Laserfiche® Volumes:

Introduction and Best Practices

White Paper

November 2005

Laserfiche®

Table of Contents

What Are Volumes?

A volume is a storage location in a Laserfiche repository. The volumes associated with a repository contains the actual TIFF images, text files and electronic documents in that repository.

Why Use Volumes?

- As different volumes can be placed in different physical locations, volumes are a good way to manage large numbers of files, which can take up a lot of hard drive space. By storing data in different physical volumes, you can spread the files over several different hard disks or computers.
- Breaking data up among volumes allows you to easily back the volumes up, and, since volumes support size constraints, you can help ensure that the data will be of the right size to store on removable backup media.
- Volumes allow a secondary level of information categorization. For instance, your folder structure might break files up by department, while your volumes categorized them by date of creation – or vice versa. As folders and volumes each have distinct security which can be set independently of one another, you can use these categories to determine who has access to what documents.

Types of Volumes

Physical Volume

A *physical volume* is a volume which has a single location in the Windows file structure. Any volume which has not been set to automatically create a new volume when its size limit is reached is a physical volume. The individual volumes within a logical volume are also physical volumes.

Physical volumes can be configured both with and without size limits. If a physical volume has a size limit, new documents can be added to that volume until the size limit is reached; when the size limit is reached, any attempt to add new documents will produce an error, and no new documents will be created or added in that volume. (However, pages and metadata may be added to the volume; see Volume Size Limits for more information.) If a physical volume has no size limit, it can continue to grow indefinitely, although it may be constrained by the size of the hard disk on which the volume files are placed.

Example: Bill is setting up his repository and wants to create two volumes, one for the Accounting department and one for the Marketing department. He is not concerned about size, and so he sets them both up as physical volumes with no size limit. He configures them with the paths C:\AccountingVolume and C:\MarketingVolume respectively. They will continue to grow until either the maximum number of pages possible in Laserfiche has been reached (2,147,483,648) or until the C drive is completely full. Later, Bill decides that it would be useful for the Human Resources department to have a volume. Since Human Resources is a small department at his company, he decides that he can restrict the size of the Human Resources volume to keep more space free for other uses. He creates a HumanResources volume with a size limit of 3 GB and sets its path to C:\HRVolume. When the HumanResources volume reaches 3 GB, it will display an error message, and no more documents will be able to be added to it.

Logical Volume

A *logical volume* is a collection of physical volumes. When you create a new volume in the Laserfiche Administration Console, you have the option to select "Automatically create a new volume when size limit is reached." If you select this option, the volume you have created will be a logical volume, containing one or more physical volumes. When a new logical volume is created, the first physical volume associated with that logical volume is also created. When the volume's size limit has been reached, a new physical volume is created and new documents are placed in the new physical volume; no new documents can be created or added to the old physical volume. (Pages and metadata may be still added to the volume; see Volume Size Limits for more information.)

To the user, the logical volume acts like a physical volume. When importing documents, importing a document into the logical volume will automatically place it in the correct physical volume. When searching, searching for volumes within a logical volume will return all the documents in that logical volume, regardless of what physical volume they are in. However, the individual physical volumes can be backed up to separate drives or burned to separate CDs or DVDs; they can be renamed or their paths changed in the Laserfiche Administration Console.

> **Example**: Sarah wants to be able to back up her repository's volumes easily on DVDs, so she doesn't want the volumes to become too large and unwieldy. However, she also doesn't want to confuse the repository's users with a large number of similar volumes. She decides to use logical volumes. She creates a volume called CityCouncil, gives it a size limit, and sets it to create a new volume when its size limit is

reached. She points sets the volume path to E:\CouncilVolume. When the CityCouncil logical volume is created, a CityCouncil1 physical volume is also created and placed at E:\CouncilVolume\CityCouncil1\. From the user's perspective, there is only one volume: CityCouncil. The users can add files to CityCouncil without worrying about size restrictions. The Laserfice server will create new volumes -- CityCouncil2, CityCouncil3, et cetera -- when the size limit is reached; the change is invisible to the users, but Sarah can easily burn the smaller physical volumes to DVD to back them up.

