IMAGE STORAGE –
A Question Of Access And Volume.

In a traditional optical workflow, most image storage issues could be addressed with a simple visit to the local hardware store. With sufficient shelving and an effective negative cataloging/order tracking system, storage and retrieval typically poses manageable obstacles for a well-organized lab.

However, in a digital workflow, image storage quickly can become a challenging and expensive proposition involving significant media, device, storage management software, and networking costs. Storage decisions made in the planning process will, to a large degree, determine the long-term productivity and effectiveness of the digital workflow.

Generally speaking, the nature of the jobs in the individual workflow for weddings, school photography, sports, etc. has a major impact on the following key drivers of image storage decisions:

- Size and volume of images.
- Lead time for accessing images.
- Long-term image use.

In analyzing storage volume requirements, labs must determine how many photographers they're dealing with; how many jobs they will handle per photographer; and how many images accompany each job. In addition, the lab must establish an image compression strategy, while understanding how the stored files will need to be accessed (and for what period of time) before the final print order arrives. School photography finishing also creates the need to establish long-term archiving requirements.

■ Categories of Storage.

Working storage (also known as online storage) is designed to provide fast access to image files sourced from scanners, as well as from digital files directly received from the photographer.

In the scanning stage, there are two distinct workflow options – one pass and two pass.

The two-pass system is designed to conserve storage space. Low-resolution image files for work in progress (WIP) reside on an internal hard drive or external (RAID) hard disk array to generate proofs for the end customer. When the final order is placed, the selected frame is re-scanned at a higher resolution to create the customer order.

By design, a one-pass system is comparatively more efficient. Labs can make a relatively substantial investment in their computer infrastructure to accommodate online storage needs for a one-pass system, or they have the option of writing the image files to CDs, DVDs and next-generation writeable media; and then appropriately store and catalog them as a entry-level alternative.

■ Resolution's Impact On Storage.

A typical wedding image, for example, when output as an 8" x 10" print may require approximately 15 MB to 38 MB of storage – depending on the resolution of the output device and the file format being created (15 MB for a 250 ppi output device, 38 MB for a 400 ppi output device).

Often, image file size requirements also are dependent on subject matter. For example, a group or wedding party shot would undoubtedly require a larger file size than a single head/shoulder portrait to ensure optimal print quality.

Image resolution serves to establish the smallest file size that delivers an acceptable print. The "right" answer is determined by comparing test prints at selected resolutions against the photographer's standard of visual quality.

Step 1

- 400 is the output device resolution in pixels per inch.
- 8" is the first dimension of the output requested.
- 3,200 is the total number of pixels in the 8" dimension.

400 x 8" = 3,200

Step 2

- 400 is the output device resolution in pixels per inch.
- 10" is the second dimension of the output requested.
- 4,000 is the total number of pixels in the 10" dimension.

400 x 10" = 4,000

Step 3

- Multiply results from previous calculations for the total amount of pixels needed (for one plane of color).

3,200 x 4,000 = 12,800,000

Step 4

- Multiply results from previous calculations by 3 (for the three colors: red, green and blue) for the total amount of pixels needed for an 8" x 10" print using an output device of 400 ppi.

12,800,000 x 3 = 38,400,000

Step 5

- To obtain approximate file size, take previous calculation and divide by one million (to get the answer in mega bytes (MB) for the total file size required to "perfectly match to the output device".*

38,400,000/1,000,000 = 38.4 MB

** Five steps to calculating files sizes required for 1:1 scan-to-print match; these calculations are approximations*

Online storage systems can vary from 500 gigabytes to more than 3 terabytes, depending on current and anticipated lab volume.

How many image files will your storage solution hold?

Drive Storage	File (20 MB tiff)	File (20 MB jpg)	File (20 MB jpg)
1 GB	51	512	102
1 TB	52,429	524,288	104,858

1 Kilobyte (KB) = 1024 bytes
1 Megabyte (MB) = 1024 KB
1 Gigabyte (GB) = 1024 MB
1 Terabyte (TB) = 1024 GB

Assume a lab scans a 20 MB file
20 MB (tiff) = 20 MB
20 MB (jpg) = 2 MB *(depending on compression)*
20 MB (jpg) = 10 MB *(depending on compression)*

After the proofs have been made, image files often are transferred to a near-line storage device (CD and/or DVD jukebox) to await retrieval once the final order has been placed. Practically speaking, it is not always feasible to maintain large image files online for the intervening weeks it often takes before the final customer order arrives. Jukeboxes with 700 slots and more provide an optimal combination of storage capacity and access to handle most high-volume digital workflow requirements.

