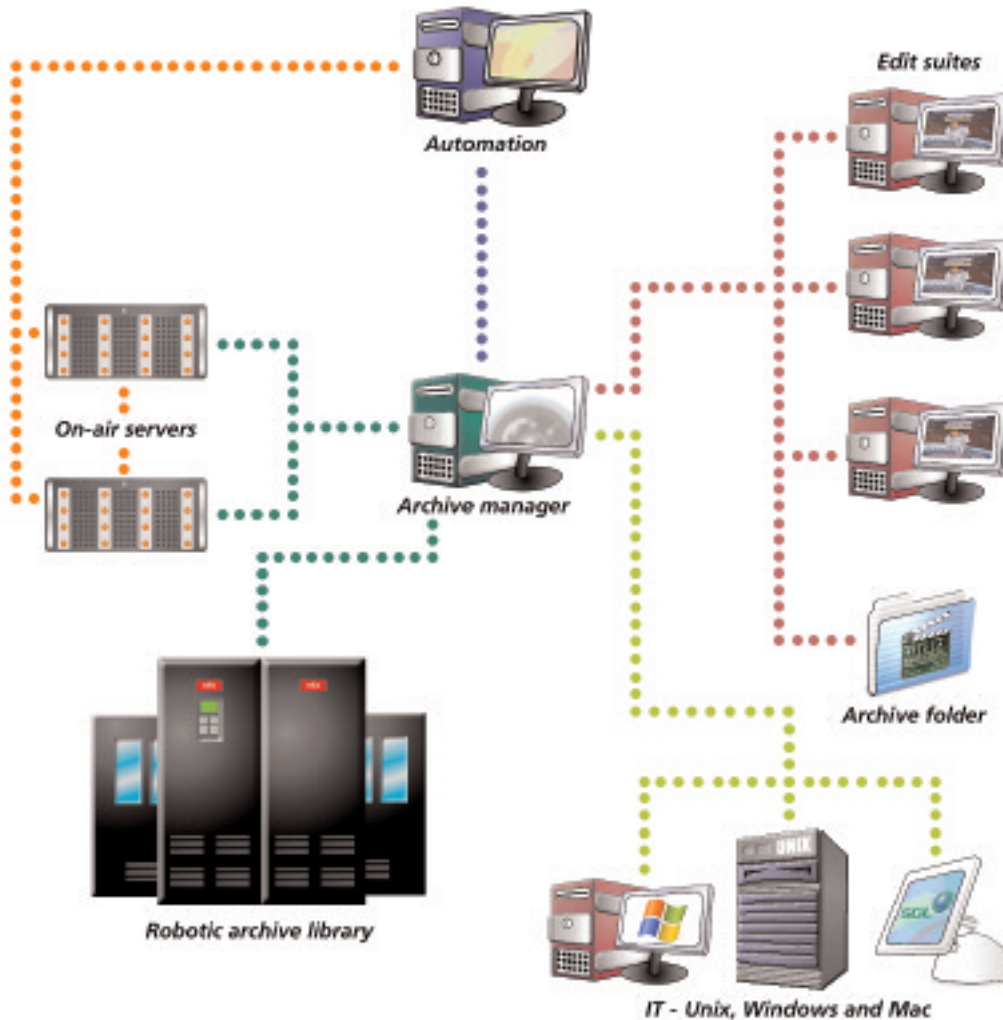


Figure 5

a specific group makes the most sense in these situations. On-air can have a set of drives for their continuous tasks, news can have a set of drives for their daily peak usage period, and IT can have one or more drives for nightly backups which are not time critical.

The penalty paid for such a grouping scheme with strict delineation is under-utilization of resources. A drive which could have been used to accelerate the backup of the IT machines in the middle of the night may not be available even though news is not using it. A middle ground approach is to have a pool of common resources available to all groups while specific resources are available only to one group. This allows on-air applications to always have a drive available for retrieval even if news is in the middle of a developing story and IT is trying to back up their machines.



# FlashNet

Lock up your data™



Exploiting grouping capabilities to optimize archive performance...

## Summary

The considerations presented here are by no means the only ones that should be evaluated before designing, purchasing and implementing an archive system. Each broadcast operation and department within that operation should look at these and other criteria unique to their operation. The point is that these considerations should be brought to the forefront prior to commissioning an archive system rather than trying to find one's way afterwards.