Volume Size and Volume Rollover

Volume Size Limits and Volume Rollover

When a physical volume within a logical volume reaches the maximum size for that logical volume, it undergoes a process called *volume rollover*. A new physical volume is created, and any new documents assigned to the logical volume will be placed within the new physical volume -- until that physical volume reaches the maximum size, at which point the process repeats. When a document is added to a given logical volume, the server performs a check to determine if the current physical volume has exceeded its maximum size. If the volume has not already exceeded its size limit, the document is added. If the volume has exceeded the size limit, a new physical volume is created for the document.

However, this size check is only performed when a new document is created, imported or migrated into the volume. Since pages from a single document cannot be split among multiple volumes, new pages can still be added to existing documents, electronic documents can be printed using Snapshot, and documents can be OCRed or have text extracted from them, even if the volume has already reached its maximum size. This can cause the volume to exceed its maximum size. If this is not accounted for, it can cause problems – for instance, a volume designed to be backed up on DVD-ROM may become too large to burn to the disc.

> **Example**: Tony is scanning a document into his logical volume FinancialRecords, which is set with a maximum size of 30 GB. When he scans it, the server checks the size of the document and determines whether FinancialRecords1 (the most recent physical volume within the FinancialRecords logical volume) has reached its size limit. The server therefore automatically creates FinancialRecords2 and places the new document in it. Tony does not need to change any settings in order for this to take place; from his perspective, he is simply continuing to scan into FinancialRecords. Next, Tony scans some pages into an existing document which has already been stored in

FinancialRecords. This document was originally stored in FinancialRecords1. Since no new document is being created, however, the server does not check the size of FinancialRecords1. The new pages are added to the document, pushing the size of FinancialRecords1 to 30.1 GB in size.

Best Practices for Volume Size

In order to allow for possible expansion of a physical volume after the volume has rolled over, it is recommended that volumes be created with a maximum size lower than the desired maximum size. Therefore, if the volume grows beyond the configured maximum size as new pages or text is added, it will still remain within the desired size range.

Exactly how much margin for error you should leave when configuring volume size will depend on how the volume is used. Since volumes expand when pages or text are added to them, you will need to leave more space in the volume for a site that frequently scans into existing documents than one that rarely or never adds pages to an existing document. Additionally, volumes published with Laserfiche Plus include metadata from the included files and may therefore be larger than the size of the volume on the disk. If you are not sure by how much your volumes are likely to expand, it is a good idea to be very conservative at first. Once the physical volumes have been in use for some time after rollover, you can get a better idea for how much space to leave, and can adjust the size of future volumes accordingly.

> **Example**: Sarah will be backing up her volumes to DVD, so she wants to make sure they won't exceed 2 GB, the size of her DVDs. She knows that the city council rarely adds new pages to existing documents, and generally OCRs and Snapshots documents shortly after adding them to the repository. They also tend to scan documents in small batches. Since most documents will not get much larger after they have been added to the repository, she knows that she does not have to leave a large margin for error. However, she also does not want to cut it too close. She therefore decides to set the volumes with a maximum size of 1.7 GB, which gives her a 15% margin for scanning into existing documents, adding to electronic documents, OCRing and printing Snapshot images to this volume.

> **Example**: Tony will be backing up his volumes to removable hard drives that hold 100 GB of data. His company handles large financial records, and frequently must add many pages of additional material to existing documents over a long period of time. They also have a high scan volume. Because of this, documents can double or even triple in size. In order to ensure that he will be able to back up his volumes, Tony sets the maximum size of his volumes to 30 GB. That way, if the

number of pages are doubled or tripled, and the additional pages are OCRed, he has plenty of room to add that material to the volume without exceeding his 100 GB limit.

Example: Since Melissa is creating a fresh new repository, she isn't sure how much material will be added to existing documents, but she knows that she wants her final volumes to be no larger than 5 GB. She decides to initially sets up her logical volume to roll over when it reaches 2.5 GB; this gives her a margin of error of 50%. When the repository has been in use for some time, she takes a look at finished physical volumes and sees that they average 4 GB in size. Therefore, in the future, she can create volumes with a smaller margin for error based on this usage that she has been able to observe. By making an overly conservative estimate at first, she can refine the rollover size of her volumes without needing to worry about exceeding the desired size.

If your volumes do exceed the desired size, there are two ways to remedy the problem. If you want to preserve the order of documents in the current set of volumes, you can create a new logical volume, set it with your new maximum size, and migrate the documents from the old volume into the new one. Documents are migrated in the order that they have been sorted in the Client; for instance, if you searched for all items in a volume and sorted them by creation date, and then migrated all items, they would be migrated into the new volumes in creation date order. If you sorted them by name, they would be migrated in alphabetical order by name, and so on. Once all the documents have been migrated, you can delete the old version; if you so desire, once the old volume is deleted, you can also rename the new logical volume with the old volume's name.