Meanwhile, expectations for long-term storage greatly depend on the lab's customer base (school, weddings, sports), and how quickly the end users make their re-order decisions. In an optical workflow environment, most labs simply return the film to the photographers. Similarly, in a digital workflow, files residing on the near-line CD/DVD storage system are written to CD and may be returned to the photographer for long-term archiving. (For optimal efficiencies, the CD should be written from the disk [RAID] – not the near-line CD/DVD system.)

Labs may also choose to warehouse digital files off-line on DVD, CD or magnetic tape systems as a service to its customers. However, at this point in the development of digital workflows, photographers are not accustomed to paying for long-term storage of image files; and few labs, if any, could be expected to offer this costly service for free.

Labs looking to become a image archiving service bureau for their professional photographers must also keep in mind that digitally stored files require digital backup (and all of its associated costs). If the processed film is to be retained by the lab, a system must be established to identify and retrieve the corresponding negative, in the event the digital file is corrupted or lost.

A Word About Compression.

File compression is a common way to enhance the efficiency of image storage and distribution over an internal or external network. In its simplest terms, compression employs mathematical algorithms to remove portions of image data (reduce its size) with the expectation that much of this data can be reconstituted when the image file is decompressed.

Although JPEG files can be compressed at a ratio of 10:1, they are generally compressed at a 2:1 or 3:1 "lossless" ratio. An emerging compression option, JPEG 2000 offers a 9:1 "lossless" compression ratio. Depending on the capacity of their storage subsystems, labs must decide what compression scheme will be used and where compression will occur in the workflow (online, near-line, off-line) given image quality standards and expectations.

How Much Storage Is Needed?

Here are the basic factors to consider in determining how much storage is required:

Working Storage – RAID

- Determine the primary type of business (wedding, school, portrait, retail, sports).
- Number of frames scanned per week (peak). Consider the size of files scanned at production resolution, not thumbnail resolution, as well as compressed file size.
- Number of digital camera images per week (peak). Consider image size and compressed file size.
- Additional customer-driven weekly online storage such as customized templates.
- Number of weeks storage is required online.

Near-line Storage – CD and/or DVD-RAM

- Consider the primary type of business (wedding, school, portrait, retail, sports).
- Longer-term storage period required (working weeks).
- Longer-term storage capacity required.

What To Look For In Storage Management Software.

Once the storage solution has been selected, labs must determine an appropriate strategy for managing the storage.

Certain storage devices such as CD/DVD are not natively supported by the operating system of the PC or workstation, and consequently require software to both manage the internal workings of the device and communicate with the host operating system.

Other storage management software optimizes the movement of images/data to and from the storage devices, or backs up large amounts of a data on a daily basis.

Depending on the complexity of the solution, storage management software should be considered as an asset protection tool. Here, the software can provide a secure means to catalog and track the importation of digital camera files and scanned images into the workflow – helping to assure the photographer's pictures can be accessed on demand at any time in the online, near-line and off-line storage cycles of the image file.

Storage management software also eliminates the random component of human error and the need for consistent human intervention by placing repetitive, time-sensitive functions such as file back up and file migration (load capacity management) into an automated mode.

Labs looking to select an integrated hardware/software solution should be mindful of the dynamic nature of the industry. As with many decisions, caution should be exercised; select branded products from well-known, reputable firms who can be counted on to support their products over an extended period. Since most suppliers do not sell directly to the lab, it is just as important to seek out value-added resellers or integrators who have strong working relationships with the vendors, and who have the resources to consult directly with the lab to configure systems that are best matched to current and future workflows.

Storage Management Software Checklist – Recommended Features And Capabilities.

- Robust support for file filtering on a per-task basis.
- Multiple tier migration – concurrently supports many tasks.
- E-mail notification of events (problems).
- Networked-hardened file transfer – detects, corrects and reports damaged file systems.
- Integration with production software.
- Customization enabled – special capabilities such as adding thumbnails to imported images while changing file name to lab nomenclature, or providing interface to production management database to detect shipping status or reorder events.
- Licensing options available.