It is also possible to 'lock down' a physical volume such that no more data can be added to it, and therefore so that it will not grow. However, this will 'freeze' the volume such that pages cannot be added to it and it cannot be OCRed or have text extracted to it. There are two ways to lock down a physical volume in this manner. First, you can set the volume as read-only. This will lock the volume for all users. To set a volume as read-only, select the volume in the Laserfiche Administration Console and choose Properties from the Action menu. Check the Read-only volume box. If you want to lock the volume only for certain users – for instance, if you wish to allow administrators to add to the documents – you can lock down the volume with volume security. Select the physical volume in the Laserfiche Administration Console and select Security from the Action menu. You will need to remove the Append Data, Modify Contents and Create Documents right from all users; this will prevent users from adding pages, OCRing or extracting text from a document, and will therefore prevent the volume from growing. However, it will also prevent users from deleting pages from that volume, moving

documents out of the volume, OCRing or extracting text. Furthermore, since there is no way to split pages in a single document between different volumes, it will be impossible to add more pages to a document either by importing images from another file or by scanning into the document.

> **Example**: John's logical volume, Marketing, contains a set of physical volumes that are coming very close to exceeding the capacity of the drives he wishes to keep them on. He wants to make sure that he does not exceed the drive size, so he considers his options. Since his users frequently scan into existing documents, locking down the volumes is not a viable option. He decides that he wants to keep the documents in order by creation date, so he will migrate the documents into a new volume. In the Administration Console, he creates a new volume called Temporary; because he underestimated the amount by which his volumes were growing, he sets its maximum size to 8 GB instead of 10 GB, which was his previous maximum size. In the Laserfiche Client, he uses a Within Volume search to locate all documents within the Marketing volume. He sorts the documents by Creation Date, then selects them all and chooses Migrate Document(s) from the Action menu. He chooses to put them in the Temporary volume. Once the documents have all been migrated, he makes sure that the Marketing volume has 0 documents in it by checking its properties in the Administration Console, then deletes the now-empty Marketing volume. In order to make the transition as smooth as possible, he then renames the Temporary volume to Marketing and changes the physical volume names from Temporary1 and Temporary2 to Marketing1 and Marketing2, and so on.

Volume Location and Moving Volumes

Volume Location

When you create a volume, either physical or logical, you will be prompted to provide a location for that volume. For a physical volume, this is the actual place on the hard disk that the TIFF, text, and electronic document files will be placed. In a logical volume, the physical volumes associated with that volume will be created as sub-folders of the specified folder.

> **Example**: The physical volume Loans is configured with the path C:\LoansVolume. Volume files will be placed directly within that folder. The logical volume StockInformation is configured with the path C:\StocksVolume. Its physical volumes will be placed within that folder -- the first physical volume will be located at C:\StocksVolume\StockInformation1, the second at C:\StocksVolume\StockInformation2, et cetera.

In order to reduce the load on the server, you may wish to keep your volumes on a different machine than the server. If the volumes are large, it may also be useful to spread them out over several computers. The volumes can be placed on any drive on the network to which the Laserfiche Server service has access. You will not be able to browse to the network volume location when setting the path, but you can type the UNC path or network drive directly into the path box.

Fixed and Removable Paths

The path for a volume can be either fixed or removable. The primary difference is that if the Server is not able to locate a fixed volume, it will display an error message; if the Server is not able to locate a removable volume, it will prompt you to mount the volume.

If a single volume has both a fixed and a removable path, you can assign different types of files (image files, text files and electronic documents) to the different paths – for instance, you could place the image files for a volume at the fixed path location and the text files for the same volume at the removable path location.

> **Tip**: Although the removable path is intended for removable media, it can be assigned to a fixed location as well. This allows you to spread image, text and/or electronic documents from a single volume across several drives.

Moving Volumes

You can change the location of volumes by selecting the volume in the Administration Console and selecting the Path tab from the Properties dialog. (You will not be able to browse to network drives or UNC paths, but you can type them manually into the path box.) It is important to note that changing the path does not move the volume files themselves, it simply sets the location on the disk or network where the repository will look for the volume. This path must match the actual location of the volume files on the disk. After the path has been changed, you will need to move the volume files to the new location as well.

Changing the volume path for a logical volume will not change the path for the physical volumes associated with that volume. Instead, it will cause new physical volumes within that logical volume to be created under the new path. You can also change the paths for individual physical volumes. This allows you to spread the physical volumes for large logical volumes over several drives or computers in order to reduce the load on a single drive.

Detaching, Exporting and Attaching Volumes

Volume exporting is a way of transferring the contents of a volume to another repository or backing up a volume. When a volume is exported, its contents are copied into a folder specified at the time of volume export. Additionally, a set of .xml files containing the metadata for those documents is created. It is these .xml files that differentiate an exported volume from a volume that has not been exported; it contains the necessary information for reconstructing the documents and folder structure of the files in a Laserfiche repository. If an external volume does not have an associated .xml file, you will not be able to attach it to a repository.

When an exported volume is attached to a repository, the documents contained in that volume, as well as their folder structure, templates, document links, tags and version information, will be added to the repository. The documents must be re-indexed in order to make them searchable in the new repository, however. If a volume is being exported for backup purposes, it is a good idea to verify that the volume has successfully exported by attaching it to a repository and checking that the documents are in place and can be successfully opened.

Volume detaching is much like volume exporting, except that the volume is deleted from the originating repository at the time of export. As volume detaching does not give you the opportunity to verify that the volume can be successfully attached before it is removed from the repository, it is usually a good idea to export a volume, verify it, and then delete it rather than detaching it.

Note: While much of the metadata for an exported volume is contained in the volume's files, some data will be lost. Specifically, all information related to users – including security – as well as all Records Management and Auditing data will be left out, as it is not possible for the server to incorporate this information into a pre-existing repository.

Volume Security

You can apply volume security to grant or restrict access to certain volumes. This can be particularly useful if you have set up your volume scheme such that certain volumes are assigned to certain departments; by granting or removing rights to that volume, you can restrict access to the departments assigned to the volumes.

The following rights are associated with Laserfiche Volumes:

Read: A user with the Read right to a volume is able to view the images and text pages associated with documents in that volume. It does not affect the

user's ability to view metadata such as template information.

Append Data: A user with the Append Data right can add or create new pages in an existing document. This right is necessary, for instance, to scan pages into a document.

Modify Contents: A user with the Modify Contents right can modify or delete images, text and electronic files associated with documents in that volume.

Note: The **Modify Contents** entry access right was called the **Write** entry access right in versions of Laserfiche prior to 7.1.1.

Create Document: A user with the Create right to a volume is able to create a new document in a volume. This right does not allow users to add pages to documents.

Note: Volume security works with Entry Access rights – neither overrides the other. A user must have both the appropriate Volume and Entry Access rights to a document to be able to open, create or modify a document.

Backing Up Volumes

In addition to backing up a repository's database files, it is important to back up the volume files as well. The entire volume directory – including the VOLUME.ID file and all subfolders – must be backed up. In the event of the loss of the original volumes, the backup volumes can then be restored by either moving them to the location of the original volumes, or by indicating their new locations in the Administration Console. Volumes backed up in this fashion must be attached to the same repository with which they were originally associated; they cannot be attached to a new repository. The backups must also be updated at the same time as the database; otherwise, documents may be present in the volume files that are absent in the database or vice versa, which can result in errors or lost information.

You can also back up volumes by exporting them and moving or burning the exported volume files to your backup medium. This is a good use for logical volumes; for instance, if you want the volumes to be backed up to DVDs, you could set your logical volumes to create new physical volumes when the volumes approach the capacity the DVD.   (See "Volume Size and Volume Rollover," above, for more information on logical volume and recommended volume size.) You can attach an exported volume exported to a new repository, if it becomes necessary.

Publishing by Volume

With Laserfiche Client 7.1.2 and Plus 7.1.2, you can publish the contents of a volume. Plus publishing copies a set of Laserfiche documents – including much of the information associated with them, such as templates and document links – into a new, read-only repository that can be accessed with a Plus viewer. This is useful for archiving, as the volumes can then be accessed straight from the published repository without needing to attach them to a repository. You can also publish multiple volumes at once. If multiple volumes are published, the volume distinctions – that is, which files are in which volume – will be preserved in the published repository.

Laserfiche®